Proceedings of the

Thirteenth Amsterdam Colloquium

December 17 — 19, 2001

Proceedings of the Thirteenth Amsterdam Colloquium

December 17 — 19, 2001

Robert van Rooy
and Martin Stokhof (eds.)

ILLC/Department of Philosophy
University of Amsterdam

Printed by Grafisch Cemtrum Amsterdam
Cover design by Crasborn Grafisch Ontwerpers bno

Preface

The 2001 edition of the Amsterdam Colloquium is the Thirteenth in a series which started in 1976. Originally, the Amsterdam Colloquium was an initiative of the Department of Philosophy. Since 1984 the Colloquium is organized by the Institute for Logic, Language and Computation (ILLC) of the University of Amsterdam.

These proceedings contain the abstracts of all papers to be presented at the colloquium. The copyright resides with the individual authors.

For the organization of the Thirteenth Amsterdam Colloquium financial support is received from:

- the Royal Dutch Academy of Sciences (KNAW)
- the Institute for Logic, Language and Computation (ILLC)
- the Dutch Graduate School in Logic (OZSL)
- Johan van Benthem's Spinoza project 'Logic in Action'
- Bronstee.com Software & Services B.V.

which is gratefully acknowledged.

Thanks to Paul Dekker, Michiel van Lambalgen, Frank Veltman, and Yde Vennema for organizing the workshops and to all the invited speakers for their willingness to accept our invitation to participate.

The organizers would also like to thank the authors for their contribution and the members of the program committee for the great job they have done:

≫ Reinhard Blutner, Gennaro Chierchia, Angelika Kratzer, and Ruth Kempson

≫ Renate Bartsch, Johan van Benthem, Jeroen Groenendijk, and Frank Veltman

Robert van Rooy and Martin Stokhof
Amsterdam, November 2001

Contents

Optimality Theory and Natural Language Interpretation

Reinhard Blutner
Humboldt University & ZAS, Berlin

Optimality Theory as developed by Prince, Smolensky and others has been fruitfully applied to phonology, morphology, and syntax. The attempt to apply the ideas developed in these fields to the semantics/pragmatics interface necessitates the conception of a bidirectional OT. The main reason for assuming the idea of bidirectionality (combining expressive and interpretive optimization) derives from a number of phenomena which, on the one hand, demand the treatment of preferred interpretations, and, on the other hand, suggest to take into account the existence of blocking effects. In this talk, some general motivation for the idea of bidirectionality is presented, the importance of this idea for the field of pragmatics (conversational implicature) is investigated, and some applications of bidirectional OT are outlined.

A central discussion point concerns the distinction between two different conceptions of bidirection. Strong bidirection is the simultaneous realization of expressive and interpretive optimization. Weak bidirection, on the other hand, describes a more sensible interaction of the two modes of optimization - one that realizes Horn's division of pragmatic labor: (un)marked expressions typically get an (un)marked interpretation). It is argued that strong bidirection describes the state of language after the OT learning algorithm has fully realized the equilibrium between expressive and interpretive optimization. The notion accounts for both unacceptability and ineffability, it is computationally tractable (to the same extent as the unidirectional optimizations are), and it is able to deal with aspects of online processing. Weak bidirection, on the other hand, should best be considered as a principle describing the direction of language change (super-optimal pairs are tentatively realized in language change). In order to install this idea, recent work by Robert van Rooy may be helpful adopting an evolutionary setting. Finally, I consider it an exciting challenge for bidirectional OT to make the term 'grammaticalization' more precise (discussing Hyman's proposal to apply the term for "the harnessing of pragmatics by a grammar").

A Unified theory of referential NPs

Gennaro Chierchia
Milan

Dynamic Syntax: the Growth of Logical Form

Ruth Kempson
King's College London
ruth.kempson@kcl.ac.uk

Given widespread recognition of the importance of context-dependence in natural language interpretation, there is tension between the avoidance of any reference to the dynamics of natural language processing in syntactic formalisms and the modelling of such dynamics in semantics. I shall argue that this tension can be resolved by adopting a representationalist input-system perspective on interpretation (broadly following assumptions of Fodor 1981, Sperber and Wilson 1986/1995), and that a model of the process of building up interpretation on a left-right basis from a sequence of words can be shown to be a sufficient vehicle for expressing syntactic generalisations. The proposed framework is the Dynamic Syntax framework (DS, Kempson et al 2001): the particular case study is the interaction between long-distance dependency and anaphoric processes in relative clause construal. The result, I shall argue, is a new concept of grammar formalism in which the dynamics of language processing is central, and the only level of structure posited is the construction of logical forms representing content as assigned in context. I conclude by reflecting on the new research questions which this shift in perspective opens up.

The sequence of arguments is as follows. First I show how, given the DS assumption that NL parsing involves a monotonic left-to-right process of building up logical forms representing interpretation with logical forms represented as trees, the construal of long-distance dependency and resolution of anaphora can both be characterised in terms of tree growth. (Trees are defined using the modal logic of finite trees, LOFT, Blackburn and Meyerviol 1994.) On the one hand, "left-dislocated" expressions are taken to project a term whose role in building up the logical form is underspecified when introduced into the emergent tree and is established subsequently in the construction process. On the other hand, pronouns are defined to involve the lexical projection of a metavariable as decoration of some node in a tree, for which a pragmatic process of substitution of some contextually provided value yields a fixed term. Such trees may be projected in tandem, with one partial tree providing the context for the subsequent construction of a second partial tree, a process which is defined as the basis for incrementally building up relative clause construal. By definition, sequences of words are defined to be well-formed iff at least one logical form can be established from the words in sequence, with no requirements on tree growth as introduced in the construction process left outstanding.

Having set out the basic framework, I show how the interaction of these processes can be straightforwardly modelled, with resumptive use of pronouns emerging as epiphenomenal, an immediate consequence of the interaction of processes of tree growth and the interpretation of anaphoric expressions in the provided context.

I then show how quantifier construal can be expressed in terms of the same left-to-right dynamics, with quantifying expressions taken to project incomplete variable-binding term operators whose specifications and any accompanying scope constraints are collected incrementally providing input to globally defined rules of evaluation. Amongst these are indefinites, which project incomplete epsilon terms. Finally, I take the parsing processes for relative clauses, anaphora and quantified expressions together, and show how they combine to provide the basis for an integrated typology of relative clause constructions according to different possible forms of interaction between anaphoric and other tree growth processes.

In the light of these preliminary results, I close by reflecting on new research questions opened up by this dynamic representationalist perspective. The question of the last decades, 'Do we need a level of semantic representation in NL grammar formalisms in addition to any syntactically motivated level(s) of representation to express semantic generalisations about natural language?' (addressed by Kamp and Reyle 1993 and many others), is, I suggest, to be replaced by the question 'Do we need any level other than that of semantic representation in NL grammar formalisms to express syntactic generalisations about natural language?'. Answering this leads to the more general question 'What does it mean to know a language, given a perspective in which parsing is taken as basic?'

References

Blackburn, P and Meyer-Viol, W.. 1994. Linguistics, logic and finite trees. The Bulletin of the Interest Group of Pure and Applied Logics, 2, pp.2-39.

Fodor, J.A. 1981. *Modularity of Mind.* MIT Press.

Kamp, H. and Reyle, U. 1993. *From Discourse to Logic.* Reidel.

Kempson, R., Meyer-Viol, W., and Gabbay, D. 2001. *Dynamic Syntax: The Flow of Language Understanding.* Blackwell.

Sperber, D. and Wilson, D. 1986/1995. *Relevance: Comunication and Cognition.* Blackwell.

Indeterminate Pronouns

Angelika Kratzer
University of Massachusetts at Amherst

The term 'indeterminate pronoun' is commonly used by scholars of Japanese to refer to the following class of pronouns:

(1)
dare 'who'
nani 'what'
dore 'which (one)'
dono 'which'(Det)
doko 'where'
itu 'when'
naze 'why'
doo 'how'

Depending on the operator they are associated with, indeterminate phrases in Japanese take on existential, universal, interrogative, negative, negative polarity, or free choice interpretations. The Japanese situation has pushed semanticists to look for a unified account of those different uses. Most recently, Junko Shimoyama has proposed part of such an account by extending Hamblin's alternative semantics for questions to universally quantified sentences in Japanese[1]. The most surprising result of her analysis is that it automatically accounts for the apparently anomalous scope properties of indeterminate phrases:

(2) **[[Dono hon-o yonda] kodomo]-mo yoku nemutta.**
Which book-Acc read child -MO well slept
'For every book x, the child who read x slept well.'
(Shimoyama 2001)

(3) **Taro-wa [[dare-ga ∅ katta] mochi]-o tabemasita ka?**
Tar-Top who-nom bought rice cake-Acc ate Q
'Who is the x such that Taro ate race cakes that x bought?'
(Shimoyama 2001)

Working within a very different framework of assumptions, Martin Haspelmath shows in his typological survey of indefinite pronouns, that 'indeterminate pronouns' in the Japanese sense constitute a unified class cross-linguistically. Here is an arbitrary example picked from Haspelmath 1996, displaying the inventory of indeterminate pronouns in Latvian[2]:

---

1. Shimoyama, Junko (2001): Wh-Constructions in Japanese. UMass Dissertation.
2. Haspelmath, p. 277, diacritics omitted. Haspelmath, Martin (1996): Indefinite Pronouns. Oxford (Oxford University Press).

	Interr.	kaut-	ne-	jeb-
person	kas	kaut kas, kads	ne-viens	jeb-kads
thing	kas	kaut kas	ne-kas	jeb-kas
place	kur	kaut kur	ne-kur	jeb-kur
time	kad	kaut kad	ne-kad	jeb-kad
manner	ka	kaut ka	ne-ka	
determiner	kads, kurs	kaut kads	ne-kads	jeb-kads, jeb-

The **kaut**- series has existentials that have specific and non-specific uses and can also occur in conditionals and questions. The **ne**-series appears under the direct scope of negation, and the **jeb**-series is found in indirect negation contexts, in comparatives, and also with a free choice interpretation.

If indeterminate phrases are a natural class cross-linguistically, the question arises what it is that makes Japanese quantifier constructions and questions look so different from the well-known quantifier constructions and questions in Indo-European languages. In my talk, I will make an initial and very tentative attempt towards answering this question by looking at some understudied uses of indeterminate pronouns in German from the point of view of Shimoyama's analysis of indeterminate phrases in Japanese. One class of examples I might examine has apparent wide-scope uses of 'jeder':

(4) **Die Blumen, die jeder jedem damals schenkte,**
The flowers that everybody everybody-dat. then gave

sind längst verwelkt.
are long faded.
'The flowers that everybody gave to everybody at the time have long since faded.'

A second class of examples is constructed around the indefinite pronoun 'irgendein', which has external (epistemic) and internal 'free choice' (?) uses, as illustrated in (5):

(5) **Sie hat an irgendeine Tür geklopft.**
She has at some or other door knocked.
'She knocked on some door or other- the speaker doesn't know or doesn't care about which door it was.'
'She knocked on some door or other -she didn't care about which one it was.'

I will explore the possibility that a Hamblin-style alternative semantics might be appropriate for wide-scope 'jeder' and 'irgendein', and possibly for the whole class of indeterminate phrases.

Semantics and Cognition

Coevolution of Languages and the Language Faculty

Ted Briscoe
Computer Laboratory
University of Cambridge

Human language acquisition, and in particular the acquisition of grammar, is a partially-canalized, strongly-biased but robust and efficient procedure. A variety of explanations have been offered for the emergence of a partially-innate language acquisition device (LAD), such as exaption of a spandrel (Gould, 1987), biological saltation (Bickerton, 1990) or genetic assimilation (Pinker and Bloom, 1990). But none provide a coherent account of both the emergence and maintenance of a LAD in an evolving population of language users exposed to a variety of linguistic systems during the period of adaptation for the LAD. I will argue that a coevolutionary approach is the only coherent account, and that computational simulation suggests that genetic assimilation of grammatical information will occur despite epistasis and pleitropy and even in circumstances of rapid concurrent linguistic evolution (pace Deacon, 1997).

In, On, Over and Between: Toward a functional geometry of spatial prepositions

Simon Garrod
University of Glasgow

Locative expressions are few in number but allow for a wide range of uses (Landau Jackendoff, 1993). This discrepancy between the small number of apparently simple spatial distinctions being made in language and the frequency and variety of uses to which locative expressions are put presents a challenge for semantic analysis. In particular, it makes it difficult to give a straightforward geometric definition for any of these expressions (Garrod Sanford,?89; Herskovits, ?86; Vandeloise, ?91).

The talk describes a series of experiments to show how locatives such as IN, ON, and BETWEEN denote location control relations defined in terms of a functional geometry(Garrod Sanford, ?89; Garrod, Ferrier and Campbell,? 99). Such relations capture the way in which objects are seen to control each other?s location by virtue of their spatial arrangement. For example, there is a control relation, which we refer to as fcontainment, by which the ground (i.e., a container) is seen to control the location of the figure (i.e., its content) by virtue of some degree of geometric enclosure. The first series of experiments I will discuss relate to the prepositions IN and ON. They show two things: (1) Confidence in descriptions containing IN and ON relate directly to judgements of the degree to which the ground controls the location of the figure, and, (2) Introducing dynamic information into scenes affects use of descriptions containing IN: Dynamic information consistent with location control of the figure by the ground enhances confidence in IN descriptions for scenes portraying equivalent geometric configurations; By contrast, inconsistent dynamic information reduces confidence in IN descriptions. The other experiment relates to the preposition BETWEEN. Here the control relation is more complicated. If X is BETWEEN Y and Z, then X is seen either to keep Y and Z separate or conversely to hold Y and Z together. This experiment tests the fconnection control relation. Scenes portraying the same geometric configuration of three objects are manipulated to include an alternative connector. Using the rationale behind Garrod et al.s? (?99) Expt. 2 the prediction is that the presence of an alternative separator should reduce confidence in BETWEEN descriptions. The results supported the prediction for two levels of alternative control analogous to Garrod et al.? (?99) findings with the preposition ON and alternative sources of support.

I will discuss the results of these studies in relation to a functional geometric account for the semantic representation of these locatives. The account has the virtue of offering a simple quasi-geometric definition of the prepositions that can accommodate the wide range of spatial situations to which they can apply. It is also consistent with recent accounts of the perception of complex spatial relations in terms of the so-called ?what?? and ?where?? systems (Landau Jackendoff, 1993) and with some accounts of physical imagery (Schwartz, 1999).

References:

- Garrod, S.C. & Sanford, A.J. (1989). Discourse models as interfaces between language and the spatial world. Journal of Semantics, 6, 147-160.
- Garrod, S., Ferrier, G. & Campbell, S. (1999) In and On: Investigating the functional geometry of spatial prepositions. Cognition, 72, 167-189.
- Herskovits, A. (1986). Language and Spatial Cognition. Cambridge: Cambridge University Press.
- Landau, B. & Jackendoff, R. (1993). "What" and "where" in spatial language and cognition. Behavioural and Brain Sciences, 16(2), 217-265.
- Schwartz, D.L. (1999) Physical imagery: Kinematic versus dynamic models. Cognitive Psychology,38, 433-464.
- Vandeloise, C. (1991). Spatial Prepositions: A Case Study from French. Translated by Bosch, A.R.K. Chicago/London: The University of Chicago Press.

The Emergence of Metacommunicative Meaning

Jonathan Ginzburg

The issue of how human language evolved is commonly taken to reduce to the issue of how *grammar* has evolved. *Grammar* is taken to be, essentially, a disembodied system of expressions with a compositional semantics. In this talk I will argue for the need to consider a somewhat different perspective on evolution and ontogeny, one where grammar is taken to be a system of types of spatio-temporally located speech events, embedded within a system of information states of interacting agents. I will adduce a number of motivations for this: first, certain semantic phenomena characteristic of mature adult linguistic competence can only be described within the latter type of approach. A striking example of this are the readings displayed by fragments used to acknowledge understanding or request clarification of an utterance made by a previous interlocuter. These constitute an instance of semantic complexity that arises without concomitant syntactic complexity. Second, by considering language as a property of a communicative interaction system one can distinguish between the contingently existing communicative system of a given species of agents and the potential competence of this species. This is clearly true for human neonates, but is also relevant with respect to evidence of (limited) language learning among primates. Third, such an approach is required in order to account for some fundamental characteristics of how certain metacommunicative meaning actually emerges, for instance the ubiquity of partial repetition by novices of competent speakers' utterances.

MODULE INTERACTION, LANGUAGE ORIGINS, LANGUAGE DIVERSITY

Pieter Muysken, Katholieke Universiteit Nijmegen
postbus 9103, 6500 HD Nijmegen
email: p.muysken@let.kun.nl

Linguists must deal with the study of language origin from the perspective of the nature of language itself. What is it about language that might inform us about its possible genesis? The most serious problem with most structural approaches to language, particularly those within the structuralist and generative traditions, is that they have tended to view the language capacity as a single monolithic whole. This monolithic view stands in the way of the gradualist perspective of language evolution that is required if we bring the origin of language in line with the genesis of other human capacities, both cognitive and more generally neurological.

A more promising approach takes the modular organization of the language capacity as its starting point. In human language, (at least) four essentially different modules intersect: (a) the structure-building and -processing capacity (*syntax*), (b) the sign forming and using capacity (*semiotics*), (c) the capacity to engage in sustained exchange of information (*interaction*), and (d) the capacity to form complex representations of information (*cognition*). This claim of four different interacting modules remains empty unless we manage to isolate the formal properties of these modules. These include:

a. *syntax:* endocentricity, `movement'

b. *semiotics*: distinctiveness, transparency, elementarity, analogy

c. *interaction*: sequentiality, cohesion

d. *cognition*: embeddedness, recursion, opposition, displacement

Only the first module is specific to language. The other three modules play a role in many different aspects of human behaviour.

A. Two crucial features of language are part of the module of *syntax*. These are not found outside of language: *endocentricity* plays a role in sentence grammar (through X-bar theory), in word formation (headedness), and in phonology (e.g. in syllable structure). The property sometimes confusingly labelled displacement (by Chomsky) I will term *`movement'* here (without any of the derivational claims often associated with this term): the fact that in language elements do not always appear in the place in the sequence where they are interpreted (where do you live < you live what place).

B. The module of *semiotics* contributes a number of properties to language; these principles or properties are however also found in non-linguistic semiotic systems. The first principle is that of *distinctiveness*: lexical elements must be sufficiently distinctive to contrast with other

elements. A second principle is *transparency*: new lexical elements ideally are transparently derived from existing elements. A third principle, *elementarity*, refers to the requirement that a lexical element ideally functions as a coherent whole, as an atom which can be combined with other elements. This principle is often referred to as the lexical integrity principle. Fourth, the principle of *analogy*, which causes new forms to be built parallel to already existing forms. The principle of analogy produces lexical subsystems characterized by paradigmaticity.

C. The human *interaction* capacities contribute several crucial properties to language. However, these are also found outside of language. *Sequentiality* is a central property of interaction, both linguistic and non-linguistic. Through a sequential patterning information is structured and made processable. Furthermore, these sequences are marked by *cohesion* in the way elements are linked: interaction systems contain a number of cross-referencing devices to maintain the information structure throughout the sequence.

D. Finally, properties of our general *cognition* play a central role in language as well: first, there is *embeddedness*, through which one cognitive unit is part of another one, and structures with internal hierarchy emerge. Specifically the embedded units can be characterized by *recursion*, through which units of the same time can embedded in one another. Our cognitive systems function in terms of contrasts or *opposition* between different feature specifications. Finally, there is *displacement*: cognitive structures exist independently of immediate experience.

What we now consider to be the set of the unique design features of human language actually is the result of the complex interaction between properties of entirely separate cognitive modules: syntax, semiotics, interaction, cognition. In human evolution these modules with their properties developed independently from one another, allowing a gradualist account.

This modular view of the language system has another crucial property: it allows us to account for the fact of language diversity. Why is there diversity at all, and what are its limits? This question is less interesting perhaps in a purely culturalist or semiotic approach to language, in which there is no claim made for a biologically conditioned human language capacity. However, when we view language from a biological perspective, the diversity encountered is a bit of a mystery. Why don't human languages resemble each other much more than they appear to do? Diversity can emerge because there are different ways the different processing systems involved in language can interact. These differences result from their different formal properties. Differences between languages are due to differential access to features defined in other modules. I will focus here on one subtheory, that of *grammaticalization*: semiotic, cognitive or interactional properties and oppositions become `visible' to syntactic operations through feature sharing at the interface. The relevant metalanguage involves notions such as *visibility of features*, *compatibility of representations*, and *optimalization of matching*.
I will illustrate this account by briefly discussing phenomena of nominal classification in some Amazonian languages.

Games in Language and Logic

Some Uses of Games in Logic

Rohit Parikh
CUNY, New York

We will survey some uses of games to understand propositional and first order logics. The topics will include Ehrenfeucht-Fraisse games, D-structures, and Ehrenfeucht's *-semantics as well as IF-logic and the more recent FI-logics (where FI stands for 'finite information'.)

D-structures generalize the games used in conventional first order logic and can be used to explain classical as well as *-semantics. Most of this work is old, done jointly with Mayberry, de Jongh and Goodman. But it suddenly has contemporary relevance.

FI-logics similarly are a variant of the IF-logics studied by Hintikka and Sandu as well as Hodges, Janssen and Vaananen, but seem to correspond to procedures occuring in ordinary life and possess both the finite model property and decidability.

Models for Games

Robert Stalnaker

Game theory provides a conceptual laboratory at just the right level of abstraction for clarifying the relationships between a range of modal, temporal, causal and epistemic concepts. Possible worlds semantics provides the right framework for charting those relationships. In this survey talk, I will discuss how semantic models of games can be used to bring out the epistemic and causal assumptions underlying different solution concepts for games, and to clarify some patterns of strategic reasoning. Specifically I will consider how assumptions of both theorists and players in a game about belief revision, about the relation between knowledge and belief, and about counterfactual independence interact in determining the rational outcome of a game.

The role of salience in the emergence of signalling conventions

Robert Sugden

In Convention, David Lewis argues that languages can be understood as conventions. His argument works by showing that conventions by which particular signs indicate particular meanings can emerge as equilibrium solutions to signalling games (a sub-class of coordination games), and by arguing that such conventions are rudimentary languages. A possible criticism of this argument is that Lewis's theory assumes that the players of signalling games share common conceptions of salience; if the existence of such common conceptions depends on the prior existence of a language community, Lewis has not explained language 'from outside'. In Evolution of the Social Contract, Brian Skyrms claims to resolve this problem by reconstructing Lewis's model in a way that dispenses with salience; conventions evolve within a population of inductive learners by the amplication of initially random variations. I argue that Skyrms fails to take account of the problem that led Lewis to invoke salience: the unlimited number of potential signals. If there is an infinity of conceivable regularities in experience, inductive learning requires the prior privileging of a small number of these regularities as ones that, if observed, would count as patterns and not as random noise. Thus, inductive learning depends on a form of salience. However, the common conceptions of salience necessary for the emergence of signalling conventions do not require the prior existence of a language community.

Contributed Papers

Pragmatics for Propositional Attitudes

Maria Aloni

Background and Motivations

Propositional attitude reports are analyzed in the framework of modal predicate logic (see [4]). Consider the following classical example from Quine which illustrates a difficulty arising for the standard version of this logic:

(1) a. Philip believes that Cicero denounced Catiline.
 b. Philip does *not* believe that Tully denounced Catiline.
 c. Philip believes that x denounced Catiline.

Suppose sentences (1a) and (1b) are true. What is the truth value of (1c) under the assignment that maps x to the individual which is Cicero *and* Tully? As Quine observes, our ordinary notion of belief seems to require that although (1c) holds when x is specified in one way, namely as Cicero, it may yet fail when the same x is specified in some other way, namely as Tully. Classical modal predicate logic, in which variables are taken to range over plain individuals, fails to account for this ordinary sense of belief.

In a possible world semantics, these 'ways of specifying objects' can be characterized by means of the notion of an (individual) *concept*, i.e. a function from the set of worlds to the set of individuals. Many authors have observed that if we let variables range over concepts rather than plain individuals, we manage to account for Quine's intuition about example (1).[1] An obvious problem arises though if we let variables range over *all* concepts. The following classical example due to Kaplan illustrates why. Suppose Ralph believes there are spies, but does not believe of anyone that (s)he is a spy. Believing that spies differ in height, Ralph believes that one among them is shortest. Ortcutt happens to be the shortest spy. The *de re* readings represented in (2b) and (3b) of the following belief reports are deviant in this situation.

(2) a. Ralph believes that someone is a spy.
 b. $\exists x \Box S(x)$

(3) a. Ralph believes that Ortcutt is a spy.
 b. $\exists x(x = o \wedge \Box S(x))$

But, if we let variables range over all concepts – among them the concept *the shortest spy* –, (2b) and (3b) are wrongly predicted to be true.

Not all concepts seem to be suitable to serve as a value for a variable. If we want to solve Kaplan's problem without automatically regenerating Quine's double vision difficulties we must somehow select the set of suitable ones. This is Quine's and Kaplan's diagnosis of these cases. The latter, in an influential analysis, has proposed a concrete characterization of the notion of a suitable representation with respect to a specific subject and an object. The following example due to Andrea Bonomi shows that a solution of this kind which attempts to characterize the set of suitable concepts as a function of the mental state of the subject disregarding the circumstances of the utterance is condemned to failure. Consider the following situation. Swann knows that his wife Odette has a lover, but he has no idea who his rival is. He knows that this person is going to meet Odette the following day

[1] Other non-classical views on trans-world identification, for example Lewis's counterpart theory, might be adopted to solve this problem (see [1] for comparison). Also these alternative analyses would be in need of the pragmatic theory I present in this article.

at the Opera. He decides to kill him there, and he tells his plan to a friend, Theo. Odette's lover is Forcheville, the chief of the army, and Theo is a member of the staff which must protect him. During a meeting of this staff, Theo (who knows all the relevant details) says (4). A murder is avoided.

(4) Swann believes that Forcheville is going to the Opera tomorrow and wants to kill him there.

Consider now sentence (5) uttered by Theo in the same situation.

(5) Swann believes that Forcheville is Odette's lover.

Sentence (5) strikes us as quite out of place. While (4) is acceptable, (5) would be deviant in this situation. On their *de dicto* readings the two sentences are false for obvious reasons. On their *de re* reading they are true only if the concept *Odette's lover* counts as a suitable representation for Swann of Forcheville. So if we follow Kaplan's strategy we are faced with a dilemma: if in order to explain the inadequacy of (5), we rule out the concept 'Odette's lover', we are unable to account for the intuitive acceptability of (4). A natural way out of this impasse would be to accept that one and the same representation can be suitable on one occasion and not on another. Many authors (Crimmins & Perry, vFraassen, Stalnaker, vRooy, etc) have observed the crucial role played by *context* in the selection of the set of suitable representations. Not much however has been said about *how* context operates such selection. This is the topic of the present article. First it makes formally precise how contextual information contributes to our evaluation of propositional attitude reports. Second it investigates the question how one arrives to select the intended meaning in actual interpretations of these constructions in a given context.

Proposal

The proposed analysis takes variables to range over sets of *separated* concepts. Two concepts are separated if their values never coincide. Different sets of concepts can be selected on different occasions. Although variables always range over the same sort of individuals, these may be differently identified. This style of quantification is adopted in modal predicate logic (see [1] for a sound and complete axiomatization of the obtained semantics). *De re* attitude reports are analyzed as quantified modal sentences $\exists x_n \Box\phi(x_n)$ which receive the standard interpretation with the only exception that x_n is taken to range over the set of concepts pragmatically selected as value for n, rather than over the domain of individuals. In this way the interpretation of these sentences may vary relative to the conceptualization of the universe of discourse which is contextually operative.

The question arise how addressees arrive to select the intended domain of quantification while interpreting these constructions in a given context. In what follows, we will address this question in the framework of O(ptimality) T(heory).

Optimal Theoretic Interpretation In OT semantics, interpretation is ruled by a relatively small number of violable principles ranked according to their relative strength (see [5]). These *soft* constraints help us in selecting a set of optimal candidates from a larger set.[2] The aim is the formulation of these constraints and their ranking in such a way that the actual interpretation of a sentence in a context is the optimal interpretation according to these constraints.

I propose the following constraints as the principles that guide our interpretations of quantified (modal) sentences in a context:

[2]In OT a certain candidate can be optimal even if it violates a constraint provided all alternative candidates lead to more severe constraint violations. A single violation of a higher ranked constraint counts as more severe than many violations of lower ranked constraints.

ANCHOR (vFraassen, Stalnaker, Zeevat, others) Interpretation should be anchored to the context.

CONV MAXIMS (Grice)

BE CONSISTENT (Stalnaker, vdSandt, Zeevat)

BE INFORMATIVE

BE RELEVANT (Horn, vRooy)

***SHIFT** (Williams, Hoop & Hendriks) Do not shift domain of quantification.

***ACC** (vdSandt, Blutner, Zeevat) Do not accommodate.

ANCH, CONV MAX > *SHIFT, *ACC

In what follow I will show that by means of these principles, which are not new and find independent motivation in the literature, we can explain the examples of *de re* belief I have discussed in the previous section. I will start by assuming a *one-dimensional* notion of optimality in which the set of candidates are potential meanings of a single syntactic form (see [5]). Meanings are here identified with subsets of the relevant alternatives (worlds or other) given by the context.

We start with the double vision example in (1). We have two possible interpretations for Quine's question *Does Philip believe that* x_n *denounced Catiline?* Either (i) concept *Cicero* is in n or (ii) concept *Tully*.[3] Only in the first case, *yes* would truly answer the question. Our principles do not select a unique optimal candidate in this case, and this explains the never ending puzzling effect of Quine's example. If (1a) alone had been mentioned, or (1b), then our principles would have selected possibility (i) or (ii) respectively, since the alternative interpretation would have violated ANCHOR or *ACC.

Let us turn now to Kaplan's problem of the shortest spy. Assume that the common ground contains the following information: Ralph does not believe anyone to be a spy, Ortcutt is the shortest spy and Ralph would assent to the sentence 'Ortcutt is thin'. Consider the following sentences:

(6) Ralph believes that Ortcutt is a spy.

(7) Ralph believes that Ortcutt is fat.

Each of these two sentences has two possible meanings in the described situation represented in (a) and (b):

(a) $\Box\phi(ortcutt)$ (*de dicto*) equivalent to

$\lambda x[\Box\phi(x)](ortcutt)$ (*de re* - identification by name)

(b) $\lambda x[\Box\phi(x)](the\ shortest\ spy)$ (*de re* - identification by description)

Interpretation (a) can be paraphrased as *Ralph would assent to 'Ortcutt is a spy/fat'*, interpretation (b) as *Ralph would assent to 'the shortest spy is a spy/fat'.*

The following diagrams summarize our OT-analysis of these two examples. I use (*) to indicate that the interpretation violates the corresponding constraint, and !(*) to indicate a crucial violation. Optimal interpretations are those which do not involve any crucial violation.

(6)	INF,	CONS	*SH
(a)		(*)	
(b)	(*)		!(*)

(7)	INF,	CONS	*SH
(a)		!(*)	
(b)			(*)

[3]Not both because the two concepts are not separated.

Interpretation (a) is correctly predicted to be optimal for sentence (6). Although it violates one constraint, namely BE CONSISTENT, the alternative candidate leads to more severe constraint violations. Indeed, content (b) violates BE INFORMATIVE – the sentence would be trivialized –, and the weaker principle *SHIFT – our description of the context suggests as active a domain which does *not* contain the problematic concept. Since BE CONSISTENT and BE INFORMATIVE are assumed not to be ranked in any way, the violation of this lower constraint becomes fatal in this case.

Interpretation (b) results however optimal for sentence (7). Our principles wrongly predict that the unnatural concept *the shortest spy* is part of our domain of quantification in this case. Let us try a first diagnosis of this problem.

An intuitive explanation of why content (b) is not assigned to a sentence like (7) in such a situation, is that a speaker who would have used such a sentence to express such a content would not have been cooperative. Indeed, (b) could have been conveyed in a much more efficient way by means of the following sentence:

(8) Ralph believes that *the shortest spy* is fat.

It is the existence of this alternative more efficient expression which seems to block the selection of (b) as preferred for (7). If a speaker had wanted to communicate (b), she would have chosen (8). She chose (7), thus she did not mean (b) (see [2] which discusses similar blocking effects). Note however that this sort of reasoning cannot be formulated in the one-dimensional OT analysis we have used so far in which inputs are given by single sentences and no reference is made to alternative forms that the speaker might have used. In order to account for this case we need a *bi-dimensional* notion of optimality (see [2] again), in which optimal solutions are searched along two dimensions, rather than one: the one of the addressee and the one of the speaker. In the following section I follow [3] and recast bi-dimensional OT interpretation processes using notions from Game Theory. The pragmatics of attitude reports is formalized in terms of *optimization games* in which speaker and addressee – whose preferences are determined by OT preferences in combination with particular goal-directed preferences– coordinate their choice towards optimal form-meaning pairs.

Interpretations as games An *optimization game* is a strategic game $(N, (A_i)_{i\in N}, (\succ_i)_{i\in n})$ involving two players, S(peaker) and H(earer), $N = \{S, H\}$. S can choose from a set $A_S = \{F1, F2, ...\}$ of alternative syntactic forms. H can choose from a set $A_H = \{C1, C2, ...\}$ of possible semantic *contents*. Optimality theoretic preferences are used in combination with particular goal-directed preferences to define the preference relations of the two players.

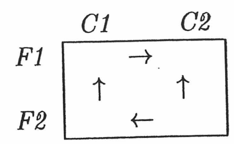
In an optimization game, language users are seen as decision makers. Speakers must decide a suitable form for a content to be communicated. Addressees must choose suitable interpretations for a given representation. Since structural *and* pragmatic factors contribute to determine the players' preference relations, contextual information determine the output of these games just like syntactic and semantics rules.

In order to make predictions about the outcome of these games we use the game-theoretic counterpart of Blutner's notion of weak optimality, namely bj-optimality. Here is the algorithm proposed in [3] to compute the bj-optimal solutions of a given game.

1. Profiles which point to a Nash equilibrium are blocked.
2. Remove preferences for blocked profiles until you reach a fixed point.
3. The Nash equilibria of the fixed point are the *bj-optimal* solutions in the original game.

As an illustration consider the simple game depicted in the following two diagrams. The matrix on the left expresses preferences in terms of payoff function, the one on the right represents them by means of vertical and horizontal arrows.

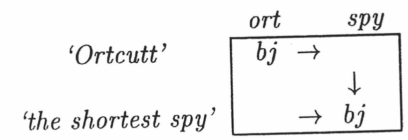
	C1	*C2*
F1	(4,1)	(4,5)
F2	(3,2)	(1,1)



This game has one Nash equilibrium, namely the profile (F1, C2). Consequently profiles (F1, C1) and (F2, C2) are blocked. We can then remove the arrow pointing to the former profile. The resulting game has two Nash equilibria, profiles (F1, C2) and (F2, C1). Since there are no more arrows to be removed on the next step, these two profiles are the bj-optimal solutions of our original game.

Let us see now how the problematic example discussed in the previous section can be analyzed in this framework.

I propose to characterize the interpretation problem posed by the example of the shortest spy by means of the following optimization game:



Speaker chooses the row to be played and hearer chooses the column. S can choose between the two forms in (9) and (10).

(9) Ralph believes that *Ortcutt* is fat.

(10) Ralph believes that *the shortest spy* is fat.

Hearer can choose between the two contents represented in (11) and (12).

(11) *ort* $\mapsto \lambda x[\Box F(x)](ortcutt)$

(12) *spy* $\mapsto \lambda x[\Box F(x)](the\ shortest\ spy)$

The preference relations of our players are given by the following OT tables.

'Ortcutt'	CONS	*SHIFT
ort	(*)	
spy		(*)

'the shortest spy'	CONS	*SHIFT
ort	(*)	
spy		

Hearer strictly prefers content *spy* over content *ort* given any syntactic input because *ort* crucially violates BE CONSISTENT. Consistent interpretations are preferred over inconsistent interpretations.

Speaker crucially prefers profile ('the shortest spy', *spy*) over ('Ortcutt', *spy*) because, as it is clear from the OT analysis, while the latter pair violates two constraints, content *spy* can be conveyed by form 'the shortest spy' without any constraint violation.

Profile ('the shortest spy', *spy*) is indeed Nash-optimal in this game. This implies that profile ('Ortcutt', *spy*) is crucially blocked, and, therefore, profile ('Ortcutt', *ort*) results bj-optimal in this situation. The problematic content *spy* is no longer optimal for sentence (9). We correctly predict that the unnatural concept *the shortest spy* is not taken to be part of our domain of quantification in that situation.

A last question arises. Consider again the case of Odette's lover discussed in the first section. In that situation Theo could have used sentence (14) to say what he wanted to say rather than (13) and he would have been more cooperative.

(13) Swann believes that *Forcheville* is going to the Opera tomorrow.

(14) Swann believes that *Odette's lover* is going to the Opera tomorrow.

In this framework the existence of more efficient forms for one content can block its selection as preferred interpretation. Why is Theo's intended meaning not blocked in this situation?

Blutner in [2] observes that *blocking* effects are not absolute and can be canceled under special contextual conditions. It seems that we have here an example of *de-blocking* triggered by contextual factors. Theo's intended meaning for (13) is not blocked by the existence of (14) because, although (14) is better than (13) according to our principles, its use in the described situation is much less effective given Theo's immediate goal of protecting Forcheville's life.

I propose to represent this case by means of the following game:

	forch	*lover*
'Forcheville'		→ *bj*
	↑	
'Odette's lover'		→ *bj*

S must choose between (13) and (14). H must choose between the following two interpretations:

(15) $forch \mapsto \lambda x[\Box\phi(x)](forcheville)$

(16) $lover \mapsto \lambda x[\Box\phi(x)](odette's\ lover)$

The horizontal arrows of Hearer are determined as in the previous case by our OT preferences depicted in the following two tableaus.

'Forcheville'	CONS	*SHIFT/*ACC
forch	(*)	
lover		(*)

'Odette's lover'	CONS	*SHIFT/*ACC
forch	(*)	
lover		

Note however that the vertical arrows representing Speaker's preferences are crucially reversed because of Theo's specific goals. Profile ('Forcheville', *lover*) is a Nash-equilibrium in this game. Theo's intended meaning for sentence (13) is optimal in the described situation.

References

[1] Maria Aloni. *Quantification under Conceptual Covers.* PhD thesis, University of Amsterdam, Amsterdam, 2000.

[2] Reinhard Blutner. Some aspects of optimality in natural language interpretation. To appear in *Journal of Semantics.*

[3] Paul Dekker and Robert van Rooy. Bi-directional optimality theory: An application of game theory. To appear in *Journal of Semantics.*

[4] Jaakko Hintikka. Semantics for propositional attitudes. In Davis, Hockney, and Wilson, editors, *Philosophical Logic.* Reidel, Dordrecht, 1969.

[5] Helen de Hoop and Petra Hendriks. Optimality theoretic semantics. *Linguistics and Philosophy*, 24(1), 2001.

Monotonicity and Relative Scope Entailments[1]

Alon Altman, Technion Edward L. Keenan, UCLA

Yoad Winter, Technion/UiL OTS

Summary This paper explores the hypothesis that simple monotonicity properties of quantifiers in natural language determine to a large extent the entailment relations between their wide/narrow scope readings. We prove that the *disjunctive normal form* of upward monotone quantifiers using principal ultrafilters correlates with whether an object narrow scope reading entails an object wide scope reading. This result naturally extends the familiar entailment relations between ∃∀ and ∀∃ quantification in first order logic into arbitrary "finitely based" upward monotone determiners (over possibly infinite models), which are precisely defined.

Given a simple sentence of the form *Subject-Verb-Object*, we are interested in the logical relations between the *object narrow scope* (ONS) and the *object wide scope* (OWS) readings of the sentence. In [4], Zimmermann (1993) fully characterizes the class of "scopeless" object quantifiers – those for which the ONS and OWS readings are equivalent for any subject. Zimmermann shows that this class is closely related to the class of (principal) ultrafilters (names). In [3], Westerståhl (1996) fully characterizes the class of "self-commuting" quantifiers, i.e. the quantifiers Q for which ONS and OWS readings are equivalent when Q is substituted for both subject and object. However, as far as we know, the more general problem of characterizing (possibly one-way) entailment relations between ONS and OWS readings has not been given serious attention.

Global determiners (see [2]) are functors that map any non-empty domain E to a binary relation over $\wp(E)$. Any set $Q_E \subseteq \wp(E)$ is called a *generalized quantifier* (GQ) *over* E. Thus, a determiner D_E over E maps any $A \subseteq E$ to the generalized quantifier $D_E(A)$ over E.

Let Q_1 and Q_2 be the GQs over E that the subject and object respectively denote in a given model. The ONS and OWS readings of the sentence in this model with respect to a binary relation $R \subseteq E \times E$, are defined using the following polyadic GQs over $E \times E$:

$$(1)\quad \begin{array}{lcll} Q_1\text{-}Q_2(R) & \stackrel{def}{=} & Q_1(\{x \in E : Q_2(\{y \in E : R(y)(x)\})\}) & \text{(ONS reading)} \\ Q_1{\sim}Q_2(R) & \stackrel{def}{=} & Q_2(\{y \in E : Q_1(\{x \in E : R(y)(x)\})\}) & \text{(OWS reading)} \end{array}$$

Let D_1 and D_2 be global determiners that correspond to the subject and object determiners respectively. The polyadic determiners $D_1\text{-}D_2$ and $D_1{\sim}D_2$, which give rise to ONS and OWS readings respectively, are defined as ternary relations between $A, B \subseteq E$ and $R \subseteq E \times E$.

$$(2)\quad \begin{array}{lcl} D_1\text{-}D_2(A)(B)(R) & \stackrel{def}{=} & ((D_1(A))\text{-}(D_2(B))(R) \\ D_1{\sim}D_2(A)(B)(R) & \stackrel{def}{=} & ((D_1(A)){\sim}(D_2(B))(R) \end{array}$$

[1] A full version of this abstract, including proofs of the propositions, appears at `http://www.technion.ac.il/~alon_a` This research was partly supported by grant no. 1999210 ("Extensions and Implementations of Natural Logic") from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. The third author was also partly supported by the fund for the promotion of research at the Technion, research no. 120-042, and by an NWO grant for visiting the UiL OTS of the Utrecht University in summer 2001.

A quantifier Q_E is called *upward monotone* (MON↑) if it is closed under supersets. A global determiner D is called *upward right monotone* (MON↑) if for every $A \subseteq E$, the quantifier $D_E(A)$ is upward monotone.

We would like to characterize whether the relation D_1-D_2 is contained in the relation $D_1{\sim}D_2$. When D_1 is *every* (*some*) and D_2 is *some* (*every*), it is well-known that the answer is negative (positive) respectively. For instance, the ONS reading of the sentence *some student saw every teacher* entails, but is not entailed by, its OWS reading. We show that in fact, in sentences with upward monotone subjects and objects, the *existential* determiner is the basis for the class of *subject* determiners that guarantee entailment from the ONS reading to the OWS reading. Symmetrically, for upward monotone subjects and objects, the *universal* determiner is the basis for the class of *object* determiners that guarantee entailment from the ONS reading to the OWS reading.

We use the fact (cf. [1]) that any upward monotone quantifier Q_E can be represented as a union of intersections of principal ultrafilters.

Fact 1 *Let Q_E be an upward monotone GQ over E. Then $Q_E = \cup_{N \in M} \cap_{x \in N} I_x$, for some subset M of $\wp(E)$, where I_y is the principal ultrafilter $\{A \subseteq E : y \in A\}$ generated by $y \in E$.*

We call M the *signature* of a *disjunctive normal form* (DNF) of an upward monotone quantifier. We define a hierarchy of the upward monotone quantifiers by requiring a DNF for a quantifier $Q_E \in$ MON↑, with a signature M that satisfies certain conditions. The classes in the hierarchy, with the respective conditions on M that define them, are listed below.

TRIV$_0$: $M = \emptyset$: Q_E is empty

TRIV$_1$: $\emptyset \in M$: Q_E is equal to $\wp(E)$

PUF: $M = \{\{a\}\}$ for some $a \in E$:
Q_E is the *principal ultrafilter I_a generated by a*;

PUF$_\cap$: $M = \{A\}$ for some (possibly empty) $A \subseteq E$:
Q_E is an intersection of PUFs: the *principal filter F_A generated by A*

PUF$_\cup$: M is a (possibly empty) collection of singletons in $\wp(E)$:
Q_E is a *union of PUFs*.

Obviously, the following relations hold between these classes of GQs: PUF $\subset$ PUF$_\cap$ $\subset$ MON↑; PUF $\subset$ PUF$_\cup$ $\subset$ MON↑; TRIV$_0$ $\subset$ PUF$_\cup$; TRIV$_1$ $\subset$ PUF$_\cap$.

Further, observe the following simple facts.

Fact 2 *A quantifier Q is in* PUF$_\cup$ *iff* $Q = \cup_{\{x\} \in Q} I_x$.

Fact 3 *A quantifier Q is in* PUF$_\cap$ *iff* $Q = \cap_{x \in \cap Q} I_x$ $(= F_{\cap Q})$.

Consider now the following simple relation between the above hierarchy and scope entailments.

Lemma 4 *Let $Q_1, Q_2 \subseteq \wp(E)$ be upward monotone GQs over E. If $Q_1 \in$ PUF$_\cup$ or $Q_2 \in$ PUF$_\cap$ then Q_1-$Q_2 \subseteq Q_1{\sim}Q_2$.*

We use our classification of MON↑ local quantifiers in order to classify MON↑ global determiners as follows. For any global determiner D:

D is PUF$^1_\cup$ iff for all $A \subseteq E$: $D_E(A)$ is in PUF$_\cup$ $\cup$ TRIV$_1$.

D is PUF$^0_\cap$ iff for all $A \subseteq E$: $D_E(A)$ is in PUF$_\cap$ $\cup$ TRIV$_0$.

D is TRIV$_0$ (TRIV$_1$) iff for all $A \subseteq E$: $D_E(A)$ is in TRIV$_0$ (TRIV$_1$).
When D is TRIV$_0$ or TRIV$_1$ we say that D is *trivial*.

D is TRIV$^\exists_0$ (TRIV$^\exists_1$) iff there exist $A \subseteq E$ s.t. $D_E(A)$ is in TRIV$_0$ (TRIV$_1$).

Thus, a determiner is called PUF$^1_\cup$ (PUF$^0_\cap$) when it generates only PUF$_\cup$ (PUF$_\cap$) and trivial quantifiers. Note that a determiner is classified as PUF$^1_\cup$ (PUF$^0_\cap$) or TRIV$_0$ (TRIV$_1$) according to its behavior on *all* domains and arguments. By contrast, for classifying a determiner as TRIV$^\exists_0$ (TRIV$^\exists_1$), it is sufficient to find *one* domain and one argument for which it is TRIV$_0$ (TRIV$_1$). The usefulness of both "universal" and "existential" classifications of determiners will be clarified as we go along.

Our main claim is that this typology of determiners allows us to determine in which cases of MON↑ determiners D_1 and D_2, the ONS reading D_1-D_2 entails (or is entailed by) the OWS reading $D_1 \sim D_2$. Before proving that, there is one qualification concerning this result that we should explain. We will assume that both D_1 and D_2 are *finitely based*, in a sense that is defined below. This restriction is needed because MON↑ determiners such as *infinitely many* behave with respect to relative scope entailments differently than MON↑ determiners such as *at least three*. Consider the following examples.

(3) a. Infinitely many students saw John or Mary.
b. At least three students saw John or Mary.

(4) a. Infinitely many students saw at least one of the two students.
b. At least three students saw at least one of the two students.

In (3a), the ONS reading entails the OWS reading: if there are infinitely many students that have the property *saw John or saw Mary*, then either John or Mary has the property *was seen by infinitely many students*. But this is obviously not the case in (3b). A similar contrast is observed between (4a) and (4b), under a Russellian treatment of the definite article. For instance:

(5) **at_least_one_of_the_n**$'(A)(B) = 1 \Leftrightarrow |A| = n \wedge A \cap B \neq \emptyset$

(6) **each_of_the_n**$'(A)(B) = 1 \Leftrightarrow |A| = n \wedge A \subseteq B$

We have seen that MON↑ determiners such as *infinitely many* show scope entailments that are different than those of similar "finite" determiners. Such "infinite" determiners, which are common in the mathematical jargon, are much less common – and have a much less defined meaning – in everyday speech. This is in contrast to more ordinary determiners such *at least three* or *every*, which English speakers use by and large with the same meaning as logicians do. The formal distinction between determiners that is held responsible for this difference is defined as follows.

Definition 1 (FB **quantifiers**) *Let E be a non-empty domain. A sequence $A_i|_{i=1}^{\infty}$ of subsets of E is called* properly monotone *if $A_i \subset A_{i+1}$ for every $i \geq 1$, or $A_i \supset A_{i+1}$ for every $i \geq 1$.*
Two properly monotone sequences $A_i|_{i=1}^{\infty}$ and $B_j|_{j=1}^{\infty}$ are called mutually monotone *if $A_i \subset B_j$ for all $i, j \geq 1$, or $A_i \supset B_j$ for all $i, j \geq 1$.*
A quantifier Q_E over E is called finitely based (FB) *iff for any two mutually monotone sequences $A_i|_{i=1}^{\infty}$ and $B_j|_{j=1}^{\infty}$ s.t. Q_E is constant on both sequences, Q_E sends both sequences to the same value.*

By "constancy" of a quantifier Q_E on a set $\mathcal{X} \subseteq \wp(E)$, we of course mean: $\mathcal{X} \subseteq Q_E$ or $Q_E \cap \mathcal{X} = \emptyset$. In the first case say we say that Q_E sends $\mathcal{X}$ to 1. In the second case say we say that Q_E sends $\mathcal{X}$ to 0.
The definition of FB determiners is derived from the definition of FB quantifiers.

Definition 2 (FB **determiners**) *A global determiner D is* FB *iff for any domain E, $D_E(E)$ is an* FB *quantifier.*

Note that this definition pays attention only to the behavior of D_E on the whole E domain, and does not take into account proper subsets of E. Thus, a determiner such as *all* is provably FB, even though on the domain of natural numbers, the quantifier *all odd natural numbers* is *not* FB. This is in accordance with the intuition that the determiner *all* does not inherently pertain to infinite sets. By contrast, the determiner *all but finitely many* provably maps any infinite domain to a non-FB quantifier, hence it is not FB itself.

Let us consider an example for a pair of FB/non-FB determiners that belong in the same class of the above hierarchy. Consider first the determiner *infinitely many*. Let N by the set of natural numbers, with $N_O \subset N$ the set of odd natural numbers. Consider two sequences $(N_O \cap [1..2i])|_{i=1}^{\infty}$ – the increasing sequence of sets of odd numbers; and $(N_O \cup [1..2i])|_{i=1}^{\infty}$ – the unions of the odd numbers with elements in the increasing sequence of sets of even numbers. These two sequences are mutually monotone, but the denotation of *infinitely many natural numbers* on the domain $E = N$ is constantly false on the first sequence but constantly true on the second sequence. Consequently, the determiner *infinitely many* is *not* FB. It is impossible to find two such sequences for the determiner *at least three*: trivially, for any domain E, the quantifier $\mathbf{at\_least\_3}'_E(E)$ cannot be false over an infinite properly monotone sequence. Consequently, the determiner *at least three* is FB. Note however that, for each of the determiners *infinitely many* and *at least three*, there are quantifiers that the determiner forms that belong in the class MON↑ \(PUF$_\cup$ ∪ PUF$_\cap$). Hence, both determiners are in the class MON↑ \(PUF$^1_\cup$ ∪ PUF$^0_\cap$). Some more examples for FB and non-FB determiners are given below. We note without proof that the class of FB determiners is closed under complements and finite intersections and unions.

FB Determiners: at least/at most/exactly 3; all; all but at least/most 3.
Non-FB Determiners: (in)finitely many; all but (in)finitely many.

We observe the following fact about upward monotone FB quantifiers.

Lemma 5 *Let Q be an FB upward monotone quantifier over a denumerable domain E. If $C_1 \supset C_2 \supset \ldots$ is a properly decreasing infinite sequence of sets in Q, then there is a finite set $A \subseteq C_1$ in Q.*

For the statement of our main claim, recall the following definitions, which are standard in GQ theory ([2]). For any global determiner D:
D satisfies *extension* (EXT) iff for all $A, B \subseteq E \subseteq E'$:$D_E(A)(B)$=$D_{E'}(A)(B)$.
D is *isomorphism invariant* (ISOM) iff for all bijections $\pi : E \to E'$, for all $A, B \subseteq E$: $D_{E'}(\{\pi(x) : x \in A\})(\{\pi(y) : y \in B\}) = D_E(A)(B)$.
D is *conservative* (CONS) iff for all $A, B \subseteq E$: $D_E(A)(B) = D_E(A)(A \cap B)$.
As in other works on GQ theory, we restrict our attention to determiners in natural language that are EXT, ISOM and CONS.

It is now possible to move on to our main claim.

Theorem 6 *Let D_1 and D_2 be two global* MON↑ *determiners that satisfy* FB, EXT *and* CONS. *Then D_1-$D_2 \subseteq D_1$~D_2 for any domain E iff both following conditions hold: (1) D_1 is* PUF$^1_\cup$ *or D_2 is* PUF$^0_\cap$*; and (2) D_1 is not* TRIV$^\exists_1$ *or D_2 is not* TRIV$^\exists_0$.

The proof of the "if" direction is quite direct. To prove the "only if" direction, we make use of the following two lemmas, which rely on the FB property.

Lemma 7 *Let D be an* FB *determiner in* MON↑ \PUF$^1_\cup$ *that satisfies* EXT *and* CONS. *Then there are $A \subseteq E$, for which there is $B \in D_E(A)$ s.t. $|B| \geq 2$ and for every $X \subset B$: $X \notin D_E(A)$.*

Lemma 8 *Let D be an* FB *determiner in* MON↑ \PUF$^0_\cap$ *that satisfies* EXT *and* CONS. *Then there is a domain E and $A \subseteq E$, for which there are $B_1, B_2 \in D_E(A)$ s.t. $B_1 \cap B_2 \notin D_E(A)$.*

Theorem 6 characterizes all the FB logical cases of upward right-monotone subject and object determiners that make the ONS reading entail (or be entailed by) the OWS reading. Simple cases like that are when the subject determiner is PUF$^1_\cup$\TRIV$^\exists_1$ or when the object determiner is PUF$^0_\cap$\TRIV$^\exists_0$. That is: when the subject always denotes a PUF$_\cup$ quantifier or the object always denotes a PUF$_\cap$ quantifier. This is the case in the following sentences.

(7) a. Some student saw every/most/at least two teachers.
b. Every/most/at least two student(s) saw every teacher.

However, to characterize completely the cases of MON↑ logical determiners for which the ONS reading entails the OWS reading, we have also considered some more complex cases of global determiners. An example for a member in PUF$^1_\cup$ ∩ TRIV$^\exists_1$ is the determiner *some or every*. Examples for members in PUF$^0_\cap$ ∩ TRIV$^\exists_0$ are the determiner *some and every* and the determiner *each of the five* (cf. definition (6)). These determiners show entailments from the ONS reading to the OWS reading in sentences such as the following.

(8) Some or (perhaps even) every student saw some or (perhaps even) every teacher.

(9) a. At least two teachers saw some and (in fact) every student.
b. At least two teachers saw each of the five students.

The complete characterization of scope entailments with MON↑ determiners explains why there is no entailment from the ONS reading to the OWS reading in simple cases such as the following.

(10) Every/most/at least two student(s) saw some/most/at least two teacher(s).

Also in more complex cases such as the following, there is no entailment from the ONS reading to the OWS reading, as theorem 6 expects.

(11) Some or (perhaps even) every student saw some teacher.

(12) Every student saw some and (in fact) every teacher.

In both cases, when there are no students and no teachers, the ONS reading is true but the OWS reading is false.

Another result concerns the following fact that is mentioned in [3] about *local* quantifiers. Westerståhl calls two quantifiers Q_1 and Q_2 over E *independent* when Q_1-$Q_2 = Q_1{\sim}Q_2$. Then he makes the following claim.

Proposition 9 (Westerståhl) *Let Q_1 and Q_2 be two quantifiers over E that are* MON↑, *non-trivial and* ISOM. *Then Q_1 and Q_2 are independent iff $Q_1 = Q_2 =$* $\mathbf{every}'_E(E)$ $(= \{E\})$ *or* $Q_1 = Q_2 = \mathbf{some}'_E(E)$ $(= \wp(E) \setminus \emptyset)$.

When we consider global determiners, we call D_1 and D_2 *independent* if D_1-$D_2 = D_1{\sim}D_2$ for any domain E. Theorem 6 entails the following fact about independent determiners. Note the ISOM requirement (as in Westerståhl's proposition), in addition to the requirements in theorem 6.

Corollary 10 *Let D_1 and D_2 be two global* MON↑ *determiners that satisfy* FB, EXT, ISOM *and* CONS.
Then D_1 and D_2 are independent iff both of the following conditions hold:

1. *At least one of the following holds: (a) D_1 and D_2 are both $\mathrm{PUF}^1_{\cup}$; or (b) D_1 and D_2 are both $\mathrm{PUF}^0_{\cap}$; or (c) D_1 is trivial; or (d) D_2 is trivial.*
2. *At least one of the following holds: (a) Neither D_1 nor D_2 are $\mathrm{TRIV}^{\exists}_0$; or (b) Neither D_1 nor D_2 are $\mathrm{TRIV}^{\exists}_1$.*

Examples for identical D_1 and D_2 that are independent are the following cases: $D_1 = D_2 =$ *some, every, some-or-every, some-and-every*. However, independent determiners do not have to be identical. For instance: *each of the two* and *each of the five* are independent determiners, since according to the Russellian definition in (6), they are both in $\mathrm{PUF}^0_{\cap} \setminus \mathrm{TRIV}^{\exists}_1$.

References

[1] E. Keenan and J. Stavi. A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9:253–326, 1986.

[2] E. Keenan and D. Westerståhl. Generalized quantifiers in linguistics and logic. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*. Elsevier, Amsterdam, 1996.

[3] D. Westerståhl. Self-commuting quantifiers. *The Journal of Symbolic Logic*, 61:212–224, 1996.

[4] T. E. Zimmermann. Scopeless quantifiers and operators. *Journal of Philosophical Logic*, 22:545–561, 1993.

Game Theoretic Foundations for Gricean Constraints

Nicholas Asher, Itai Sher and Madison Williams
Department of Philosophy
University of Texas at Austin
Austin, TX, 78712, USA
email: nasher@bertie.la.utexas.edu

1 Introduction

Gricean maxims of quality, quantity and relevance have an intuitive appeal and would appear to be a part of any adequate pragmatic theory. Previous work in formal pragmatics has formalized some of these principles; e.g., using the conditional operator $>$ of Asher and Morreau (1991), we can formalize some Gricean intuitions in the two defaults, Competence and Sincerity:

- Competence (for the hearer): $\mathcal{B}_S\phi > \mathcal{B}_H\phi$
- Sincerity (for the speaker): $Says_S\phi > \mathcal{B}_S\phi$

This allows us to ask and to begin to answer a question fundamental to the notion of pragmatics as a theory of *rational* action in discourse: do such defaults have a deep connection to established models of rational behavior? While Sperber and Wilson (1986) *inter alia* have criticized Gricean approaches and denied any important connection, we disagree and give here an affirmative answer to this question for the defaults, Competence and Sincerity. Given certain assumptions, we show that these defaults describe

a unique, strict Nash Equilibrium for speakers and hearers in a particular game. We will also exploit an equivalence between strict Nash Equilibria and stabilization points in evolutionary game theory to give an argument for how such defaults could come about.

Our approach generalizes, we believe, to other pairs of defaults that formalize Grice's intuitions about conversational behavior.

2 Pragmatics and Static Games

The sort of games we are interested in involve two players, a speaker (S) and a hearer (H). We will assume that S has some information or misinformation he wishes to convey to H. The basic strategies open to a speaker are to tell the truth (T) or to lie (F), and for H they are to believe (B)or to not believe ($\neg B$) what S said. Thus, at first glance, it appears that we may represent the game with a simple 2×2 matrix where the rows represent S's strategies T and F while the columns represent H's strategies. The strategic aspects of this game are as follows. H would rather

believe what he is told if S tells him the truth; on the other hand, he would rather disbelieve what he is told if S lies to him. Presumably it is H's goal in this exchange to gain some new information, with the proviso that this information be accurate (hence useful to him). From S's perspective, it is better to be believed than disbelieved for the simple reason that, whether his information is true or not, if he is disbelieved, he will have failed to convey her information to player 2.

In trying to get a decent game theoretic representation of this complex epistemic situation in dialogue, we run into a difficulty in completely specifying the strategic aspects of S. Regardless of whether it is in his interest to tell the truth or not, it seems reasonable to suppose that he will be indifferent between telling H the truth and lying to him if H will not believe him in any case. But if it is in S's interest to tell the truth, then he will rather tell H the truth if H will believe him. And if it is in S's interest to lie, then he will prefer to lie to the hearer if the latter will believe him. Thus, a simple matrix is insufficient to represent the interplay of strategies needed to support or to undermine the defaults of

Competence and Sincerity. This distinguishes us from Lewis (1969) and Axelrod (1984).

A more adequate representation of the game is a 4×4 matrix in which the simple strategies for H and S now depend on a parameter α. α stands for the proposition that it is in S's interest to tell the truth in his current situation, whereas $\neg\alpha$ means that it is in S's interest to lie. The pure-strategy best reply correspondence for this game for the figure below is as follows: $b_{44} > b_{4i}, i \neq 4; a_{i4} = a_{j4}, \forall i, j; a_{21} > a_{2i}, i \neq 1; b_{11} > b_{1j}, j \neq 1; b_{22} > b_{2j}, j \neq 2; b_{33} > b_{3j}, j \neq 3; a_{33} > a_{i3}, i \neq 3$.

	$\alpha\to B$, $\neg\alpha\to B$	$\alpha\to B$, $\neg\alpha\to\neg B$	$\alpha\to\neg B$, $\neg\alpha\to B$	$\alpha\to\neg B$, $\neg\alpha\to\neg B$
$\alpha\to T$, $\neg\alpha\to T$	a_{11}, b_{11}	a_{12}, b_{12}	a_{13}, b_{13}	a_{14}, b_{14}
$\alpha\to T$, $\neg\alpha\to F$	a_{21}, b_{21}	a_{22}, b_{22}	a_{23}, b_{23}	a_{24}, b_{24}
$\alpha\to F$, $\neg\alpha\to T$	a_{31}, b_{31}	a_{32}, b_{32}	a_{33}, b_{33}	a_{34}, b_{34}
$\alpha\to F$, $\neg\alpha\to F$	a_{41}, b_{41}	a_{42}, b_{42}	a_{43}, b_{43}	a_{44}, b_{44}

In this larger game, both players condition their behavior on whether α or $\neg\alpha$ is the case. If H believes S no matter what, then S will adopt T given α and F given $\neg\alpha$. If H believes S given α but not given $\neg\alpha$, S will tell the truth given α and be indifferent between T and F given $\neg\alpha$. If H disbelieves S no matter what, then S will be indifferent among his strategies. H's best response correspondence to any strategy is that he will believe S in any circumstance in which S adopts T and disbelieve S whenever he adopts F. This game has two equilibria: $\langle(\alpha\to F, \neg\alpha\to F), (\alpha\to\neg B, \neg\alpha\to\neg B)\rangle$ (abbreviated as $\langle(F,F)(\neg B,\neg B)\rangle$, and $\langle(T,F),(B,\neg B)\rangle$, neither of which support the defaults we are interested in. They are also both weak equilibria, in the sense that they are not strictly preferred by both S and H to all other strategies. In particular the pair that supports our defaults $\langle(T,T)(B,B)\rangle$ is not an equilibrium; for, if H always believes S, then S will be able to take advantage and optimize with respect to α and $\neg\alpha$.

To support Competence and Sincerity, we must construct a situation in which S tells the truth despite the fact that it is in his interest to lie. Given our set up this means we must assume that what S does in the situation in which it is in his interest to lie has an impact on his payoff in the situation in which it is in his interest to tell the truth. One method of achieving this might be to study reputation effects in repeated games; another, which is the avenue we pursue, is to take into account the costs of employing a more complex strategy. To illustrate, let $C_S : \{T, F\} \to \mathcal{R}$ be a cost function for the speaker. We can reasonably assume:

$$X \neq Y \to C_S(X,Y) > C_S(X,X) = C_S(Y,Y) \quad (1)$$

since it is more expensive to choose a strategy dependent upon α than one that is independent of α. For now we only factor in the costs of strategies for S in computing the best response correspondence. So the H's best responses remain as before, but S's preferences will now partially depend on the costs of the strategies. Consider how the two previous equilibria fare. If H disbelieves S no matter what, then S will be indifferent between truth telling and lying. But to save costs he will either choose (T,T) or (F,F). Now consider the case when H believes S given

α but not when $\neg\alpha$. Then S will be indifferent between T and F given $\neg\alpha$, but he will prefer a simple strategy that maximizes his gain at α; i.e. (T,T). Now what happens where H adopts (B,B)? We can immediately rule out (F,T) since it is complex and assigns the least appropriate response in either α or $\neg\alpha$. To choose between the two simple strategies for S, we must assume something about the subjective probability of α. If we assume

$$p(\alpha) > 1/2 \quad (2)$$

, which seems reasonable if a basic function of language is information exchange to help speakers attain their needs, we can rule out (F,F).

We now need only to figure out how S distinguishes between (T,T) and (T,F). In order to do this, we must make some additional assumptions about the utility function of S. To this end, we decompose the speaker's utility function u into two parts:

$$u(a,(X,Y)) = \overline{u}(a,Z) - C(X,Y) \quad (3)$$

where $Z = X$ if $a = \alpha, Z = Y$ if $a = \neg\alpha$. C is the cost function discussed above; $\overline{u}$ is a function that gives the utility the speaker derives from lying or telling the truth for a given value of *a given that the hearer believes the speaker*. Therefore both the functions u and $\overline{u}$ are calculated on the assumption that the hearer believes the speaker. To proceed further, we assume about $\overline{u}$ that:

$$\overline{u}(\alpha,T) = \overline{u}(\neg\alpha,F) > \overline{u}(\alpha,F) = \overline{u}(\neg\alpha,T) \quad (4)$$

The inequality in (4) holds, if it is more worthwhile for the speaker to do what it is in his interest to do than to do what it is not in his interest. The equalities are simplifying assumptions; we are not interested in considerations that arise from a differences in penalties imposed on the speaker for going against his interests in α and $\neg\alpha$. Now what are the conditions under which

$$E_a[u(a,(T,T)] > E_a[u(a,(T,F)] \quad (5)$$

That is, what are the conditions under which it is in the speaker's interest to employ a uniformly truthful strategy, rather than a strategy which is truthful depending on the circumstances, on the assumption that the hearer will
believe the speaker no matter what? (5) holds iff

$$p(\alpha)[\overline{u}(\alpha,T) - C(T,T)] + (1-p(\alpha))[\overline{u}(\neg\alpha,T) - C(T,T)]$$
$$> p(\alpha)[\overline{u}(\alpha,T) - C(T,F)] + (1-p(\alpha))[\overline{u}(\neg\alpha,F) - C(T,F)]$$

iff by algebra:

$$C(T,F) - C(T,T) > (1-p(\alpha))[\overline{u}(\neg\alpha,F) - \overline{u}(\neg\alpha,T)] \quad (6)$$

By (4) $\overline{u}(\neg\alpha,F) - \overline{u}(\neg\alpha,T) > 0$ so that (6) holds iff :

$$p(\alpha) > 1 - \frac{C(T,F) - C(T,T)}{\overline{u}(\neg\alpha,F) - \overline{u}(\neg\alpha,T)} \quad (7)$$

Note that according to (7) if the difference in cost between (T,F) and $((T,T)$ rises, then the probability of α needed to make (T,T) preferable to (T,F) shrinks. This makes sense because when $C(T,F) - C(T,T)$ rises, this will deter the speaker from choosing the more complex strategy so that he will need less motivation in terms of probability of α, the state which is favorable to his course of action, to motivate him to choose (T,T). On the other hand (7) also says that if $\overline{u}(\neg\alpha,F) - \overline{u}(\neg\alpha,T)$

rises, then so must $p(\alpha)$ in order to justify (T,T). This also makes sense because the bigger the difference in utilities between F and T given that $\neg\alpha$, the bigger the penalty the speaker pays in $\neg\alpha$, and so the higher the probability of α he needs to justify (T,T).

Combining (2) and (7), we have that

$$p(\alpha) > \max\{1/2, 1 - \frac{C(T,F) - C(T,T)}{\overline{u}(\neg\alpha, F) - \overline{u}(\neg\alpha, T)}\} \quad (8)$$

In order to insure that $p(\alpha)$ is a probability, we must be sure that $1 - \frac{C(T,F)-C(T,T)}{\overline{u}(\neg\alpha,F)-\overline{u}(\neg\alpha,T)} \leq 1$, which is equivalent to the assumption that $1 - \frac{C(T,F)-C(T,T)}{\overline{u}(\neg\alpha,F)-\overline{u}(\neg\alpha,T)} > 0$. However this follows from the fact that both the numerator and denominator are positive by (1) and (4).

If we assume (8), we can show that this matrix there has two equilibria, $\langle(F,F),(\neg B,\neg B)\rangle$ which is weak and $\langle(T,T),(B,B)\rangle$ which is strict. Even if we were to take equlibria in mixed strategies into account, $\langle(T,T),(B,B)\rangle$ would be the only strict equilibrium, since no mixed strategy equilibrium can be strict. Note also that the assumption that the hearer does not take costs into account may now be lifted without changing the equilibria. If the hearer were to take costs into account, then his best response correspondence would only change in the second and third rows of the matrix, since these are the only rows in which his best response correspondence selects a complex strategy. Thus, in this case, $((T,T),(B,B))$ is the unique strict Nash Equilibrium of the game.

Factoring in the costs of H doesn't change the outcome of our game. If we assume that more complex strategies are more costly than simple ones for both H as well as S, then H's best response correspondence will only change in the second and third rows of the matrix, since these are the only rows in which his best response correspondence selects a complex strategy. However, from inspection of the matrix, it is obvious that it does not matter which strategy the hearer's best response correspondence assigns to the hearer in the second and third rows, since there is no strategy in these rows which is the output of the *speaker's* best response correspondence in those rows. So regardless of what the hearer's best response correspondence does in these two rows, no new pure strategy equilibrium will be added or deleted.

Two important questions remain about the interpretation of our game. The strict Nash Equilibrium described above involves a particular choice of strategies for S and H. How does this exactly support the defaults of Competence and Sincerity? We have shown something universal in a simple setting. Real world communication, however, is more complex and information α and the costs of the strategies is more partial and more uncertain. Calculations concerning the costs of various strategies for the speaker and whether α obtains or not might well be in principle impossible for the hearer to do, since he may lack information relevant to discovering whether α; and while the speaker may have access to the information relevant to determine α, it may be a very difficult calculation for him to

do under the constraints of real time conversation. In more complex situations, the conditions under which the simple strategies are chosen become part of the "normal" conditions of discourse situations, and thus we arrive at a justification of the defaults themselves. The defaults of Charity and Sincerity are, as Mill (1957) argued in another area of philosophy, shorthand approximations of the ideally rational situation; they have the advantage that they bypass complex calculations about costs and whether α holds or not.

3 Pragmatics and Evolutionary Games

The second question is, how such defaults might come about? While semantics and pragmatics have almost nothing to say about the issue, it appears that evolutionary approaches to game theory might give us an answer. Our result above translates easily into an evolutionary framework in which populations of players of the game above are distributed across various strategies. These populations' selection of strategies "evolves" according to a "selection dynamic", a set of time differential equations of the form $\dot{x} = \varphi(x)$, where $\varphi : X \subseteq \mathcal{R}^k \to \mathcal{R}^k$ describes changes in the state space (population shares of strategies). The goal in this framework is to see where the evolutionary dynamics of the system stabilizes; asymptotic stability is the standard stability condept in these games. The theorem (see Weibull 1995) that a population state is asymptotically stable at x (a choice of strategies) in the standard replicator dynamic iff x is a strict Nash equilibrium to transfer our earlier result into the dynamic framework. This entails that $((T,T),(B,B))$ is the only asymptotically stable population state, and that other weak equilibria in the static game are not asymtotically stable.

We can broaden the class of selection dynamics that would support our defaults. We want a close connection between the payoffs and the growth rate. We need two concepts, payoff positivity and payoff monotonicity, which are defined as follows: g is payoff monotonic in a game Θ with a set of players $I = \{H, S\}$, and a set of strategies $S_i, i \in I$, iff $\forall$ population distributions $x \in \Theta, i \in I, h, k \in S_i, u_i(e_i^h, x_{-i}) > u_i(e_i^k, x_{-i}) \Leftrightarrow g_{ih}(x) > g_{ik}(x)$, where e_i^h is the payoff to player i playing pure strategy h, u_i is the utility function for i and x_{-i} is the population distribution $-i$. G^m is the class of monotonic growth rate functions. In a payoff-monotonic selection dynamic, a pure strategy with a higher payoff always has a higher growth rate. Now for payoff positivity: g is payoff positive if and only if $\forall x \in \Theta, i \in I, h \in S_i, sgn[g_{ih}(x)] = sgn[u_i(e_i^h, x_{-i}) - u_i(x)]$. G^p is the subclass of payoff positive growth rate functions. Intuitively, a selection

dynamics is payoff positive when a pure strategy has a positive growth rate if and only if its payoff is above the population average. An imitation selection dynamic, which might more plausibly characterize the evolution appropriate for our default, is payoff linear (and so in $G^p \cap G^m$), while the standard replicator dynamic is not (though it's monotonic). But we can make do with less:

Theorem 1 *A profile $x \in \Theta$ is asymptotically stable in a selection dynamic $g \in G^m \cup G^p$ if and only if x is a strict Nash Equilibrium.*

Proof: Let x be an asymptotically stable profile in a dynamic $g \in G^m \cup G^p$. But for any $g \in G^m \cup G^p$, if $x \in \Theta^{NE}$ is a pure but not a strict Nash equilibrium, or x belongs to to a non-singleton component of Θ^{NE}, then x is not asymptotically stable in $\dot{x}_{ih} = g_{ih}(x)x_{ih} \forall i \in I, h \in S_i, x \in \Theta$ (Weibull, Proposition 5.12). So x is a strict Nash Equilibrium. Now let $x \in \Theta^{NE}$ be strict. Every strict Nash equilibrium $x \in \Theta$ is asymptotically stable in all weakly payoff-positive selection dynamics $\dot{x}_{ih} = g_{ih}(x)x_{ih} \forall i \in I, h \in S_i, x \in \Theta$ (Weibull, Proposition 5.11). So x is stable (since $G^m \cup G^p \subseteq G^w$).

Evolutionarily, we can show that a population will converge on the population state corresponding to our norm x_{TT}, y_{BB} from any point in the interior of the game space Θ, provided the payoff structure is as described above. The interior of Θ, $int(\Theta) = \times_{i \in I} int(\Delta_i) = \times_{i \in I}\{x_i \in \Delta_i | \forall h \in S_i, x_{ih} > 0\}$. Δ_i is a mixed strategy space of population i, or, for our purposes, the population mixture of a particular strategy. $bd(\Theta) = \times_{i \in I} bd(\Delta_i) = \times_{i \in I}\{x \in \Delta_i | x_i \notin int(\Delta_i)\}$. By showing that, given a certain plausible payoff structure, the only pure Nash equilibrium is at x_{TT}, y_{BB}, we are able to demonstrate that any $x \in int(\Theta)$ will converge to the population state corresponding to our norm. More formally, we have shown that the vector $x = \langle 1,0,0,0,1,0,0,0\rangle \in \Theta$, the strategy space assigning population share 1 to x_{TT} and y_{BB} and 0 to all others, is the only asymptotically stable strategy and converges everywhere in $int(\Theta)$. We now turn to an example with the replicator dynamic to show how this works. Since this dynamic is intended to simulate a biological replication it has the feature that, if $x_i \in \Theta = 0$ then $\dot{x}_i = 0$: a population that does not exist cannot grow. It follows that an $x \in bd(\Theta)$ will not necessarily converge to our norm. For instance if we start at $x_{FF}, y_{DD} = 1$ we will be stuck there. However, a dynamic with any possibility of perturbation will, at some point, put the population into $int(\Theta)$ and so will always go to x_{TT}, y_{BB}. Take the following values for a_{ij} and b_{kl} in the 4×4 matrix above now understood as fitness values,, where the speaker and hearer each get a point for "doing the right thing" but

where the speaker loses half a point for choosing a complex strategy: $a_{11} = a_{41} = a_{12} = a_{43} = b_{12} = b_{13} = b_{24} =$
$b_{34} = b_{42} = b_{43} = 1; a_{22} = a_{23} = a_{24} = .5; a_{3i} = -.5; b_{44} = b_{33} = b_{22} = b_{11} = 2.$

As usual, each population share of a strategy $x_h \in [0,1]$ and $\sum_{h \in S} x_h = \sum_{k \in H} y_k = 1$. Let $A = \{x\} = \langle 1,0,0,0,1,0,0,0\rangle$. To prove our claim, we need to show that there is a neighborhood, $D = int(\Theta)$, of A and a continuous function $v : D \to \mathcal{R}_+$ such that $v(x) = 0$ if and only if $x \in A$ and $v(\xi(t,x)) < v(x)$ if $x \notin A, t > 0$ and $\xi(s,x) \in D, \forall s \in [0,t]$. (c.f. Theorem 6.3, Weibull p. 246.) Trivially, A is closed. Let D be a neighborhood of A, and define $v(x) : D \to \mathcal{R}_+ = \dot{x}_{TT}(x) + \dot{y}_{BB}(x)$ restricted to D, where ẋ is the change in over time with respect to the row component and ẏ is the change with respect to the column component. First, let us check what happens when $v(x) = 0$ (i.e. when $\dot{x}_{TT}(x) = \dot{y}_{BB}(x)$.) Thus $[\sum_{k \in S_2} a_{TT,k} y_k - \sum_{j \in S} \sum_{k \in H} x_j a_{jk} y_k] x_{TT} = [\sum_{h \in S} b_{h,BB} y_h - \sum_{j \in H} \sum_{k \in S_1} x_j a_{hj} y_h] y_{BB}$ Since everything in D must be ε-close to A, $x_{TT}, y_{BB} \leq 1 - \varepsilon$ with the remaining population distributed over the other three strategies. So we have $\dot{x}_{TT} = (1-\varepsilon)(\frac{7\varepsilon}{3}) = (1-\varepsilon)(\frac{-75\varepsilon^2}{9}) = \dot{y}_{BB}$ So $v(x) = 0$ when $\varepsilon = 0$ (i.e. at A) or $\varepsilon = 1$ (i.e. everywhere else along the boundary). This establishes the first clause.

For the second, assume $x \notin A, t > 0$, and $\xi(s,x) \in D, \forall s \in [0,t]$ (hence $x \in D$). Since $\xi(t,x) = x + (\dot{x}_{TT} + \dot{y}_{BB})^t$ and $v(x) = (1-\varepsilon)(\varepsilon - 6\varepsilon^2), v(\xi(t,x))$ (which we can think of as the t^{th} moment after x) $= ((1-\varepsilon)(\varepsilon - 6\varepsilon^2))^t$ If we take $\varepsilon < 0.1$ we have $(1-\varepsilon)(\varepsilon - 6\varepsilon^2) < 1$ hence $((1-\varepsilon)(\varepsilon-6\varepsilon^2))^t < ((1-\varepsilon)(\varepsilon - 6\varepsilon^2))$ for $t > 1$ which establishes the second clause. Hence x is asymptotically stable. Further, because we have shown fixed points of the dynamic exist only in $bd(\Theta), D = int(\Theta)$. Finally, every point in $D = int(\Theta)$ converges to our norm. If it did not, since v is continuous, there would have to be a point, $x \in D$ where $v(x) = 0$ but we have shown that there is not. Hence, x converges to the norm across $D = int(\Theta)$.

4 References

Asher, N. and M. Morreau. 'Commonsense Entailment: A Modal, Nonmonotonic Theory of Reasoning', *Proceedings of IJCAI 91.*

Axelrod, R. *The Evolution of Cooperation.* New York: Basic.

Lewis, David. 1969. *Convention.* Cambridge: Harvard University Press.

Mill, John Stuart. 1957. *Utilitarianism.* Oskar Priest, ed. Indianapolis: Bobbs-Merrill.

Sperber, Dan and Deirdre Wilson. *Relevance.* Oxford: Oxford University Press

Weibull, Jorgen W. 1995. Evolutionary Game Theory. Cambridge, MA: The MIT Press.

Types for linguistic typologies. A case study: Polarity Items

Raffaella Bernardi, UiL OTS, University of Utrecht

Abstract

This paper describes how in categorial type logics (CTL) derivability patterns among types give a precise way to (i) gain a deeper understanding of the typological classifications proposed in the literature of formal linguistics, (ii) carry out cross-linguistic comparisons, (iii) clarify the consequences predicted by the typologies opening the way to further investigations, and (iv) establish new dependency between linguistic phenomena. In particular, the base logic studied in [2] ($\mathsf{NL}(\diamond, \cdot^0)$) is used to elucidate the typology of Greek polarity items discussed in [3] and compare it with Italian data and types. The picture obtained predicts the existence of (i) non veridical contexts which do not license polarity items of the classes considered so far and which could license other sort of polarity items instead; and of (ii) contexts shared by 'negative' and positive polarity items. Finally, it sheds light on a connection between dynamic Montague grammar (DMG) and CTL giving new intuitions about the interpretations of the unary operators of $\mathsf{NL}(\diamond, \cdot^0)$.

1 Derivability relations in categorial type logic

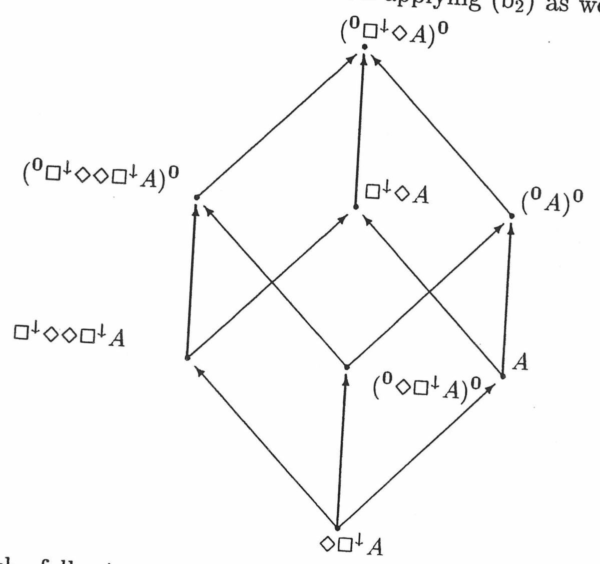
In formal linguistic literature, one finds examples of theories based on classifications of items which belong to the same syntactic category but which differ in some respect. For example, generalized quantifiers have been classified considering the different ways of distributing with respect to negation [1]; a typology of wh-phrases can be given considering their sensitivity to different weak-islands strength [6]; adverbs differ in their order relations [5]; similarly, polarity items have been distinguished by the sort of licensors they require for grammaticality [7, 3].

Our aim is to show how categorial type logic can contribute to the study of linguistic typologies. In particular, we employ $\mathsf{NL}(\diamond, \cdot^0)$ to account for the classification of polarity items based on the distinction between veridical and non-veridical expressions proposed in [3]. In this paper we concentrate on the linguistic applications and present only the essential ingredients of the logic system. The reader interested in the logic details of $\mathsf{NL}(\diamond, \cdot^0)$ is referred to [2].

Besides functional and compositional connectives $(\backslash, \bullet, /)$, $\mathsf{NL}(\diamond, \cdot^0)$ has unary operators $(\diamond, \Box^{\downarrow}, {}^0\cdot, \cdot^0)$, which can be used to encode the different (semantic) features which characterize items of the same (syntactic) category, e.g. non-veridicality vs. veridicality. Let us introduce this logic briefly. $\mathsf{NL}(\diamond, \cdot^0)$ is the base logic characterized by two simple algebraic principles closely connected, i.e. its (unary and binary) operators form either a residuated or a Galois pair. In particular, $\diamond, \Box^{\downarrow}$ form a residuated pair of functions, whereas the other ones ${}^0\cdot, \cdot^0$ are Galois connected operators. More specifically, the former are order-preserving with respect to derivability ($\longrightarrow$), whereas the latter are order-reversing and they are governed by the relations below:

(a_1)	$\diamond\Box^{\downarrow}A \longrightarrow A$	(b_1)	$A \longrightarrow ({}^0A)^0$
(a_2)	$A \longrightarrow \Box^{\downarrow}\diamond A$	(b_2)	$A \longrightarrow {}^0(A^0)$

From these simple relations an interesting net of derivability patterns can be derived. The ones we have explored so far are summarized in the picture below which can be extended with a third level applying (b_2) as well.



Moreover, the following equivalences hold. Let (f_1, f_2) be either the residuated or the Galois pair above, $f_1 f_2 f_1 A \longleftrightarrow f_1 A$, and similarly $f_2 f_1 f_2 A \longleftrightarrow f_2 A$. For example,

$$\Diamond\Box^{\downarrow}\Diamond A \longleftrightarrow \Diamond A \quad \text{and} \quad \Box^{\downarrow}\Diamond\Box^{\downarrow}A \longleftrightarrow \Box^{\downarrow}A.$$

Let $F \in \{\Box^{\downarrow}\Diamond\cdot, ({}^0\cdot)^0, {}^0(\cdot^0)\}$, $FA \longleftrightarrow FFA$, viz. F is a closure operation. For example,

$$\Box^{\downarrow}\Diamond A \longleftrightarrow \Box^{\downarrow}\Diamond\Box^{\downarrow}\Diamond A \quad \text{and} \quad ({}^0A)^0 \longleftrightarrow ({}^0(({}^0A)^0))^0.$$

Before going to look at the linguistic applications of this system few comments are in order. Categorial type logic (CTL) fremowork provides a modular architecture to study constants and cross-linguistic variations of grammatical composition, viz. the former are accounted for in the base logic and the latter by means of the structural module. The logic we employ, $\mathsf{NL}(\Diamond, \cdot^0)$, is the base logic where no extra structural postulates are considered. In the next section we are going to show how its derivability relations displayed in the cube above give the right expressivity to account for the typology of polarity items (PIs) described in [3]. Note that in this case cross-linguistic differences between Greek and Italian PIs distribution are accounted for simply by means of lexical type assignments. A question which stays open is whether the logical type equivalences we have just shown will play a role when extending the fragment of the linguistic items we have considered so far.

2 Typology of Polarity Items

When speaking of PIs a classical reference is the work by Ladusaw [4] where negative polarity items (NPIs) are characterized as expressions requiring to be in the scope of a downward monotonic context (DM). In [7] it is shown

	Veridical	Non-veridical	Anti-veridical
$f_{(t,t)}$	now/yesterday	perhaps/usually	It is not the case
$f_{(t,(e,t))}$	weak int. vp	too-clauses	
$f_{((e,t),(e,t))}$		will/can	doesn't/without
$f_{((a,t),((c,t),t))}$: 1^{st} ar.	a	no/every	
$f_{((a,t),((c,t),t))}$: 2^{nd} ar.	a		no
$f_{(a,(a,a))}$: 1^{st} ar.	and	or/if	neither ... nor
$f_{(a,(a,a))}$: 2^{nd} ar.	and	or/if	neither ... nor
$f_{(a,(t,t))}$: 1^{st} ar.	before/without		
$f_{(a,(t,t))}$: 2^{nd} ar.	after	before	without

Table 1: (Non-)veridical contexts.

that downward monotonic functions can be classified into subclasses and this classification is reflected on the NPIs which differ in the class of licensor they require. In [3] Giannakidou shows that this classification based on a hierarchy of downward monotonic functions, though on the right track, is not general enough and proposes to enlarge the focus of the lens used to classify the licensors.

In particular, she gives evidence for a typology of PIs described in terms of sensitivity to (non-)veridicality – where quite intuitively a non-veridical expression (NV) is such that it does not entail the truth of the proposition it takes in its scope. The set of non-veridical functions contains as subset the one formed by anti-veridical expressions (AV) which are negation-like operators. We will refer to these two sets as NV and AV, respectevely. Notice that Ginnakidou's analysis does not refuse the previous proposals since in [8] it is proved that downward monotonic functions form a proper subset of the non-verdical ones.

Before going to present the typology of PIs and their account into categorial type logic, let us give a closer look to the classification of the NV which determines the one of PIs. In [8] non-veridicality is defined for expressions denoted by monadic and dyadic functions. Having assumed a categorial type logic perspective offers a deeper understanding of the typological classifications proposed in the literature. It gives a way to extend this definition to n-adic functions and prove the inclusion relation DM $\subseteq$ NV considering such extension. For reasons of space, we cannot go into the detail of such aspect and we limit ourselves to give a classification of some (non-)veridical contexts with the corresponding typed functions (Table 1). Note that since a subset relation holds between anti-veridical and non-verdical contexts, all the items which occur in the rightmost column should be repeated in the middle one as well. We can now move to see how PIs interact with these contexts.

2.1 PIs and non-veridical contexts

A first distinction among PIs is given by considering whether they can or cannot occur in veridical contexts (V). The ones which are felicitous in such contexts are considered positive polarity items (PPIs); the ones which are not, are named affective polarity items (APIs). Among the latter a further classification can be made by distinguishing the items which require to be in the scope of an anti-verdical expression, the label of negative polarity items (NPIs) is reserved for them.

An important aspect of this theory is that when an item is sensitive to a certain property A and this property is found to be in a subset relation with a more specific one B (i.e. B $\subseteq$ A), the item will then be expected to occur in contexts characterized by A as well as by B. This is the case of APIs which are felicitous both in non-veridical and anti-veridical contexts. On the other hand, since NPIs are sensitive to anti-verdicality and AV is a subset of NV, NPIs will not be grammatical in non-veridical contexts.

Finally, the relations between an item and the semantic property it depends on can be of different nature. This opens the way to a further classification of expressions belonging to the same sort of sensitivity dependency besides the one given by the set-relation holding between properties we have just considered. Giannakidou defines two possible relations which can be summarized as below[1]. Let SET be either NV, or AV, or V (or any set defined by a semantic property)

1. An item A is *licensed* by an expression $B \in$ SET iff $\forall C \in$ SET, A is grammatical in the (immediate) scope of C.

2. An item A is *anti-licensend* by an expression $B \in$ SET, if A is ungrammatical in the (immediate) scope of B.

This distinction creates the possibility of having for example PIs which are ungrammatical in veridical contexts, but which do not have to be felicitous in all non-veridical ones. A case in point is *any* which is in an anti-licensing relation with veridicality. If we consider the non-veridical contexts introduced in Figure 1, we see that *any* though forbidden in the veridical contexts is not grammatical, for example, in the scope of perhaps-clause **Perhaps Paul talked to anybody*, whereas it is in most of the other non-veridical contexts.

2.2 Categorial typle logic analysis

In categorial type logic (CTL) the assembly of linguistic expressions corresponds to functional application and therefore to type-matching. In particular, due to the logic property of the functional connectives a structure of type A/B (or $B\backslash A$) will compose with a structure of type B or of any other type C such that $C \longrightarrow B$. This will be the main property we are going to exploit in our analysis of polarity items.

In this light, Giannakidou' s anlysis of NPIs and APIs interaction with NV contexts can be summarized as below, where $\circ$ is the composition operator, Δ[X] means that X is in the structure Δ, and * marks ungrammatical composition.

$$\begin{array}{ll} \text{AV} \circ \Delta[\text{NPI}] & {}^*\text{NV} \circ \Delta[\text{NPI}], \\ \text{AV} \circ \Delta[\text{API}] & \text{NV} \circ \Delta[\text{API}], \\ {}^*\text{V} \circ \Delta[\text{NPI}] & {}^*\text{V} \circ \Delta[\text{API}]. \end{array}$$

Since the polarity must be in the immediate scope of its licensor, we can consider Δ[NPI] to be of type *npi*, viz. the NPI has scope over the whole structure and determines its type, and similarly Δ[API] is of type *api*. Combining this and the observation about functional composition it follows that (i) the type assigned

[1]Besides the licensing vs. anti-licensing condition, Giannakidou distinguishes some items which are licensed indirectly due to the negative implicature given rise by the sentence in which they occur. However, this third case is not relevant for our investigation which mainly concerns the relation between the types of an overt licensor and the sensitive item.

by APIs to Δ[API] derives the one assigned by NPIs to Δ[NPI], (ii) AVs have *npi* as argument, whereas NVs have *api*, and (iii) neither *npi* nor *api* derives the argument type of veridical expressions. Summing up,

$$\text{AV} \in A/npi \quad \text{NV} \in A/api, \quad \text{V} \in A/ppi$$

$$api \longrightarrow npi \qquad npi \not\longrightarrow ppi \qquad api \not\longrightarrow ppi.$$

We leave a general formula A on the value of the licesors'type, since it is not relevant for the understanding of the main idea. Notice that the inclusion relation AV $\subseteq$ NV at the heart of the linguistic theory, follows as a consequence from these types, viz. $A/npi \longrightarrow A/api$. Let's make things more concrete by means of an example. In Table 1 we have seen that *yesterday*, *usually* and *it is not the case* are all denoted in the domain $D_t^{D_t}$, hence their (syntatic) category is s/s. In order to account for PIs distribution we need to differentiate these types as explained. In Section 1 we have seen that unary operators give us the right expressivity to account for such distinction. For example, we can consider *npi*, *api* and *ppi* as ${}^0(\Diamond\Box^{\downarrow}s))^0$, $({}^0s)^0$ and $\Box^{\downarrow}\Diamond s$, respectevely. Consequentely, *it is not the case* $\in s/({}^0s)^0$, *usually* $\in s/({}^0(\Diamond\Box^{\downarrow}s))^0$ and *yesterday* $\in s/\Box^{\downarrow}\Diamond s$. In this way, the derivability relations stipulated above follow from the logical properties of the unary operators. These types will correctly block a structure contaning a NPI or API (having wide scope) to compose with the verdical expression *yesterday* (see 1(a)-(b) below), and will predict the difference between NPIs and APIs with respect to NV contexts, like *usually* (2(a) vs. 2(b)).

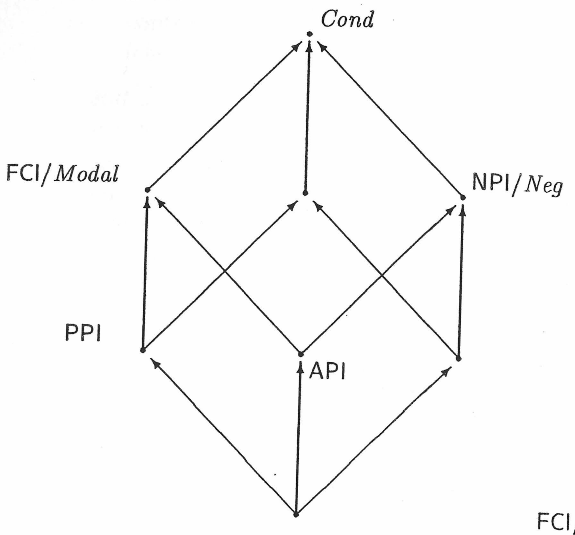
1. (a) *Yesterday I spoke with *anybody* I met.
   (b) *Yesterday I said *a word.*
2. (a) Usually I speak with *anybody* I meet.
   (b) *Usually I say *a word.*

The cube of derivability relations given above offers a much richer hierarchy of types which will allow us to make more fine-grained distinctions among polarity items as actually is required by the linguistic data. In the remainder of this section we will present the results given by the application of this CTL analysis to PIs in Greek and Italian showing how it helps carry out cross-linguistic comparisons. Again, for reason of space we summarize (part of) the data in the tables below and the needed types in the two cubes to be compared with the previous one. The items we have considered are (i) NPI:*ipe leksi*, API: *kanenan*, free choice item (FCI): *opjondhipote*, for Greek; and (ii) NPI: *nessuno*, API: *mai*, FCI: *chiunque*, for Italian.

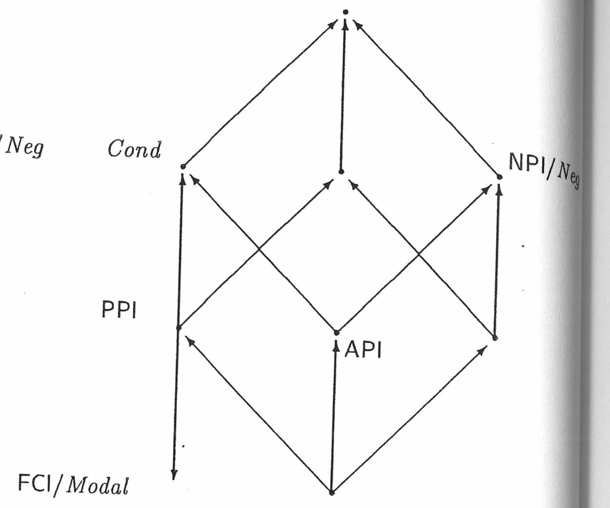
Greek	FCI	API	NPI
Veridical	*	*	*
Negation	*	Yes	Yes
Modal verb	Yes	Yes	*
Conditional	Yes	Yes	Yes

Italian	FCI	API	NPI
Veridical	*	*	*
Negation	*	Yes	Yes
Modal verb	Yes	*	*
Conditional	*	Yes	*

From this comparison the following conclusions can be drawn: (i) both in Greek and Italian, there could be contexts where both PPI and API are felicitous (e.g. conditionals); (ii) in Italian there could be other sorts of polarity items sensitive to non-veridical contexts like modal verb, and ungrammatical in the scope of conditional and negation. Note that the last point shows how the difference between the two levels of non-veridicality given by $({}^0\cdot)^0$ and ${}^0(\cdot^0)$ allows us to account for anti-licensing and licensing conditions. The small framgent of items we have considered is intended to give an example of how the type logic perspective helps clarify the consequences predicted by the typologies and open the way to further investigations.



Greek



Italian

We conclude by pointing out another potential advantage of CTL analysis which could be explored in further research, namely it helps establish dependency between linguistic phenomena and linguistic theories. In particular, our search for the right lexical type assignments suggests a possible connection between non-verdicality and $({}^{0}\cdot, \cdot^{0})$, and veridicality and $(\Box^{\downarrow}, \Diamond)$. Due to the close relation between (non-)veridical contexts and the (anti-)licensing of discourse anaphora the proposed account sheds light on a connections between the operators $\uparrow, \downarrow$ introduced in dynamic Montague grammar (DMG) and the unary operators of $\mathsf{NL}(\Diamond, \cdot^{0})$. If an expression is in the scope of ${}^{0}(\cdot^{0})$ (and $({}^{0}\cdot)^{0}$) it is closed; if it is in the scope of $\Box^{\downarrow}\Diamond\cdot$ anaphoric links are allowed. Which translated into DMG terms means that $\Box^{\downarrow}\Diamond$ corresponds to $\uparrow$ where $\uparrow \phi =_{def} \lambda p.(\phi \wedge {}^{\vee}p)$ and ${}^{0}(\cdot^{0})$ (and $({}^{0}\cdot)^{0}$) to $\downarrow$ where $\downarrow \psi =_{def} \psi({}^{\wedge}\mathsf{true})$.

References

[1] F. Beghelli and T. Stowell. Distributivity and negation: The syntax of each and every. In A. Szabolcsi, editor, *Ways of Scope Taking*, pages 72–107. Kluwer, 1997.

[2] R. Bernardi C. Areces and M. Moortgat. Galois connections in categorial type logic. In R. Oherle and L. Moss, editors, *Electronic notes in theoretical Computer Science. Proceedings of FGMOL'01*, volume 47. Elsevier Science B.V., 2001.

[3] A. Giannikidou. *The landscape of polarity items.* PhD thesis, Groningen, 1997.

[4] W. Ladusaw. *Polarity sensitivity as inherent scope relations.* PhD thesis, University of Texas at Austin, 1979.

[5] Ø. Nilsen. *Adverb order in Type Logical Grammar.* In this volume.

[6] A. Szabolcsi and F. Zwarts. Weak islands and a algebraic semantics for scope taking. In A. Szabolcsi, editor, *Ways of Scope Taking*, pages 217–262. Kluwer, 1997.

[7] T. van der Wouden. *Negative Contexts.* PhD thesis, University of Groningen, 1994.

[8] F. Zwart. Nonveridical contexts. *Linguistic Analysis*, 25:286–312, 1995.

Quantified Hybrid Logic and Natural Language

Patrick Blackburn — INRIA Lorraine

Maarten Marx — University of Amsterdam

Introduction Quantified Hybrid Logic (*QHL*) is an extension of first-order modal logic in which the modal component is augmented with names for states, binders over states, and a mechanism for asserting that a proposition holds at a particular state (by 'state' we mean 'world' or 'time' or 'location' or whatever it is that the elements of the Kripke model are meant to be). Recent technical work has shown that the added machinery of state reference solves most of the notorious problems (such as incompleteness and interpolation failure) that have haunted first-order modal logic since its conception, and solves them in a general way (see [2]). Moreover, no matter what assumptions on the first-order domains are made, complete analytic tableaux procedures exist that can be used to compute interpolants (see [5]).

Such positive results are a good indication that *QHL* is well-designed, and in particular that there is a good fit between its syntax and semantics. In this paper we argue that *QHL* is also well-motivated from the perspective of natural language semantics. We do so by looking at the interaction of tense, temporal reference, and quantification.

The work presented here is a stepping stone on the way to our ultimate goal: the creation of a higher-order hybrid language, satisfying certain technical criteria, that can handle a significant portion of natural language semantics. We briefly discuss this goal at the end of the paper.

Quantified Hybrid logic To make a language of *QHL*, take a first-order modal logic, and add a set *NOM* of new propositional variables called *nominals*, and also modal operators $@_i$ and binders $\downarrow i$ for every nominal i. In this paper we assume we have carried out this process (called *hybridization*) to the language of first-order tense logic, thus the modalities at our disposal are F (sometime in the future), and P (sometime in the past). Models for the hybridized language are simply the (first-order) Kripke models for the underlying tense logic, with the crucial proviso that each nominal is true at exactly one time. Thus nominals, although they are proposition symbols and not terms, are a referential mechanism: they 'name' the unique time they are true at. This increases the expressive power at our disposal. For example, $F(i \wedge \varphi) \wedge F(i \wedge \psi) \rightarrow F(\varphi \wedge \psi)$ is valid precisely because i can only be true at one time (this schema would not be valid if we replaced i by an arbitrary propositional variable p). The modalities of the form $@_i$ are *holds* or *satisfaction* operators, and they have a very natural interpretation:

$@_i\varphi$ says that at the time named i, φ holds. As for the downarrow binder, in effect this lets us create names for the current time on the fly: $\downarrow i$ binds all occurrences of i in its scope and insists that they refer to the current time. That is, $\downarrow i.\varphi$ is true at a time t iff φ is true at t when all occurrences of i in φ bound by $\downarrow i$ are regarded as true at t.

Nominals, satisfaction operators, and downarrow are standard tools in contemporary hybrid logic (for more information on hybrid logic, see the hybrid logic homepage at `http://www.hylo.net`, the hybrid logic 'manifesto' [4], or Chapter 7 of [6]). In what follows, however, we shall also need a mechanism to handle non-rigid designators. Assume we have a set of first-order constant symbols denoting individual concepts. Then if i is a nominal, and q an individual concept, then $@_i q$ is a first-order *term* that picks out the denotation of q at the world named i.

Prior's legacy Hybrid logic dates back to work of Arthur Prior [8] in the mid 1960s, and played a fundamental role in his work. It is possible to think of time in terms of the A-series of past, present and future (to do so is to take a 'situated', 'internal' or 'deictic' view of time) or in terms of the B-series of earlier and later (to do so is to take 'external' or 'classical' view of time). Prior believed the A-series to be fundamental: we live *in* time, and the internal perspective is imprinted on natural language in various ways.

Prior designed two logical calculi for talking about time. The T-calculus (tense logic) was designed to mirror A-series talk. The U-calculus (essentially the first-order language of earlier-than and later-that) was designed to model B-series talk. But then Prior faced a problem. He viewed A-series talk as fundamental, thus he believed the T-calculus should be able to express anything expressible in the U-calculus. Unfortunately the reverse was obviously true: T-calculus was far weaker than U-calculus.

Hybridization was Priors solution. His key insight was that *propositions could be used as terms*, and this led him to nominals (or as he called them, 'world propositions'). He also added two other ingredients: binders of the form $\forall i$, for every nominal i, together with what is now called the global modality A ($A\varphi$ means φ is true at all times in the model). This apparatus, as Prior showed, enabled the T-calculus to 'swallow' the U-calculus. Hybridization thus saved Prior's philosophical program.

Contemporary hybrid logic is motivated by rather different concerns. There is less interest in showing that modal approaches can do everything first-order approaches can do, rather the emphasis is on designing languages that can do enough, and no more. Thus the use of the $\forall i$ tends to be regarded as overkill; the more restrained $\downarrow i$, which only binds to the current

state, is considered more interesting (see [1] for some reasons why). Similarly, satisfaction operators can be considered to be more restricted forms of the global modality; a little thought shows that $A(i \rightarrow \varphi)$, which is the way Prior typically used the global modality, means $@_i\varphi$. But while the tools used in *QHL* are weaker than those Prior used, they enable us to deal with a surprisingly wide range of semantic phenomena, as we shall now see.

Using nominals and satisfaction operators Given that he invented the nominal, it is ironic that Prior never fully appreciated the fundamental role temporal reference plays in natural language. For Prior, nominals were essentially a tool that enabled him to carry out a philosophical program. But they also offer something more down-to-earth: a unified approach to the semantics of tense and temporal reference.

Temporal reference is central to the semantics of tense. For example, the simple past tense sentence "Vincent squeezed the trigger" means that at some particular, contextually determined, past time, Vincent squeezed the trigger. $P(\textit{vincent-squeeze-the-trigger})$, its Priorean representation, does not capture this contextual determination (it simply locates the trigger squeezing sometime in the past). But with the aid of a nominal, localization to a particular past time is easily achieved:

$$P(i \wedge \textit{vincent-squeeze-the-trigger}).$$

This can be glossed as: "in the past is the time named i, and at that time *vincent-squeeze-the-trigger* is true". Similarly, the past perfect sentence "Vincent had squeezed the trigger" can be represented by:

$$P(i \wedge P(\textit{vincent-squeeze-the-trigger})).$$

The link with Reichenbach's ideas should be obvious: in essence, the nominal i marks the point of reference. Indeed, with the aid of nominals, it is easy to write down schemas capturing all of Reichenbach's analyses, and doing so overcomes the chief weakness of Reichenbach's diagrammatic representations, namely their overspecificity (see [3] for further details). Priorean and Reichenbachian approaches to tense are *not* mutually antagonistic as Prior himself (and most commentators since) have supposed. Prior's key innovation, the nominal, enables them to be brought together in a mutually beneficial way.

Temporal reference is also crucial to the way tense functions in text. Consider the text "Vincent woke up. Something felt very wrong. Vincent reached under his pillow for his Uzi." Note the contrast between the temporal locations of the second and third sentences: the second is to be true at the same time at the first (because it is a state sentence), whereas the third sentence is to be true sometime after (because it is an event sentence).

Satisfaction operators give us a natural way of handling this. (The analysis sketched here improves on the one given in [3], which does not make use of satisfaction operators.) The key observation is this: satisfaction operators let us talk about the relationship between different points of time. For example, $@_i j$ means that i and j name the same time, $@_i Fj$ means that the point j lies to the future of point i, and $@_i Pj$ means that the point j lies to the past of point i. So we can represent our little text as follows:

$$\begin{array}{lll} & P(i \wedge \mathit{vincent\text{-}wake\text{-}up}) & \\ \wedge & P(j \wedge \mathit{something\text{-}feel\text{-}very\text{-}wrong}) & \wedge\ @_j i \\ \wedge & P(k \wedge \mathit{vincent\text{-}reach\text{-}under\text{-}pillow\text{-}for\text{-}uzi}) & \wedge\ @_k Pj \end{array}$$

This analysis has a natural computational interpretation, one that factors the (mechanical) computation of the core representation from the temporal reference resolution step (which requires inference). The idea is this. When the semantic representation of the first sentence is built, a previously unused nominal (here i) is introduced to represent its temporal location. The same process occurs when the second sentence is analysed (here the introduced nominal is j), but something additional happens: a second conjunct, $@_j i$, is introduced. This is added by the temporal inference component, which we assume contains some theory of tense in text. (For illustrative purposes, we have here assumed a rather simplistic theory: state verbs identify reference times, while event verbs advance them.) Thus the identity of the times named by i and j follows, and the conjunct $@_i j$ states it in the object language. Much the same thing occurs in the third sentence: a representation is built and then the inference component (because the main verb of the third sentence is the event verb "reached") 'annotates' the representation with the additional information $@_k Pj$, which advances the reference time as required.

Using downarrow Downarrow offers us even more power (though far less power than Prior, with his $\forall i$ binders allowed himself). Because it lets us create names for times on the fly, names which can later be used by the satisfaction operators, downarrow is a natural tool for controlling the interaction between tense, temporal reference, and quantification, even in the presence of non-rigid designators.

Let q denote "the queen of Holland" and *Tall* be the one place predicate meaning that someone is tall. Then the phrase "The present queen of Holland is tall" can be represented as $\downarrow s.\mathit{Tall}(@_s q)$, which can be paraphrased as "at this moment, call it s, the queen of Holland at time s is tall." (Note how,

just like $\downarrow s.$, the phrase *call it s* does not add to the meaning of the sentence. It is not even a declarative statement, but an imperative one relating to the structuring of the ongoing discourse.)

But now consider sentence "The queen of Holland will be tall." With our machinery we can distinguish a number of different readings:

$$\downarrow s.F\,\mathit{Tall}(@_s q) \tag{1}$$

$$F\downarrow t.\,\mathit{Tall}(@_t q) \tag{2}$$

$$\downarrow s.F\downarrow t.@_s\,\mathit{Tall}(@_t q) \tag{3}$$

$$\downarrow sF\downarrow t@_s\,\mathit{Tall}(@_s q) \tag{4}$$

Reading (1) says that some day in the future, the present queen will be tall, according to the future notion of tallness. The reading with temporal wide scope is (2). This means that the *then* present queen will be tall. In the third reading, however, we have something more subtle: (3) might be paraphrased as "Some day, the then present queen will be tall according to the present notion of tallness". In a situation in which the present queen consults a fortune-teller about the crown princess' future body length, (3) seems the intended meaning of the answer. Finally (4) is, as the reader may verify, just a complicated way of just saying $\downarrow s.\mathit{Tall}(@_s q)$.

Concluding remarks The advantage hybrid logics have over orthodox modal logics is that they enable states to be named and formulas to be evaluated at named states. For many semantic phenomena this ability is crucial, but our discussion has only scratched the tip of the iceberg. For a start, hybridization can be used in conjunction with richer temporal logics, such as interval based or event based logics, and this enables more interesting phenomena to be tackled. Moreover, hybridization is relevant to non-temporal semantic phenomena. For example the binding approach to the semantics of "actually" advocated in [9] can be captured with the aid of $\downarrow$.

Our goal is to develop a higher-order hybrid logic capable of handling a significant portion of natural language semantics. We believe that hybridization may make Montague-style type theories (by which we mean type theories which offer both classical quantification and the use of modalities) better behaved technically, and more relevant to natural language semantics.

There are problems with Montague-style type theories. Muskens's [7] criticizes them on technical and conceptual grounds, and advocates dispensing with modal syntax in favour of classical quantification over worlds and times. His criticisms are sound, but we believe that hybridization may provide a cure. First, hybridization should enable us to overcome the technical shortcomings that Muskens cites (such as failure of the Church-Rosser property). Moreover, it should be possible to develop usable inference systems for such logics. Finally (as we hope this paper his shown) we believe that the representations hybridization makes possible are simple and natural. And in terrain as complex as natural language semantics, simplicity and naturalness are not optional extras, they are essential.

References

[1] C. Areces, P. Blackburn, and M. Marx. Hybrid logics. Characterization, interpolation and complexity. *Journal of Symbolic Logic*, 66(3): 977-1010, 2001.

[2] C. Areces, P. Blackburn, and M. Marx. Repairing the interpolation theorem in quantified modal logic. To appear in *Annals of Pure and Applied Logic*. Available at `www.hylo.net`.

[3] P. Blackburn. Tense, temporal reference and tense logic. *Journal of Semantics*, 11: 83-101, 1994.

[4] P. Blackburn. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL*, 8(3): 339-365, 2000.

[5] P. Blackburn and M Marx. Constructive interpolants for every bounded fragment definable modal logic. Submitted for journal publication.

[6] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge UK, 2001.

[7] R. Muskens. *Meaning and Partiality*. CSLI Publications, Stanford, 1995.

[8] A.N. Prior. *Past, Present and Future*. Oxford University Press, 1967.

[9] Y. Stephanou. Indexed actuality. *Journal of Philosophical Logic*, 30(4): 355-393, 2001.

Duality and Anaphora

António Branco
University of Lisbon

Abstract *The four constraints on sentential anaphoric binding, known as binding principles, are observed to form a square of oppositions. With the formal tools of phase quantification, these constraints are analysed as the effect of phase quantifiers over reference markers in grammatical obliqueness hierarchies. The four quantifiers are shown to be organized in a square of duality. The impact of this result on the distinction quantificational vs. non quantificational NPs and on the semantics of nominals in general is discussed.*

1 Quantification and Duality

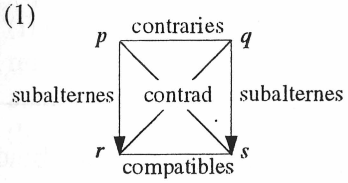
Logical duality has been a key issue in natural language semantics. It is a pattern noticed in many phenomena, ranging from the realm of determiners to the realm of temporality and modality, including topics such as the adverbials *still/already* or the conjunctions *because/although*, etc. ([6], [7], [5], [9] *i.a.*).

While noting that the ubiquity of the square of duality may be the sign of a semantic universal, [1],p.23 highlighted its heuristic value for research on quantification inasmuch as "it suggests a systematic point of view from which to search for comparative facts"— a hint we explore in this paper.

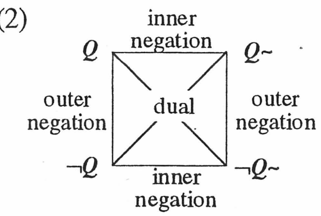
2 Anaphoric Binding Constraints

Given our purpose here, it is of note that the square of duality in (2) is different from the classical square of oppositions in (1).

(1)



(2)



The difference lies in the fact that *duality*, *inner negation* and *outer negation* are third order concepts, while *compatibility*, *contrariness* and *implication* are second order concepts. There are instantiations of the square of oppositions without corresponding squares of duality, and vice-versa ([6],p.56 for discussion).

Although the two squares are logically independent, the empirical emergence of a square of oppositions naturally raises the question about the possible existence of an associated square of duality. This is where we get focussed into our research topic, given the emergence of a square of oppositions with the four constraints on sentential anaphoric binding, also known as binding principles.

Binding constraints capture empirical generalizations concerning the relative positioning of anaphors with respect to their antecedents in the grammatical geometry of sentences. We follow here the definition proposed in [8] for these constraints, and subsequent extension in [10], [2]:

Principle A: A locally o-commanded short-distance reflexive must be locally o-bound.

Lee$_i$ thinks [Max$_j$ saw himself$_{*i/j}$].

Principle Z: An o-commanded long-distance reflexive must be o-bound.

Zhangsan$_i$ cong Lisi$_j$ chu tingshuo [Wangwu$_k$ bu xihuan ziji$_{i/*j/k}$]. [10]:ex(2)

Zhangsan$_i$ heard from Lisi$_j$ [Wangwu$_k$ doesn't like "himself"$_{i/*j/k}$].

Principle B: A pronoun must be locally o-free.

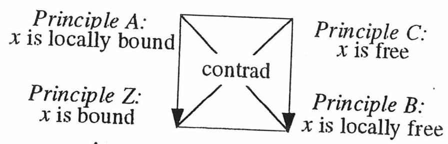
Lee$_i$ thinks [Max$_j$ saw him$_{i/*j}$].

Principle C: A non-pronoun must be o-free.

[Kim$_i$'s friend]$_j$ thinks [Lee saw Kim$_{i/*j}$].

X *o-binds* Y iff X o-commands Y and X is the antecedent of Y. *O-commands* is a partial order under which, in a clause, the Subject o-commands the Direct Object, the Direct Object o-commands the Indirect Object, and so on, following the obliqueness hierarchy of grammatical functions; in multiclausal sentences, the upward arguments o-command the embedded arguments, etc. [8],p.279. The *local domain* is, roughly, the subcategorization domain of the predicator selecting the anaphor (details in [3]).

When stripped away from procedural phrasing and non-exemption requirements, these generalizations instantiate the following square of oppositions ([2] for detailed discussion):

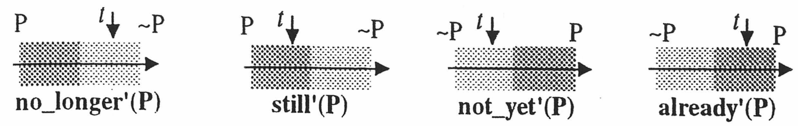


Given this square, the question to pursue is whether this is a sign that binding principles are the effect of some underlying quantificational structure, i.e. whether there is a square of duality associated with the constraints on anaphoric binding.

3 Phase Quantification

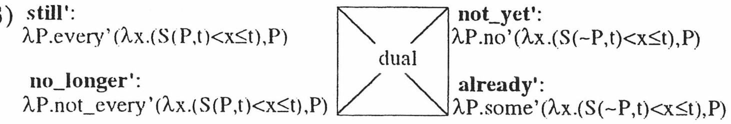
We argue that the answer to this question is affirmative. Before this result may be worked out, some analytical tools are to be introduced first.

We resort to the notion of phase quantification, introduced in [6] to study the semantics of aspectual adverbials and shown to be extended to characterize quantification in general [6],p.74. For the sake of concreteness, consider a diagrammatic display of the semantics of such adverbials:



Very briefly, phase quantification requires the following ingredients: (i) an order over the domain of quantification; (ii) a parameter point t; (iii) a property P defining a positive phase in a sequence of two opposite phases; and (iv) the starting point of the relevant semiphase given the presupposition about the linear order between P and $\sim P$.

For aspectual adverbials, (i) the order is the time axis; (ii) t is the reference time of the utterance; (iii) P denotes the instants where the proposition containing the adverbial holds; (iv) the starting point $S(R,t)$ is the infimum of the set of the closest predecessors of t which form an uninterrupted sequence in R — e.g. the adverbials *no longer* and *still* bear the presupposition that phase P precedes $\sim P$. These adverbials express the following quantifiers:

(3)


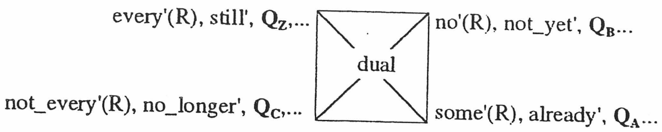
4 Quantificational Anaphors

With this in place, the empirical generalizations captured in the binding principles can be argued to be the visible effect of the phase quantificational nature of the corresponding nominals: Below, anaphors are shown to express one of four quantifiers acting on the grammatical obliqueness order.

Phase quantification here is assumed to unfold over entities in grammatical representations, vz. reference markers [4], and its ingredients are as follows: (i) *Order:* reference markers are ordered according to the o-command relation; (ii) *Parameter point:* t is a here, the marker of the antecedent for the anaphoric nominal at stake; (iii) *Phase property:* P is D here, which denotes the set of markers in the grammatical domain of the anaphor: For a nominal anaphor N, D is determined by the relative position of N in the obliqueness order which N enters. Given m, the reference marker of N, semiphase D_m is a stretch containing m and elements that are less than m in the obliqueness order, i.e. markers of o-commanders of N; if semiphase D_m is presupposed to precede $\sim D_m$, D_m is such that the last successor in it is local wrt to m; and if $\sim D_m$ precedes D_m, the first predecessor in D_m is local wrt to m, however locality for binding may be parameterised in each language [3]. In both cases the closest D_m neighbour of semiphase $\sim D_m$ is local wrt m: $D_m(x)$ iff

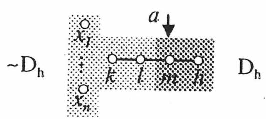
$x \leq r \wedge \forall y[(\sim D_m(y) \wedge (x \prec y \vee y \prec x)) \rightarrow (x \text{ is local wrt } m \wedge y \text{ is not local wrt } m)]$

With this replacements in (3), one gets four phase quantifiers — we termed $\mathbf{Q_Z}$, $\mathbf{Q_B}$, $\mathbf{Q_C}$ and $\mathbf{Q_A}$ — entering the square of duality and aligning with other quantifiers of similar quantificational force at each of the corners:



These four phase quantifiers ensure the same empirical predictions as secured by the four binding principles, as we can briefly check out below.

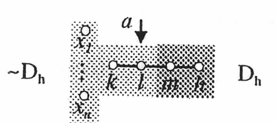
A: The quantifier expressed by short-distance reflexives is associated with the presupposition that $\sim D.D$. It receives the following definition, which is easily interpreted against the diagram corresponding to the example sentence, *Kim said Lee thinks Max$_i$ hit himself$_i$* —k, l, m and h are resp. the markers of *Kim*, *Lee*, *Max* and *himself*, and $x_1,...,x_n$ are markers not in the obliqueness relation of h, possibly introduced in other sentences of the discourse or available in the context (Hasse diagrams displayed with a turn of 90° right):



Q_A:
λP.some'(λx.(S(~P,a)<x≤a,P)

$Q_A(D_h)$ is satisfied iff between the bottom of the uninterrupted sequence $\sim D_h$ most close to the antecedent a and a inclusive, there is at least one reference marker in D_h. As $\sim D_h$ precedes D_h, this amounts to requiring that a be in D_h, the local domain of h here, and consequently that a be a local o-commander of h, which matches the requirement in Principle A. Binding phase quantifier $\mathbf{Q_A}$ shows positive existential force and short-distance reflexives align in the square of duality with items like *some N*, *already*, *possibly*, etc.

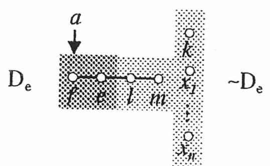
B: The phase quantifier expressed by pronouns, in turn, lies at the same corner as the quantifiers no'(R) or not_yet' in (3). The presupposition conveyed by these anaphors is also that $\sim D.D$, and $\mathbf{Q_B}$ is easily understood when considering the diagrammatic description of an example like *Kim said Lee$_i$ thinks Max hit him$_i$*:



Q_B:
λP.no'(λx.(S(~P,a)<x≤a,P)

$Q_B(D)$ is satisfied iff no reference marker between the bottom of $\sim D$ and the antecedent a inclusive is in D, which implies that a has to be in $\sim D$, i.e. it has to be outside the local domain of the pronoun, as required in Principle B.

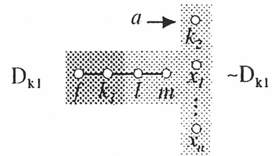
Z: Turning to long-distance reflexives, we consider an example from Portuguese (*[O amigo de Kim]$_i$ disse que [ele próprio]$_i$ acha que Lee viu Max.*) *[Kim's friend]$_i$ said "ele próprio"$_i$ thinks Lee saw Max*:



Q_Z:
λP.every'(λx.(S(P,a)<x≤a,P)

As with short-distance reflexives, a is here required to occur in D_e though the presupposition conveyed now is that semiphase D is followed by semiphase $\sim D$. Taking into account the definition of D_e above, the antecedent is required to be an o-commander (local or not) of e. The semantics of the phase quantifier $\mathbf{Q_Z}$ is such that, for $Q_Z(D_e)$ to be satisfied, between the bottom of the uninterrupted sequence D_e closest to the antecedent a and a inclusive, every reference marker is in D_e. This amounts to requiring a to be in D_e, i.e. to requiring it to be an o-commander of e, as predicted by Principle Z.

C: While long-distance reflexives show positive universal force, the quantifier expressed by non-pronouns appears at the same corner as quantifiers like not_every'(R), no_longer', etc. Let us consider a first version of the diagram of *Kim(2)'s friend said Kim(1)$_i$ thinks Lee saw Max*:

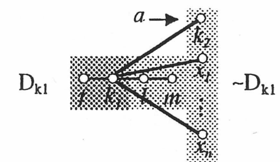


Q_C:
λP.not_every'(λx.(S(P,a)<x≤a,P)

The antecedent a should be required to occur in $\sim D_{k1}$, which means that a cannot be an o-commander of k_1: This renders the same constraint as expressed by principle C, that non-pronouns are free. As in previous diagrams, $\sim D$ is taken as the complement set of D. Correct empirical prediction requires this to be refined and a more accurate definition of $\sim D$ to be given for phase quantification in non-linear orders — as the one under consideration — where not all elements are comparable.

For $Q_C(D_{k1})$ to be satisfied, between the bottom of D_{k1} and the antecedent a inclusive, not every reference marker is in D_{k1}. In examples as the one above, λx.(S(D_{k1},a)<x≤a), the restrictor of $\mathbf{Q_C}$, is always empty: It is not the case that S(D_{k1},a)≤a because a=k_2 (or a=x_i for any i) is not comparable to any element of D_{k1}, including its bottom. Hence, not_every'(λx.(S(D_{k1},a)<x≤a), D_{k1}) is false whatever reference marker k_2 or x_i is taken as the antecedent for k_1. The specific anaphor resolution in our example would be incorrectly ruled out.

This suggests that when phase quantification operates on non-linear orders, negation of semiphase P may be slightly more sophisticated than simple Boolean negation rendering its complement set. We are taught that negation of P also involves the lifting of the complement set, $\overline{P}\bot$, with $\bot$ equal to the top of P when P.~P (k_1 in our example). We can check that this specification of ~P makes it possible to satisfy $Q_C(D_{k1})$ in the correct anaphoric links for non pronouns:.[1]



[1] For the sake of formal uniformity, when ~P.P, the order-theoretic dual of this definition for ~P can also be assumed.

5 The Semantics of Nominals

These results may shed new light over a number of interesting issues. For instance, given their parameterised validity across natural languages [3], the universal character of binding principles has been seen as a striking feature: When envisaged as a set (so-called binding theory), they appear as one of the best candidates to be a module of universal grammar. Given the universality of quantification, if binding principles are the noticeable effect of quantifiers, it is not surprising that they are universally operative across natural languages.

Second, not all languages have anaphors of each of the four binding types. In English, there is no long-distance reflexives. This is in line with the well known fact that not every corner of a duality square may be "lexicalised", as Löbner puts it: In some squares, there may not exist a single expression for a given corner, which is then expressed by some other means (e.g. a complex expression, such as *not every*, etc. — [6],p.65 for a fully-fledged discussion).

Finally, it is interesting to notice the inverted analogy between referential and quantificational NPs. Nominals with "genuine" quantificational force (*every man, most students,...*) have a somewhat "secondary" referential force, as revealed in e-type anaphora: Although they introduce markers in the grammatical representation that can be picked as antecedents by anaphors (vd. Σ-abstraction [4]Ch4), they cannot be used to refer to "outside world" entities.

Conversely, NPs with "genuine" referential force (*he, the book, John,...*), we can consider it now, have a somewhat "secondary" quantificational force: They introduce quantificational requirements over grammatical entities, but cannot be used to directly quantify over "outside world" entities.

If the results reported here are meaningful, and taking aside indefinites, every NP, quantificational or referential, has a dual nature by making a contribution in both dimensions of quantification and reference, but with respect to different universes.

References

[1] van Benthem, 1991. "Linguistic Universals in Semantics". In [11], 17-36.
[2] Branco and Marrafa, 1999, "Long-Distance Reflexives and the Binding Square of Opposition", in Webelhuth, Koenig and Kathol, eds., Lexical and Constructional Aspects of Linguistic Explanation. CSLI, Stanford, Ch. 11.
[3] Dalrymple, 1993. *The Syntax of Anaphoric Binding*. CSLI, Stanford.
[4] Kamp and Reyle, 1993. *From Discourse to Logic*. Kluwer, Dordrecht.
[5] König, 1991. "Concessive Relations as the Dual of Causal Relations". In [11], 190-209.
[6] Löbner, 1987, "Quantification as a Major Module of NL Semantics", in Groenendijk, de Jongh and Stokhof, eds., *Studies in DRT and the Theory of Generalized Quantifiers*. Foris.
[7] Meulen, 1988, "The Semantic Properties of English Aspectual Verbs", NELS 21. University of Massachusetts, Amherst.
[8] Pollard, and Sag. 1994. *Head-driven Phrase Structure Grammar*. CSLI, Stanford.
[9] Smessaert, 1997. "Aspectual Duality Regained". In Dekker *et al.*. eds. *Proceedings of the 11th Amsterdam Colloquium*, ILLC, Amsterdam, 271-276.
[10] Xue, Pollard and Sag, 1994, "A New Perspective on Chinese *Ziji*". *Proceeedings of the West Coast Conference on Formal Linguistics*, 13. CSLI, Stanford.
[11] Zäfferer, ed. 1991. *Semantic Universals and Universal Semantics*, Foris, Berlin.

How pronominal are expletive pronouns?

Ronnie Cann, University of Edinburgh

1. Introduction

Following a persistent tendency in linguistic theory to equate differences in syntactic or semantic behaviour with different construction types, expletive uses of pronouns are invariably analysed as unrelated to anaphoric ones, cf. the two instances of *it* in (1) and *there* in (2).

(1) It's likely that it's been sick on the carpet.

(2) There are no buskers there this year.

In this paper, I propose a theory of *it* and *there* which assumes that expletive uses follow directly from a minimal modification of the properties of a full pronoun and that expletive pronouns are identical to their non-expletive counterparts.

The framework of *Dynamic Syntax* (*DS*, Kempson et al 2001) defines the process of natural language understanding as a monotonic tree growth process over the left-right sequence of words, with the goal of establishing some propositional formula as interpretation. Taking information from words, pragmatic processes and general rules, the theory derives partial semantic tree structures that represent the content of a string as interpreted in context up to the current point in the parse. Intrinsic to this process are concepts of underspecification whose resolution is driven by requirements which determine the process of tree growth, having to be satisfied for a parse to be successful.

All nodes in the semantic trees constructed during a parse are introduced with requirements to be fulfilled, folowing the universal initial requirement to build a representation of the propositional content of a string, expressed as $?Ty(t)$. To satisfy such requirements, a parse relies on information from various sources. General processes of construction modify partial trees: e.g. a tree rooted in $?Ty(Y)$ may be expanded to one with argument daughter $?Ty(X)$ and functor daughter $?Ty(X \rightarrow Y)$. Other construction rules introduce nodes that are underspecified with respect to position within the current tree, associated with the requirement to find such a position.[1] This is the basis for analysing long distance dependencies, analysed in terms of unfixed nodes whose position in the emergent tree structure is fixed at some later stage in the parsing process. Additional information comes from actions induced by parsing words in sequence, which provide formula values of

[1] This requirement is of the form $?\exists x Tn(x)$, where Tn is the treenode label which provides the address of some node in terms of its relative position with respect to the topnode of the tree.

appropriate types and may build and/or annotate other nodes with labels or requirements.

Interacting with these tree growth processes is the context-dependent processing of anaphoric expressions. This phenomenon is modelled in *DS* as the projection of a metavariable, $Fo(\mathbf{U})$, with an associated requirement to find a fixed value, $?\exists\mathbf{x}.Fo(\mathbf{x})$, to be provided by some term through the process of SUBSTITUTION. This process is pragmatic and system-external, but may be constrained by additional restrictions (e.g. gender on pronouns) that lead to additional presuppositions derivable from the containing propositional formula (e.g. that something is neither male nor female). Thus, the pronoun in *It is possible* projects a metavariable which must be substituted by a formula of type t provided by the context (e.g. $Happy(Ronnie)$) to yield a complete propositional formula (e.g. $Possible(Happy(Ronnie))$), representing the content of this string in context.

2. *It* extraposition

All tree-theoretic characterisations of the structure of natural languages assume that contentive expressions provide labels for the terminal nodes of the trees they inhabit, reflecting compositionality of meaning. The *DS* analogue of this is provided by the so-called 'bottom restriction', resulting from the parse of a contentive item. Represented as the modality $[\downarrow]\bot$, the effect of this restriction is to prevent further tree growth below the current point, because no node may lack any annotation.[2]

One of the differences between functional and contentive expressions may, however, be ascribed to the absence of this restriction. In English, most pronouns are like fully contentive expressions in being associated with the bottom restriction. Amongst other things, this prevents unfixed trees (the *DS* analogue of movement) from merging with a node annotated by a pronoun, so accounting for the ungrammaticality of, e.g. **Which students did John not like them?*. If an expression does not impose the bottom restriction, it will allow further development of that node. This property is characteristic of an expletive and so if a pronoun loses the bottom restriction, it may act as an expletive by allowing substitution not just by some *formula* but by some structural representation provided (e.g.) by an unfixed node, as in:

(3) It is possible that I am wrong.

This example may be analysed as involving the normal annotation by *it* of

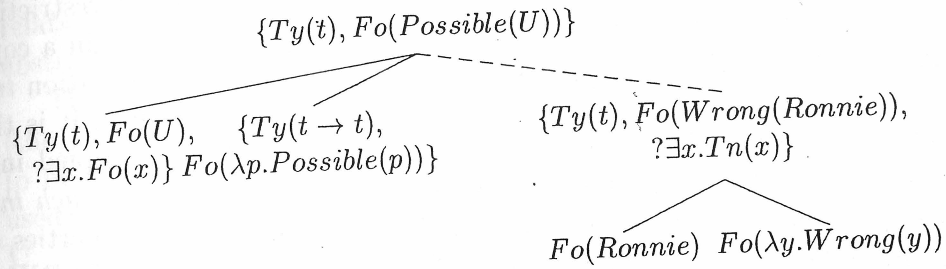
[2]The formula exemplifies the Logic of Finite Trees (LOFT, Blackburn and Meyer-Viol 1994) which forms the backbone of *DS* by defining relations between nodes in trees using modal operators such as: $\langle\downarrow\rangle$ the general daughter relation; $\langle\downarrow_0\rangle$ and $\langle\downarrow_1\rangle$ the argument and functor daughter relations, respectively; $\langle\downarrow_*\rangle$ the dominance relation (the reflexive, transitive closure of the daughter relation); and the inverses of these using the mother relation, $\uparrow$. $\langle.\rangle$ is the existential modality of the system while $[.]$ is the universal.

a (propositional) node in subject position with a metavariable and its associated formula requirement. If SUBSTITUTION does not apply at this point, the predicate is then parsed and the resulting propositional tree is completed to yield an incomplete formula value: $Fo(Possible(\mathbf{U}))$. At this point, a construction rule may be invoked which provides an unfixed node:[3]

*Final*Adjunction*

$$\frac{\{\{Tn(a),\ldots Ty(t),\Diamond\}\}}{\{\{Tn(a),\ldots,Ty(t)\},\quad \{\langle\uparrow_*\rangle Tn(a),\ldots,?Ty(X),\Diamond\}\}}$$

In parsing (3), an application of Final*Adjunction permits the construction of an unfixed node of $Ty(t)$ that allows the the string final clause to be analysed. This unfixed tree carries a requirement that a fixed position is to be found within the larger tree and so must MERGE with some node in this structure, as illustrated in Figure 1:



The only node with which the final unfixed node can merge successfully is that decorated by the metavariable, yielding a final formula value for (3) as $Possible(Wrong(Ronnie))$.

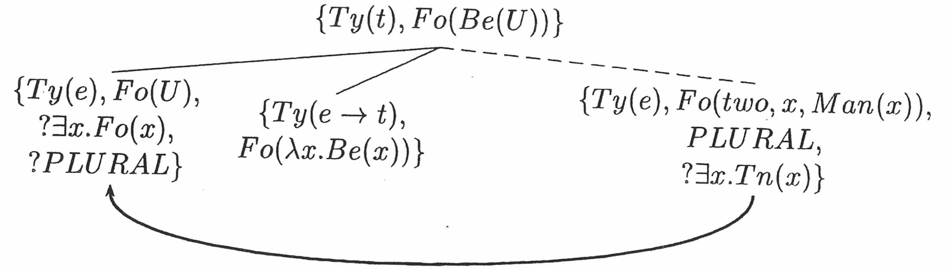
Apart from the loss of the bottom restriction nothing distinguishes this use of *it* from others and the pragmatic process of SUBSTITUTION could have applied at the point at which the pronoun is parsed, e.g. instantiating $\mathbf{U}$ as $Happy(Ronnie)$. In such a circumstance, however, the tree analysing the string final clause would have no resolution to its requirement $?\exists\mathbf{x}.Tn(\mathbf{x})$, precluding all such derivations. Hence, the 'expletive' characterisation of this construction.

3. *There* Constructions

There constructions can be analysed in the same way. *There* projects a metavariable like any other pronoun, but, like *it*, without a bottom restriction, thus permitting merger by a later unfixed node. The analysis of a sentence like *There are two men* thus involves the projection of a metavariable by the subject pronoun and completion of the main predicate provided

[3]See Cann et al. 2001 for a justification of this rule.

by parsing *are*, which I take to be the *be* of existence, i.e. a one-place predicate (see also McNally 1998). An application of Final*Adjunction licenses the parsing of the ASSOCIATE, *two men*. This unfixed tree must merge with the node decorated by the metavariable to satisfy the requirements associated with both nodes. Figure 2 shows that all oustanding requirements are (only) satisfied by the merger of the unfixed node with the subject.[4]



The theory allows a predicate to follow the associate, as in *There are two men sitting on a wall*, with *sitting on a wall* analysable either as a restrictive relative where the structure projected from it is required to contain a copy of the variable introduced by the nominal *men*, giving an interpretation *two men who are sitting on a wall exist*; or as a correlative, in which it is the epsilon term projected from *two men* which is required to be copied into the associated structure, yielding an interpretation *two men exist such that they are sitting on a wall*. On either analysis, the structural properties of the postposed modifying phrases emerge from general properties of LINKed structures (Kempson et. al. 2001 ch 4).

The fact that expletive *there* may be analysed as a pronoun without a bottom restriction does not, however, show that the locative and expletive uses of the expression are closely related. Under the assumption espoused here, that there is no essential difference between expletive pronouns and their non-expletive counterparts and so only a single representation of *there* in English, the properties of the expletive construction should be derivable from those associated with the adverbial, i.e. from its locative and demonstrative characteristics. In fact, these properties do account in a large part for the properties of the expletive construction under the analysis given above.

In the first place, that *there* is a locative accounts for the fact that the associate may determine the form of the main verb. In Figure 2, the requirement for a plural subject is projected by the copula but is not satisfied by *there*, as this expression, not being nominative, does not project agreement

[4]All noun phrases are taken to project expressions of type e with quantification expressed as variable-binding term operators using the epsilon calculus (Kempson et al, ch. 7). The $PLURAL$ requirement on the subject is imposed by the agreement morphology of the verb.

information (only nominative expressions inducing full agreement on finite verbs in Indo-European generally). Because the associate merges with the subject node, its agreement properties can (indeed must) satisfy the requirements imposed by the verb.[5] Thus, one of the aspects of this construction that proves most puzzling for other accounts is directly captured (see e.g. Chomsky 1995).

It is again the locative dimension of the pronoun that accounts for both locational and existential readings of the *there* construction. I suggest that *there* does not assert a locative relation as part of the content (formula value) of some node, but imposes it instead as a condition on substitution of the metavariable it projects. This treats the locative information like gender on personal pronouns as a presupposition of the resulting formula.[6] Representing such presuppositions as subscripts on metavariables, *there* in subject position projects as its formula value, $\mathbf{U}_{LOC(\mathbf{U},\mathbf{V})}$. Anything that substitutes for $\mathbf{U}$ must, therefore, truthfully satisfy the first argument of some locative relation (which I treat as a weak semantic relation, interpretable more or less abstractly depending on context).[7]

The locative relation is itself underspecified since the second argument (the x where something is located) is itself a metavariable which requires to be substituted in order to provide some full presupposition that can be used in determining the appropriateness of the substitution operation for $\mathbf{U}$. This metavariable must itself (by the interpretation of the locative relation) be substituted by some expression denoting some spatial or spatio-temporal entity. With some spatial term substituting $\mathbf{V}$, we derive typical locative readings of the pronoun such as in *There are those two men outside (again)*. In parsing such a string we derive a presupposed locative context, $LOC((Two, x, Man(x)), \mathbf{V})$ where $\mathbf{V}$ is substituted by some formula value describable as *outside*. The metavariable may otherwise be substituted by the index of the proposition under construction, S_i. This gives rise to existential readings, since what is being constructed is a statement, $Be(Two, x, Man(x))$ which has a presupposition that 'locates' the two men at the current index, $LOC((Two, x, Man(x)), S_i)$, thus yielding a reading for the sentence: *Two men are (exist) at the current index*, a necessarily existential statement. No-

[5]Where the singular copula appears with a plural associate, the assumption must be that the latter is not 'nominative' and so does not project agreement properties. This can most clearly seen in the contrast between *There am/*is I with my hair in curlers* and *There *am/is me with my hair in curlers*.

[6]See Kempson et al 2001:236, for a characterisation of the anaphoric properties of definite NPs in similar terms.

[7]As a locative adverbial, *there* projects $\mathbf{V}_{LOC(\mathbf{U},\mathbf{V})}$, a distinction that is probably ultimately reducible to the properties of different lexical entailments.

tice further that where **V** is substituted by an index, the dependence of the term expressed by the associate has narrow scope with respect to that index.

Finally, it is the demonstrative force of the pronoun that gives rise to notions of new information connected with the associate (see e.g. Ward and Birner 1995). I assume a relevance-theoretic approach to demonstratives whereby the use of such expressions by a speaker makes manifest to the hearer that some contextually accessible referent has current relevance. The merging of the information provided by the associate with that of *there* signals the relevance of the former: where the associate is definite, there is a foregrounding of a contextually accessible referent, re-presenting it for further consideration; where the associate is indefinite, it is the existence of some object denoted by the associate that is foregrounded, buttressing the existential force derived from the existential verb and the locative property.

Although the hypothesis that adverbial and subject *there* are identical does not provide a complete account of the expletive constructions (e.g. why some quantifiers are excluded from existential readings), it does allow a uniform account of expletive and non-expletive uses of the same pronoun. The demonstrative and locative nature of the pronoun yielding the multiple ways of construing *there* constructions according to: the salience of some spatial location; the definiteness or otherwise of the associate; and the relevance of the latter's being made manifest within the current context. This analysis thus allows the treatment of expletives as a regular, though extended, anaphoric device. The sole difference between pronouns that can be used expletively and those that cannot is that the former have lost their 'bottom restriction' and so may be substituted by unfixed trees. This simple difference provides a means of accounting for expletive uses of pronouns without assuming homonymy, reducing the distinction to an epiphenomenon.

Blackburn, P and W Meyer-Viol (1994) 'Linguistics, logic and finite trees'. *Bulletin of Interest Group of Pure and applied Logics* 2:2-39.

Cann, R, R Kempson, L Marten, M Otsuka and D Swinburne (2001) 'On left and right dislocation: A Dynamic Perspective'. MS University of Edinburgh/King's College London/SOAS.

Chomsky, N (1995) *The Minimalist Program.* Cambridge, Mass., MIT Press.

Kempson, R, W Meyer-Viol and D Gabbay (2001) *Dynamic Syntax.* Oxford, Basil Blackwell.

McNally, L (1998) 'Existential sentences without existential quantification'. *Linguistics and Philosophy* 21:353-392.

Ward, G and B Birner (1995) 'Definiteness and the English existential'. *Language* 71:722-742.

Expressivity of extensions of dynamic first-order logic

Balder ten Cate[1], Jan van Eijck[2] and Juan Heguiabehere[3]

Abstract

Dynamic predicate logic (DPL), presented in [5] as a formalism for representing anaphoric linking in natural language, can be viewed as a fragment of a well known formalism for reasoning about imperative programming [6]. An interesting difference from other forms of dynamic logic is that the distinction between formulas and programs gets dropped: DPL formulas can be viewed as programs. In this paper we show that DPL is in fact the basis of a hierarchy of formulas-as-programs languages.

1 Formulas-as-Programs in DPL and its Extensions

In this paper we investigate the landscape of extensions of DPL with the six operations $\cup$, ${}^{\smile}$, σ, $\breve{\sigma}$, $\cap$, $\exists\!\!\exists$. All 64 combinations are classified with respect to their expressive power. Extensions of DPL with $\cap$ (relation intersection) and $\exists\!\!\exists$ (local variable declaration) are studied in [8], while in [1], an extension of DPL with $\cup$ (relation union) and σ (simultaneous substitution) is axiomatized, and ω-completeness is proved for the extension of DPL with $\cup$, σ and $*$ (Kleene star). In a dynamic setting, left-to-right substitutions σ have right-to-left counterparts $\breve{\sigma}$ (converse substitutions). For pre- and postcondition reasoning with extensions of DPL, converse substitution and relation converse ${}^{\smile}$ are attractive.

Some of these operators have been studied from a linguistic perspective. In particular, [4] discusses potential linguistic applications of the $\cup$ and $\cap$ operator. The relevance of $\cap$ operator is mentioned there for the analysis of *plurals*, and it is argued that $\cup$ is needed for a proper analysis of sequences such as: *"A professor or an assistant professor will attend the meeting of the university board. She will report to the faculty."*

Unlike [5], we allow function symbols, so DPL terms are given by $t ::= x \mid c \mid f(t_1,\dots,t_n)$. DPL formulas are given by $\phi ::= \exists x \mid Rt_1 \dots t_n \mid t_1 \approx t_2 \mid \neg\phi \mid \phi_1;\phi_2$. Terms are interpreted as usual. We use $t^{\mathcal{M},g}$ for the interpretation of term t in model $\mathcal{M}$ under valuation g.

A substitution is a finite set of bindings $x := t$, with the usual conditions that no binding is trivial (of the form $x := x$) and that every x in the set has at most one binding (substitutions are functional). Examples are $\{x := f(x)\}$ ("set new x equal to f-value of old x"), $\{x := y, y := x\}$ ("swap values of x and y"). If a substitution contains just a single binding we omit the curly brackets and write just the *assignment statement* $x := t$. A converse substitution is a finite set of converse bindings $(x := t)^{\smile}$, with the same conditions as those for substitutions. An example is $(x := f(x))^{\smile}$ ("set old x equal to f-value of new x", or "look at all inputs g that differ from the output h only in x, and that satisfy $f(g(x)) = h(x)$").

A variable state in a model $\mathcal{M} = \langle D, I\rangle$ is a member of D^{Var}, where Var is the set of variables of the language. Letting g, h, k range over variable states, the interpretation of DPL and the extensions that we study, in a model $\mathcal{M} = \langle D, I\rangle$, is given by the following definition:

$$
\begin{array}{rcl}
g[\![\exists x]\!]^{\mathcal{M}}h & \text{iff} & h = g^x_d \text{ for some } d \in D \\
g[\![Rt_1 \dots t_n]\!]^{\mathcal{M}}h & \text{iff} & h = g \text{ and } \langle t_1^{\mathcal{M},g},\dots,t_n^{\mathcal{M},g}\rangle \in R^{\mathcal{M}} \\
g[\![t_1 \approx t_2]\!]^{\mathcal{M}}h & \text{iff} & h = g \text{ and } t_1^{\mathcal{M},g} = t_2^{\mathcal{M},g} \\
g[\![\neg\phi]\!]^{\mathcal{M}}h & \text{iff} & h = g \text{ and there is no } k \text{ such that } g[\![\phi]\!]^{\mathcal{M}}k \\
g[\![\phi;\psi]\!]^{\mathcal{M}}h & \text{iff} & g[\![\phi]\!]^{\mathcal{M}}k \text{ and } k[\![\psi]\!]^{\mathcal{M}}h \text{ for some } k \in D^{\mathrm{Var}} \\
g[\![\sigma]\!]^{\mathcal{M}}h & \text{iff} & h = g^{x_1\cdots x_n}_{d_1\cdots d_n} \text{ where } \{x_1,\dots,x_n\} = \mathrm{dom}(\sigma) \text{ and } d_i = \sigma(x_i)^{\mathcal{M},g} \\
g[\![\breve{\sigma}]\!]^{\mathcal{M}}h & \text{iff} & g = h^{x_1\cdots x_n}_{d_1\cdots d_n} \text{ where } \{x_1,\dots,x_n\} = \mathrm{dom}(\sigma) \text{ and } d_i = \sigma(x_i)^{\mathcal{M},h} \\
g[\![\exists\!\!\exists x(\phi)]\!]^{\mathcal{M}}h & \text{iff} & (g^x_d)[\![\phi]\!]^{\mathcal{M}}k \text{ and } h = k^x_{g(x)} \text{ for some } d \in D, k \in D^{\mathrm{Var}} \\
g[\![\phi \cup \psi]\!]^{\mathcal{M}}h & \text{iff} & g[\![\phi]\!]^{\mathcal{M}}h \text{ or } g[\![\psi]\!]^{\mathcal{M}}h \\
g[\![\phi \cap \psi]\!]^{\mathcal{M}}h & \text{iff} & g[\![\phi]\!]^{\mathcal{M}}h \text{ and } g[\![\psi]\!]^{\mathcal{M}}h \\
g[\![\phi^{\smile}]\!]^{\mathcal{M}}h & \text{iff} & h[\![\phi]\!]^{\mathcal{M}}g.
\end{array}
$$

[1]ILLC, Amsterdam

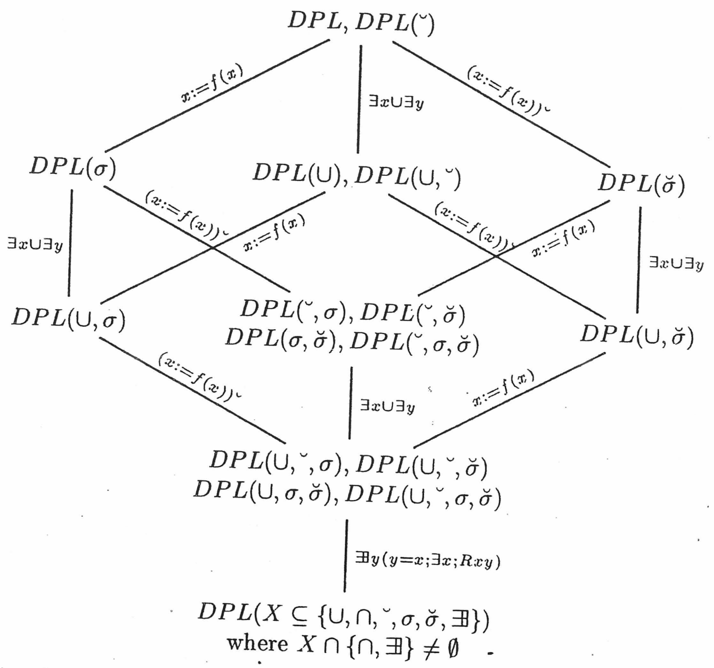
[2]CWI and ILLC, Amsterdam, Uil-OTS, Utrecht

[3]ILLC, Amsterdam

With DPL we will denote the basic language. Extensions will be indicated with DPL(X), where X is a se[t] of operators. For instance, DPL(σ, $\smile$) denotes the extension of DPL with simultaneous substitutions and converse.

2 The Lattices of DPL and DPL(*) Extensions

The following figure represents the lattice of all possible combinations of DPL with operators from $\{\cup, \cap, \smile, \sigma, \breve{\sigma}, \exists\}$ (union, intersection, converse, simultaneous substitution, converse substitution, hiding). It indicates which operators can be defined in terms of which; the labels on the arrows indicate counterexamples to equal expressivity, i.e., formulas from the lower language that don't have a counterpart in the upper language.



Note that all 64 combinations of the six operators are present in the diagram. The diagram makes immediately clear which extensions of DPL are closed under converse: precisely those which are in the same node of the lattice as the corresponding version of DPL *with* converse operator. Adding Kleene star gives an isomorphic lattice for DPL(*) and its extensions: none of the distinctions collapse because the same counter-examples to equal expressivity still work. The justification of the lattices is in Section 5.

3 Expressivity and Programming Constructs

Most of the operators we consider for extending DPL have obvious uses in programming [2, 6]. Kleene star allows the implementation of *while* loops.

$$\exists s; \exists i; s \approx 0; i \approx -1; (i \not\approx n; i := i + 1; s := s + b(i))^*; i \approx n \qquad (1)$$

Program (1) computes the sum of the $n + 1$ elements of an array b and places the result in the variable s. If the value for n is not known beforehand (at 'compile time'), it is not possible to write an equivalent program without the use of *.

$$(x \approx 1;\ \exists y; y \approx t_1) \cup (x \approx 0; \exists z; z \approx t_2) \qquad (2)$$

Nondeterminism ($\cup$) allows conditional execution: program (2) changes the value of either y or z, depending on the value of x. This cannot be done without $\cup$ (or stronger operators like $\cap$ or $\exists$).

The $\exists$ operator allows for the declaration of local variables. Simultaneous substitution permits performing certain computations without the use of auxiliary variables, and finally converse and converse simultaneous substitution are useful for pre- and postcondition reasoning, as they allow us to define the inverses of programs under certain conditions (Chapter 21 in [3]).

4 Left-to-Right and Right-to-Left Substitution

Because the semantics of DPL formulas is completely symmetric, performing a substitution in a DPL formula can be done in two directions: left-to-right and right-to-left [8] (see also [7], where substitutions for DPL with a stack semantics are studied). Left-to-right substitutions affect the left-free variable occurrences, right-to-left substitutions the right-free (or 'actively bound') variable occurrences.

The substitution lemma for first order logic has the form $\mathcal{M} \models_g \sigma(\phi)$ iff $\mathcal{M} \models_{g_\sigma} \phi$, given an obvious safety condition on σ, to the effect that no variables in the range of σ should get accidentally captured by quantifiers in ϕ.

One would like to prove a left-to-right substitution lemma for DPL to the effect that $g[\![\sigma(\phi)]\!]h$ iff $g_\sigma[\![\phi]\!]h$ (suppressing the parameter $\mathcal{M}$ for readability). Viewing the substitution itself as a state change, we can decompose this into $g[\![\sigma]\!]g'[\![\phi]\!]h$. This uses $g[\![\sigma]\!]k$ iff $k = g_\sigma$. Since, in general, σ is not expressible in DPL, we need to extend the language with substitutions [1].

The right-to-left substitution lemma for DPL that one would like to prove says that $g[\![\breve{\sigma}(\phi)]\!]h$ iff $g[\![\phi]\!]h_\sigma$. Viewing the right-to-left substitution itself as a state change, we can decompose this into $g[\![\phi]\!]h'[\![\breve{\sigma}]\!]h$. This uses $k[\![\breve{\sigma}]\!]h$ iff $k = h_\sigma$. Again, since in general $\breve{\sigma}$ is not expressible in DPL, we have a motivation to extend the language with converse substitutions.

Use $\circ$ for relational composition of substitution expressions, defined in the standard way. Note that $\sigma; \rho$ is equivalent to $\sigma(\rho)$, which is in turn equivalent to $\rho \circ \sigma$. E.g., $x := x + 1; y := x$ is equivalent to $\{x := x + 1, y := x + 1\}$.

Every DPL($\cup, \sigma$) formula can be written with ; associating to the right, as a list of predicates, quantifications, negations, choices and substitutions, with a substitution ρ at the end (possibly the empty substitution). Left-to-right substitution in DPL($\cup, \sigma$) is defined by:

$$\begin{array}{rcl}
\sigma(\rho) & := & \rho \circ \sigma \\
\sigma(\rho; \phi) & := & \rho \circ \sigma; \phi \\
\sigma(\exists v; \phi) & := & \exists v; \sigma'\phi \text{ where } \sigma' = \sigma \backslash \{v := t \mid t \in T\} \\
\sigma(P\bar{t}; \phi) & := & P\sigma\bar{t}; \sigma\phi \\
\sigma(t_1 \approx t_2; \phi) & := & \sigma t_1 \approx \sigma t_2; \sigma\phi \\
\sigma(\neg(\phi_1); \phi_2) & := & \neg(\sigma\phi_1); \sigma\phi_2 \\
\sigma((\phi_1 \cup \phi_2); \phi_3) & := & \sigma(\phi_1; \phi_3) \cup \sigma(\phi_2; \phi_3)
\end{array}$$

A term t is left-to-right free for v in ϕ if all variables in t are input-constrained in all positions of the left-free occurrences of v in ϕ. A substitution σ is safe for ϕ if all bindings $v := t$ of σ are such that t is left-to-right free for v in ϕ. This allows us to prove:

Lemma 1 (Left-to-Right Substitution) *If σ is safe for ϕ then $g[\![\sigma(\phi)]\!]h$ iff $g_\sigma[\![\phi]\!]h$.*

Right-to-left substitution is defined in a symmetric fashion, now reading the formulas in a left-associative manner, with a converse substitution at the front, and overloading the notation by also using $\circ$ for the relational composition of converse substitutions (defined as one would expect, to get $\breve{\sigma} \circ \breve{\rho} = (\rho \circ \sigma)^\smile$):

$$\begin{array}{rcl}
\breve{\sigma}(\breve{\rho}) & := & \breve{\sigma} \circ \breve{\rho} \\
\breve{\sigma}(\phi; \breve{\rho}) & := & \phi; \breve{\sigma} \circ \breve{\rho} \\
\breve{\sigma}(\phi; \exists v) & := & \breve{\sigma}'\phi; \exists v \text{ where } \breve{\sigma}' = \breve{\sigma} \backslash \{(v := t)^\smile \mid t \in T\} \\
\breve{\sigma}(\phi; P\bar{t}) & := & \breve{\sigma}\phi; P\sigma\bar{t} \\
\breve{\sigma}(\phi; t_1 \approx t_2) & := & \breve{\sigma}\phi; \sigma t_1 \approx \sigma t_2 \\
\breve{\sigma}(\phi_1; \neg(\phi_2)) & := & \breve{\sigma}\phi_1; \neg(\breve{\sigma}\phi_2); \\
\breve{\sigma}(\phi_1; (\phi_2 \cup \phi_3)) & := & \breve{\sigma}(\phi_1; \phi_2) \cup \breve{\sigma}(\phi_1; \phi_3)
\end{array}$$

A term t is right-to-left free for v in ϕ if all variables in t are output-constrained in all positions of the right-free (actively bound) occurrences of v in ϕ. A converse substitution $\breve{\sigma}$ is safe for ϕ if all converse bindings $(v := t)^\smile$ of $\breve{\sigma}$ are such that t is right-to-left free for v in ϕ. This allows us to prove:

Lemma 2 (Right-to-Left Substitution) *If $\breve{\sigma}$ is safe for ϕ then $g[\![\breve{\sigma}(\phi)]\!]h$ iff $g[\![\phi]\!]h_\sigma$.*

5 The Justification of the DPL Extension Lattices

This section contains the proofs of the theorems that justify the lattices of DPL and DPL(*) extensions. Many of the proofs are generalizations of, or adaptations of, proofs given in [8].

Theorem 1 *$DPL(\exists)$ is equally expressive as $DPL(\cup, \cap, {}^{\smile}, \sigma, \breve{\sigma}, \exists)$.*
Proof: *Let a formula ϕ be given, and let V be the set of variables occurring in ϕ. Furthermore, let V' and V'' be sets of variables, such that V, V' and V'' are mutually disjoint and of equal cardinality. Let $V = \{x_1, \dots, x_n\}$, $V' = \{x'_1, \dots, x'_n\}$, and $V'' = \{x''_1, \dots x''_n\}$. The following function C translates a formula from $DPL(\cup, \cap, {}^{\smile}, \sigma, \breve{\sigma}, \exists)$ into a test from DPL.*

$$
\begin{array}{lcl}
C(\exists y) & = & \bigwedge_{x \in V \setminus \{y\}} x' \approx x \\
C(Rt_1 \dots t_n) & = & \bigwedge_{x \in V} x' \approx x; Rt_1 \dots t_n \\
C(t_1 \approx t_2) & = & \bigwedge_{x \in V} x' \approx x; t_1 \approx t_2 \\
C(\neg\phi) & = & \bigwedge_{x \in V} x' \approx x; \neg(\exists x'_1; \dots; \exists x'_n; C(\phi)) \\
C(\phi;\psi) & = & \neg\neg(\exists x''_1; \dots; \exists x''_n; C(\phi)^{[x'_1/x''_1, \dots, x'_n/x''_n]}; C(\psi)^{[x_1/x''_1, \dots, x_n/x''_n]}) \\
C(\phi \cap \psi) & = & C(\phi); C(\psi) \\
C(\phi \cup \psi) & = & \neg(\neg C(\phi); \neg C(\psi)) \\
C(\phi^{\smile}) & = & C(\phi)^{[x_1/x'_1, \dots, x_n/x'_n, x'_1/x_1, \dots, x'_n/x_n]} \\
C(\sigma) & = & \bigwedge_{x \in dom(\sigma)} x' \approx \sigma(x); \bigwedge_{x \in V \setminus dom(\sigma)} x' \approx x \\
C(\breve{\sigma}) & = & \bigwedge_{x \in dom(\sigma)} x \approx \sigma(x)^{[x_1/x'_1, \dots, x_n/x'_n]}; \bigwedge_{x \in V \setminus dom(\sigma)} x' \approx x \\
C(\exists x.\phi) & = & \neg\neg(\exists x; \exists x'; C(\phi)); x' \approx x
\end{array}
$$

Here, $\bigwedge$ is used as a shorthand for a long composition, which is non-ambiguous because the order of the particular sentences involved doesn't matter. By induction, it can be shown that every ϕ containing only variables in V, is equivalent to $\exists x'_1 \dots x'_n(C(\phi); x_1 := x'_1; \dots x_n := x'_n)$. □

Theorem 2 *$DPL(^*, \exists)$ is equally expressive as $DPL(^*, \cup, \cap, {}^{\smile}, \sigma, \breve{\sigma}, \exists)$*
Proof: *As the proof of Theorem 1, now adding the following clause to the definition of C.*

$$C(\phi^*) \quad = \quad \neg\neg(\exists x''_1; \dots; \exists x''_n; (C(\phi)^{[x'_i/x''_i]}; \bigwedge_{x \in V} x := x'')^*; \bigwedge_{x \in V} x \approx x')$$

□

It follows immediately that every formula $\phi \in DPL(\cup, \cap, {}^{\smile}, \sigma, \breve{\sigma}, \exists)$ is equivalent to a first order logic formula, in the sense that ϕ can be executed in $\mathcal{M}$ with input assignment g iff the first order translation of ϕ is true in $\mathcal{M}$ under g.

Theorem 3 (Visser) *$DPL(\exists)$ can be embedded into $DPL(\cap)$.*
Proof: *Let ϕ be of the form $\exists x(\psi)$, and let $\{y_1, \dots, y_n\} = B(\phi) \setminus \{x\}$, where $B(\phi)$ are the blocking variables of ϕ, i.e., the variables y such that ϕ contains an $\exists y$ not in the scope of a negation [8]. Then ϕ is equivalent to $(\exists x; \psi; \exists x) \cap (\exists y_1; \dots; \exists y_n)$* □

In a similar way, the following can be proved:

Theorem 4 *$DPL(^*, \exists)$ can be embedded into $DPL(^*, \cap)$.*

It is also easy to show that * gets us beyond first order expressivity:

Theorem 5 *The formula*

$$\neg(\exists y; y \approx 0; (\exists z; z \approx f(x); \exists y; z \approx f(y))^*; x \approx y)$$

cannot be expressed in $DPL(\cup, \cap, {}^{\smile}, \sigma, \breve{\sigma}, \exists)$.
Proof: *On the natural numbers (interpreting f as the successor relation), this formula defines the odd numbers. Oddness on the natural numbers cannot be captured in a first order formula with only successor.*
□

Definition 1 *A substitution $\{x_1 := t_1, \dots, x_n := t_n\}$ is* full *if every x_i occurs in some t_j and every t_i contains some x_j.*

Examples of full substitutions are $x := f(x)$ and $\{x := y, y := x\}$, while the substitution $x := y$ is not full. It is easy to see that full substitutions are closed under composition. Note that a substitution without function symbols is full iff it is a renaming. Also, note that any formula of $DPL(\sigma)$ or any of its extensions can be transformed into a formula in the same language containing only full substitutions, by replacing bindings of the form $x := t$, where t does not contain variables, by $\exists x; x \approx t$.

Lemma 3 *Every formula $\phi \in DPL(\sigma)$ is equivalent to a formula of one of the following forms (for some ψ, x, χ, σ, where σ is full):*

1. *$\neg\neg\chi; \sigma$.*
2. *$\psi; \exists x; \neg\neg\chi; \sigma$.*

Proof: *First rewrite ϕ into a formula that contains only full substitutions. After that, the only non-trivial case in the translation instruction is the case of $\tau; \psi$, where τ is full and ψ is of the first form, i.e., where ψ is equivalent to $\neg\neg\chi; \sigma$, for some χ, σ, with σ full. In this case, $\tau; \psi$ is equivalent to $\neg\neg(\tau; \chi); \sigma \circ \tau$, where $\sigma \circ \tau$ is full because σ and τ are.* □

Theorem 6 *$(\exists x \cup \exists y)$ cannot be expressed in $DPL(\sigma)$.*
Proof: *Suppose $\phi \in DPL(\sigma)$ is equivalent to $(\exists x \cup \exists y)$. Take a model with as domain the natural numbers, and let R be the interpretation of ϕ. By Lemma 3, it follows that ϕ is equivalent to $\psi; \exists z; \neg\neg\chi; \sigma$, for some formulas ψ, χ, some variable z and some full substitution σ (otherwise, ϕ would be deterministic). Two cases can be distinguished.*

1. *z does not occur in σ. Without loss of generality, assume that $z \neq x$. Take any pair of assignments g, h such that $g \neq h$ and $g \sim_x h$. Then gRh. Take any $k \neq h$ such that $k \sim_z h$. Then gRk, but g and k differ with respect to two variables (x and z), which is in contradiction with the fact that ϕ is equivalent to $(\exists x \cup \exists y)$.*
2. *z occurs in σ. By the fact that there are no function symbols involved, and by the fact that σ is full, there must be exactly one binding in σ of the form $u := z$. We can apply the same argument as before, now using u instead of z, and again we arrive at a contradiction.*

□

Since every substitution is equivalent to a DPL formula containing only full substitutions, and since every full substitution without function symbols is a renaming, and therefore has a converse that is also a renaming, we get:

Lemma 4 *Every converse substitution containing no function symbols is equivalent to a formula in $DPL(\sigma)$.*

This immediately gives:

Corollary 1 *$(\exists x \cup \exists y)$ cannot be expressed in $DPL(\sigma, {}^{\smile})$.*

By a straightforward induction we get:

Lemma 5 *Every formula in $DPL(\sigma, \cup)$ is equivalent to a formula of the form $\phi_1 \cup \dots \cup \phi_n$ $(n \geq 1)$ where each $\phi_i \in DPL(\sigma)$.*

Theorem 7 *$(x := f(x))^{\smile}$ cannot be expressed in $DPL(\sigma, \cup)$.*
Proof: *Suppose $\phi \in DPL(\sigma, \cup)$ is equivalent to $(x := f(x))^{\smile}$. By Lemma 5, we can assume that ϕ is of the form $\phi_1 \cup \dots \cup \phi_n$, where each $\phi_i \in DPL(\sigma)$. Consider the model with as domain $\{0, \dots, n\}$, and where f is interpreted as the "successor modulo $n+1$" function.*

Let us say that a relation R fixes a variable x if for $\forall gh \in cod(R)$: $g \sim_x h$ implies that $h = g$. Analyzing each ϕ_i, we can distinguish the following two cases.

- *ϕ_i is equivalent to $\neg\neg\chi; \sigma$, with σ full. Then $[\![\phi_i]\!]$ fixes x.*
- *ϕ_i is equivalent to $\psi; \exists y; \neg\neg\chi; \sigma$, again with σ full. If y occurs in σ, then let z_i be the (unique) variable such that σ contains a binding of the form $z_i := f^k(y)$. If σ does not contain y then let $z_i = y$. Then it must be the case that $[\![\phi_i]\!]$ fixes z_i, for otherwise $[\![\phi_i]\!]$ is not injective.*

Thus, we have that every ϕ_i fixes some variable z_i. Let $\{z_1, \dots, z_m\}$ be all variables that are fixed by some ϕ_i (where $m \leq n$).

Consider all possible ways of assigning objects from the domain to the variables $z_1, \dots z_m$ (assigning to all other variables). This gives us $(n+1)^m$ assignments, each of which is in the co-domain of ϕ. Now, of this space of assignments, each ϕ_i can cover only a small part: at most $(n+1)^{m-1}$ (since one variable is fixed). So, together, $\phi_1, \dots, \phi_n$ can cover at most $n * (n+1)^{m-1} = (n+1)^m - (n+1)^{m-1} < (n+1)^m$ assignments, which means that some assignments are not in the co-domain of ϕ. This is in contradiction with the fact that ϕ is equivalent to $(x := f(x))^\smile$. □

By symmetry, we get the following

Corollary 2 *$x := f(x)$ cannot be expressed in DPL($\breve{\sigma}$, ∪).*

Theorem 8 *$\exists y(y \approx x; \exists x; Rxy)$ cannot be expressed in DPL(∪, σ, ˘).*
Proof: *The same proof as for Theorem 7 can be used. Assume a signature without function symbols. Let the domain of the model be the set $\{0, \dots, n\}$. Let R be interpreted as "successor modulo $n+1$". Then R is interpreted in the same way as f was in the proof of Theorem 7. Notice that, under this interpretation, $\exists y(y \approx x; \exists x; Rxy)$ means the same as $(x := f(x))^\smile$ did in the proof of Theorem 7. It follows that $\exists y(y = x; \exists x; Rxy)$ cannot be expressed in DPL(∪, σ). Since the signature contains no function symbols, it follows by Lemma 4 that this formula cannot be expressed in DPL(∪, σ, ˘) either.* □

6 Conclusion and Future Work

Our study of extensions of DPL started from an observation about the process of substitution in a dynamic setting. The substitution lemmas shed some further light on how variables are handled in DPL, and thus indirectly on what the processes of anaphoric reference and anaphora resolution in natural language analysis and variable declaration and variable use in programming have in common.

The formulas-as-programs languages that extend DPL hold promises for correctness reasoning and program analysis. Tableau style reasoning for DPL(σ, ∪, *) could be used for generating counterexamples from incorrect programs. The formulas-as-programs perspective also simplifies the process of program inversion. Gries's example of the inversion of $\{x \approx 3\}$ $x := 1$ to $\{x \approx 1\}$ $x := 3$ involved a magical transformation of a command into an assertion and vice versa. In the formulas-as-programs perspective this becomes the mapping of $x \approx 3; \exists x; x \approx 1$ to $x \approx 1; \exists x; x \approx 3$, which is just a matter of two applications of the inversion law $(R \circ S)^\smile = S^\smile \circ R^\smile$. Inversion of substitutions can always be done by means of ˘, but sometimes $\breve{\sigma}$ can again be written as a substitution: converses of renamings are themselves renamings, $(x := x+1)^\smile$ is equivalent to $x := x - 1$, and so on. Further analysis of these matters is future work.

References

[1] J. van Eijck, J. Heguiabehere, and B. Ó Nualláin. Tableau reasoning and programming with dynamic first order logic. *Logic Journal of the IGPL*, 9(3), May 2001.

[2] R. Goldblatt. *Logics of Time and Computation, Second Edition, Revised and Expanded*, volume 7 of *CSLI Lecture Notes*. CSLI, Stanford, 1992 (first edition 1987). Distributed by University of Chicago Press.

[3] D. Gries. *The Science of Programming*. Springer, Berlin, 1981.

[4] J. Groenendijk and M. Stokhof. Dynamic Montague Grammar. In L. Kalman and L. Polos, editors, *Papers from the Second Symposium on Logic and Language*, pages 3–48. Akademiai Kiadoo, Budapest, 1990.

[5] J. Groenendijk and M. Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14:39–100, 1991.

[6] D. Harel. *First-Order Dynamic Logic*. Number 68 in Lecture Notes in Computer Science. Springer, 1979.

[7] C. Vermeulen. A calculus of substitutions for DPL. *Studia Logica*, 68:357–387, 2001.

[8] A. Visser. Contexts in dynamic predicate logic. *Journal of Logic, Language and Information*, 7(1):21–52, 1998.

Counting Concepts

Cleo Condoravdi Dick Crouch Xerox PARC

Martin van den Berg FXPAL

In [2] we argued that the object argument of *prevent* must be concept-denoting rather than individual-denoting. The singular indefinite NP *a strike* in (1) existentially quantifies over sub-concepts of the concept Strike, including the concept Strike itself.

(1) Negotiations prevented a strike.

(1) entails that Strike or some sub-concept of it is uninstantiated in the actual world, where negotiations occur. But in a counterfactual world, identical apart from the absence of negotiations, the concept is instantiated.

This paper discusses how this analysis can be extended to deal with concept denoting indefinite NPs with numeral determiners, as in (2).

(2) Negotiations prevented three strikes.

If the NP *three strikes* quantifies over concepts, what is counted by the numeral *three*? We argue that *three* corresponds to a generalized determiner counting appropriately identified sub-concepts of Strike. More direct attempts to count the number of individuals instantiating Strike fail to predict the correct entailments for (2).

Section 1 reviews some of the entailment patterns of sentences like (1). Section 2 points out the difficulties these patterns pose for standard treatments of intensional verbs, and reviews the arguments in [2] that taking NPs to quantify over concepts rather than individuals accounts for these patterns. Section 3 discusses various ways of analyzing plural concept-denoting NPs, and explains why some of the more obvious possibilities fail to do justice to examples like (2).

1 General and Specific Prevention

Perhaps the most striking feature of (1) is its negative existential entailment: the prevented strike does not come into existence. Less immediately obvious is that (1) is ambiguous between a general reading (no strikes occurred, paraphrasable as *Negotiations prevented any strike*) and a specific reading (some particular potential strike did not occur). This ambiguity is brought out more clearly by

(3) Safety procedures at Chernobyl prevented a serious accident.

As a general statement (3) is notoriously false. But as a specific statement (3) might well be true: perhaps there were other occasions when the safety procedures really did prevent an accident.[1]

General and specific interpretations of prevent statement also license distinct monotonicity entailments.

(4) a. Negotiations prevented a (specific) long strike. $\not\dashv\;\models$
Negotiations prevented a (specific) strike.

b. Negotiations prevented a (any) long strike. $\not\dashv\;\not\models$
Negotiations prevented a (any) strike.

[1] Non-indefinite NPs are restricted to a specific reading. For instance, *Negotiations prevented every strike/most strikes* implies that there is a set of particular potential strikes quantified over.

Under the specific interpretation (4a), *prevent* is upward monotone in its second argument: if negotiations prevented some particular long strike, then, clearly, they prevented some particular strike. Under the general interpretation (4b), however, *prevent* is is neither upward nor downward monotone. Upward monotonicity fails as follows: even if negotiations prevented any long strikes, short ones may still have occurred. Downward monotonicity fails in a subtler way. Preventing any strikes clearly entails that no long strikes occur. But this does not necessarily mean that any long strikes have been prevented: maybe none of the prevented strikes would have been long to begin with.

Specific prevention is thus upward monotone in its factual and its counterfactual entailments. General prevention is downward monotone in its factual entailments but upward monotone in its counterfactual entailments, and hence neither upward nor downward monotone overall.

2 Prevention & Quantification over Concepts

Montague's classical account of intensional verbs like *want* does not properly account for existence entailments and general/specific readings when extended to verbs like *prevent*[2]:

(5) a. General, *de dicto*:
$\mathsf{prevent}(n, ^\wedge[\lambda P.\exists x.\ \mathsf{strike}(x) \wedge P(x)])$ — no existential commitment

b. Specific, *de re*:
$\exists x.\ \mathsf{strike}(x) \wedge \mathsf{prevent}(n, ^\wedge[\lambda P.P(x)]))$ — existential commitment

The general, *de dicto* reading (5a) is too weak for a predicate like *prevent*: *prevent* requires an entailment of non-existence rather than a mere non-entailment of existence. This can be rectified by a meaning postulate stating that the set of properties serving as the second argument must be empty in the world of evaluation. However, this does not help with the specific, *de re* reading (5b), which carries an (undesired) entailment of existence, regardless of any meaning postulates. Indeed the proposed postulate for *prevent* makes the specific reading self-contradictory: how can an individual sublimation be empty? Put another way, Montague's analysis ensures that all opaque verbs carry existential commitment under specific, *de re* readings, regardless of any lexico-semantic differences between the verbs.

A possible fix that can quickly be dismissed (dismissed at greater length in [2]) would be for quantifiers to range over possible rather than actual individuals. The *de re* reading would then merely entail the presence of a possible individual rather than the actual existence of a real individual. However, technical problems aside, the identity criteria for (un-named) possible individuals are extremely murky.

The core difficulty with Montague's analysis is that while *prevent* takes a concept as an argument, NPs quantify over / refer to individuals. Zimmermann [4] argues for an ambiguity between (i) individual NPs, which quantify over or refer to individuals, and (ii) concept NPs, which are non-quantificational, referential NPs denoting concepts. General readings arise from concept denoting NPs interpreted in situ (6a). Specific readings arise from an individual, quantificational NPs taking scope over the opaque predicate, plus type raising of the individual to an individual concept (6b).

(6) a. General: $\mathsf{prevent}(n, \mathsf{Strike})$ — no existential commitment

b. Specific: $\exists x.\ \mathsf{strike}(x) \wedge \mathsf{prevent}(n, ^\wedge\lambda y.y = x)$ — existential commitment

[2]The second argument to *prevent* is a concept, i.e., the intension of a set of properties. The first is some constant n referring to a particular set of negotiations.

However, Zimmermann's specific reading suffers exactly the same problem as Montague's: an existential commitment that arises from as a logical form rather than lexical entailment.

2.1 Singular Reference to and Quantification over Concepts

In [2] we extended Zimmermann's referential concept NPs to also allow quantificational concept NPs. The object of *prevent* is uniformly a concept NP, with a distinction between an NP referring to the nominal concept (general reading) or quantifying over sub-concepts of it (specific reading). This leads to the following (here somewhat simplified) analyses of (1):

(7) a. General (referential): $\exists Y.Y \sqsubseteq \mathsf{Negotiation} \wedge \mathsf{prevent}(Y, \mathsf{Strike})$

b. Specific (quantificational): $\exists X.X \sqsubseteq \mathsf{Strike} \wedge \exists Y.Y \sqsubseteq \mathsf{Negotiation} \wedge \mathsf{prevent}(Y, X)$

Here Strike and Negotiation refer to the concept of strike-events and negotiations respectively, and $\sqsubseteq$ means 'sub-concept'.

To spell out the entailments of (7a,b), we need to say something about the lexical entailments of the *prevent* predicate. Informally, $\mathsf{prevent}(Y, X)$ is true if the following two conditions hold. (i) In the actual world, w_a, the preventor concept Y has an instance but the preventee concept X has no instance; for example, there is an instance of a negotiation and no instance of a strike. (ii) There is a counterfactual world w_c as similar as possible to w_a apart from not having an instance of of the preventor Y[3], and in w_c there *is* an instance of X; for example, in the counterfactual world there are no instances of a negotiation but there is an instance of a strike. Stating this more formally

(8) $w_a \models \mathsf{prevent}(Y, X)$ iff
$\mathsf{inst}(w_a, Y)$, $\neg\mathsf{inst}(w_a, X)$, and for all w_c $\neg\mathsf{inst}(w_c, Y)$, and $\mathsf{inst}(w_c, X)$

where $\mathsf{inst}(w, C)$ is true if the concept C has a non-null extension in world w, and where we leave the precise relation between w_a and w_c open.

Given that an instance of a specific concept also counts as an instance of a concept that generalizes it:

(9) $\forall X, Y, w.\ (X \sqsubseteq Y \wedge \mathsf{inst}(w, X)) \rightarrow \mathsf{inst}(w, Y)$

and abbreviating the concepts of *strike* and *long strike* as S and LS, and where $LS \sqsubseteq S$, we obtain the following factual and counterfactual entailments:

General	Specific
Prevent long strike $\not\models_a$ Prevent strike $\neg\mathsf{inst}(w_a, LS) \not\models \neg\mathsf{inst}(w_a, S)$	Prevent long strike $\models_a$ Prevent strike $\exists X.X \sqsubseteq LS \wedge \neg\mathsf{inst}(w_a, X) \models \exists X.X \sqsubseteq S \wedge \neg\mathsf{inst}(w_a, X)$
Prevent long strike $\models_c$ Prevent strike $\mathsf{inst}(w_c, LS) \models \mathsf{inst}(w_c, S)$	Prevent long strike $\models_c$ Prevent strike $\exists X.X \sqsubseteq LS \wedge \mathsf{inst}(w_c, X) \models \exists X.X \sqsubseteq S \wedge \mathsf{inst}(w_c, X)$
Prevent strike $\models_a$ Prevent long strike $\neg\mathsf{inst}(w_a, S) \models \neg\mathsf{inst}(w_a, LS)$	Prevent strike $\not\models_a$ Prevent long strike $\exists X.X \sqsubseteq S \wedge \neg\mathsf{inst}(w_a, X) \not\models \exists X.X \sqsubseteq LS \wedge \neg\mathsf{inst}(w_a, X)$
Prevent strike $\not\models_c$ Prevent long strike $\mathsf{inst}(w_c, S) \not\models \mathsf{inst}(w_c, LS)$	Prevent strike $\not\models_c$ Prevent long strike $\exists X.X \sqsubseteq S \wedge \mathsf{inst}(w_c, X) \not\models \exists X.X \sqsubseteq LS \wedge \mathsf{inst}(w_c, X)$

[3]For the purposes of this paper we fortunately do not need to commit ourselves on the vexed issue of how to identify the most similar counterfactual world in which there is no instance of Y, or indeed whether there is a unique most similar world.

This correctly predicts the monotonicity properties of general and specific prevention discussed in section 1. Specific prevention is compatible with the existence of other strikes, whereas general prevention says there were no strikes at all.[4]

3 Plural Concept NPs

We have argued here, and at greater length in [2], that the object NP of the verb *prevent* must either refer to or quantify over concepts, rather than individuals. We now turn to the interpretation of plural sentences like (2). At issue is integrating our treatment of NPs as quantifiying over concepts and sub-concepts with a treatment of plural quantifiers. In this section we consider three broad approaches to plural quantification which involve, respectively: (i) counting the number of objects instantiating a concept, (ii) employing plural and singular concepts, and (iii) counting the number of (sub)concepts. Only the last option, where the NP is a generalized quantifier over concepts, gives sensible results when applied to (2).

3.1 Counting Instances

One possible way of capturing the cardinality of the concept NP *three strikes* (2) in would be to say that the concept / subconcept has three instances. This might lead to NP meanings along the lines of

(10) a. $\lambda P.\ \#\mathsf{inst}(\mathsf{Strike}, w, 3) \wedge P(X)$

b. $\lambda P.\ \exists X.\ [X \sqsubseteq \mathsf{Strike} \wedge \#\mathsf{inst}(X, w, 3)] \wedge P(X)$

where $\#\mathsf{inst}(X, w, 3)$ says that the number of individuals instantiating the concept X is world w is (at least) three.

It is important to relativize the counting of concept instances to a world. As we have already seen, the same concept can have different instantiations in different worlds. However, this relativization to worlds poses severe problems for sentences like (2). We might represent one reading of (2) as

(11) $\exists X.\ [X \sqsubseteq \mathsf{Strike} \wedge \#\mathsf{inst}(X, w, 3)] \wedge \mathsf{prevent}(\mathsf{Negotiation}, X)$

Which world should w be? In the actual world, the concept X should have no objects instantiating it. It is only in the counterfactual world implicit in *prevent* that X has three instances. But (i) how does the numeral quantification obtain access to this implicitly defined counterfactual world? And (ii) how can the numeral quantification also be set up to entail that there are zero instances in the actual world, while apparently it quantifies over three strikes?

The root of the problem is that the instantiation (or otherwise) of a concept is the result of the predicates that apply to the concept. In the case of *prevent*, the predication leads to different instantiation claims for different worlds. But the proposed treatment of cardinal NPs also makes instantiation part of the meaning of the quantifier over concepts. This is wrong. The NP itself should not be making any existence / instantiation claims; only the predicates that apply to it should.

3.2 Plural Concepts

The previous approach took counting instances to be part of the NP quantifier, and this led to problems. Suppose instead that counting instances is incorporated into

[4] [2] discusses various ways in which the general assertion that there are no strikes at all can be relativized to mean no strikes meeting a contextually salient description of relevant possible strikes.

the concepts over which the NP quantifies. For example, suppose that we take the numeral 3 to be an adjective, so that the NP *three strikes* might be represented as

(12) a. $\lambda P.\ P(3\sqcap\mathsf{Strike})$

b. $\lambda P.\ \exists X.\ X \sqsubseteq 3\sqcap\mathsf{Strike} \wedge P(X)$

where $3\sqcap\mathsf{Strike}$ is the combination of the concept of a strike and the concept of a threesome. Instances of this concept would be groups of (at least) three things, all of which are strikes.

We are now in the improved position where *prevent* takes a concept of three-strikes and evaluates it at different worlds. In the actual world the concept is uninstantiated: there are not three strikes. In the counterfactual world the concept is instantiated: there are three strikes.

Unfortunately, this analysis makes implausible predictions. Sentence (2) is interpreted to claim (generally) that the concept *three strikes* is uninstantiated in the actual world, or that (specifically) some sub-concept of *three strikes* is uninstantiated. What does it mean for such concepts to be uninstantiated? It means that there are two or fewer strikes. This is clearly wrong: under a general interpretation of (2), there should be no strikes in the actual world. Similarly under a specific interpretation there should be no instances of strikes meeting the more restricted sub-concept. In both cases there should be no instances, not just less than three. Put another way, this analysis treats (2) as synonymous with

(13) Negotiations prevented there being three strikes.

but the two sentences differ in meaning.

3.3 Counting Concepts

Suppose that we treat *three* as a generalized determiner quantifying over concepts

(14) $3(\lambda X.X \sqsubseteq \mathsf{Strike}, \lambda X.\mathsf{Prevent}(X))$

We must be careful about the way that we count concepts. Concepts are ordered by specificity in a way that individuals are not, and this needs to be accounted for when counting concepts. Monotonicity of instantiation (9) means that one individual can lead to the instantiation of multiple concepts. Simply counting all instantiated concepts will not lead to a count of the instantiating individuals. But counting maximally specific instantiated concepts will lead to an accurate count of individuals. The difficult part is to define what is meant by a maximally specific instantiated concept.

Assuming that the determiner is conservative (14) is equivalent to (15).

(15) $3(\lambda X.X \sqsubseteq \mathsf{Strike}, \lambda X.X \sqsubseteq \mathsf{Strike} \wedge \mathsf{Prevent}(X))$

We need to define a way of counting that will retrieve three maximally specific concepts from the set of prevented strikes, $\lambda X.X \sqsubseteq \mathsf{Strike} \wedge \mathsf{Prevent}(X)$.

Let's start with a simple case first. Suppose there are only two worlds: a world w_0, in which there are no strikes, and a world w_c in which there are exactly three strikes (a, b, c). The set $\lambda X.X \sqsubseteq \mathsf{Strike} \wedge \mathsf{Prevent}(X)$ is a partially ordered set of concepts, all of which hold of one or more of the individuals a, b or c. If a, b and c are distinct individuals, then there have to be incomparable concepts that distinguish them. Two concepts are incomparable iff they do not share any sub-concepts. This means that we can always find a triple of incomparable concepts A B and C (any three that distinguish a, b and c). But you will not be able to find a quadruple of incomparable concepts. Since there are only three strikes, two of the four concepts must be comparable, as they will apply to the same strike. We can thus count the number of individual strikes by finding the maximum n for which there is an n-tuple of incomparable concepts in $\lambda X.X \sqsubseteq \mathsf{Strike} \wedge \mathsf{Prevent}(X)$.

The general case is slightly more complicated. We find the maximum n for which there is an n-tuple of incomparable concepts in $\lambda X.X \sqsubseteq \mathsf{Strike} \wedge \mathsf{Prevent}(X)$. But we have to be careful what it means for two concepts X, Y to be incomparable. We know that X, Y are instantiated in every alternative world. And because they are from a maximal set, they will be instatiated by only one individual. But we have to ensure that there is no world where X and Y are instantiated by the same individual. We can do this by demanding that there is no sub-concept of strike, prevented or not, that is more specific than both X and Y. In particular, if X and Y were instantiated by the same individual a in a world w, the concept instantiated by a in w, and uninstantiated anywhere else would be more specific than X and Y. This leads to the following way of counting.

Let $\mathsf{xc}_{\mathcal{X}}(X_1, X_2, \ldots, X_n)$ mean that every X_i, X_j is mutually exclusive in the sense that there is no concept in $\mathcal{X}$ that is more specific than both X_i and X_j.

(16) $3(\mathcal{X}, \mathcal{Y}) \iff \exists X, Y, Z \in \mathcal{Y} \,.\, \mathsf{xc}_{\mathcal{X}}(X, Y, Z) \wedge \neg\exists X, Y, Z, P \in \mathcal{Y} \,.\, \mathsf{xc}_{\mathcal{X}}(X, Y, Z, P)$

With this meaning for the generalized quantifier 3, this does result in the right reading for (14). The maximal set of mutually exclusive concepts that are strike-concepts and satisfy the conditions of prevent has a cardinality 3, which means that in all accessible worlds, they are instantiated by 3 strikes.[5]

References

[1] van den Berg, M. *Some Aspects of the Internal Structure of Discourse* ILLC Dissertation Series 1996-3, Amsterdam, 1996.

[2] Condoravdi, C., R. Crouch, M. van den Berg, J. Everett, V. Paiva, R. Stolle, D. Bobrow. Preventing Existence *Formal Ontology in Information Systems.* ACM Press, 2001.

[3] van der Does, J. *Applied Quantifier Logics: Collectives & Naked Infinitives.* PhD thesis, University of Amsterdam, 1992.

[4] Zimmermann, E. On the proper treatment of opacity in certain verbs. *Natural Language Semantics* 1, 149–179, 1993.

[5] There may still be worlds with more strikes, but at least some worlds should have exactly three.

Meanwhile, Within the Frege Boundary

Paul Dekker*
ILLC/Department of Philosophy

dekker@hum.uva.nl

1 Introduction

Edward Keenan (Keenan 1992) has shown that type ⟨2⟩ quantifiers (properties of binary relations) which are reducible to two type ⟨1⟩ quantifiers (properties of unary relations) are identical if they behave the same on relations which are products. This is remarkable because it allows us to draw universal conclusions about two predicates (over a domain of relations) from their behavior over a highly restricted domain of relations (products, that is). Normally, knowing that two predicates behave uniformly over a small domain (that the nice students are the good students, for instance), does not generalize to larger domains (that nice humans are good humans, a non-sequitur). So what is the rationale behind Keenan's result?

Keenan's result is useful because it allows us to actually prove quite a few type ⟨2⟩ quantifiers to be *not* reducible to two type ⟨1⟩ quantifiers, merely by inspecting their behaviour on relations which are products. However, the result is not entirely satisfying since it leaves a few questions unanswered. Firstly, Keenan himself already realized that we can not use this result to show, for any irreducible type ⟨2⟩ quantifier, that it is irredubile. Secondly, it does not give us a method for deciding, given the behaviour of a type ⟨2⟩ quantifier on relations which are products, what its possible reduction to two type ⟨1⟩ quantifiers could be. Thirdly, it has so far been unclear if, or how, Keenan's result generalizes to type ⟨n⟩ quantifiers, properties of n-ary relations.

In this squib I will generalize Keenan's result about type ⟨2⟩ quantifiers to one about type ⟨n⟩ quantifiers and I will show that if we are given the behaviour of a reducible type ⟨n⟩ quantifier on products, we can determine the n type ⟨1⟩ quantifiers the composition of which is the type ⟨n⟩ one. I will first state the setting and terminology, restate Keenan's main result and its impact (section 2). In section 3 I will give a more general version of Keenan's result which, I think, provides more insight into what is really at stake when we consider reducibility. Section 4 winds up the results.

2 Keenan on type ⟨2⟩ Quantifiers

Let E be our universe of at least two individuals. A type ⟨1⟩ quantifier f is a property of sets of individuals: $f \in \mathcal{P}(\mathcal{P}(E))$, a type ⟨2⟩ quantifier F^2 is a property of binary relations between

*. I thank Marcus Kracht for directing me towards a simplification of Keenan's results which goes further than the one published in a previous version of this paper. The research for this work is supported by a grant from the Netherlands Organization for Scientific Research (NWO), which is gratefully acknowledged.

individuals: $F^2 \in \mathcal{P}(\mathcal{P}(E^2))$ and, more generally, a type $\langle n \rangle$ quantifier F^n is a property of n-ary relations over individuals: $F^n \in \mathcal{P}(\mathcal{P}(E^n))$. For a type $\langle n \rangle$ quantifier F^n and $R^n \in \mathcal{P}(E^n)$, I will write $F^n(R^n) = 1$ if $R^n \in F^n$ and $F^n(R^n) = 0$ otherwise.

By means of a rule of division (Geach) a type $\langle m \rangle$ quantifier can also be applied to a relation R^{n+m} of arbitrary arity $n+m$, yielding an n-ary relation $F^m(R^{n+m})$ as the result:

- $F^m(R^{n+m}) = \{\langle d_1, \ldots, d_n \rangle \mid F^m(\{\langle d_{n+1}, \ldots, d_{n+m} \rangle \mid \langle d_1, \ldots, d_{n+m} \rangle \in R\}) = 1\}$

(Notice that if $n = 0$, indeed $F^m(R^{0+m})$ is either $\{\langle\rangle\}$, the truth value 1, or $\emptyset$, the truth value 0.) Using the rule of division type $\langle 1 \rangle$ quantifiers f and g can be composed to produce a type $\langle 2 \rangle$ quantifier. Thus, $\forall R \in \mathcal{P}(E^2)$:

- $(f \circ g)(R) = f(g(R)) = f(\{d \mid g(\{d' \mid \langle d, d' \rangle \in R\}) = 1\})$

For readers familiar with montague grammar, $f \circ g$ is indeed the property of relations R satisfying: $T(\lambda x\ T'(\lambda y\ S(x)(y)))$, with T interpreted as f, T' as g and S as R. For example, consider the composition given by "every cat — a mouse":

- $[\![$*every cat*$]\!] \circ [\![$*a mouse*$]\!] = \{R \mid \forall d \in [\![$*cat*$]\!]$: $\exists d' \in [\![$*mouse*$]\!]$: $\langle d, d' \rangle \in R\}$

This type $\langle 2 \rangle$ quantifier holds of any relation R (such as chase, for instance) iff every cat Rs (chases) a mouse.

A both philosophically and linguistically interesting question is concerned with the possibility of characterizing type $\langle 2 \rangle$ quantifiers by means of the composition of two type $\langle 1 \rangle$ quantifiers. (Keenan 1992) presents a number of natural language examples which can not, and he actually proves they are not. The key concept is that of reducibility:

Definition 1 (Reducibility) a type $\langle 2 \rangle$ quantifier F^2 is reducible iff there are type $\langle 1 \rangle$ quantifiers f and g: $F^2 = f \circ g$

If a type $\langle 2 \rangle$ quantifier is not reducible, Keenan has it that it lives beyond the Frege boundary. Keenan's observations are backed up by two theorems, the one of which we focus upon here:

Theorem 1 (Reducibility Equivalence) if F^2 and G^n are reducible type $\langle 2 \rangle$ quantifiers:

- $F^2 = G^2$ iff $\forall P, Q \in \mathcal{P}(E)$: $F^2(P \times Q) = G^2(P \times Q)$

Reducible quantifiers have the special property that if they behave the same on relations which are products, they behave the same on all relations. (A product $(P \times Q)$ is of course the relation $\{\langle d, d' \rangle \mid d \in P\ \&\ d' \in Q\}$ holding between all members of P and Q, respectively.) This is an intriguing and remarkable result which can be used to show certain type $\langle 2 \rangle$ quantifiers to be not reducible.

Consider an arbitrary type $\langle 2 \rangle$ quantifier F^2 and the question whether it is reducible or not. Of course, if we take $[\![$*every cat*$]\!] \circ [\![$*a mouse*$]\!]$ we know it is reducible because the type $\langle 2 \rangle$ quantifier is defined in terms of two type $\langle 1 \rangle$ quantifiers. But then consider a property like that of transitivity or reflexivity. Transitivity and reflexivity are (contingent) properties of relations so they are type $\langle 2 \rangle$ quantifiers as defined above. Can we define these properties using two type $\langle 1 \rangle$ quantifiers? Now you may try to do this, and you may fail to succeed in reducing these quantifiers but this does not need to show that they are not reducible. Maybe you have not tried hard enough! Keenan here offers an ingenious method to establish that these quantifiers are indeed not reducible. Consider what transitivity and reflexivity say about product relations. It turns out that:

- $\text{TRANS}(P \times Q) = 1$ $(\forall P, Q \in \mathcal{P}(E))$

$\text{REFL}(P \times Q) = 1$ iff $P = Q = E$

This means that transitivity and reflexivity display precisely the same truth value pattern on product relations as the type $\langle 2 \rangle$ quantifiers $(\top \circ \top)$ and $(\text{ALL} \circ \text{ALL})$, respectively. (Here, $\top$ is the type $\langle 1 \rangle$ quantifier true of all sets of individuals.) Notice that the latter two type $\langle 2 \rangle$ quantifiers are reducible, because they are each defined in terms of two type $\langle 1 \rangle$ quantifiers. With Keenan's theorem (1) we now know that if transitivity and reflexivity are reducible then $\text{TRANS} = (\top \circ \top)$ and $\text{REFL} = (\text{ALL} \circ \text{ALL})$. But since the latter two equations are definitely false, the assumptions that transitivity and reflexivity are reducible must be false as well. A proof of the non-reducibility of a type $\langle 2 \rangle$ quantifier F thus consists in defining a type $\langle 2 \rangle$ quantifier $f \circ g$ which behaves the same as F on product relations. If $f \circ g$ is not in general equal to F we know F to be non-reducible.

Before we proceed, let us look at three natural language examples.

(1) Lois and Clark posed the same two stupid questions.

(2) Every student criticised himself.

(3) A sum total of five theories handled a sum total of five sentences.

If we only look at models where "posed" may denote relations which are products $P \times Q$, then (1) is true iff Lois and Clark are in P, and there are exactly two questions in Q. But these are precisely the same products for which "Lois and Clark posed exactly two stupid questions" is true. Since $([\![$*Lo+Cl*$]\!] \circ [\![$*ex2stqu*$]\!])$ is reducible and not equal to $\{R \mid [\![$*Lo+Cl V sa2stqu*$]\!]_{V/R}\}$, the latter is not reducible. Similarly, only looking at product interpretations of "criticised", example (2) is true iff "Every student criticised every student" is true, but certainly the two sentences are not generally equivalent. The same finally goes for example (3) (on the cumulative reading) and "Exactly five theories handled exactly five sentences." See (Keenan 1992) for more discussion.

Keenan's Reducibility Equivalence thus is a truly interesting result, but it leaves us with a couple of questions. Firstly, it is not quite clear exactly why reducible type $\langle 2 \rangle$ quantifiers behave as Keenan's theorem says they do. What makes it that their behaviour on the full domain $\mathcal{P}(E^2)$ is, in a sense, determined by their behaviour on $\mathcal{P}(E) \times \mathcal{P}(E)$? I must submit that, although I could follow Keenan's own proof of theorem (1), it did not give me the feeling I could see what is at stake. (Notice that it is certainly not the case that $(f \circ g)(R) = (f(d(R)) \wedge g(r(R)))$, where $d(R)$ indicates the domain of R and $r(R)$ its range.) Secondly, Keenan's theorem is only partly helpful in proving non-reducibility. For to prove type $\langle 2 \rangle$ F^2 not reducible we still have to find a (different) quantifier $(f \circ g)$ which behaves the same as F^2 on products. But if we do not find such a composition of two type $\langle 1 \rangle$ quantifiers it at best shows that F^2 is not reducible or, again, we have not tried hard enough! Besides, as we will see below, there are type $\langle 2 \rangle$ quantifiers, viz., the property of being a symmetric relation, the behaviour of which on products can not be characterized by any reducible quantifier. Thirdly, it has so far been an open question whether Keenan's reducibility equivalence generalizes to type $\langle n \rangle$ quantifiers. The following section is devoted to answer these questions.

3 Generalizing Keenan's Result

Let us first generalize our notion of reducibility:

Definition 2 (Type $\langle n \rangle$ Reducibility) a type $\langle n \rangle$ quantifier F^n is (n)-reducible iff there are n type $\langle 1 \rangle$ quantifiers $f_1, \ldots, f_n$: $F^n = f_1 \circ \ldots \circ f_n$

One of the key concepts which Keenan also uses is that of a quantifier which is 'positive'. A quantifier F^n (of arbitrary type $\langle n\rangle$) is positive iff $F^n(\emptyset) = 0$. Our observations in this squib will be stated for the most part with respect to positive quantifiers and with respect to type $\langle n\rangle$ quantifiers which are reducible to n positive type $\langle 1\rangle$ quantifiers, without loss of generalization. For:

Observation 1 if F^n is an (n)-reducible type $\langle n\rangle$ quantifier then there are n positive type $\langle 1\rangle$ quantifiers $f_1, \ldots, f_n$ such that $F = f_1 \circ \ldots \circ f_n$ or $\neg F^n = f_1 \circ \ldots \circ f_n$

Proof. Suppose F^n is (n)-reducible so that $F^n = f_1 \circ \ldots \circ f_n$. Starting from $i = n-1$ up to $i = 1$, if f_{i+1} is not positive, use $\neg f_{i+1}$ instead, which is positive, and use $f_i\neg$ in stead of f_i. Obviously, $f_i\neg \circ \neg f_{i+1} = f_i \circ f_{i+1}$. This, thus, is a recipe for characterizing an (n)-reducible type $\langle n\rangle$ quantifier F^n or $\neg F^n$ by means of n positive type $\langle 1\rangle$ quantifiers.

We will also use a generalization of the following observation from Keenan:

Observation 2 for a positive type $\langle 1\rangle$ quantifier f and any $P, Q \in \mathcal{P}(E)$:

- $f(P \times Q) = P$ if $f(Q) = 1$
  $= \emptyset$ otherwise

Proof. If $d \notin P$, $d \notin f(P \times Q)$, since f is positive; if $d \in P$, $d \in f(P \times Q)$ iff $f(Q) = 1$. The generalization we use is this:

Observation 3 if $F^n = f_1 \circ \ldots \circ f_n$ and the f_i are positive, then

- $F^n(Q_1 \times \ldots \times Q_n) = 1$ iff $f_1(Q_1) = \ldots = f_n(Q_n) = 1$

Proof. Assuming that $f_1(Q_1) = \ldots = f_n(Q_n) = 1$, $n-1$ applications of observation (2) give us that $F^n(Q_1 \times \ldots \times Q_n) = (f_1 \circ \ldots \circ f_n)(Q_1 \times \ldots \times Q_n) = (f_1 \circ \ldots \circ f_{n-1})(Q_1 \times \ldots \times Q_{n-1}) = \ldots = f_1(Q_1) = 1$. Furthermore, if, for any i $(1 < i \leq n)$ $f_i(Q_i) = 0$, $F^n(Q_1 \times \ldots \times Q_n) = (f_1 \circ \ldots \circ f_n)(Q_1 \times \ldots \times Q_n) = f_1(\emptyset) = 0$ (because the f_i are positive), and otherwise, if only $f_1(Q_1) = 0$, $(f_1 \circ \ldots \circ f_n)(Q_1 \times \ldots \times Q_n) = f_1(Q_1) = 0$ as well.

Now suppose F^n and G^n are (n)-reducible type $\langle n\rangle$ quantifiers. We can for the sake of convenience assume that $F^n = f_1 \circ \ldots \circ f_n$ and $G^n = g_1 \circ \ldots \circ g_n$, with all of the f_i and g_j positive (cf., observation 1). Keenan's theorem is now easily generalized:

Theorem 2 (Type $\langle n\rangle$ Reducibility Equivalence) if F and G are (n)-reducible type $\langle n\rangle$ quantifiers:

- $F = G$ iff $\forall Q_1, \ldots Q_n \in \mathcal{P}(E)$: $F(Q_1 \times \ldots \times Q_n) = G(Q_1 \times \ldots \times Q_n)$

Proof. Let F^n be reducible so that $F^n = f_1 \circ \ldots \circ f_n$ and assume all of the f_i to be positive. This means $F^n(Q_1 \times \ldots \times Q_n) = 1$ iff $f_i(Q_i) = 1$ for all i: $1 \leq i \leq n$. The same goes for $G^n = g_1 \circ \ldots \circ g_n$. If F^n and G^n behave the same on products, and the f_i and g_j are positive, the f_i must be identical to the g_i so that $F^n = G^n$.

Keenan's findings about (2)-reducible type $\langle 2\rangle$ quantifiers are thus generalized to type $\langle n\rangle$. The behaviour of n-reducible type $\langle n\rangle$ quantifier on arbitrary n-ary relations is somehow determined by their behaviour on relations which are products of n sets of individuals. An obvious next question is this. Given the behaviour of a quantifier F^n on products, can we determine what, if any, are type $\langle 1\rangle$ quantifiers $f_1, \ldots, f_n$ such that $F^n = f_1 \circ \ldots \circ f_n$? We can, if F^n shows a certain invariance, defined as follows:

Definition 3 (Invarience) a type $\langle n\rangle$ quantifier F^n is invariant for sets in products in F^n iff $\forall Q_1, \ldots, Q_n, Q'_1, \ldots, Q'_n$ (all non-empty) and for any i $(1 \leq i \leq n)$:

- if $F^n(Q_1 \times \ldots \times Q_i \times \ldots \times Q_n) = F^n(Q'_1 \times \ldots \times Q'_i \times \ldots \times Q'_n) = 1$ then $F^n(Q_1 \times \ldots \times Q'_i \times \ldots \times Q_n) = 1$

If we are given the behaviour of a type $\langle n\rangle$ quantifier on products we can now determine whether that behaviour can be generated by n type $\langle 1\rangle$ quantifiers. The point is not that F^n is invariant iff F^n is reducible, but the idea comes close:

Theorem 3 (Reducible Product Equivalents) a type $\langle n\rangle$ quantifier F^n or $\neg F^n$ is invariant for sets in products in $(\neg)F^n$ iff there is a product equivalent (n)-reducible correlate G^n of $(\neg)F^n$

Proof, Only if. Suppose F^n is invariant for sets in products in F^n. Define, for non-empty Q_i: $g_1(Q_1) = \ldots = g_n(Q_n) = 1$ iff $F^n(Q_1 \times \ldots \times Q_n) = 1$, $g_2(\emptyset) = \ldots = g_n(\emptyset) = 0$ and $g_1(\emptyset) = F^n(\emptyset)$. By F^n's invariance this is well-defined. Take $G^n = g_1 \circ \ldots \circ g_n$. G^n is construed to be equivalent with F^n on product relations and it is (n)-reducible by definition. If. Suppose F^n is positive (otherwise, take $\neg F^n$). Suppose F^n is not invariant, so there are $Q_1, \ldots, Q_n, Q'_1, \ldots, Q'_n$ (all non-empty): $F^n(Q_1 \times \ldots \times Q_i \times \ldots \times Q_n) = F^n(Q'_1 \times \ldots \times Q'_i \times \ldots \times Q'_n) = 1$ and $F^n(Q_1 \times \ldots \times Q'_i \times \ldots \times Q_n) = 0$. Now consider any arbitrary reducible $G^n = g_1 \circ \ldots \circ g_n$ with all of the g_i positive. Suppose $G^n(Q_1 \times \ldots \times Q_i \times \ldots \times Q_n) = 1$ and $G^n(Q_1 \times \ldots \times Q'_i \times \ldots \times Q_n) = 0$. Then $g_i(Q_i) = 1$ and $g_i(Q'_i) = 0$. So $G^n(Q_1 \times \ldots \times Q'_i \times \ldots \times Q_n) = G^n(Q'_1 \times \ldots \times Q'_i \times \ldots \times Q'_n) = G^n(\emptyset) = 0$. Hence, G^n is not product equivalent with F^n.

Theorem (3) tells us, when we know the behaviour of a type $\langle n\rangle$ quantifier $(\neg)F^n$ on products, we know whether there is an (n)-reducible quantifier which has that behaviour. Moreover, the proof gives us a method for defining this reducible quantifier as the composition of n constructively defined type $\langle 1\rangle$ quantifiers:

Corollary 1 (Non-invariance) If a type $\langle n\rangle$ quantifier F^n and its negation are not invariant for sets in products in $(\neg)F^n$, then $(\neg)F^n$ is not (n)-reducible

Proof, by contraposition. Suppose F^n is (n)-reducible. Then F^n has a product equivalent correlate G^n, namely F^n itself, which is (n)-reducible by supposition. So, by theorem (3), we find that F^n or $\neg F^n$ is invariant for sets in products in $(\neg)F^n$. Besides, using theorem (2):

Corollary 2 (Decomposition) If a type $\langle n\rangle$ quantifier F^n is invariant for sets in products in F^n, then F^n is (n)-reducible iff $F^n = G^n$, with G^n defined as in the proof of theorem (3)

Proof. If F^n is invariant it has a reducible product equivalent correlate G^n (theorem 3) and if F^n is reducible as well it must be identical to G^n (theorem 2).

The last result may help us understand the impact of the behaviour of reducible quantifiers on products. The sequences of sets that matter for a reducible quantifier F^n are the key, almost literally, for the quantifier to work. If each of the sets Q_i in the sequence fit the corresponding g_i (i.e., if $g_i(Q_i) = 1$), then F^n is true of the sequence. In case a set in the sequence does not matter for F^n (if it does not fit the lock, so to speak), then it does not matter which other Q_i you use, for in that case $F^n(Q_1 \times \ldots \times Q_n) = F^n(\emptyset) = 0$, if F^n is positive.

4 Conclusion

With this squib I have hoped to contribute to understanding Keenan's result from (Keenan 1992). Reducible type $\langle 2 \rangle$ quantifiers that behave the same on product relations are the same. I have given an alternative proof of this result, which applies to type $\langle n \rangle$ quantifiers in general. Not only is this a new and welcome generalization, it also gives—or at least it has given me—insight into the intimate relation between (n)-reducible type $\langle n \rangle$ quantifiers and n-ary product relations. If type $\langle n \rangle$ quantifier F^n is (n)-reducible, that is, if $F^n = f_1 \circ \ldots \circ f_n$ (with the f_i positive), then F^n is satisfied by $Q_1 \times \ldots \times Q_n$ iff each composing f_i is satisfied by Q_i. All of the f_i's do their own job (inspect Q_i), but the overall (positive) outcome depends on the (positive) joint outcome of all of them. For non-product relations they do essentially the same job, but then in a hierarchical manner, dependent on the relative scope of the f_i.

Corollary (2) and the construction used in the proof of theorem (3) here prove very useful if we want to apply Keenan's theorem (1) to establish non-reducibility results. Take transitivity and reflexivity again. Transitivity holds on all products. The construction used in the proof of theorem (3) automatically gives us ($\top \circ$ SOME) as a specification of the one and only (2)-reducible type $\langle 2 \rangle$ quantifier which behaves thus on products. (Of course, ($\top \circ$ SOME) = ($\top \circ g$) for arbitrary g.) Since TRANS $\neq$ ($\top \circ$ SOME), we know transitivity is not reducible. Reflexivity holds only on the product $E \times E$. The only reducible quantifier which has this property is $(f_1 \circ f_2)$, with $f_1(Q) = f_2(Q') = 1$ iff $Q = Q' = E$. This is indeed the reducible quantifier (ALL $\circ$ ALL), different from REFL. So reflexivity is not reducible. In general, if we are given the behaviour of a type $\langle n \rangle$ quantifier F^n on products, we now know what is the reducible type $\langle n \rangle$ quantifier G^n (if any) with which F^n has to be compared to establish reducibility of F^n.

Theorem (3) also helps us settle the matter about type $\langle 2 \rangle$ quantifiers such as SYMM. Products $(P \times Q)$ are symmetric iff $P = Q$ or one of them is empty. But certainly SYMM is not invariant for sets in products in SYMM: SYMM$(P \times P)$ = SYMM$(Q \times Q)$ = 1 while SYMM$(P \times Q)$ = SYMM$(Q \times P)$ = 0 if $\emptyset \neq P \neq Q \neq \emptyset$. Thus, SYMM and $\neg$SYMM are not invariant and theorem (3) tells us that there is no (2)-reducible type $\langle 2 \rangle$ quantifier with the same behaviour as (non-)symmetry on products. This explains why we can not use Keenan's theorem (1) to show symmetry not to be reducible. And it also explains why we do not at all need theorem (1) for that purpose. Theorem (3), or, rather, corollary (1), already tells us that it is not reducible. The generalization of Keenan's theorem presented in this squib may, thus, not only improve our understanding of it, but it also extends its application to other questions about natural language quantifiers.

Bibliography

Keenan, E. L.: 1992, Beyond the Frege boundary, *Linguistics and Philosophy 15*, 199–221

Learning Categorial Grammars from Semantic Types[1]

Daniela Dudău-Sofronie, Isabelle Tellier, Marc Tommasi
LIFL-Grappa Team

Abstract

This paper investigates the inference of Categorial Grammars from a new perspective. To learn such grammars, Kanazawa's approach consists in providing, as input, information about the structure of derivation trees in the target grammar. But this information is hardly arguable as relevant data from a psycholinguistic point of view. We propose instead to provide information about the semantic type associated with the words used. These types are considered as general semantic knowledge and their availability is argued. A new learning algorithm from types is given and discussed.

1 Introduction

Learning a Categorial Grammar from sentences it generates consists in identifying the categories assigned to each word of its vocabulary, when the rewriting schemas (i.e. Forward Application, Backward Application, etc) are supposed to be known.

The formal learnability of AB-Categorial Grammars in Gold's model has been deeply studied in [Kan96], [Kan98]. Kanazawa has proved that the class G_k of Categorial Grammars assigning at most k different categories with any given word is learnable in Gold's model from positive Structural Examples. When $k = 1$, the grammars are called "rigid" and they can be efficiently learned. When $k > 1$, the learnability is NP-hard [C.01]. But structural examples are very informative and are difficult to justify from a psycholinguistic point of view as input of a learning algorithm. The learnability of G_k also holds from texts (i.e. positive sequences of words) but is also intractable. We believe that one cannot avoid giving more knowledge to a learning process to get practical learning algorithms.

In the meantime we should emphasize that Categorial Grammars are well known for their formalized connection with semantics [Mon74], [DWP81], providing a good compromise between formalism and linguistic expressivity [OBW88]. The idea of "semantic bootstrapping" [Bre96] (semantic information inserted to help syntax reconstruction) seems a path not yet exploited enough in formal learnability researches. Links between Kanazawa's learning strategy and semantic information have been shown in [Tel99]. This first approach is still not satisfactory as it does not avoid combinatorial explosion.

Our purpose in this article is also to learn Categorial Grammars, but from a new kind of input data, more informative than texts and more relevant than Structural Examples. These data will consist of sequences of types associated with words and can be interpreted as semantic information, allowing to distinguish facts from entities and from properties satisfied by entities.

Section 2 contains preliminaries introducing the nature of the information admitted as input to the learning process and its links with logical semantics. Section 3 illustrates our learning strategy with a detailed example. The conclusion criticizes its limitations and argues about the cognitive plausibility of the data provided to this algorithm, in comparison with the data usually used in other learning algorithms. Perspective issues are also evoked.

2 Semantic information

2.1 A typing system

We will not define here a full semantic language, as types can be associated with very different kinds of semantic representation [Mon74]. For us, the only semantic information needed will be a typing system making a basic distinction between entities and facts and allowing to express the types of functors, among which are the predicates. This typing system is a an-intensional version of Montague's intensional logic. The set Θ of possible types is defined by:

- elementary types : $e \in \Theta$ (type of entities) and $t \in \Theta$ (type of truth values) are the elementary types of Θ;
- Θ is the smallest set including every elementary type and satisfying : if $u \in \Theta$ and $v \in \Theta$ then $\langle u, v \rangle \in \Theta$ (the composed type $\langle u, v \rangle$ is the type of functors taking an argument of type u and providing a result of type v).

[1] this work was supported by the ARC INRIA project GRACQ

These types combine following rewriting rules similar with the ones for classical Categorial Grammars.

- **TF** (Type Forward) : $\langle u, v \rangle . u \to v$;
- **TB** (Type Backward) : $u . \langle u, v \rangle \to v$.

Example 1. *For example, the types associated with some words (representing respectively an entity, one-place predicates, and a predicate modifier) may be:*

- *John:* e ;
- *man, woman, runs, walks:* $\langle e, t \rangle$;
- *fast:* $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$.

The main difference between syntactic categories of a Categorial Grammar and types in Θ is that the *direction* of a functor-argument application in categories is indicated by the operator used (/ or \), whereas this direction is lost in types (as both are replaced by a ","). The functor-argument application of types is commutative, whereas it is not for categories. Note that for a given type containing n elementary types, there are 2^n possible ways to replace each of its $n-1$ "," by either / or \.

Note also that, although we consider only two elementary types, Categorial Grammars include a non bounded number of elementary syntactic categories.

2.2 From logic to types

Before describing the learning process, let us slightly deepen the links between semantics and types. As a matter of fact it seems possible, under conditions, to deduce the type corresponding with a word from its meaning representation in a Montague-like system.

The typing system we use is well adapted to the language of first order predicate logic extended with typed-lambda calculus (corresponding with a simplified un-intensional version of Montague's intensional logic, without the temporal and modal operators).

Two classical syntactic rules that contribute to the construction of the set of meaningful well-formed expressions of type $\tau \in \Theta$, ME_τ are the following:

1. *λ-abstraction*: if $\alpha \in ME_\tau$ and u is a variable of type $b \in \Theta$ then $\lambda u.\alpha \in ME_{\langle b, \tau \rangle}$;
2. *application*: if $\alpha \in ME_{\langle a, b \rangle}$ and $\beta \in ME_a$ then $\alpha(\beta) \in ME_b$.

Example 2. *The semantic translations of some words in this logical language may be:*

- *John: John′ (the prime symbol distinguishes local constants from words)*
- *man:* $\lambda x.man'(x)$; *woman:* $\lambda x.woman'(x)$
- *runs:* $\lambda x.runs'(x)$; *walks:* $\lambda x.walk'(x)$
- *a:* $\lambda P.\lambda Q.\exists x[P(x) \wedge Q(x)]$
- *fast:* $\lambda P.\lambda x.(fast'(P))(x)$

Nevertheless, the reader should note that we need to impose some syntactic conditions in order to derive types from logical formulas. For the language evoked they include the following:

- symbols denoting predicates must be used *in extension*, i.e. displaying all their arguments: for example, a two-place predicate *loves′* will appear as $\lambda x.\lambda y.loves(x)(y)$;
- formulas associated with words must be closed, to avoid free variables and all lambdas and all quantifiers must bind variables that have at least a free occurrence in the sub-formula that follows. For example we forbid formulas like $\lambda P.\lambda y.\lambda x.\neg P(x)$ and also formulas like $\lambda x.\exists y[man'(x) \wedge runs'(x)]$ (where the variable y is not in the sub-formula that follows its binding by a lambda or a quantifier).

These two constraints refer in a certain manner to the "structure of the type"(the organization of couples of $\langle \rangle$). The number of λ symbols ahead a formula (in fact the number of application of a λ- *abstraction* rule) determinates the inner parenthesized description of the final type.

Example 3. *One λ symbol determines:* $\langle \Box, \Box \rangle$; *two λ symbols determines:* $\langle \Box, \langle \Box, \Box \rangle \rangle$, *etc.*

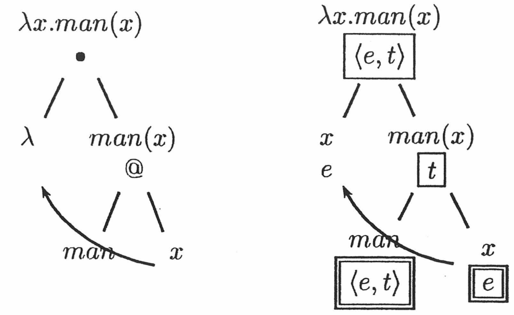


Figure 1: Type deduction example

As soon as the structure is discovered it remains to decorate it with types. Types of variables and constants are supposed to be known and thus one can label the leaves of the tree structure that represents the formula. Then the whole type of a word is evaluated by propagating contraints over types.

The first tree in the Figure 1 illustrates the decomposition of the semantic translation of the word *man* where • is a *λ-abstraction* rule and @ is a *application* rule. The type of the leaves in the double boxes are supposed to be known. We infer that $man(x)$ has the type t and the entire formula has the type $\langle e, t \rangle$.

3 A new learning algorithm from types

We adopt here *a semantic-based theory of syntax learning*, i.e. we consider that the capacity of acquiring a grammar is conditioned by the ability to build a representation of the situation described. In previous semantic-based methods of learning [HW75], [Hil83], [Fel98] word meanings are supposed to be already known when the grammatical inference mechanism starts. We make here the smoother hypothesis that the crucial information to be extracted from the environment is the semantic type of words. These types may be inferred from semantic information (as shown previously) or directly extracted from the environment.

The input of our learning algorithm is a sample of couples composed of a syntactically correct sentence and a corresponding sequence of types. The purpose is to build a rigid Categorial Grammar able to generate this sample. The missing information is the direction of the functor/argument applications, as each type of the form $\langle u, v \rangle$ can combine with a type u placed either on its left (with a **TB** rule), or on its right (with a **TF** rule). The identification of the operators / and \ will be processed in two steps: **a parsing step** and a step to **get categories from types.**

Parsing types We assume that every sentence given as input is syntactically correct and thus that the associated sequence of types can be reduced using the TB and TF rules to the type of truth values t. The first step of the algorithm is to find such a reduction.

Example 4.

a	*man*	*runs*
$\langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle$	$\langle e, t \rangle$	$\langle e, t \rangle$

This reduction can be viewed as a parse in a context free grammar. **TF** and **TB** are interpreted as schemes of rules instantiated by types in the input sequence. As every instantiated rule fitting one of these schemes is in Chomsky Normal Form, it is possible to adapt a standard parsing algorithm to find the parse on types. In this paper, we modify the Cocke-Younger-Kasami (CYK) algorithm. The whole process is the search for replacing the "," in type expressions to get categories. Possible values for "," are / , \ or "left unchanged" (a "," is left unchanged when the corresponding type is never in a functor place). To this aim, we identify every "," with a distinct variable. A reduction via **TF** or **TB** implies some equalities between categories and sub-categories that must be propagated.

Hence, type reductions translate into variable equalities and a given parse leads to a system of equalities.

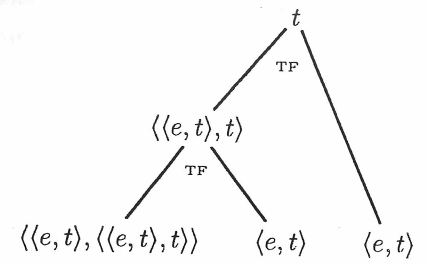


Figure 2: A parse tree for types

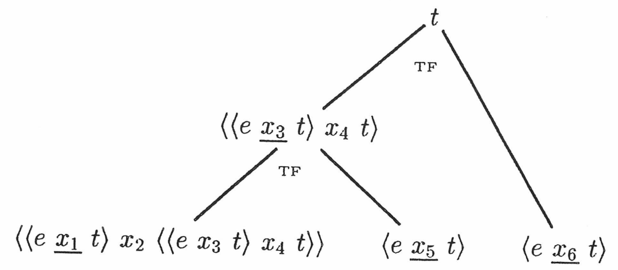


Figure 3: A parse tree underling equality between variables

Note that there could be more than one parse for a given input. The outcome of the parser is thus a set of parses each of which being described by a set of variable equalities.

Getting categories Now given such constraints on types, we easily deduce categories and hence a Categorial Grammar that accepts the sequence of words as input. To this aim, we consider that **TB** and **TF** have the classical rules in the categorial grammar formalism as a counterpart. The association between types and categories is done in the following way:

- t is associated with S.
- any distinct type that never occurs at a functor place in the parse is associated with a new variable category.
- If **TF**: $\langle ux_iv\rangle.u \to v$ and u (resp. v) is associated with the category A (resp. B) then $x_i = /$ and $\langle u\ x_i\ v\rangle$ is associated with the category B/A.
- If **TB**: $u.\langle ux_iv\rangle \to v$ and u (resp. v) is associated with the category A (resp. B) then $x_i = \backslash$ and $\langle u\ x_i\ v\rangle$ is associated with the category $A\backslash B$.

Thus, while getting categories, we add equality constraints of the form $x_i = /$ or $x_i = \backslash$ for some i . The process is always guaranteed to give a set of categories because of the properties of the sequences in input.

Learning process An on-line and incremental learning algorithm can be designed using the strategy describe above. It consists in inferring equality constraints between variables and equality constraints fixing the value of variables from each input couple (in the example below, both are inferred at the same time). We also take into account that the target Categorial Grammar is rigid, that is every word is associated with at most one category.

Example 5. *A first input is provided ...*

a	man	runs
$\langle\langle e\ x_1\ t\rangle x_2\langle\langle e\ x_3\ t\rangle x_4 t\rangle\rangle$	$\langle e\ x_5\ t\rangle$	$\langle e\ x_6\ t\rangle$

$$x_1 = x_5$$
$$x_2 = /$$
$$x_3 = x_6$$
$$x_4 = /$$

At this point, the Categorial Grammar inferred consists in the associations: a: $(S/B)/A$; man: A; runs: B.

Then the second input is provided and the set of variable equalities is enriched. Since every word has at most one category we consider the same variables in the type of words already encountered in a previous sentence.

a	woman	walks
$\langle\langle e\ x_1\ t\rangle x_2\langle\langle e\ x_3\ t\rangle x_4 t\rangle\rangle$	$\langle e\ x_7\ t\rangle$	$\langle e\ x_8\ t\rangle$

$$x_1 = x_7 = x_5$$
$$x_2 = /$$
$$x_3 = x_8 = x_6$$
$$x_4 = /$$

Consider a third input. The set of variable equalities becomes:

a	woman	walks	fast
$\langle\langle e\ x_1\ t\rangle x_2\langle\langle e\ x_3\ t\rangle x_4 t\rangle\rangle$	$\langle ex_7t\rangle$	$\langle ex_8t\rangle$	$\langle\langle e\ x_9\ t\rangle x_{10}\langle e\ x_{11}\ t\rangle\rangle$

$$x_1 = x_7 = x_5$$
$$x_2 = /$$
$$x_3 = x_8 = x_6 = x_{11} = x_9$$
$$x_4 = /$$
$$x_{10} = \backslash$$

Finally, if the last sentence is "John walks", equality constraints propagate x_8 to several categories.

John	walks
e	$\langle e\ x_8\ t\rangle$

.... $x_3 = x_8 = x_6 = x_{11} = x_9 = \backslash$ ...

To summarize, after renaming e in T and $\langle ex_1t\rangle$ in CN, the Categorial Grammar obtained consists in *John* : T; *man, woman* : CN; *runs, walks* : $T\backslash S$; *a* : $(S/(T\backslash S))/CN$; *fast* : $(T\backslash S)\backslash(S\backslash T)$. This grammar is able to recognize sentences like "John runs fast" and some generalization has thus occurred. In more complicated cases where more than one parse is possible on a given input, the algorithm must maintain a set of sets of constraints representing a set of candidate grammars.

4 Conclusion

We will discuss here the complexity of our approach and after that we will try to evidence our point in comparison with other works. In the end we will furnish the future directions of our research.

A parse in a context free grammar in Chomsky Normal Form can be computed in polynomial time in the size of the input. This complexity result is stated for a given context free grammar. In our setting, we do not have a context free grammar but a "set of possible ones" defined by schemes of rules. This changes a lot the complexity issues as illustrated by the following example:

Example 6. *Let us consider the following input:*

a ... a	b	a ... a
e ... e	$\langle e, \langle e, \ldots, \langle e, t\rangle\rangle\rangle$	e ... e

One can parse such a string in a context free grammar in polynomial time with respect to the size of the input sequence. But there exist C_n^{2n} different context free grammars compatible with this input. As a matter of fact the type associated with b expects $2n$ arguments among which n are on its right and n are on its left. Since our algorithm builds all possible grammars, it will run in exponential time with respect to n. Nonetheless, note that this n does not depend on the length of the input. It is related to the lexicon and the size of types associated with words.

A first idea to circumvent this combinatorial explosion is to try to build only one grammar among the set of possible ones. But something has to be done to get an incremental learning procedure and handle mind changes that may arise with new inputs. A second way of research is to find an economic and efficient manner to handle the representation of a set of grammars compatible with a given set of input sentences.

Learning a Categorial Grammar means associating categories with words. Categories are built from basic categories and operators. Learning categories from strings of words seems impossible in reasonable time. So, richer input data need to be provided.

The approach developed so far by Buskowsky & Penn and Kanazawa consisted in providing Structural Examples, i.e. giving the nature of the operators / and \ and letting the basic categories to be learnt. Our approach is dual, as it consists in providing the nature of the basic categories (under the form of types) and letting the operators to be learnt. But, we must emphasize that the mapping between categories and types is not perfect. For example, there exist some words like common nouns and intransitive verbs that initially receive the same type $\langle e, t\rangle$ corresponding with a one place predicate and that won't receive the same syntactic category at the end of the learning process because they have different combinatorial properties in English.

The algorithm we propose is able to identify in the limit any rigid Categorial Grammar. It still needs to be extended to k-valued Categorial Grammars. In fact, two sources of polymorphism should be distinguished: the case where a word must be associated with different semantic types (for example words like "and"), and the case where a word must be associated with different syntactic categories corresponding with the same semantic type (which must be the case for words like "a"). The first case should be handled as a natural extension of our learning strategy, where each semantically distinct instance of the word is treated as a new word, but the second case should be harder to deal with. Another extension to be explored is the case when only some of the types associated with words are known (for example, those associated with lexical words), whereas some others (those associated with grammatical words) remain unknown.

References

[Bre96] Michael R. Brent. *Computational Approaches to Language Acquisition.* MIT Press, 1996.

[C.01] Costa Florncio C. Consistent identification in the limit of any of the classes k-valued is np-hard. In *Logical Aspects of Computational Linguistics*, volume 2099 of *Lecture Notes in Artificial Intelligence*, pages 125–134. Springer Verlag, 2001.

[DWP81] D. R. Dowty, R. E. Wall, and S. Peters. *Introduction to Montague Semantics.* Linguistics and Philosophy. Reidel, 1981.

[Fel98] J. A. Feldman. Real language learning. In *ICGI'98, 4th International Colloquium in Grammatical Inference*, pages 114–125, 1998.

[Hil83] J. C. Hill. A computational model of language acquisition in the two-year-old. *Cognition and Brain Theory*, 3(6):287–317, 1983.

[HW75] H. Hamburger and K. Wexler. A mathematical theory of learning transformational grammar. *Journal of Mathematical Psychology*, 12:137–177, 1975.

[Kan96] Makoto Kanazawa. Identification in the limit of categorial grammars. *Journal of Logic, Language, and Information*, 5(2):115–155, 1996.

[Kan98] M. Kanazawa. *Learnable Classes of Categorial Grammars.* The European Association for Logic, Language and Information. CLSI Publications, 1998.

[Mon74] R. Montague. *Formal Philosophy; Selected papers of Richard Montague.* Yale University Press, 1974.

[OBW88] Richard T. Oehrle, Emmon Bach, and Deirdre Wheeler, editors. *Categorial Grammars and Natural Language Structures.* D. Reidel Publishing Company, Dordrecht, 1988.

[Tel99] I. Tellier. Towards a semantic-based theory of language learning. In *12th Amsterdam Colloquium*, pages 217–222, 1999.

VP Ellipsis by Tree Surgery

Katrin Erk and Alexander Koller
Saarland University, Saarbrücken, Germany

Abstract

We present *jigsaw parallelism constraints*, a flexible formal tool for replacing parts of trees with other trees. Jigsaw constraints extend the Constraint Language for Lambda Structures, a language used in underspecified semantics to declaratively describe scope, ellipsis, and their interaction. They can be used, among other things, to model certain classes of ellipses where existing approaches run into overgeneration problems.

1 Introduction

This paper defines *jigsaw parallelism constraints*, a generalization of the parallelism constraints found in the Constraint Language for Lambda Structures (CLLS [4]). They state that two trees must be isomorphic, except for regions exempt from parallelism, which can be specified very flexibly. They are a powerful tool for tree surgery: We can replace parts of trees by other trees, within the framework of a completely declarative formalism.

We apply jigsaw parallelism to VP ellipsis. Jigsaw parallelism can handle cases where the contrasting elements are complex parts of the sentence meanings (and not just NPs). Some of these cases are problematic for existing theories because the precise position of deleted material must be specified, as in the sentence

(1) Dan left, but George didn't.

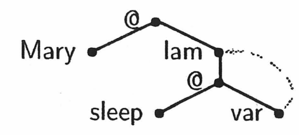
At the same time, our analysis of ellipsis can be used for the underspecified analysis of ellipses where contrasting elements partake in scope ambiguities. The fact that we may not know where they end up in the sentence semantics is a serious problem for traditional approaches, but is handled very naturally by ours.

We believe jigsaw parallelism has a wide range of further possible applications that we do not look at here; for instance, it can easily model the tree operations of TAG. For the purposes of processing, jigsaw parallelism can be compiled into *group parallelism constraints* [1], whose computational aspects are well investigated [6].

Plan of the paper. In Section 2, we sketch the definition of CLLS and the traditional analysis of ellipsis with parallelism constraints. In Section 3, we look more closely at the ellipsis problems we want to solve. Section 4 defines jigsaw parallelism constraints, and Section 5 shows how jigsaw constraints can be applied to ellipsis. We conclude in Section 6 and present some lines of future research.

2 The Constraint Language for Lambda Structures

In this section, we give the briefest possible overview of the Constraint Language for Lambda Structures, CLLS; for a more careful introduction (and clean definitions), see [4]. Generally speaking, CLLS is a logic for the partial description of *lambda structures* which can be used in an underspecified analysis of scope, ellipsis, and anaphora.



Lambda structures are node-labelled trees, enriched by two partial functions that map "bound" nodes to "binding" nodes: λ, which models variable binding, and ante, which models anaphoric reference. An example, encoding the λ-term $\mathsf{Mary}(\lambda x.\mathsf{sleep}(x))$, is shown above. Application is represented as the binary label @. Abstraction and bound variables are represented using the labels lam and var, and the λ-binding function (indicated by the dashed arrow) is used to indicate which variable is bound by which binder.

The syntax of CLLS is defined as follows. X, Y, etc. are variables which denote nodes in a lambda structure; A and B are explained below.

$$\begin{array}{lcl}\varphi & ::= & X{:}f(X_1,\ldots,X_n) \mid X\triangleleft^* Y \\ & \mid & \lambda(X){=}Y \mid \mathsf{ante}(X)=Y \\ & \mid & A \sim B \mid \varphi \wedge \varphi'\end{array}$$

A CLLS formula or *constraint* is a conjunction of atomic literals; it is satisfied by a lambda structure and a variable assignment iff all its literals are satisfied

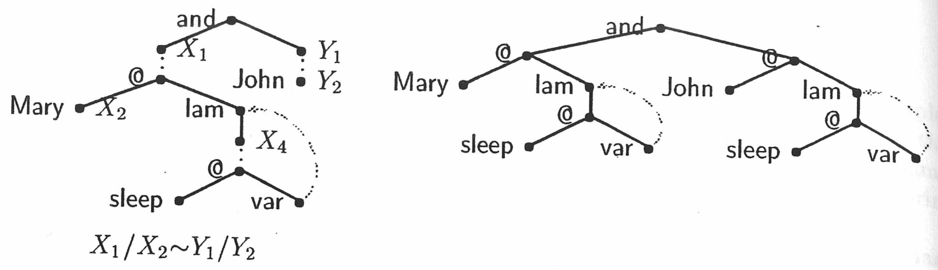
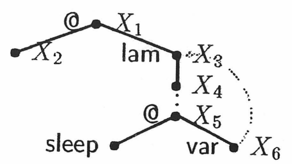


Figure 1: A parallelism constraint for (2) and a lambda structure that satisfies it.

in the following sense. A *labeling literal* $X{:}f(X_1, \ldots, X_n)$ is satisfied iff X denotes a node with label f and children that are denoted by $X_1, \ldots, X_n$. A *dominance literal* $X \lhd^* Y$ is satisfied iff X denotes a (reflexive, transitive) ancestor of Y in the lambda structure. The *binding literals* $\lambda(X){=}Y$ and $\mathsf{ante}(X){=}Y$ are satisfied iff the respective binding functions map the denotation of X to the denotation of Y. In the application, dominance constraints are used to model the outscoping relation of terms.

We draw constraints as *constraint graphs*, as shown to the right. The nodes in this graph stand for variables in a constraint and the edges and labels represent different types of literals. This particular graph represents a constraint that starts $X_1{:}@(X_2, X_3) \wedge X_4 \lhd^* X_5 \wedge \lambda(X_6){=}X_3 \wedge \ldots$ and, incidentally, is satisfied by the example lambda structure above.



The most complex, and for this paper the most important, literal is *parallelism* $A \sim B$. A and B are terms of the form $X/Y_1, \ldots, Y_n$, where $n \geq 0$. This term denotes a *segment* of the lambda structure, i.e. a subtree from which some further subtrees have been cut out. Formally it is a pair $\alpha = u/v_1, \ldots, v_n$ of a *root* $r(\alpha) = u$ and a sequence $hs(\alpha) = \{v_1, \ldots, v_n\}$ of *holes*, where u and the v_i are nodes of the lambda structure. We must have $u \lhd^* v_i$, but not $v_i \lhd^* v_j$ for all i, j. Segments specify parts of the lambda structure as sets of nodes *between* the root and the holes:

$$\begin{aligned} \mathsf{b}^-(\alpha) &:= \{w \mid r(\alpha) \lhd^* w, \text{ but not } v \lhd^* w \text{ for any } v \in hs(\alpha)\} \\ \mathsf{b}(\alpha) &:= \mathsf{b}^-(\alpha) \cup hs(\alpha) \\ \mathsf{i}(\alpha) &:= \mathsf{b}^-(\alpha) - \{r(\alpha)\} \end{aligned}$$

The difference is that $\mathsf{b}(\alpha)$ contains both the root and the holes; $\mathsf{b}^-(\alpha)$ contains the root but not the holes; and $\mathsf{i}(\alpha)$ (i for "interior") contains neither. Note that a segment with no holes specifies a subtree, and that for segments $\alpha = u/u$ ("singletons"), $\mathsf{b}^-(\alpha) = \emptyset$ and $\mathsf{b}(\alpha) = \{u\}$.

Two segments are considered parallel iff a *correspondence function* between them exists and certain conditions on binding hold, which we cannot go into here due to lack of space. A correspondence function relates pairs of nodes in parallel positions, requiring that corresponding nodes have the same labels and corresponding children. Thus they enforce structural identity of the two segments.

Definition 1 *A correspondence function between two tree segments α and α' is a bijective mapping $c : \mathsf{b}(\alpha) \to \mathsf{b}(\alpha')$ such that for all nodes $w \in \mathsf{b}^-(\alpha)$ and every label f of arity n,*

$$w{:}f(w_1, \ldots, w_n) \Leftrightarrow c(w){:}f(c(w_1), \ldots c(w_n)).$$

Parallelism constraints can be used in a semantic account of VP ellipsis, inspired by [3], but with tighter control over the reconstruction process and with an analysis of strict/sloppy ambiguities using anaphoric link chains [8]. By way of example, consider the sentence (2), which is analyzed as the left-hand graph in Fig. 1. A lambda structure satisfying this constraint is shown on the right.

(2) Mary sleeps. John does too.

The semantics of the source sentence is specified by the constraint graph below X_1; we want to constrain the substructure below Y_1 to the target semantics. Initially, we specify the contrasting element in the target sentence ("John" at Y_2). Then the two substructures are related by a parallelism constraint. This means there must

be a correspondence between the two segments, which can only exist if all nodes below X_1 have copies below Y_1. Except for X_2 (the hole), each copy must have the same label as the original. X_2 must correspond to Y_2, which has the label John.

Thus, parallelism constraints allow a very tight control over the copying process; different occurrences of Mary-labeled nodes would be kept strictly apart, and no equivalent of a primary occurrence restriction is needed. The analysis can deal with a large class of examples from the literature [4], e.g. interactions with scope and antecedent-contained deletion. Computationally, parallelism constraints are equivalent to context unification, whose decidability is open. Semi-decision procedures exist [6], and a linguistically useful decidable fragment has been identified [7].

3 The Problems

Now let us look at the problems we want to solve. For the first, consider Dalrymple et al.'s [3] (DSP, for short) account of VP ellipsis, which employs higher-order unification (HOU) and still has to count as the gold standard for semantical approaches to the problem. Among many other phenomena, they discuss sentence (3), which contains contrasting elements that are not simply NPs. Its semantics is shown (4); reconstruction of the ellipsis amounts to instantiating the higher-order variable P according to the unification problem (5).

(3) Dan didn't leave, but George did.

(4) $neg(left(dan)) \wedge P(george)(\lambda x.x)$

(5) $P(dan)(neg) = neg(left(dan))$

However, a problem occurs in the following slight variant.

(6) Dan left, but George didn't.

(7) $left(dan) \wedge P(george)(neg)$

(8) $P(dan)(\lambda x.x) = left(dan)$

The DSP account can indeed compute the correct meaning of the sentence, corresponding to the solution $P = \lambda x \lambda Q.Q(left(x))$ of the unification problem. But their approach does not preclude wrong solutions such as $P = \lambda x \lambda Q.left(Q(x))$. The core of the problem is that HOU performs silent β-conversions, to the effect that it can no longer easily distinguish between the different *occurrences* of $\lambda x.x$ in $P(dan)(\lambda x.x)$. The particular example could be saved by imposing well-typedness restrictions, but the general problem remains.

The current analysis of ellipsis in the CLLS framework does not have this problem, but it cannot provide a satisfactory analysis of examples like (6) either: It is limited to equality of complete segments, but sentences of this type require that parts of the parallel segments are exempt from parallelism. The problem could in principle be solved by using *group parallelism constraints* [1] and manually specifying all the parallel material explicitly, in a way similar to [9]. But this is certainly quite awkward, and can probably not be automated easily.

Another limitation (the second problem mentioned in the introduction) is that the segments must be known beforehand, which causes problems when the contrasting elements are involved in scope ambiguities:

(9) John went to the station, and every student did too, on a bike.

Fig. 4 shows a CLLS constraint graph which is an underspecified description of the meaning of (9). The target clause has two readings (of admittedly different availability): either "every student" outscopes "a bike" ($\forall\exists$), or the other way round ($\exists\forall$). As outscoping in a term is the same as dominance in a tree, the $\forall\exists$ reading corresponds to a solution that additionally satisfies $Y_2'' \lhd^* Y_3$; then the parallel segments in the target clause must be $Y_1/Y_2, Y_3$ and $Y_4'/$. In the $\exists\forall$ reading we have $Y_4' \lhd^* Y_2'$, so now the parallel segments are Y_1/Y_3 and Y_4'/Y_2. This shows that it may be necessary to enumerate all the (potentially many) readings of an ambiguous sentence before one can even specify the parallelism constraint. It would be much preferable to state and process the parallelism on an underspecified description such as Fig. 4, deriving an underspecified description of the target sentence efficiently.

4 Jigsaw Parallelism

To represent these examples, we generalize ordinary CLLS parallelism to *jigsaw parallelism.* It inverts the perspective on complex ellipses, or rather, sets it right: We specify a global parallelism, plus the parts that are *not* parallel, instead of listing all the separate pieces that are. We will now define jigsaw parallelism in three steps by lifting the notions of segment, correspondence, and finally parallelism to the new setting.

Jigsaw segments are sets of ordinary segments that are specified as subtractions of ordinary segments; intuitively, the jigsaw segment $\alpha - (\gamma_1, \dots, \gamma_n)$ is the result of "cutting out" the γ_i segments from α, leaving behind a set of smaller segments. (Hence the name "jigsaw segment".) This intuition is illustrated by Fig. 2 (a,b).

The simplest case is to cut out only a single segment γ from a bigger segment α. In this case, we get a set of segments with roots that are either $r(\alpha)$ or in $hs(\gamma)$, and holes that are either $r(\gamma)$ or in $hs(\alpha)$. Each root forms a segment with all the holes it dominates, unless there is another root in between. We say that segments α, β *overlap properly* iff $\mathsf{b}^-(\alpha) \cap \mathsf{b}^-(\beta) \neq \emptyset$ and write $u \triangleleft^+ v$ for *proper dominance*, i.e. the intersection of dominance and inequality.

Definition 2 *Let α, γ be segments of the same lambda structure. Then $\alpha - \gamma$ is a set of segments defined as follows.*
1. $\alpha - \gamma = \{\}$ if $\mathsf{b}(\alpha) \subseteq \mathsf{b}(\gamma)$.
2. $\alpha - \gamma = \{\alpha\}$ if either α and γ do not overlap properly, or α is a singleton with $\mathsf{b}(\alpha) \not\subseteq \mathsf{b}(\gamma)$.
3. For non-singleton α to which the first two cases do not apply, let

$$\mathsf{ro}(\alpha - \gamma) = (\{r(\alpha)\} - \mathsf{b}^-(\gamma)) \cup (hs(\gamma) \cap \mathsf{i}(\alpha))$$
$$\mathsf{ho}(\alpha - \gamma) = (hs(\alpha) - \mathsf{i}(\gamma)) \cup (\{r(\gamma)\} \cap \mathsf{i}(\alpha))$$
$$\mathsf{ch}(u, \alpha - \gamma) = \{v \in \mathsf{ho}(\alpha - \gamma) \mid u \triangleleft^+ v \text{ and } \nexists u' \in \mathsf{ro}(\alpha - \gamma) \text{ such that } (u \triangleleft^+ u' \triangleleft^+ v)\}$$

Then
$$\alpha - \gamma = \{u_0/u_1, \dots, u_n \mid u_0 \in \mathsf{ro}(\alpha - \gamma), u_1, \dots, u_n \text{ are the members of } \mathsf{ch}(u_0, \alpha - \gamma) \text{ ordered left to right}\}.$$

The definition can be extended to cut out multiple segments from α. Let $\alpha_1, \dots, \alpha_n, \gamma$ be segments of the same λ-structure such that for all $1 \leq i < j \leq n$, α_i and α_j do not properly overlap. Then $\{\alpha_1, \dots, \alpha_n\} - \gamma := \bigcup_{i=1}^{n} (\alpha_i - \gamma)$.

Definition 3 *Let $\alpha, \gamma_1, \dots, \gamma_n$ be segments of the same λ-structure. Then $\omega = \alpha - (\gamma_1, \dots, \gamma_n) := \Big(\big((\alpha - \gamma_1) - \gamma_2\big) \dots\Big) - \gamma_n$ is a jigsaw segment.*

We call the elements of ω *alpha segments* for short, and we call the excluded segments *gamma segments*. Also, we write $\mathsf{b}(\omega) = \bigcup_{\alpha' \in \omega} \mathsf{b}(\alpha')$. It can be shown that no two alpha segments of a jigsaw segment overlap properly, all are contained in α, and together with the γ_i, they cover all of α. Furthermore, the order in which gamma segments are subtracted does not matter.

The process of cutting out gamma segments is illustrated in Fig. 2. (a) is a schematic diagram of a segment α with two holes, from which segments $\gamma_1, \gamma_2, \gamma_3$ are being cut out. γ_1 overlaps only partially with α, and γ_3 is a singleton segment. If we compute the set $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \alpha - (\gamma_1, \gamma_2, \gamma_3)$ according to Def. 3, we obtain a picture as in (b): α is cut along the gammas. Note that adjacent segments generally share a node, which is the root of one and a hole of the other segment.

The way that the alpha and gamma segments are plugged into each other is represented in the *alpha-gamma tree* shown in (c). An alpha-gamma tree is a tree which contains exactly one node with label α_i for each i, at most one node with label γ_i for each i, and nodes with label $\circ$ for the holes of α. The children of each node are the segments plugged into the holes of the corresponding segment, in the correct left-to-right order. It can be shown that if the gamma segments do not overlap properly, such an alpha-gamma tree always exists. If they exist, alpha-gamma trees are unique up to permutations of equal singleton gamma segments.

Now we can use alpha-gamma trees to define *jigsaw correspondence functions*, correspondence functions between jigsaw segments. Two jigsaw segments correspond iff first, gamma segments with the same index are in the same positions in the alpha-gamma trees, and second, alpha segments in the same positions correspond in the ordinary sense.

Definition 4 *A* jigsaw correspondence function *between jigsaw segments $\omega = \alpha - (\gamma_1, \dots, \gamma_n)$ and $\omega' = \alpha' - (\gamma_1', \dots, \gamma_n')$ is a bijective mapping $c : \mathsf{b}(\omega) \to \mathsf{b}(\omega')$ which satisfies the following conditions:*

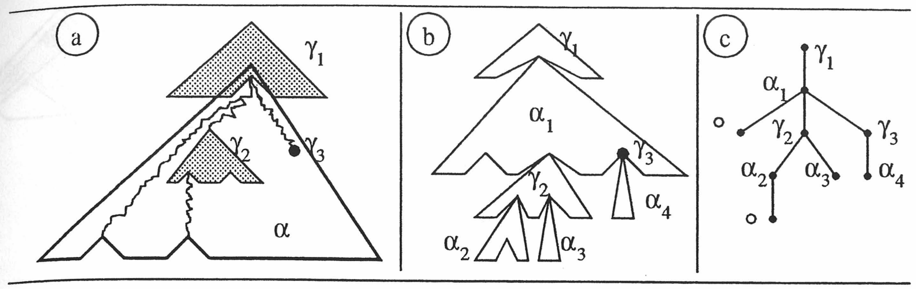


Figure 2: Jigsaw segments and alpha-gamma trees.

1. There are isomorphic alpha-gamma trees θ, θ' of ω and ω'; call the (tree) isomorphism h.

2. Every node u of θ is labeled γ_j iff $h(u)$ is labeled γ_j'; u is an alpha node iff $h(u)$ is; and u is labeled $\circ$ iff $h(u)$ is.

3. Let $\omega = \{\alpha_1, \dots, \alpha_m\}$. Then for every i, the restriction of c to $\mathsf{b}(\alpha_i)$ is an ordinary correspondence function between α_i and $h(\alpha_i)$.

Jigsaw parallelism is obtained by simply replacing the words "segment" and "correspondence function" in the definition of ordinary parallelism by "jigsaw segment" and "jigsaw correspondence function". We extend the syntax of CLLS by *jigsaw parallelism literals* in the straightforward way.

5 VP Ellipsis Using Jigsaw Parallelism

We will now apply jigsaw parallelism constraints to describe the meanings of (6) and (9).

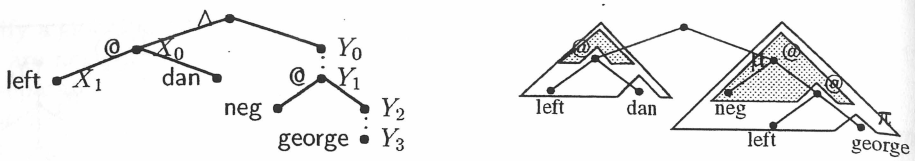
A constraint describing (6) is shown on the left-hand side of Fig. 3; the right-hand side of the picture displays the unique lambda structure satisfying this constraint. The constraint works in complete analogy to the introductory example in Fig. 1: We specify the source clause semantics below X_0 and the contrasting elements of the target clause (*neg* and *george*) below Y_0. Then we add a (jigsaw) parallelism constraint to force the subtree below Y_0 to represent the rest of the target semantics as well. This constraint expresses that the segment X_0/X_1 must have the same structure as Y_0/Y_3, with the exception of the two segments X_0/X_0 and Y_1/Y_2, which must occupy the same positions in X_0/X_1 and Y_0/Y_3, respectively, but can contain different material. Note that X_0/X_0 is a singleton segment, which contains no nodes, but still has a defined position.

In the solution, Y_0 and Y_1 have been mapped to the node μ, Y_3 has been mapped to π, etc. The two white segments contain the same nodes with the same labels. The gray segments may have different contents because they are exempt from the parallelism, but are at the same positions (namely, at the top) within the white segments. The nodes labeled *dan* and *george* are allowed to have different labels because they are not part of the parallel segments. Thus the position of the negation in the target sentence is fixed, and the overgeneration problem is avoided.

A constraint for (9) is shown in Fig. 4. As before, the constraint consists of an explicit description of the source clause, descriptions of the parallel elements in the target clause, and a (jigsaw) parallelism constraint relating source and target semantics. But now jigsaw parallelism allows us to exempt the two segments that constitute the meaning of "on a bike" from the parallelism directly, instead of waiting for the random placement of the quantifiers and exempting whichever ends up higher. This is done by subtracting the segment terms Y_3/Y_3' and Y_4/Y_4' on the right-hand side. The corresponding gamma segments on the left-hand side are simply new singletons.

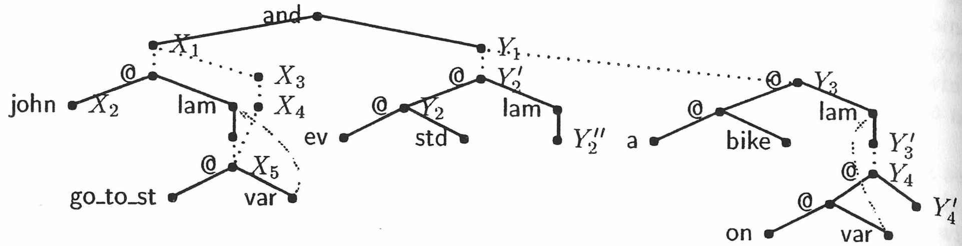
6 Conclusion

We have defined jigsaw parallelism constraints, a generalization of the original parallelism in CLLS which expresses that two segments of a lambda structure must be equal, except for subsegments which may be different but must occur at parallel positions. They were useful in an analysis of VP ellipsis because we could specify in the constraint where added material had to be inserted in the target semantics. In addition, the ability to exempt complete



$$X_0/X_1 - X_0/X_0 \sim Y_0/Y_3 - Y_1/Y_2$$

Figure 3: Constraint for sentence (6) and its solution.



$$X_1/X_2 - (X_3/X_3, X_4/X_4) \sim Y_1/Y_2 - (Y_3/Y_3', Y_4/Y_4')$$

Figure 4: Constraint for sentence (9).

subsegments from parallelism was useful when the relative positions of these subsegments was unknown due to underspecification.

We expect that jigsaw parallelism has uses outside of ellipsis. For instance, the TAG operations of substitution, adjunction, and sister adjunction can all be represented using jigsaw parallelism constraints. Adjunction, for example, is expressed by the constraint $\alpha - \gamma \sim \alpha' - \gamma'$ where α, α' are the complete trees before and after the adjunction, γ is a singleton segment at the adjunction site, and γ' is the tree that is to be adjoined.

The next question that needs to be considered is how to process jigsaw parallelism constraints. Jigsaw parallelism can be almost completely reduced to a *group parallelism constraint* [1] of the alpha segments; but there are some subtleties with binding which we have glossed over above. This should make it easy to obtain a sound and complete solution procedure from known procedures from group parallelism. But jigsaw parallelism avoids some hard instances of group parallelism and may thus be amenable to more efficient specialized techniques, similar to [2].

Finally, while we have shown how to *represent* some cases of ellipsis, we have said nothing about how to *obtain* these representations from a syntactic analysis. First steps towards this goal in the CLLS framework have been taken in [5]; we expect that the increased flexibility of jigsaw parallelism will simplify the interface design.

Acknowledgments. The authors are indebted to Markus Egg, Geert-Jan Kruijff, David Milward, and Bonnie Webber for fruitful discussions and nasty examples.

References

[1] M. Bodirsky, K. Erk, A. Koller, and J. Niehren. Beta reduction constraints. In *Proceedings of RTA*, Utrecht, 2001.

[2] M. Bodirsky, K. Erk, A. Koller, and J. Niehren. Underspecified beta reduction. In *Proc. 39th ACL*, Toulouse, 2001.

[3] M. Dalrymple, S. Shieber, and F. Pereira. Ellipsis and higher-order unification. *Linguistics & Philosophy*, 14:399–452, 1991.

[4] M. Egg, A. Koller, and J. Niehren. The constraint language for lambda structures. *J. Logic, Language, and Information*, 10(4), 2001.

[5] Markus Egg and Katrin Erk. A compositional account of VP ellipsis. In *8th Intl. Conference on HPSG*, Trondheim, 2001.

[6] K. Erk, A. Koller, and J. Niehren. Processing underspecified semantic representations in the constraint language for lambda structures. *Journal of Language and Computation*, 2001. To appear.

[7] Katrin Erk. *Parallelism Constraints in Underspecified Semantics.* PhD thesis, Saarland University, 2001. Forthcoming.

[8] Andrew Kehler. A discourse copying algorithm for ellipsis and anaphora resolution. In *Proceedings of EACL*, 1993.

[9] M. Schiehlen. *Semantikkonstruktion.* PhD thesis, University of Stuttgart, 1999.

Foundationalist Belief Revision in Update Semantics

Anthony S. Gillies
Department of Philosophy
The University of Texas at Austin

1. A Dilemma in Belief Revision

I want to introduce two intuitive distinctions in belief revision. First, we can differentiate theories of epistemic change by whether they are *semantic models* or *syntactic models*. Second there is a distinction between *foundations* or *foundationalist* belief change and *coherence* or *coherentist* belief change.

Syntactic models are models in which epistemic states are represented in part by a set of sentences in some language ("belief sets"), and belief revision is some kind of operation on such states. The AGM theory is a syntactic model since it takes epistemic states to be, at least in part, constituted by Cn-closed subsets of the language $\mathcal{L}$ of classical propositional logic (Alchourrón, *et al.*, 1985). Semantic models, on the other hand, are models in which epistemic states are represented as semantic objects (sets of possible worlds plus an ordering relation of plausibility, for example), and revision is some kind of operation on such states. The idea is that semantic models are models for changing an agent's representation of the world. Spohn's OCF model is a semantic model in this sense (Spohn, 1988). Semantic models have a certain *prima facie* plausibility, owing to both their rich structure (it is easy to represent modal properties of beliefs, even when beliefs themselves are restricted to plain facts) and their transparency.

Foundations theories of belief dynamics take seriously the idea that some beliefs are held on the basis of others—some beliefs (so-called "derived" beliefs) are held "just because" others are. Moreover, and more importantly, foundations theories insist that the structure of rational belief change is affected by the fact that some beliefs are held on the basis of others. The main task left for a foundations theory, then, is to give philosophical and formal analyses of the nature of the locution "just because", and say how it plays a role in how agents ought to revise their picture of the world. *Coherence* theories of belief change deny that a proper subset of an agent's beliefs are afforded any special status. The standard AGM theory is, again, a paradigm example of a coherence theory of belief dynamics (Alchourrón, *et al.*, 1985; Gärdenfors, 1992). Typically, foundations theories of belief change have taken the form of *belief-base* models (Hansson, 1989; Fuhrmann, 1991). Here the idea is that epistemic change operations are performed on a belief base—i.e., on a non-logically closed subset of an agent's beliefs.

From the point of view of someone who wants to explore rational belief change from within a semantic framework, the trouble is that, owing to an example originally due to Hansson (1989), we also want to have a foundations theory of belief change. And it is far from obvious how to get that in a semantic model. Suppose you walk down the main street in a town on a public holiday. You know that the town has just two cafés, A and B. In one scenario, imagine that you see Jones enjoying a coffee, from which you infer that at least one of the cafés is open—i.e., you believe $(A \vee B)$. In the distance you see evidence that in particular A is open (e.g., you see the lights on). Your belief set, call it K, is $\mathrm{Cn}(\{A, (A \vee B)\})$. Coming closer to A, you see a sign in the window: CLOSED ALL DAY FOR CLEANING. You must now revise with the fact that $\neg A$. In the resulting state, however, you should still believe $A \vee B$ since you had independent reason to think that one café was open. Contrast this with a scenario in which you do not see Jones (or anyone else) with a coffee. Then, in the distance, you see evidence that A is open. Of course, since A entails $(A \vee B)$, you also believe $(A \vee B)$. Your belief set, call it K', is $\mathrm{Cn}(\{A\})$. Coming closer to A, you see a sign in the window: CLOSED ALL DAY FOR CLEANING. You must now revise with the fact that $\neg A$. In the resulting state, you should not still believe $(A \vee B)$ since your only reason for believing $(A \vee B)$ was your belief in A. Clearly $K = K'$, but the two belief sets have different intuitive revision policies. The disjunctive belief, in the first scenario, is in some sense a "non-derived" belief whereas in the second it is "derived"—i.e., you believe $(A \vee B)$ *just because* you believe A.

The point of the example is that it suggests that this pre-theoretic difference makes a difference to the contours of rational belief change. The most common way of formalizing the derived–non-derived belief distinction is in terms of belief bases. For any given AGM belief set K, A is a base for K iff $\mathrm{Cn}(A) = K$. The idea in belief-base models is that epistemic change is a process of making changes directly to belief bases. Those changes then percolate out to the belief set at large. The belief-base approach has the advantage of offering a completely transparent analysis of the derived–non-derived belief distinction: a belief $\phi \in K$ is a derived belief relative to base A iff $\phi \in \mathrm{Cn}(A) \setminus A$. Revision of a belief set K, then, is always relative to a contextually salient base for K. Change the base and only the base to change the belief set (Hansson, 1989; Fuhrmann, 1991). It is easy to why such belief-base models get the facts right about Hansson-style café cases since in the first scenario the disjunctive belief is in the base, and in the second scenario it is not.

It is also no accident that such belief-base models are syntactic models. It is hard to see how semantic models could adequately represent the difference between "derived" and "non-derived" beliefs, and how these differ in belief change. Possible worlds are logically closed structures, so how is it that they could be used to represent the difference between believing, say, $\phi \vee \psi$ in a non-derived sense and believing it in a derived sense?

We seemed to be faced with a choice between either denying our semantic intuitions or else denying our foundationalist intuitions. But we would really prefer to dodge the choice. My goal here is to explore how we can dodge this choice by trying to accommodate our foundationalist intuitions in a semantic theory.

2. Update Semantics

Not just any dodge will do. Semantic theories of belief change tend to assume that possibilities are totally ordered by a relation of plausibility. Revision with respect to ϕ, then, amounts to adopting as new beliefs those formulas that are true at the most plausible ϕ-worlds. These measures of plausibility are the linchpin for semantic models of epistemic change. Assuming, rather than explaining, the existence and properties of such an ordering is on that basis undesirable. The accounts of belief change are driven by this structure over possibilities, and so it needs to be explained not assumed. Stipulating that possibilities are so ordered is

tantamount to hypothesizing a ghost in the machine of belief change. A satisfactory theory ought not rely on such ghosts.[1]

Intuitively, an agent judges one possibility as more plausible than another if the former obeys the sorts of commonsense inferences we routinely make and the second does not. Plausibility is a function of the default rules that an agent has accepted, and the comparative preferences thus acquired are encoded by an agent's expectation pattern. Veltman (1996) provides a dynamic semantic theory for representing default reasoning in update semantics (US) which has just this structure. My proposal is to adopt, plus or minus a bit, that basic semantic picture and use it as the basis for a theory of epistemic change.[2]

For simplicity, let us restrict attention to the case of general defaults—i.e., defaults of the form normally ϕ, for $\phi \in \mathcal{L}$. Let $\natural$ symbolize the normally operator, and let $\mathcal{L}_1$ be the smallest superset of $\mathcal{L}$ such that if $\phi \in \mathcal{L}$ then $\natural\phi \in \mathcal{L}_1$. Given a space W of possible worlds (functions from atoms to truth-values), we can then define epistemic states as follows:

Definition 1 (Epistemic States). $\kappa = (s, \varepsilon)$ is an epistemic state iff: (1) $s \subseteq W$; and (2) ε is a reflexive, transitive ordering over W such that some world is ε-minimal.[3]

If an agent is in state $\kappa = (s, \varepsilon)$, then s is her *acceptance base* and ε is her *expectation pattern*.[4] It is sometimes more convenient to write $w \preceq_\varepsilon v$ for $\langle w, v\rangle \in \varepsilon$, and to adopt the usual conventions about the asymmetric (strict) and symmetric (equivalent) parts.

There are two kinds of information that our agents can decide to accept. A model of epistemic change details in a precise way how epistemic states should change under the influence of the new information. First, an agent can come to accept a new default. In this case, revision should reduce to a refinement of her expectation pattern, just as in Veltman's US.

Definition 2. The refinement of ε by ϕ, $\varepsilon \bullet \phi$, is the pattern $\varepsilon' \subseteq \varepsilon$ such that $\langle w, v\rangle \in \varepsilon'$ iff: $\langle w, v\rangle \in \varepsilon$ and if $v \in [\![\phi]\!]$ then $w \in [\![\phi]\!]$.

Given an epistemic state κ, we would like a canonical way of assessing what an agent in κ is rationally committed to believing. The basic idea is simple enough: she ought believe ϕ just in case ϕ is the case in the most plausible worlds compatible with what she accepts. The formal definition adds the small wrinkle that we want an agent's commitments to be "skeptical" if she has competing defaults.

Definition 3. Let $\kappa = (s, \varepsilon)$ be an epistemic state. Then: (1) w is *optimal* in κ iff $w \in s$ and there is no $v \in s$ such that $v \prec_\varepsilon w$. (2) X is an *optimal set* in κ iff $X = \{v \in s : w \equiv_\varepsilon v$ for some w optimal in $\kappa\}$. (3) $\mathbf{M}_\kappa$ is the set of optimal sets in κ. (4) K is the *objective belief set based on* κ iff $K = \{\phi \in \mathcal{L} : \bigcup \mathbf{M}_\kappa \subseteq [\![\phi]\!]\}$.

[1] The force of this criticism applies equally to syntactic models of belief change, including belief-base models.

[2] It turns out that Veltman's information states for representing defaults can be given a purely epistemological motivation (Gillies, 2001).

[3] This last requirement is just to ensure that our agents don't accept an incoherent set of defaults, which just underscores our assumption that the sort of genuine revision we are interested in here is with respect to plain facts.

[4] Agents are assumed to start with the trivial expectation pattern $W \times W$; as they learn defaults they gain more structure via the update rules.

3. General Default Epistemic Change

The basic idea is that changes to an epistemic state can be neatly partitioned into those changes which bear on the expectation pattern, and those that bear on the acceptance base. With respect to the latter, a revision should be decomposable into a first a weakening of the state (a *downdate*), followed by a simple-minded strengthening (an *update*). Updating is trivial:

Definition 4 (Base Updating). Let s be an acceptance base and ϕ be any wff of $\mathcal{L}$. Then $s \uparrow \phi = s \cap [\![\phi]\!]$. If $\kappa = (s, \varepsilon)$, let $\kappa + \phi$ abbreviate $(s \uparrow \phi, \varepsilon)$.

What is needed is a construction of downdates. The basic idea for downdates—epistemic weakenings—is that they should in some sense be the minimal changes necessary to effect the desired change in the epistemic state. We can do this by considering slightly more fine-grained constituents of acceptance bases, and then exploit relations between them to construct a base downdate operator.

Definition 5. A *situation* is a non-empty partial function from atoms to truth-values. Given a situation x, the possible worlds based on x, $\mathbf{pw}(x)$, is the set $\{w \in W : x \subseteq w\}$.

Definition 6 (Remainder Sets). Let s be an acceptance base, and consider $\phi \in \mathcal{L}$. The *remainder set of situations* $s \bot \phi$ is the set of $\subseteq$-maximal situations x such that: (1) there is a $w \in s$ such that $x \subseteq w$; and (2) $\mathbf{pw}(x) \not\subseteq [\![\phi]\!]$.

Definition 7 (Base Downdate). Let s be any acceptance base, and consider any $\phi \in \mathcal{L}$. Then: (1) If $s \subseteq [\![\phi]\!]$, then $s \downarrow \phi = \{w : x \subseteq w \text{ some } x \in s \bot \phi\}$; (2) Otherwise, $s \downarrow \phi = s$. If $\kappa = (s, \varepsilon)$, let $\kappa - \phi$ abbreviate $(s \downarrow \phi, \varepsilon)$.

In the limiting case, your acceptance base is not included in $[\![\phi]\!]$. So, it is consistent with what you take yourself to know that $\neg\phi$. In this case, intuitively, the minimal change to your acceptance base s that does not commit you to accepting ϕ is to leave s as it is. That is what the second clause guarantees. Now, consider the principle case. Suppose you accept that ϕ. Then to downdate your acceptance base with respect to ϕ, you gather together all the possibilities which extend some situation in $s \bot \phi$.

Revision, then, comes to the composition of a downdate with an update (unless the agent is accepting a new rule).

Definition 8 (Revision). Let $\kappa = (s, \varepsilon)$ be an epistemic state. Then the general default revision function $\circ$ taking states and formulas of $\mathcal{L}_1$ to states is defined as follows: (1) If $\phi \in \mathcal{L}$, then $\kappa \circ \phi = (\kappa - \neg\phi) + \phi$; (2) Otherwise (i.e., if $\phi = \natural\psi$, some $\psi \in \mathcal{L}$), $\kappa \circ \phi = (s, \varepsilon \bullet \psi)$.

4. Basic Properties

Consider Hansson's café example. It is important, for the example to work as it was intended (namely, as a counterexample to the AGM theory), that you do *not* form the belief Jones is enjoying a coffee. To make sure that the belief sets are the same in both cases, we have to stipulate that you do not have any beliefs about the grounds or origins for your belief in $(A \vee B)$. The force of the example is that we have a clear intuitive judgment that the commitment to $(A \vee B)$ in the first case is *in some sense or other* basic or independent or non-derived, and that the commitment to $(A \vee B)$ in the second case is *in some sense or other* non-basic or dependent or

non-derived. The point is that this pre-theoretic difference makes a difference to the contours of rational belief change. And belief-base models are one way of cashing out how this goes. Belief-base models, however, are not unambiguously favored by Hansson's café example—one can accept the foundationalist intuition without resorting to belief bases.

There is a clear epistemic difference between *acceptances* and *expectations*. If open the blinds one morning and see that it is sunny, then I just *accept* that it is sunny. If, however, I don't open the blinds but instead rely on the default that I have learned normally it is sunny, then I still believe that it is sunny because I *expect* that it is. This intuition can be codified in a precise way as follows: in an epistemic state $\kappa = (s, \varepsilon)$ you accept that ϕ iff $s \subseteq [\![\phi]\!]$, you expect that ϕ iff it is not an acceptance but $\bigcup \mathbf{M}_\kappa \subseteq [\![\phi]\!]$. My proposal is that we can look at the difference between non-derived and derived beliefs as the difference between acceptances and expectations. If this is right, then Hansson-style cases do not unambiguously favor belief-base models (which are syntactic models of belief dynamics). And this suggests that our foundationalist intuitions about belief change are more general than many have thought.

Consider the first scenario: you walk into town and see Jones enjoying a coffee. From an epistemological standpoint, this ought not give you any *direct* or basic information about A or B being open. At most, the direct information you can get from this sort of information is that Jones is enjoying a coffee. But, as I pointed out earlier, we have to stipulate that you do not form this belief at all in order for the example to have its intended force. Pre-theoretically, seeing Jones enjoy a coffee gives you a defeasible expectation that one of the cafés is open. Since we are only allowing general defaults, this amounts to inducing the default $\natural(A \vee B)$. Let $W = \{w_1, \ldots, w_4\}$ and let A be the formula A is open and B be the formula B is open. Suppose your initial epistemic state is $\kappa_0 = (s_0, \varepsilon_0)$ where we have the following:

- $w_1(A) = w_1(B) = 1$; $w_2(A) = 1$, $w_2(B) = 0$; $w_3(A) = 0$, $w_3(B) = 1$; $w_4(A) = w_4(B) = 0$.
- $s_0 = W$
- $\varepsilon_0 = W \times W$

Observation 1. *Let κ_0 be as specified above. Then:*

(1) $\kappa_0 \circ \natural(A \vee B) = \kappa_1 = (s_0, \varepsilon_1)$ *where* $w_1 \equiv_{\varepsilon_1} w_2 \equiv_{\varepsilon_1} w_3 \prec_{\varepsilon_1} w_4$.
(2) $\bigcup \mathbf{M}_{\kappa_1} \subseteq [\![A \vee B]\!]$.
(3) $\kappa_1 \circ \neg A = (\{w_1, w_2), \varepsilon_1) = \kappa_2$.
(4) $\bigcup \mathbf{M}_{\kappa_2} \subseteq [\![A]\!]$.
(5) $\kappa_2 \circ \neg A = (\{w_3, w_4\}, \varepsilon_1) = \kappa_3$.
(6) $\bigcup \mathbf{M}_{\kappa_3} \subseteq [\![B]\!] \subseteq [\![A \vee B]\!]$.

In contrast to this, we have the following facts about the second scenario:

Observation 2. *Let κ_0 be as above. Then:*

(1) $\kappa_0 \circ A = (\{w_1, w_2\}, \varepsilon_0) = \kappa'$.
(2) $\bigcup \mathbf{M}_{\kappa'} \subseteq [\![A]\!] \subseteq [\![A \vee B]\!]$.
(3) $\kappa' \circ \neg A = (\{w_3, w_4\}, \varepsilon_0) = \kappa''$.
(4) $\bigcup \mathbf{M}_{\kappa''} \not\subseteq [\![A \vee B]\!]$.

The fact that, in the second case, you have no expectation about whether the cafés are open makes a difference. The difference is that you believed $(A \vee B)$ because you accepted as true that A was open. Giving that up forced you to give up the disjunctive belief. But in the first case, you had independent, though defeasible, grounds for thinking that $(A \vee B)$. Giving up that A is open, in the presence of this indirect information, does not force you to relinquish your belief that some café is open.[5]

Let me just end by pointing out a few of the properties of this model of revision in US. First, my revision function $\circ$ is not fully AGM compatible. Second, my downdating operator $\downarrow$ looks suspiciously like a possible worlds analogue of so-called "full meet contraction". The trouble with full-meet contraction is that it is too severe, removing everything from an agent's belief set K when contracting by ϕ excpet those items which are entailed by $\neg\phi$ (Alchourrón, *et al.*, 1985). But, happily, $\downarrow$ is not a full-meet operator in this sense. Third, $\circ$ takes full epistemic states to full epistemic states. This means that iterated revision poses no special problem. Finally, given the dynamic semantics base of the model, it is possible to introduce epistemic modalities and conditionals into the language of epistemic commitments, avoiding certain triviality results without completely re-thinking our model of epistemic change.

REFERENCES

Alchourrón, C., P. Gärdenfors, and D. Makinson (1985). "On the logic of theory change: Partial meet functions for contraction and revision," *Journal of Symbolic Logic* **50**: 510–530.

Fuhrmann, A. (1991). "Theory contraction through base contraction," *Journal of Philosophical Logic* **20**: 175–203.

Gärdenfors, P. (1992). "The dynamics of belief systems: Foundations vs. coherence theories," in C. Bicchieri and M. L. dalla Chiara (eds.), *Knowledge, Belief and Strategic Interaction* (Cambridge: Cambridge University Press).

Gillies, A. S. (2001). *Rational Belief Change*. Ph.D. dissertation, Department of Philosophy/Program in Cognitive Science: University of Arizona.

Grove, A. (1988). "Two modellings for theory change," *Journal of Philosophical Logic* **17**: 157–170.

Hansson, S. O. (1989). "New operators for theory change," *Theoria* **55**: 114–132.

Spohn, W. (1988). "Ordinal conditional functions: A dynamic theory of epistemic states," in W. L. Harper and B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics II* (Dordrecht: Kluwer Academic Publishers).

Veltman, F. (1996). "Defaults in update semantics," *Journal of Philosophical Logic* **25**: 221-261.

[5] I am not, to be clear, claiming that this is the only plausible analysis of the Hansson examples. I am rather aiming at showing that the intuitions he his pumping are in fact more general than they have been taken to be.

Type raising, continuations, and classical logic

Philippe de Groote
Inria-Lorraine

Abstract. There is a striking analogy between type raising, as introduced by Montague (1973), and the notion of continuation that has been developped in programming language theory in order to give compositional semantics to control operators (Stratchey and Wadsworth, 1974). In fact, this analogy is such that it is possible to see Montague's semantics as a continuation based semantics.

On the other hand, the notion of continuation allows classical logic to be given a Curry-Howard interpretation (Griffin 1990). In particular, the double negation law $((A \to \bot) \to \bot) \to A$ is provided with a computational content, which may be used to give a type logical interpretation of type lowering.

Putting the pieces of the picture together, it is possible to use "classical extensions" of the λ-calculus in order to express the semantic components of the lexical entries of Morrill's (1994) type logical grammars. This solution offers the advantage of not burdening the syntax by enforcing type raising to the worst case.

1 Type raising and continuations

Montague (1973) introduced *type raising* as a way of providing a compositional semantics to constructs that may give rise to scope ambiguities. Such constructs (typically, quantifiers) have semantic scopes that may be wider than their apparent syntactic scopes. Around the same time, computer scientists were trying to provide a compositional semantics to full jumps (i.e., 'goto' statements), which led to the discovery of *continuations* (Stratchey and Wadsworth, 1974).

Both problems are similar, and both solutions present striking similitudes. Montague's type raising is based on Leibniz's principle, which consists of identifying an entity with the set of its properties. Consequently, the type of entities e is replaced by $(e \to t) \to t$, where t is the type of propositions. In programming language theory, a *continuation semantics* (as opposed to a *direct semantics*) consists in providing the semantic function with the continuation of the program as an explicit parameter. Let P be a program, let $[\![-]\!]$ be the semantic function, and let s be some initial state. If we consider programs as state transformers, a direct semantics is such that $[\![P]\!]\, s \in \mathbf{State}$. On the other hand, a continuation semantics gives $[\![P]\!]\, s \in (\mathbf{State} \to \mathbf{State}) \to \mathbf{State}$. In fact, in both cases (type raising and continuation semantics), a type A is replaced by a type $(A \to O) \to O$, where O is the type of observable entities or facts.

2 Negative translations and classical logic

In the realm of the λ-calculus, the notion of continuation gave rise to the so-called CPS-transformations (Plotkin 1975). These are continuation-based syntactic transformations of the λ-terms that allow given evaluation strategies (typically, call-by-name or call-by-value) to be simulated.

For instance, Plotkin's call-by-value CPS-transformation is as follows:

$$\begin{aligned}
\overline{c} &= \lambda k.\, k\, c;\\
\overline{x} &= \lambda k.\, k\, x;\\
\overline{y} &= \lambda k.\, k\, (\lambda v.\, \lambda k.\, k\, (y\, v));\\
\overline{\lambda x.\, M} &= \lambda k.\, k\, (\lambda x.\, \overline{M});\\
\overline{M\, N} &= \lambda k.\, \overline{M}\, (\lambda m.\, \overline{N}\, (\lambda n.\, m\, n\, k))
\end{aligned}$$

Now, compare the following naive type logical grammar, where the lexical items are assigned a direct interpretation:

John	–	J	:	NP
Mary	–	M	:	NP
loves	–	$\lambda x.\, \lambda y.\, \text{LOVE}\, y\, x$	:	$(NP \setminus S) / NP$

together with the grammar, where the lexical items are assigned a Montague-like interpretation:

John	–	$\lambda k.\, k\, \text{J}$	:	NP
Mary	–	$\lambda k.\, k\, \text{M}$	:	NP
loves	–	$\lambda f.\, \lambda g.\, f\, (\lambda x.\, g\, (\lambda y.\, \text{LOVE}\, y\, x))$	:	$(NP \setminus S) / NP$

Again, the analogy between continuation and type raising is striking. The Montague-like interpretation may almost be seen as the call-by-value CPS-transform of the direct interpretation. This opens a new line of research that has been advocated in Barker's recent work (2000, 2001).

When applying a CPS-transformation to a typed λ-term, it induced another transformation at the type level (Meyer and Wand, 1985). For instance, the above CPS-transformation induces the following type transformation:

$\overline{\alpha} = (\alpha^* \rightarrow \bot) \rightarrow \bot$, where:

$$\begin{aligned}
a^* &= a, \quad \text{for } a \text{ atomic;}\\
(\alpha \rightarrow \beta)^* &= \alpha^* \rightarrow \overline{\beta}.
\end{aligned}$$

Griffin (1990) observed that these type transformations amount to double negative translations of classical logic into minimalist logic, and that it allows classical logic to be provided with a formulae-as-type interpretation. In this setting, the double negation law $((A \rightarrow \bot) \rightarrow \bot) \rightarrow A$, which corresponds to type lowering, is given a computational content by considering the absurd type $\bot$ to be the type of observable entities (this is radically different from the usual interpretation of $\bot$ as the empty type).

3 The $\lambda\mu$-calculus

Griffin's discovery gave rise to several extensions of the λ-calculus, which aim at adapting the Curry-Howard isomorphism to the case of classical logic. The $\lambda\mu$-calculus (Parigot 1992) is such a system.

The $\lambda\mu$-calculus is a strict extension of the λ-calculus. Its syntax is provided with a second alphabet of variables ($\alpha, \beta, \gamma, \ldots$ — called the μ-variables), and two additional constructs: μ-abatraction ($\mu\alpha.\, t$), and naming ($\alpha\, t$).

These constructs obey the following typing rules:

$$\frac{\alpha : \neg A \qquad t : A}{\alpha\, t : \bot} \qquad\qquad \frac{\begin{matrix}[\alpha : \neg A]\\ t : \bot\end{matrix}}{\mu\alpha.\, t : A}$$

Besides β-reduction:

$(\beta)\ (\lambda x.\, t)\, t \longrightarrow t[x := u]$

a notion of μ-reduction is defined:

$(\mu)\ (\mu\alpha.\, u)\, v \longrightarrow \mu\beta.\, u[\alpha\, t_i := \beta\, (t_i\, v)]$

where $u[\alpha\, t_i := \beta\, (t_i\, v)]$ stands for the term u where each subterm of the form $\alpha\, t_i$ has been replaced by $\beta\, (t_i\, v)$. It corresponds to the following proof-theoretic reduction:

$$\dfrac{\dfrac{\begin{matrix}\dfrac{\dfrac{}{\alpha : \neg(A \rightarrow B)}\,1 \qquad \begin{matrix}\vdots\\ t_i : A \rightarrow B\end{matrix}}{\alpha\, t_i : \bot}\\ \vdots\\ u : \bot\end{matrix}}{\mu\alpha.\, u : A \rightarrow B}\,1 \qquad \begin{matrix}\vdots\\ v : A\end{matrix}}{(\mu\alpha.\, u)\, v : B} \quad\longrightarrow\quad \dfrac{\begin{matrix}\dfrac{\dfrac{}{\beta : \neg B}\,1 \qquad \dfrac{\begin{matrix}\vdots\\ t_i : A \rightarrow B\end{matrix} \qquad \begin{matrix}\vdots\\ v : A\end{matrix}}{t_i\, v : B}}{\beta\, (t_i\, v) : \bot}\\ \vdots\\ u[\alpha\, t_i := \beta\, (t_i\, v)] : \bot\end{matrix}}{\mu\beta.\, u[\alpha\, t_i := \beta\, (t_i\, v)] : B}\,1$$

As well-known, classical logic is not naturally confluent. Consequently, there exist variants of the $\lambda\mu$-calculus that do not satisfy the Church-Rosser property (Parigot 2000). This is the case if we also consider the symmetric of the μ-reduction rule:

$(\mu')\ v\, (\mu\alpha.\, u) \longrightarrow \mu\beta.\, u[\alpha\, t_i := \beta\, (v\, t_i)]$

Finally, for the purpose of the example given in the next section, we also add the following simplification rules:

$(\sigma)\ \mu\alpha.\, u \longrightarrow u[\alpha\, t_i := t_i]$

which may be applied only to terms of type $\bot$.

4 Semantic recipes as $\lambda\mu$-terms

Dealing with a calculus that do not satisfy the Church-Rosser property is not a defect in the case of natural language semantics. Indeed, the fact that a same term may have several different normal forms allows one to deal with semantic ambiguities.

If we consider the sentential category S (or, semantically, Montague's type t) to be our domain of observable facts, the following typing judgement is derivable:

$$\dfrac{\dfrac{\dfrac{\dfrac{\text{PERSON} : e \rightarrow t \qquad \dfrac{}{x : e}\,0}{(\text{PERSON}\, x) : t} \qquad \dfrac{\dfrac{}{\alpha : e \rightarrow t}\,1 \qquad \dfrac{}{x : e}\,0}{(\alpha\, x) : t}}{(\text{PERSON}\, x) \supset (\alpha\, x) : t}}{\forall x.(\text{PERSON}\, x) \supset (\alpha\, x) : t}\,0}{\mu\alpha.\, \forall x.(\text{PERSON}\, x) \supset (\alpha\, x) : e}\,1$$

This allows the following type logical lexical entries to be defined:

everybody	–	$\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x)$	:	NP
somebody	–	$\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x)$	:	NP
loves	–	$\lambda x.\lambda y.\,\text{LOVE}\,y\,x$	:	$(NP \setminus S) / NP$

Then, the sentence

everybody loves somebody

has only one parsing, to which is associated the following semantic reading:

$$(\lambda x.\lambda y.\,\text{LOVE}\,y\,x)\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x))\,(\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x)).$$

This $\lambda\mu$-term may be considered as an underspecified representation. Indeed, its possible reductions yield two different normal forms:

$$\begin{array}{lr}
(\lambda x.\lambda y.\,\text{LOVE}\,y\,x)\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x))\,(\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x)) & \\
\longrightarrow (\lambda y.\,\text{LOVE}\,y\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x)))\,(\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x)) & (\beta) \\
\longrightarrow \text{LOVE}\,(\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x))\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x)) & (\beta) \\
\longrightarrow (\mu\beta.\forall x.(\text{PERSON}\,x) \supset (\beta\,(\text{LOVE}\,x)))\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x)) & (\mu') \\
\longrightarrow \mu\beta.\forall x.(\text{PERSON}\,x) \supset (\beta\,(\text{LOVE}\,x\,(\mu\alpha.\exists y.(\text{PERSON}\,y) \wedge (\alpha\,y)))) & (\mu) \\
\longrightarrow \forall x.(\text{PERSON}\,x) \supset (\text{LOVE}\,x\,(\mu\alpha.\exists y.(\text{PERSON}\,y) \wedge (\alpha\,y))) & (\sigma) \\
\longrightarrow \forall x.(\text{PERSON}\,x) \supset (\mu\alpha.\exists y.(\text{PERSON}\,y) \wedge (\alpha\,(\text{LOVE}\,x\,y))) & (\mu') \\
\longrightarrow \forall x.(\text{PERSON}\,x) \supset (\exists y.(\text{PERSON}\,y) \wedge (\text{LOVE}\,x\,y)) & (\sigma)
\end{array}$$

$$\begin{array}{lr}
(\lambda x.\lambda y.\,\text{LOVE}\,y\,x)\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x))\,(\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x)) & \\
\longrightarrow (\lambda y.\,\text{LOVE}\,y\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x)))\,(\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x)) & (\beta) \\
\longrightarrow \text{LOVE}\,(\mu\alpha.\forall x.(\text{PERSON}\,x) \supset (\alpha\,x))\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x)) & (\beta) \\
\longrightarrow (\mu\beta.\forall x.(\text{PERSON}\,x) \supset (\beta\,(\text{LOVE}\,x)))\,(\mu\alpha.\exists x.(\text{PERSON}\,x) \wedge (\alpha\,x)) & (\mu') \\
\longrightarrow \mu\alpha.\exists y.(\text{PERSON}\,y) \wedge (\alpha\,((\mu\beta.\forall x.(\text{PERSON}\,x) \supset (\beta\,(\text{LOVE}\,x)))\,y)) & (\mu') \\
\longrightarrow \exists y.(\text{PERSON}\,y) \wedge ((\mu\beta.\forall x.(\text{PERSON}\,x) \supset (\beta\,(\text{LOVE}\,x)))\,y) & (\sigma) \\
\longrightarrow \exists y.(\text{PERSON}\,y) \wedge (\mu\beta.\forall x.(\text{PERSON}\,x) \supset (\beta\,(\text{LOVE}\,x\,y))) & (\mu) \\
\longrightarrow \exists y.(\text{PERSON}\,y) \wedge (\forall x.(\text{PERSON}\,x) \supset (\text{LOVE}\,x\,y)) & (\sigma)
\end{array}$$

These correspond to subject and object wide scope readings, respectively.

5 conclusions

We have argued that Montague's type raising is a particular case of continuation. Consequently, continuation based formalisms, which have been developped in the context of programming language theory, may be used to deal with the sort of semantic ambiguities for which Montague invented type raising. Parigot's $\lambda\mu$-calculus is such a formalism, and we have shown how it may be used to cope with quantifier scope ambiguities. We claim that the $\lambda\mu$-calculus is particularly suitable for expressing compositional semantics of natural languages. For instance, it allows Cooper's (1983) storage to be given a type logical foundation. In fact, it allows a lot of dynamic constructs to be defined, which is of particular interest for discourse representation.

references

Barker, C. (2000) *Continuations and the nature of quantification* working paper, submitted for publication.

Barker, C. (2001) Introducing Continuations. In R. Hastings, B. Jackson, and Z. Zvolenszky, editors, *Proceedings of SALT 11*, CLC Publications, Ithaca, New York.

Cooper, R. (1983). *Quantification and Syntactic Theory*. Dordrecht: Reidel.

Griffin, T. G. (1990). A formulae-as-types notion of control. In *Conference record of the seventeenth annual ACM symposium on Principles of Programming Languages*, pages 47–58.

Meyer, A. and Wand, M. (1985). Continuation semantics in typed lambda-calculi (summary). In R. Parikh, editor, *Logics of Programs*, volume 193 of *Lecture Notes in Computer Science*, pages 219–224. Spinger Verlag.

Montague, R. (1973). The proper treatment of quantification in ordinary english. In J. Hintikka, J. Moravcsik, and P. Suppes, editors, *Approaches to natural language: proceedings of the 1970 Stanford workshop on Grammar and Semantics*, Dordrecht, Reidel. Reprinted: Montague, (1974, pages 247–270).

Montague, R. (1974). *Formal Philosophy: selected papers of Richard Montague, edited and with an introduction by Richmond Thomason.* Yale University Press.

Morrill, G. (1994). *Type Logical Grammar: Categorial Logic of Signs.* Kluwer Academic Publishers, Dordrecht.

Parigot, M. (1992). $\lambda\mu$-Calculus: an algorithmic interpretation of classical natural deduction. In A. Voronkov, editor, *Proceedings of the International Conference on Logic Programming and Automated Reasoning*, volume 624 of *Lecture Notes in Artificial Intelligence*, pages 190–201. Springer Verlag.

Parigot, M. (2000). On the computational interpretation of negation. In P.G. Clote and H. Schwichtenberg, editors, *Computer Science Logic*, volume 1862 of *Lecture Notes in Computer Science*, Springer Verlag.

Plotkin, G. D. (1975). Call-by-name, call-by-value and the λ-calculus. *Theoretical Computer Science*, 1:125–159.

Stratchey, C. and Wadsworth, C. (1974). Continuations a mathematical semantics for handling full jumps. Technical Report PRG-11, Oxford University, Computing Laboratory.

Even-NPIs in Questions

Elena Guerzoni, Department of Linguistics MIT

This paper presents an account of the rhetorical effect induced in questions by a certain subset of NPIs, the so called 'minimizers', which include idioms like *lift a finger* and *the faintest idea* but not *any*. It is known at least since Ladusaw (1979) that standard NPIs like *any* and minimizers like *lift a finger* are grammatical in questions, but affect their interpretation differently: While questions with *any/ever* can be used as unbiased requests of information, questions with minimizers are always biased towards a negative answer.

The analysis I will propose elaborates on Ladusaw's original appeal to general pragmatic principles linking the way a question is asked to the speaker's expectations concerning its answer. Specifically, I show that the rhetorical effect of questions with minimizers is a consequence of a presupposition which reduces the set of possible answers in the utterance context to the singleton containing the negative answer. The distinctive property of minimizers that accounts for this presupposition is as already proposed in Heim (1984) that minimizers contain a silent *even* (like Hindi NPIs, see Lahiri 1998[1]), *any* and *ever* do not (contra Lee & Horn 1994).

One crucial ingredient of the proposal is Wilkinson's 1996 scope theory of *even*. The paper shows that once the scope of *even* in a question is taken into account, the rhetorical effect follows from the semantics and pragmatics of questions. In addition, this view provides an account for the unexpected presupposition of questions containing minimizers and more generally of questions where *even* associates with expressions denoting the lower end-point of pragmatic scales.

1. A Brief Survey of the Facts

The questions in (1) can be used as unbiased information-seeking questions. Those in (2) come with what has been often described as a negative **rhetorical** flavor (Ladusaw 1979, Heim 1984, Wilkinson 1996, Han 1997).[2]

(1) a. Did *anybody* call?
 b. Has John *ever* been in Paris?

(2) a. Did anyone *lift a finger* to help you? — negative biased
 b. Does John have *the least bit of taste*? — negative biased
 c. Does Sue have *the faintest idea* of how hard I'm working? — neg. biased

[1] Importantly Lahiri points out that Hindi questions with NPIs are biased as well.
[2] Some scholars have identified negative rhetorical questions with negative assertions (Progovac 1993, Han & Siegel 1996). The use of this term here is meant to be a theory neutral way to indicate, following the above mentioned tradition, that the questions under consideration are felt to be biased towards the negative answer.

To the extent that questions as in (2) can be used to elicit information, they cannot be used to do so *disinterestedly* (Ladusaw 1979, Ch. 8, p.188). The presence of minimizers is felt to signal the speaker's expectation (/bias) for a negative answer. Borkin (1971) illustrates this point by showing that questions like those in (2) are infelicitous in contexts where the speaker is clearly unbiased as for what the true answer would be like, e.g. (3) below. Notice that their counterparts with *any* are, instead, fine.

(3) *Sue and I gave a party. Some friends helped organizing it. I'd like to thank all those who helped. I ask Sue*
 a. Did anyone give you any help?
 b. Did anyone lift a finger to help you?

2. *Even* in Questions

Besides involving an optionally hidden *even,* minimizers like *lift a finger* clearly denote the low end-point of the contextually relevant pragmatic scale (cf. Horn 1989, 399). Interestingly, the semantic effect of *even* in questions depends precisely on the position of its focus on the contextually relevant scale: when the focus is the lower end-point, the question has the same rhetorical flavor to it as questions with minimizers.

(4) a. Can you *even* [add 1+1]$_f$? — negative rhetorical
 b. Can you *even* [solve this very difficult equation]$_f$? — info seeking

(5) Can Sue *even* solve [problem 2]$_f$? — ambiguous
(i) < the most difficult problem, problem n,..., problem 2> — negative biased
(ii) < problem 2, problem n....,..., the easiest problem > — info seeking

Given this, an explanation of the rhetorical reading of (4a) and 5(i) will automatically extend to (2).

3. The Presuppositions of *Even*

It is uncontroversial that *even* doesn't contribute to the truth conditions of a sentence. I will take the contribution of this focus particle to be a scalar presupposition (ScalarP, henceforth).[3] Specifically, (6a) asserts (6b) and presupposes (6c):

(6) a. Mary can even answer [this difficult question]$_f$.
 b. Assertion (p): Mary can answer this difficult question.
 c. ScalarP: For any salient alternative x to this difficult question, it is MORE likely that M can answer x than that M can answer this difficult question i.e. *p is the **LEAST** likely among the alternatives*

If this scalar presupposition is a logical presupposition, the function of

[3] *Even* also introduces an existential presupposition which will be ignored here.

even in a declarative affirmative sentence like (6a) is to introduce partiality in meaning in the following way:

(7) $[\![even]\!] = \lambda C^{4}_{<st,t>}. \lambda p_{<st>}: \forall q_{<st>}[q \in C \;\&\; q?p \rightarrow q >_{likely} p]. p$

When we turn to negative sentences, however, *even* appears to introduce a different scalar presupposition. This is shown in (8).

(8) a. Sue cannot even add 1+1.
  b. Assertion (not p): Sue cannot add 1+1.
  c. ScalarP: For any alternative x to 'adding 1+1', that Sue can do x is LESS likely than that Sue can add 1+1. I.e. <u>*p is the **MOST** likely among the alternatives!*</u>

K&P (1977) and Wilkinson (1996) explain this phenomenon as a consequence of the scope of *even* with respect to negation (contra Rooth 1985, who attributes it to a lexical ambiguity of *even*): if *even* has wide scope, our lexical entry in (7) captures its presuppositions in these cases as well.

(9) a. LF: even [Sue cannot [add t $1+1]_f$]
  b. ScalarP: 'For every contextually relevant alternative x, that S canNOT do x is MORE likely than that S canNOT add 1+1, i.e. <u>*not p is the **LEAST** likely among the alternatives* ⇔ *p the **MOST** likely*</u>

In questions, finally, what presupposition *even* introduces appears to depend on the position of its focus on the contextually relevant scale (a problem first addressed in Wilkinson 1996). If the focus of *even* is the higher end-point, the question comes with a presupposition that is typical of affirmative utterances, **hardP** henceforth (this is shown in (10b)). This is expected as there is no negation in (10a). Surprisingly, when the focus of *even* is the low end-point, the presupposition is the one typically found in negative environments, **easyP** henceforth (see (10c)).

(10)a. Can Sue *even* solve [problem $2]_f$? ambiguous
  b. < problem 2, problem 5, problem 3, …,…,..,the easiest problem >
  ScalarP: For any alternative x, it is MORE likely that S can solve x than that S can solve Pr.2. <u>*p is the LEAST likely among the aternatives*</u> **(HardP)**

  c. < the most difficult problem, problem 3, problem 5, …, problem 2>
  Scalar P: For any salient alternative x to problem 2 it is LESS likely that Sue can solve x than that Sue can solve problem 2. I.e .<u>*p is the MOST likely among the alternatives*</u> **(EasyP)**

[4] C is the set of contextually salient alternative propositions (see Rooth (1996)).

Notice that there are two aspects to the 'ambiguity' of questions with *even*:

yes-no questions with *even* +	**A: Lower end-Point**	**B: Higher end-point**
Interpretation **Presuppositions**	Rhetorical **EasyP**	Info seeking **HardP**

Table 1

If the scope theory of *even* is correct the differences in table 1 should be the effect of a scope ambiguity. It's worth pointing out that stipulating a lexical **ambiguity** of *even* (Rooth1985) would predict **easyP** but not it's co-occurrence with the negative rhetorical effect. The next section shows that an explanation based on a single meaning for *even* (as in (7)) and the syntactic (scope) configurations in which it is interpreted accounts for both the presence of **easyP** and its cooccurrence with the rhetorical flavor.

4. Scope Ambiguity of Questions with *even*

In confronting the task of deriving **easyP** presuppositions in questions with *even* we can start by pointing out that **easyP** would be the presupposition of the negative answer, if *even* was present in this answer and had wide scope over negation (the opposite scope relation would generate instead **hardP**).

(11)a. Q: Can Sue even solve [problem $2]_f$
  b. A: No, Sue cannot even solve [problem $2]_f$
  c. LF1: even [NOT Mary can solve [problem $2]_f$] (even>not)
  ScalarP: not p is the **LEAST** likely among the alternatives ⇔ **easyP**
  d. LF2: NOT even [Mary can solve t_1 [problem $2]_f$] (not>even)
  ScalarP : p is the LEAST likely among the alternatives ⇔ **hardP**

The next question we need to address is how this ambiguity can be derived from the question in (11a) alone. The task is to derive this possibility from an ambiguity of the question itself.

In order to entertain this hypothesis we will need to make two assumptions here. The first assumption is that y/n questions always involve a hidden *whether,* which mirrors other *wh*-words in its syntax and semantics. In the spirit of a Karttunen-style semantics of *who*, *whether* denotes an existential quantifier (over functions of type $<st,st>$) with an implicit restrictor: the set containing the identity ($\lambda p.p$) and the negation ($\lambda p. \sim p$) functions. This amounts to saying that *whether* means *which of yes or no.*

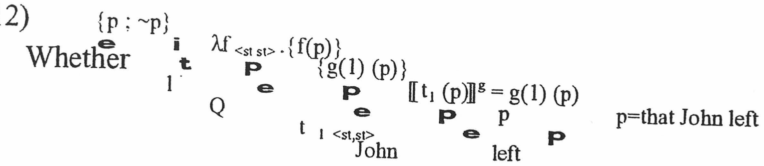
$[\![$*whether*$]\!] = \lambda f_{<<st,st>t>}. \exists h_{<st,st>} [h = \lambda p.p \text{ or } h = \lambda p.\sim p] \text{ and } f(h)=1$[5]

Whether moves over the set-creating Q morpheme, leaving a trace of

[5] Compare with the Karttunen-style meaning for *who*: $[\![who]\!] = [\![which\ person]\!] = \lambda P_{<et>}. \exists x_e$ s.t. x is a person and P(x)=1

type <st,st> in its base position:

(12)



The second assumption is that *even* can have narrow or wide scope relative to the trace of *whether*. The two LFs of (11), are thus (13a) and b.

(13) a. [Whether$_1$ [Q [**t$_1$** [*even* [Mary solved [problem 2]$_f$]] — $t_{whether}$>even

b. [Whether$_1$ [Q [*even* [**t$_1$** [Mary solved [problem 2]$_f$]] — even> $t_{whether}$

The denotations of these two LFs are different in that they contain different partial propositions (see also appendix):

(14) Let's p be 'that M. solved Pr2'   a. ⟦13a⟧= {⟦even⟧(p) ;~ ⟦even⟧ (p)}
b. ⟦13b⟧= {⟦even⟧ (p) ; ⟦even⟧ (~p)}

Returning now to the negative answers, the scope of even and the presuppositions will be as specified in (15).

(15) a. ~ ⟦even⟧ (p) is the *no* answer to (13a), since not>even → **hardP**
b. ⟦even⟧ (~p) is the *no* answer to (13b), since even>not → **easyP**

5. The Disambiguating Function of the Context

The utterance context has the important function of providing the information as to how high on a pragmatic scale the denotation of the focused expression is ranked with respect to the relevant alternatives.

We can make some speculations about how this affects the interpretation of a question containing *even* in a given context. Since answers with false presuppositions are presupposition failures, it is reasonable to assume that a speaker uttering a question in a context c is biased towards those answers whose presuppositions are satisfied in c. If Q is the Hamblin-set denoted by a question, let's call Q/c the subset of Q of the possible answers the speaker is presenting as live alternatives in a context c, i.e. those answers that are defined in c. (See Heim (2001)).

Now suppose that a question like (11a), repeated below, is uttered in a context in which Problem 2 is very easy to solve:

(11) Can Mary solve even Problem 2?
C:: < the most difficult problem, ….,…,.. problem 2 >

In C, **hardP** is false and **easyP** true, a situation that given our considerations above, has two consequences. The first consequence is that

reading (13a) (trace $_{whether}$> even) repeated here as (16a) is absent in C because all its answers would be infelicitous there.

(16)a. ⟦Whether$_1$Q **t$_1$** *even* M. solved [Pr2]$_f$]] ⟧ = {⟦even⟧(p) ; ~ ⟦even⟧(p)}
Since *Yes* presupposes **hardP**, ⟦*yes*⟧ ∉ ⟦16a⟧ /C
Since *No* also presupposes **hardP**, ⟦*no*⟧ ∉ ⟦16a⟧ /C → ⟦16a⟧/C = ∅

The second consequence is that, only the negative answer to (13b) (repeated in 16b) is possible according to the speaker's expectations.

(16)b. ⟦Whether$_1$ Q *even* **t$_1$** M. solved [Pr 2]$_f$]⟧ ={⟦even⟧(p) ; ⟦even⟧(~p)}
Since *Yes* presupposes **hardP**, ⟦*yes*⟧ ∉ ⟦16b⟧ /C
Since *No* presupposes **easyP**, ⟦no⟧ ∈ ⟦16b⟧/C → ⟦16b⟧/C ={⟦even⟧(~p)}

This is why a question containing *even* is felt to be biased towards a negative answer in contexts where the focus of *even* is the lowest scale end-point. In addition, as the singleton of the possible answers in these contexts contains the answer presupposing **easyP** the question unambiguously presupposes **easyP** in them.

In light of the proposal that minimizers contain *even*, this situation also explains, why questions with minimizers are always negative biased, since minimizers denote the lower end-point of the scale in every context.

6. Conclusions

This work provides a unified perspective on two puzzling properties of questions with minimizers: A rhetorical effect and an unusual presupposition. Adopting Heim's (1984) hypothesis that NPIs of this variety contain a hidden *even* I argued that the two above properties follow from: the scope theory of *even* and the pragmatics of questions.

Selected References:

Borkin, Ann (1971). Polarity Items in Questions. CLS
Heim, I (1984). A Note on Negative Polarity and DE-ness. In *Nels 14.*
Heim, I. (2001). Class Notes for the seminar "Topics in Semantics", fall 2001, MIT.
Han,C.&S,L(1996). Syntactic and Semantic Conditions on NPI Licensing in Questions.*WCCFL 14.*
Han, C. (1998). Deriving Interpretation of Rhetorical Questions. *Proceedings of WCCFL 16.*
Horn, L.(1989). *A Natural History of Negation.* The University of Chicago Press..
Ladusaw (1979). *Polarity Sensitivity as Inherent Scope Relation.* Dissertation, Austin.
Lahiri, U.(1998).Focus and Negative Polarity in Hindi. in *NLS 6.1.*
Progovac, L (1993). Negative Polarity and Grammatical Representation. *L&P 10.*
Wilkinson, K. (1996). The scope of Even. *NLS 4.2.*

Appendix:

(13a) Trace$_{whether}$>EVEN	(13b) Even>Trace$_{whether}$
{⟦even⟧(p) ; ~(⟦even⟧(p)}	{⟦even⟧(p) ; ⟦even⟧ ~(p)}
t y λf$_{<st,st>}$.{f (⟦even⟧ (p))}	t y λf$_{<st,st>}$.{ ⟦even⟧ (f (p))}
Whether t y {g(1) (⟦even⟧) (p)}	Whether t y {⟦even⟧ (g(1)) (p)}
1 t y g(1) (⟦even⟧) (p)	1 t y ⟦even⟧ (g(1)) (p)
Q t y ⟦even⟧ (p)	Q t y (g(1)) (p)
t$_{1<st,st>}$ t y	even t y
even p	t$_{1<st,st>}$ p

Complementary preference overruled: stressed pronouns in discourse

Helen de Hoop

The use of anaphoric elements such as pronouns has its source in balancing between minimization of effort (for the speaker's sake) and maximization of expression (for the hearer's sake). All languages in the world appear to have personal pronouns, but they come in different forms, for instance full versus reduced ones or free versus bound ones. In languages that have both reduced and nonreduced pronouns, the reduced ones are specialized for anaphoricity, the nonreduced ones have focus functions (cf. Bresnan 2001). The general correlation between reduced form and anaphoricity (or topic-continuity) is observed by Givón (1984) under the name 'referential iconicity'. Haiman (1985) argues that this correlation can be explained by economy: the effort of the speaker can be minimized by reducing expressions of anaphoric (hence, predictable) referents. In languages that do not have different types of pronouns, the interaction with prosody gives the same result: unstressed pronouns need less effort, hence they are specialized for anaphoricity, while the stressed ones have focus functions. However, stress is used for different reasons in language (new information, contrast, shift in reference) and it is not always clear what principles guide a hearer's interpretation of a stressed or unstressed pronoun in a certain context. What are the different types of constraints that play a part in (un)stressed pronoun resolution and how do these interact? In this paper the interpretation of stressed pronouns in context will be analyzed in an optimality theoretic fashion.

1. Stressed pronouns: deictic or anaphoric?

Semantically, pronouns are usually associated with variables, which are essentially free. For instance, if a speaker doesn't know the name of a person she wishes to refer to, she might point to this person and say:

(1) She/SHE doesn't have a hand-out yet.

In most cases, however, pronouns are anaphoric. Anaphoric pronouns are usually topics or part of the background. As such, the variables they are associated with refer to familiar objects in the discourse. In these cases there is usually an antecedent in the linguistic context to which the pronoun is anaphorically linked. Anaphoric pronouns are not necessarily destressed. They may be stressed, in which case the stress does not indicate deixis, but rather contrastive focus. An example is Vallduví's (1990) dialogue in (2) [stress indicated by capitals]:

(2) S1: Good morning. I am here to see Mrs. Bush again.
S2: Sure, Mr. Smith. Let's see... One of her assistents will be with you in a second.
S1: Could I see HER today? I'm always talking to her assistents.

In (2) the conversational implicature evoked by the focus is that S1 does not want to see one of Mrs. Bush's assistents again. In contrast, S1 wants to see HER/Mrs. Bush herself. Constituent focus evokes contrast within a contextually salient set of alternatives. This gives rise to the implicature that a certain predicate which holds for the focussed object does not hold for the other elements of the set of alternatives (cf. Rooth 1992).

2. Stressed pronouns: complementary preference?

Kameyama (1994) claims to have a unified account of interpretation preferences of stressed and unstressed pronouns in discourse. The central intuition is expressed as the "Complementary Preference Hypothesis" taking the interpretation preference of the unstressed pronoun as the base from which to predict the interpretation preference of the stressed pronoun in the same discourse position. For example, Kameyama discusses the following two famous sequences:

(3) Paul called Jim a Republican. Then he insulted him. (Paul insulted Jim.)
(4) Paul called Jim a Republican. Then HE insulted HIM. (Jim insulted Paul.)

As Prince (1981) notes, because HE and HIM are marked as being new, they cannot refer to Paul and Jim respectively. Being pronouns, they do, however, refer to some entities already in the discourse model, so here only Jim and Paul are left, in that order. Thus, we get a change in coreference. On the basis of these examples, Kameyama claims there to be a systematic relation between the stressed and unstressed counterparts which is of a complementary preference within a suitable subset of the domain. The assumption is that stressed and unstressed counterparts have exactly the same range of possible values. However, in the following sequence we get the same stress pattern as in (4), despite the fact that the two pronouns do not have the same range of possible values (cf. Prince 1981):

(5) Paul called Jane a Republican. Then SHE insulted HIM.

The stress on the pronouns in (3)-(5) seems to be the result of the topic-focus structure of the sentence, rather than of a shift in preferred reference. Note furthermore that if the stress is on a different constituent, the pronouns can remain unaccented:

(6) Paul called Jane a Republican. Then she HIT him.

3. Stressed pronouns: complementary preference not possible

In the novel 'Fire from heaven' by Mary Renault I found 50 examples of stressed pronouns, indicated by the author by means of italics [replaced by capitals by me, HdH]. For the vast majority of these examples, it can be argued that the stress signals contrastive focus; none of these examples can be accounted for by the Complementary Preference Hypothesis (alone). In the following two examples, the stress cannot be accounted for by the Complementary Preference Hypothesis as there are no alternative values for the first and second person pronoun:

(7) 'Gold, my boy, gold is the mother of armies. I pay my men round the year, war or no war, and they fight for ME, under my officers.' (ME = the speaker)

(8) 'Well, it teaches you to bear your wounds when you go to war.'
'War? But you're only six.'
'Of course not, I'm eight next Lion Month. You can see that.'
'So am I. But YOU don't look it, you look six.' (YOU = the addressee)

Another example where Kameyama's Complementary Preference Hypothesis (KCPH) cannot apply is (9) with a stressed third person pronoun:

(9) 'Of course,'he said. 'I shall kill Attalos as soon as I can do it. It will be best in Asia.' Hephaistion nodded; he himself, at nineteen, had long lost count of men he had already killed. 'Yes, he's your mortal enemy; you'll have to get rid of HIM. The girl's nothing then, the King will find another as soon as he's on campaign.'(HIM = Attalos)

In the dialogue in (9) there is only one male individual in the third person, namely Attalos. Yet, he is referred to by a stressed pronoun. In all the cases above the stress signals contrastive focus. In (7) the implicature is that the men work for the speaker, not for anybody else. In (8) the contrast is between the speaker and the addressee: the first one looks eight years old, but the second one only looks six. In (9), finally, the contrast is between Attalos and the girl. That is, the implicature of the sentence is that the addressee will only have to get rid of Attalos and not of the girl.

4. Stressed pronouns: complementary preference overruled

In (9) above the Complementary Preference Hypothesis cannot be applied as there is no alternative referent available for the unstressed counterpart of the pronoun. Other cases provide direct evidence against the hypothesis. Consider for instance one text fragment with two instances of stressed pronouns:

(10) 'So, think which of them can't afford to wait. Alexander can. Philip's seed tends to girls, as everyone knows. Even if Eurydike throws a boy, let the King say what he likes while he lives, but if he dies, the Macedonians won't accept an heir under fighting age; HE should know that. But Olympias, now, that's another matter. SHE can't wait.' (HE = Philip = the King; SHE = Olympias)

In (10) there are two referents available for the masculine pronoun (namely, Alexander and Philip) and two for the feminine pronoun (*viz.*, Eurydike and Olympias). The reader may verify that Kameyama's hypothesis would predict *HE* to refer to Alexander and *SHE* to Eurydike. Neither prediction is borne out. In other words, the stressed pronouns do not indicate a shift in reference. Instead, the stress evokes conversational implicatures: (i) other people might not know that the Macedonians won't accept an heir under fighting age; (ii) Alexander can wait. These are once again relevant implicatures in the context.

A theory that tries to derive the interpretation of anaphoric expressions from constraint interaction is Optimality Theoretic Semantics (cf. Hendriks and De Hoop 1997, 2001). In this theory each utterance is associated with an in principle infinite number of interpretations. Hearers arrive – as fast as they do – at one or two optimal interpretations of the utterance by evaluating the candidate interpretations with respect to a set of (conflicting) constraints. The interpretation that arises for an utterance within a certain context maximizes the degree of constraint satisfaction and is as a consequence the best alternative (hence, optimal interpretation) among the set of possible interpretations.

The optimal interpretations that are assigned to stressed pronouns in discourse can be analyzed in terms of three ranked constraints. The constraints are formulated below.

(11) *Contrastive Focus (CF)*: A stressed pronoun signals contrast within a contextually determined set of alternatives;
(12) *Continuing Topic (CT)*: A pronoun is interpreted as a continuing topic;
(13) *Kameyama's Complementary Preference Hypothesis (KCPH)*.

The optimal interpretations for the stressed pronouns in (4)-(10) follow from the ranking CF >> CT >> KCPH, as shown in the tableau.

(14) Constraint tableau for the interpretations of (4), (9) and (10).

Input		Output	Contrastive Focus	Continuing Topic	KCPH
(4)		HE=Paul; HIM=Jim	*		*
	☞	HE=Jim; HIM=Paul		*	
(9)	☞	HIM=Attalos			*
		HIM≠Attalos		*	
(10)	☞	HE=Philip			*
		HE=Alexander		*	
	☞	SHE=Olympias			*
		SHE=Eurydike	*	*	

In other words, Kameyama's Complementary Preference Hypothesis is neither sufficient nor necessary for a proper analysis of the use of stressed pronouns in discourse. In the examples discussed above, KCPH is only satisfied in the famous example (4). This is possible because in this example satisfaction of the KCPH corresponds with satisfaction of Contrastive Focus, which in turn is stronger than the constraint Continuing Topic that is violated by the winning candidate interpretation. If Paul called Jim a Republican and then Paul insulted Jim, stress on the pronouns would not be licensed by Contrastive Focus. However, if the stressed pronouns refer to Jim and Paul in the continuation, Contrastive Focus is satisfied. The contrast is between the two events: first Paul insulted Jim and then the other way around. This is in general the implicature that we get when two pronominal arguments of a predicate are stressed in the absence of further context. This is also what happens in the example (15), a translation from a Dutch newspaper fragment:

(15) SYDNEY – In the train returning to town from the Olympic Park, Dutch fans sang "Inge is OK, ole ole ole, Inge is OK, ole ole ole..." on Thursday night. Prince Willem Alexander, who congratulated Inge de Bruijn on her second olympic title with a kiss one hour earlier in the Aquatic Centre, thinks so too.
During the press conference following her victory in the 100 meter freestyle, De Bruijn spoke about the one moment in Sydney in which she did NOT take the initiative: "No, no, HE kissed ME."

In (15) the KCPH cannot explain the use of the stressed pronouns because there are no complementary discourse values available. We do get an interpretation via Contrastive Focus, however, comparable to the interpretation of (4): the implicature is that Willem-Alexander kissed Inge and not the other way around.

Clearly, an analysis in terms of complementary preference of available discourse values does not account for the subtle interaction of conditions dealing with discourse relations, prosody, and topic-focus structure. I conclude that this interaction is better analyzed in terms of constraint satisfaction (see also Beaver 2000, for an Optimality Theoretic approach to pronoun resolution in Centering Theory).

5. Semantic variation

It was shown above that in English Kameyama's principle is a rather weak one, the effect of which only show under specific circumstances. Within an optimality theoretic approach to interpretation, we predict that there may be cross-linguistic variation induced by a different ranking of the same set of constraints. Fro example, we predict that there are languages in which the KCPH outranks Continuing Topic. In order to test this prediction I will briefly consider two languages where Contrastive Focus does not interfere in the interpretation of the pronouns because stress does not have to distinguish between continuing and shifted topics. In some languages continuing and shifted topics will be morphologically distinct pronouns, in other languages these will be stressed pronouns as opposed to unstressed pronouns, and sometimes, these will be overt pronouns as opposed to null pronouns. Consider the following Chinese sentences, discussed in Huang (1991):

(16) Xiaoming yi jin wu, - jiu ba men guan shang le.
"As soon as Xiaoming$_1$ enters the house, he$_{1/2}$/I/you/we/they... close(s) the door."
(17) Xiaoming yi jin wu, **ta** jiu ba men guan shang le.
"As soon as Xiaoming$_1$ enters the house, he$_2$ closes the door."

According to Huang, the use of the pronoun *ta* in (17) where a zero anaphor could occur (as in (16)), implicates a contrast in reference. In (16) the preferred reading for the null pronoun would be the continuing topic reading; in (17) the overt pronoun must be interpreted as a shifted topic. Obviously, this paradigm shows that a generalized KCPH is satisfied in Chinese (if we view *ta* as similar to a stressed pronoun in English). Note, however, that when a topic is actually given (in topic position), both the zero pronoun as well as the overt pronoun must refer to this topic and they do not allow for a shifted topic reading anymore:

(18) Xiaohua, Xiaoming yi jin wu, - jiu ba men guan shang le.
"Xiaohua$_1$, as soon as Xiaoming$_2$ enters the house, he$_1$ closes the door."

(19) Xiaohua, Xiaoming yi jin wu, **ta** jiu ba men guan shang le.
"Xiaohua$_1$, as soon as Xiaoming$_2$ enters the house, he$_1$ closes the door."

In other words, the KCPH is overruled in Chinese by Continuing Topic (as it was in English). Chichewa differs from a language such as Chinese in this respect, as illustrated in the following paradigm (Bresnan 2001):

(20) Fîsi a-na-dyá chí-manga. Á-tá-**chí**-dya, a-na-pítá ku San Francîsco.
"The hyena ate the corn. Having eaten it, he went to San Francisco."

(21) Fîsi a-na-dyá chí-manga. Á-tá-dyá **icho**, a-na-pítá ku San Francîsco.
"The hyena ate the corn. Having eaten it (something other than the corn), he went to S.F."

By using the full pronoun *icho* in (21) the interpretation that *it* refers to a continuing topic (anaphor) is unavailable. This leads to the very odd and in fact incoherent interpretation as given by the translation in (21). Unlike in Chinese, the presence of a topic in topic position, does not overrule this interpretive effect of a full pronoun, whence the (near) ungrammaticality of (22):

(22) ?* M-kángó uwu fîsi a-na-dyá **íwo**.
"This lion, the hyena ate it."

This suggests that we may fruitfully account for the cross-linguistic patterns of pronoun resolution by making use of the reranking possibilities of constraints that the theory offers.

6. The problem of unintelligibility

At first sight, sentence (22) above seems to be a problematic example for Optimality Theory because of the clash in interpretation which in fact makes the (syntactic) *input* ill-formed. In principle, any input must lead to an optimal interpretation, no matter how many constraints are violated. Intuitively, the explanation of the 'unintelligible' syntactic input in (22) lies in the existence of an alternative (well-formed) input that blocks a coherent interpretation of this one (cf. Blutner 2000, Zeevat 2000). Blutner (2000) claims that the simplest explanation for blocking is a bidirectional Optimality Theory that takes into account the comprehension as well as the production perspective. An expression is blocked with regard to a certain interpretation if this interpretation can be generated more economically by an alternative expression. It may seem that unintelligibility is not a real problem for the theory, once we recognize that sometimes infelicitous interpretations might be optimal and hence should be part of the candidate set of interpretations. However, I have argued in De Hoop (2001) that unintelligibility differs from optimal infelicity, yet it may be solved within a bidirectional approach to optimization.

Conclusion

In general, the existence of (morphological) alternatives raises strong interpretive blocking effects. When there are two optimal lexical forms, it is economical to use them for different interpretations. Thus, when there are two pronominal forms for the third person singular, one might be optimally interpreted as a continuing topic, the other one as a shifted topic or focus. In this paper I have argued that in English stress plays only a minor role in the process of pronoun resolution (i.e., disambiguation). Rather, meaning effects of stress on pronouns are general semantic/pragmatic effects of focus. In particular, stressed pronouns indicate contrast within a set of contextually determined

alternatives. The optimal interpretations that are assigned to stressed pronouns in discourse can be analyzed in terms of three constraints and their ranking, such that Contrastive Focus >> Continuing Topic >> Kameyama's Complementary Preference Hypothesis. The possibility of reranking these constraints might be used to explain different patterns of pronoun interpretation in other languages.

References

Beaver, D. 2000. The Optimization of Discourse. Manuscript Stanford University.

Blutner, R. 2000. Some Aspects of Optimality in Natural Language Interpretation. *Journal of Semantics* 17, 189-216.

Bresnan, J. 2001. The Emergence of the Unmarked Pronoun. *Optimality Theoretic Syntax,* ed. G. Legendre, J. Grimshaw, and S. Vikner. Cambridge, MA: MIT Press.

Givón, T. 1984. *Syntax. A Functional-Typological Introduction*. Amsterdam: John Benjamins.

Haiman, J. 1985. *Natural Syntax. Iconicity and Erosion*. Cambridge: Cambridge University Press.

Hendriks, P., and H. de Hoop. 1997. On the interpretation of semantic relations in the absence of syntactic structure. *The Proceedings of the 11th Amsterdam Colloquium*, ed. Paul Dekker, Martin Stokhof, and Yde Venema. Amsterdam: ILLC/Department of Philosophy.

Hendriks, P., and H. de Hoop. 2001. Optimality Theoretic Semantics. *Linguistics and Philosophy* 24, 1-32.

Hoop, H. de. 2001. The Problem of Unintelligibility in OT Semantics. *Making Sense: from Lexeme to Discourse*, ed. G. van der Meer and A.G.B. ter Meulen. *GAGL* 44. Groningen: CLCG.

Huang, Y. 1991. A neo-Gricean pragmatic theory of anaphora. *Journal of Linguistics* 27, 301-335.

Kameyama, M. 1994. Stressed and Unstressed Pronouns: Complementary Preferences. *Focus & Natural Language Processing. Volume 3: Discourse*, ed. Peter Bosch & Rob van der Sandt. Heidelberg: IBM.

Prince, E.F. 1981. Toward a Taxonomy of Given-New Information. *Radical Pragmatics*, ed. P. Cole. New York: Academic Press.

Rooth, M. 1992. A Theory of Focus Interpretation. *Natural Language Semantics* 1, 75-116.

Vallduví, E. 1990. *The Informational Component*. PhD dissertation, University of Pennsylvania, Philadelphia.

Williams, E. 1997. Blocking and Anaphora. *Linguistic Inquiry* 28, 577-628.

Zeevat, H. 2000. The Asymmetry of Optimality Theoretic Syntax and Semantics. *Journal of Semantics* 17, 243-262.

Indefinites and Sluicing. A type logical approach

Gerhard Jäger
Zentrum für Allgemeine Sprachwissenschaft Berlin

Abstract

Following Jäger (2001a), we propose to extend the Lambek calculus with two additional implications, where the first one models anaphora and the second one indefiniteness. Both pronouns and indefinites are interpreted as (possibly partial) identity functions, but they give rise to different types and are thus subject to different interpretation strategies. The descriptive content of indefinites is interpreted as a domain restriction on the corresponding function. The resulting grammar of indefinites treats the scopal behavior of these NPs in an empirically adequate way. Furthermore it leads to a straightforward surface compositional analysis of Sluicing. The assumed division of labor between syntax and semantics in Sluicing is in accordance with the facts; Sluicing is correctly predicted to be insensitive to syntactic islands, but sensitive to morphological features of the antecedent.

1 The type logical treatment of anaphora

In Jäger (2001a) a type logical version of Jacobson's (1999) variable free grammar of anaphora is given and successfully applied to VP ellipsis. To this end, the associative Lambek calculus **L** is extended with a third implication "|", where $A|B$ is the type of an anaphoric expression of type A which requires an antecedent of type B. Semantically this corresponds to a function from B-denotations to A-denotations. The behavior of the new implication is governed by the following sequent rules (we replace the rule of use from Jäger 2001a by a somewhat stronger version, but the difference is irrelevant for all practical purposes).

$$\frac{X \Rightarrow M : A \qquad Y, x : A, Z, y : B, W \Rightarrow N : C}{Y, X, Z, w : B|A, W \Rightarrow N[M/x][wM/y] : C}[|L]$$

$$\frac{X, x : A, Y \Rightarrow M : B}{X, y : A|C, Y \Rightarrow \lambda z.M[yz/x] : B|C}[|R]$$

The first rule is responsible for anaphora resolution, while the second one ensures that anaphora slots can percolate up in larger syntactic structures. All relevant instances of Jacobson's combinators **Z** and **G** are theorems of the resulting logic. If we adopt the Jacobsonian type assignment $np|np$ and meaning assignment λxx for anaphoric pronouns, her CCG-analysis of pronominal anaphora thus mainly carries over to this type logical system.[1] In Jäger (2001a) it is furthermore shown that we obtain a straightforward account of VP ellipsis if we assume that the stranded auxiliary *did* in constructions like (1) has type $(np\backslash s)|(np\backslash s)$ and denotes the identity function on properties.

(1) John entered before Bill did.

2 A type logical grammar of indefinites

We propose that the semantics of indefinite NPs is similar to the one of anaphoric pronouns. Consider a simple minimal pair such as

[1]The main difference is that the type logical version predicts linear precedence to be the licensing structural configuration for anaphora where Jacobson assumes a version of c-command. The interested reader is referred to Jäger (2001b) for further discussion of this issue.

(2) a. It moved
b. Something moved

According to the analysis mentioned above, (2a) has the category $s|np$ (i.e. it is a clause containing one unresolved pronoun) and the meaning representation $\lambda x.\text{MOVE}'x$, i.e. it denotes the property of moving. The central claim of the present paper is that (2b) should be analyzed analogously; its meaning representation is also $\lambda x.\text{MOVE}'x$. The two sentences differ in truth conditions and in their semantic contribution to larger constructions because they belong to different syntactic categories. To implement this idea, we extend the Lambek calculus with another binary connective, $\rightsquigarrow$. The intuition here is that a sign of category $A \rightsquigarrow B$ is like a sign of category B except that it introduces a discourse referent of the type corresponding to category A. So an indefinite NP will receive the category $np \rightsquigarrow np$, and a sentence containing an indefinite has the category $np \rightsquigarrow s$. (In linguistic applications, A will always be instantiated with np.) A sign of category $A \rightsquigarrow B$ denotes a function from A-denotations to B-denotations, so

$$Dom(A \rightsquigarrow B) = Dom(B|A) = Dom(A\backslash B) = Dom(B/A) = Dom(B)^{Dom(A)}$$

A simple indefinite like *something*, having category $np \rightsquigarrow np$, thus denotes a Skolem function. This function is lexically specified to be the identity function λxx, so *something* comes out as synonymous with *it*.
The property of introducing a discourse referent can be inherited from sub-constituents to super-constituents. Formulated in type logical terms, this means that the deductive behavior of $\rightsquigarrow$ is governed by the following rule, which is entirely analogous to the rule of proof of the anaphora slash.

$$\frac{X, x : A, Y \Rightarrow M : B}{X, y : C \rightsquigarrow A, Y \Rightarrow \lambda z.M[yz/x] : C \rightsquigarrow B}[\rightsquigarrow]$$

It is easy to see that this rule, as well as the two rules for the anaphora slash, have the subformula property. Furthermore, Lambek's (1958) proof of Cut elimination smoothly carries over to the extended Lambek calculus. The resulting system is thus guaranteed to be decidable and to have the finite reading property (i.e. there are only finitely many Cut free proofs for each theorem).
Indefinites are not anaphoric; the argument slot that is created by an indefinite can thus not be filled by anaphora resolution. In type logical terms, this means that the logic of $\rightsquigarrow$ is solely determined by the rule $[\rightsquigarrow]$. There is not counterpart of the rule of use of the anaphora slash for $\rightsquigarrow$. (This can be regarded as a proof theoretic implementation of Heim's 1982 Novelty Condition.)
The syntactic derivation of (2b) is given in figure 1. If we plug in the lexical semantics λxx for *something* and MOVE' for *moved*, we obtain the sentence meaning $\lambda v.\text{MOVE}'v$ for (2b).

$$\dfrac{\dfrac{\dfrac{}{x : np \Rightarrow x : np}[id] \qquad \dfrac{}{y : s \Rightarrow y : s}[id]}{z : np, w : np\backslash s \Rightarrow wz : s}[\backslash L]}{u : np \rightsquigarrow np, w : np\backslash s \Rightarrow \lambda v.w(uv) : np \rightsquigarrow s}[\rightsquigarrow]$$

Figure 1: Derivation for (2b)

In the general case, indefinite descriptions come with a non-trivial descriptive content. We analyze the descriptive information as a restriction on the domain of the corresponding function. The denotation of the indefinite description *a cup*, for instance, would come out as the identity function on the set of cups. To express partial functions in the typed λ-calculus, we extend its syntax and semantics in the following way:

Definition 1

- If M and ϕ are terms of types σ and t respectively and v is a variable of type τ, then $\lambda v_\phi M$ is a term of type $\langle \tau, \sigma \rangle$.
- $\|\lambda v_\phi M\|_g = \{\langle a, \|M\|_{g[v \to a]} \rangle : \|\phi\|_{g[v \to a]} = 1\}$

The lexical entry for the indefinite article is given in (3b); its meaning is a function from a set to the identity function over this set. Given this, the sequent in (3c) is a theorem, and the meaning of the sentence (3a) is thus (3d), i.e. it denotes the partial truth-valued function that returns 1 for each cup that moved, 0 for each cup that didn't move, and that is undefined for all non-cups.

(3) a. A cup moved
b. a – $\lambda P x_{Px}.x : (np \rightsquigarrow np)/n$
c. $y : (np \rightsquigarrow np)/n, z : n, w : np\backslash s \Rightarrow \lambda u.w(yzu) : np \rightsquigarrow s$
d. $\lambda u.\text{MOVE'}((\lambda x_{\text{CUP'}x}.x)u)$
$(\equiv \lambda u_{\text{CUP'}u}.\text{MOVE'}u)$

3 Truth and negation

Since sentence denotations need not be truth values in the present system but may be partial truth-valued function of any arity, truth has to be defined polymorphically, and it has to be relativized to the syntactic category of the sentence in question. (This will ensure that (2a) and (b) will have different truth conditions despite their identical denotations.) Furthermore, we follow Dekker (2000) in relativizing truth to sequences of objects (that supply the referents of unbound pronouns). So truth is defined as a four place relation between a sentence denotation (expressed by the meta-variable α below), a syntactic category (where we use the meta-variable S), a model (which is suppressed in the notation below) and a sequence of individuals $\vec{e}$. ($c\vec{e}$ is the sequence that results if you add c as top element to the sequence $\vec{e}$.)

Definition 2 (Truth)

$$\begin{array}{rcl} \vec{e} \models \alpha : s & \text{iff} & \alpha = 1 \\ c\vec{e} \models \alpha : S|np & \text{iff} & \vec{e} \models (\alpha c) : S \\ \vec{e} \models \alpha : np \rightsquigarrow S & \text{iff} & \vec{e} \models (\bigcup_{\alpha c \text{ is defined}} (\alpha c)) : S \end{array}$$

Intuitively, the argument slots originating from pronouns are filled by the elements of the sequence $\vec{e}$, while the slots coming from indefinites are existentially bound. It is easy to see that the truth conditions of our examples come out as expected, i.e. (2) is true wrt. a sequence iff the first element of this sequence moved, (2b) is true if something moved, and (3a) if some cup moved.
As in Dynamic Predicate Logic or in Dekker's Predicate Logic with Anaphora, negation is an operation that operates on the truth conditions of its operand rather than on its meaning directly. In the present system this means that negation is relativized to the syntactic category of its operand. As in the truth definitions above, argument slots originating from indefinites are existentially closed, while those coming from pronouns are passed further up.

Definition 3 (Negation)

$$\begin{array}{rcl} \sim \alpha : s & = & 1 - \alpha \\ \sim \alpha : S|A & = & \lambda c. \sim (\alpha c) \\ \sim \alpha : A \rightsquigarrow S & = & \sim (\bigcup_{c \in Dom(A)} \alpha c) \end{array}$$

We assume that the arguments of other propositional operators (like conjunctions or verbs of propositional attitudes) undergo a similar operation of existential closure as well.

4 Linguistic consequences

Consider the following construction:

(4) If a cup moved, the ghost is present

Suppose that the particle *if* has the polymorphic category $S_1/S_1/S_2$ (where $S_{1/2}$ range over sentential categories) and the meaning representation $\lambda pq.(\neg p \vee q)$ (where $\neg$ is the syntactic counterpart of the semantic negation operation defined above, and $\phi \vee \psi$ abbreviates $\neg(\neg\phi \wedge \neg\psi)$). To simplify the discussion, we ignore the internal structure of the main clause *the ghost is present* and represent its meaning as GHOST_IS_PRESENT. Given this, (4) is predicted to be structurally ambiguous—depending on the stage in the derivation where the rule $[\rightsquigarrow]$ is applied—and to receive the following two semantic representations:

(5) a. $\neg\lambda u_{\text{CUP'}u}\text{MOVE'}u \vee \text{GHOST\_IS\_PRESENT}$
b. $\lambda u_{\text{CUP'}u}.\neg\text{MOVE'}u \vee \text{GHOST\_IS\_PRESENT}$

According to the truth definitions given above, (a) is true if either no cup moved or the ghost is present—i.e. the indefinite has narrow scope wrt. the conditional—and (b) is true if there is a certain cup u such that either u doesn't move or the ghost is present. This corresponds to the specific reading of the indefinite *a cup*.
This example illustrates the following noteworthy properties of the present analysis of indefinites:

A An indefinite can take scope over each clause it is contained in, and it scopally interacts with superordinate propositional operators like negation. However, the scoping mechanism for indefinites is entirely independent of the type logical scoping mechanism for genuine quantifiers (like Moortgat's 1996 *in situ* binder), and if quantifier scope is clause bounded or otherwise restricted, this has no implications for the scope of indefinites. So the unrestricted scope taking behavior of indefinites is expected.

B The present theory assumes that indefinites are interpreted *in situ*. Therefore the scope of indefinites is not subject to constraints on movement. Nonetheless the descriptive content of an indefinite—being interpreted as a domain restriction on a function—is inherited by its super-constituent after semantic composition. This ensures that the existential impact of an indefinite and its descriptive content are never unduly divorced.

C No particular problem arises if the extension of the descriptive content of an indefinite is empty. For instance, if there were no cups, both (3a) and (4) in the reading (5b) would denote the empty function and would therefore be false.
The second and third point pose problems for other *in situ* theories of indefinites like the choice function approach (see for instance the discussion in Geurts 2000), and it has been argued that some kind of movement analysis is inevitable for these reasons.

The present approach demonstrates that the empirical facts can be analyzed in a surface compositional way.

Last but not least it should be mentioned that Dekker's treatment of donkey anaphora can be incorporated into the present system without major problems by designing a polymorphic version of conjunction. Space prevents us from pursuing this issue further; the interested reader is referred to Jäger (2001b).

5 Sluicing

This grammar of indefinites, paired with the mentioned type logical treatment of anaphora, leads to a straightforward surface compositional analysis of Sluicing. This is a version of ellipsis where under certain contextual conditions, a bare *wh*-phrase stands proxy for an entire question. The source clause is typically a declarative clause which contains an indefinite NP. The target clause is interpreted as the question that is obtained if this indefinite is replaced by a *wh*-phrase. For example, the Sluicing construction in (6a) is interpreted as (6b).

(6) a. A cup moved, and Bill wonders which cup
b. A cup moved, and Bill wonders which cup moved

We assume that the *wh*-determiner *which* in an embedded question like in (6b) has the lexical entry given in (7a). (We adopt Moortgat's 1988 gap operator "$\uparrow$", i.e. $s \uparrow np$ is the type of a clause with an np gap. q is the type of embedded questions. The predicate Q^+ denotes the positive extension of the predicate Q. This means that $\|Q^+\|c = 1$ iff $\|Q\|c = 1$, and $\|Q^+\|c = 0$ otherwise.) The question (7b) thus receives the semantic representation (7c).

(7) a. which cup moved
b. which – $q/(s \uparrow np)/n : \lambda PQ?x.Px \wedge Q^+x$
c. $?x.\text{CUP'}x \wedge \text{MOVE'}x$

Now consider a Sluicing construction like (6a). The antecedent clause has the category $np \leadsto s$, i.e. it is a clause containing an indefinite. Its interpretation is $\lambda v_{\text{CUP'}v}\text{MOVE'}v$. We assume that the *wh*-word *which* has the same interpretation here as in the non-elliptical interpretation, but a different category. The second lexical entry for *which* is

(8) which – $q|(np \leadsto s)/n : \lambda PQ?x.Px \wedge Q^+x$

So if we combine *which* with a common noun like *cup*, we get an item of type $q|(np \leadsto s)$, i.e. an anaphoric embedded question which needs an antecedent of type $np \leadsto s$, i.e. a clause containing an indefinite. Plugging in the meaning of the antecedent in the example (6a), the embedded question receives the resolved meaning $?x\text{CUP'}x \wedge (\lambda x_{\text{CUP'}}\text{MOVE'}x)^+x$, which is equivalent to $?x.\text{CUP'}x \wedge \text{MOVE'}x$. This analysis handles the essential empirical characteristics of Sluicing correctly:

A It follows from the type of the *wh*-phrase in Sluicing constructions that the antecedent clause has to contain an indefinite which corresponds to the *wh*-phrase in the elliptical clause. So the oddity of (9) is predicted.

(9) *The cup moved, and Bill wonders which cup

B Sluicing is island insensitive. This means that Sluicing constructions are grammatical even if their non-elliptical counterparts are deviant due to an island violation. A typical example is (10) (taken from Chung et al. 1995), where the non-elliptical version involves a violation of the Complex NP Constraint.

(10) a. The administration has issued a statement that it is willing to meet with one of the student groups, but I'm not sure which one
b. *The administration has issued a statement that it is willing to meet with one of the student groups, but I'm not sure which one the administration has issued a statement that it is willing to meet with

Under the present approach this is predicted because the elliptical and the non-elliptical construction are not transformationally related. For the Sluicing construction to be grammatical, it is sufficient that the antecedent clause contains a wide scope indefinite, and the scope of indefinites is not subject to island constraints.

C The descriptive part of the antecedent indefinite is interpreted as an additional restriction on the interrogative operator in the ellipsis site. For example, the next two sentences are correctly predicted to be synonymous.

(11) John invited a philosopher, but I don't know {which philosopher/who}

D In case marking languages it can be observed that the indefinite in the source clause must have the same case as the corresponding *wh*-phrase in the target clause. The following German example (from Ross 1969) illustrates this point:

(12) Er will jemandem schmeicheln, aber sie wissen nicht {wem / *wen}
HE WANTS SOMEONE$_{\text{DAT}}$ FLATTER BUT THEY KNOW NOT {WHO$_{\text{DAT}}$ / WHO$_{\text{ACC}}$}
'He wants to flatter someone, but they don't know whom'

This generalization can easily be covered if we represent morphological information in the syntactic categories. Let us say simplifyingly that a name in dative case has the category $np(dat)$. An indefinite in dative case would then receive the category $np(dat) \leadsto np(dat)$, and a clause containing such an indefinite has the category $np(dat) \leadsto s$. A sluiced *wh*-phrase in dative like *wem* above has the category $q|(np(dat) \leadsto s)$, i.e. it requires a clause containing an indefinite in dative as antecedent.

References

Chung, S., W. Ladusaw, and J. McCloskey (1995). Sluicing and Logical Form. *Natural Language Semantics*, **3**, 239–282.

Dekker, P. (2000). Grounding dynamic semantics. manuscript, University of Amsterdam.

Geurts, B. (2000). Indefinites and choice functions. *Linguistic Inquiry*, **31**, 731–738.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts, Amherst.

Jacobson, P. (1999). Towards a variable-free semantics. *Linguistics and Philosophy*, **22**(2), 117–184.

Jäger, G. (2001a). Anaphora and quantification in categorial grammar. In M. Moortgat (ed.), *Logical Aspects of Computational Linguistics*, no. 2014 in Lecture Notes in Artificial Intelligence, 70–89. Springer, Berlin, Heidelberg.

— (2001b). Anaphora and type logical grammar. manuscript, Utrecht University. Available from http://www.let.uu.nl/~Gerhard.Jaeger/personal/.

Lambek, J. (1958). The mathematics of sentence structure. *American Mathematical Monthly*, **65**, 154–170.

Moortgat, M. (1996). Generalized quantification and discontinuous type constructors. In W. Sijtsma and A. von Horck (eds.), *Discontinuous Constituency*. De Gruyter, Berlin.

Ross, J. (1969). Guess who? In R. I. Binnick, A. Davison, G. M. Green, and J. L. Morgan (eds.), *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, 252–286. University of Chicago.

Implicit slashing in IF logic

Theo M.V. Janssen, ILLC, University of Amsterdam

1 Introduction

In game theoretical semantics the truth of a formula is determined by a game between two players, one who tries to check the formula, and one who tries to refute it. A version of such games, introduced by J. Hintikka, is IF logic: independence friendly logic (e.g. in Hintikka & Sandu (1997)). In IF logic the quantifier $\exists y_{/x}$ arises, which means that a value for y has to be chosen independent of the value for x, and the disjunction $\psi \vee_{/x} \theta$ where a subformula has to be chosen independent of x. Following Hodges (1997*a*) we call the players ∃loise (female) and ∀belard (male); this has the advantage that pronouns can be used without the danger of confusion. Furthermore, the names reflect the choices the players make (in the situations arising in this paper): ∃loise makes the choice for $\exists y$, $\exists y_{/x}$, $\vee$, and $\vee_{/x}$, and ∀belard for $\forall$ and $\wedge$. A formula is true if ∃loise has a winning strategy and false if ∀belard has one.

Below, some examples are given which illustrate the aims of IF logic. These, and all later examples, are interpreted on the natural numbers (**N**).

(1) $\forall x\, \exists u_{/x}\, [x = u]$

When ∀belard has chosen, ∃loise has to choose y independently of the value of x, and therefore it may happen that she selects another value. Hence ∃loise has no winning strategy, the formula is 'not true'. Also ∀belard has no winning strategy, so the formula is neither true nor false. For the variant with $\neq$ instead of $=$, the same holds, neither 'true' nor 'false'.

(2) $\forall x\, \exists y_{/x}\, [y \leq x]$

The strategy 'let y be 0' is winning for ∃loise whatever ∀belard plays, hence the formula is 'true'.

(3) $\forall x[x = 7\ (\vee/x)\ x \neq 7]$

This is 'not true', because without information on x a guaranteed correct choice between the disjuncts is not possible.

Hintikka claimed that a compositional semantics for IF logic would not be possible. However, Hodges has given a compositional interpretation (Hodges (1997*a*), Hodges (1997*b*)). By doing so, he clarified several aspects of the logic.

The main aim of this paper is to show that the semantics for IF logic, viz. game theoretical semantics, does not capture the intuitions about independence. We will focus on will focus on the implicit slashing convention, which is supposed to remedy these clashes with intuition, and it will be shown that it introduces new problems. For other conflicts with intuition, see Janssen (1999).

2 Implicit Slashing of Existential quantifiers

Consider

(4) $\forall x \exists u\ \exists v_{/x}\, [x = v]$

This examples resembles example (1), the difference is that a vacuous quantifier u is inserted. The information given by $\exists u$ that there exists a number tells us nothing new, so one might expect that this change makes no difference for the task to find a v independent of x. In the semantics for IF logic, given in Hodges (1997*a*), the surprising result holds that (4) is true. The strategy ∃loise follows for $\exists v_{/x}$ is to play $v := u$. This strategy does not mention the value for x, and therefore is allowed. The strategy for $\exists u$ is to play $u := x$; this dependence is allowed because $\exists u$ has no slash. So $u = x$, and, since $v = u$, it follows that $x = v$. By choosing these strategies, ∃loise always wins (although the formula without the empty quantification is not true).

This result is not accordance with intuitions about independence. The value of v is carefully chosen in such a way that it equals x, hence is very dependent on x.

The issue is clarified in (Hintikka 1996), which appeared after (Hodges 1997*a*) was written. Hintikka says 'the small extra specification that is needed is that moves connected with existential quantifiers are always independent of earlier moves with existential quantifiers' (Hintikka 1996, p. 63). This means that in case an existential quantifier $\exists v$ occurs within the scope of $\exists u$, then $\exists v$ should be interpreted as if it was written as $\exists v_{/u}$. In the appendix to Hintikka's book, Sandu presents a formal interpretation of IF logic in which this independence of existential variables on other ones is formalized (Hintikka 1996, p.256).

That such an independence was always intended, can be seen for instance in the examples given in one of the earlier papers on game theoretical semantics: (Hintikka 1974). Furthermore, Hintikka (1996, p. 64) gives a formula which only is true in infinite models and in the empty model, but looses this property (which is not expressible in first order logic) when the slashing convention is dropped.

We will refer in the sequel to this independence of existential quantifier as the 'slashing convention' and with 'GTS' to the semantics with this convention. It is the 'official' interpretation of IF logic and described by Sandu in Hintikka (1996). Note that, although in the language of IF logic the only variables which may occur after a slash are those which are bound by a universal quantifier, also existential variables may, due to the convention, implicitly arise after a slash.

The slashing convention means that Hodges (1997*a*) does not give a compositional version of GTS, but for a closely related interpretation. In Hodges (1997*b*) it is indicated how he would design a compositional semantics for IF logic with the slashing convention.

The slashing convention has consequences which are intuitively very strange. Consider:

(5) $\exists u \exists v[u = v]$

Due to the convention the second quantifier is implicitly slashed for the first. This means that the formula is equivalent with:

(6) $\exists u\ \exists v_{/u}\, [u = v]$

Of course, (5) is true, but in (6) it is, according to my intuition, impossible to find a v independent of u such that the two are equal, so (6) is not true. This reaction on the slashing convention is found on several places in the literature. In their review of Hintikka (1996), Cook & Shapiro (1998, p. 311) comment on Hintikka's extra specification as follows: 'However [...] then a sentence like $\exists x \exists y[x = y]$ would not be true over the natural numbers, contradicting Hintikka's claim that game theoretical semantics agrees with ordinary semantics on first-order sentences [Hintikka (1996), p. 65].' Related remarks are found in the review by de Swart, Verhoeff & Brands (1997) and in the unpublished PhD-dissertation of Pietarinen (2001, p. 31).

However, these intuitions about (5) and (6) are not reflected in their game theoretical interpretation: (6) is in game theoretical semantics a true formula. One sees this as follows. Let n be some natural number. ∃loise's strategy for the first quantifier is $u := n$, and for the second quantifier $v := n$. This second strategy does not mention u, and therefore it allowed in GTS. So, although in (6) the strategies for u and v must be the same and a change in the strategy for u has immediate consequences for the strategy for u, they are nevertheless formally independent in GTS. Note the strange fact that, due to the convention, the obvious strategy for (5) (viz. $u := v$) is not allowed, whereas in the closely related formula (7) it is allowed:

(7) $\forall x \exists u[u = x]$

Also the next example is one from classic first order logic (i.e. it has no slashes) where the slashing convention seems to prohibits the classical interpretation and where constant strategies are not possible.

(8) $\forall x \exists u[u > x \wedge \exists v[v > u]]$

We would like to follow a strategy like $v := u + 2$: how else could we achieve that $v > u$? The trick is that the strategy for v the strategy for u. If we have chosen $u := x + 17$, then we now follow the strategy $v := (x + 17) + 2$. So again, although u and v are formally independent, their strategies have a close relationship.

3 Implicit slashing of disjunctions

As was the case with existential quantifiers in example (4), also in a disjunction a dummy variable y can be used to transfer information concerning x towards the choice for $\vee$ which must be independent of x:

(9) $\forall x \exists y[x = 4 \vee_{/x} x \neq 4]$

Example (9) becomes true if chooses $y := x$ and then decides on to $\vee_{/x}$ using the value of y. To avoid this signaling the same solution can be used as for (4): state that a disjunction is implicitly slashed for the existential variables which have scope over it. Although Hintikka does not mention this variant of the slashing convention, it is incorporated in the appendix by Sandu (see Hintikka (1996, p.256)). So (9) is not true in game theoretical semantics.

The slashing convention for disjunctions has strange consequences. We will consider three examples. The first is (10) which is, due to the convention, equivalent with (11):

(10) $\exists u[u = 4 \vee u \neq 4]$

(11) $\exists u[u = 4 \vee_{/u} u \neq 4]$

Of course, (10) should be true, but according to my intuition of independence, (11) is not true because it requires to make independent of u a choice between $u = 4$ and $u \neq 4$. Nevertheless, both (10) and (11) are true in GTS. The trick is to follow a constant strategy for the existential quantifier, e.g. $u := 2$, and for the disjunction always to choose R. So instead of letting the value of u determine the choice (which is not allowed), the strategy for $\exists u$ determines the choice.

The second example has no slashes and classically it is true.

(12) $\forall x \exists u[u = x \wedge [u = 4 \vee u \neq 4]]$

Since $u = x$, a constant strategy for $\vee$ cannot be used, and in GTS the obvious strategy for the disjunction *if* $u = 4$ *then* L *else* R is not available because of implicit slashing. But (12) is never the less true because x can be used: follow the strategy *if* $x = 4$ *then* L *else* R. This is an unnatural strategy because x does not occur in the subformula $u = 4 \vee u \neq 4$.

Also in the third example a constant strategy cannot be used:

(13) $\forall x \exists u[u = x \wedge [u = 4 \vee_{/x} u \neq 4]]$

This formula is intuitively true because the choice between $u = 4$ and $\neq 4$ has nothing to do with x. In GTS neither this strategy, nor one using x is allowed. The only strategy ∃loise might play is a constant strategy, but then she will loose for at least one x. Hence (13) is not true in GTS.

It is remarkable that the following variants of (13) are true in GTS:

(14) $\forall x \exists u[u \leq x \wedge [u = 4 \vee_{/x} u \neq 4]]$

(15) $\forall x \exists u[u > x \wedge [u = 4 \vee_{/x} u \neq 4]]$

The strategy for the first example is $u := 0$, and for the second $u := x + 5$.

Example (13) is a remarkable example. In all other cases where intuition differs from GTS results, it was the case that a formula which intuitively is not true, was never the less true in GTS. Here the situation is reverse.

Hintikka (1974, p. 65) claims that IF logic is a conservative extension of classic predicate logic. The examples given in this and in the previous section illustrate that this relation is in neither direction a direct one:

1. If Tarskian semantics suggests a game strategy for (5) $\exists u \exists v[u = v]$ or (10) $\exists u[u = 4 \vee u \neq 4]$, it will be a strategy based upon the value of u, and not a constant strategy as in GTS.
2. In GTS a strategy for ψ can be based upon variables that do not occur in ψ, as was the case for $u = 4 \vee u \neq 4$ in (12) $\forall x \exists u[u = x \wedge [u = 4 \vee u \neq 4]]$. But in standard semantics the interpretation of ψ depends only on variables which occur in ψ.

4 ∀belard's choices

In section (2) and (3), we have seen how ∃loise could use her own choices to signal a value to herself that she is assumed not to know, and that the slashing convention prohibits this. But there is form of signaling that is not prohibited: she may use a variable chosen by ∀belard to get information she should not use. Consider:

(16) $\forall x \forall z[x = z \vee \exists u_{/x}[u \neq x]]$

My intuition says that (16) should not be true because the left disjunct is not always true, and the right disjunct is not true at all. However, ∃loise can use y to signal the value of x: she follows the strategy $u := y$, and the formula becomes true.

A more elaborated version is

(17) $\forall x \forall y \forall z[x = z \vee \exists v\ \exists u_{/x}[v \neq y \wedge u \neq x]$

This variant is true due signaling the value of x by z. But if we change the name of the bound variable v to z, this seems not possible any more. I do not know whether Hintikka's interpretation allows to distinguish the to occurrences of z, but in the formalization of Caicedo & Krynicki (1999) change of bound variables is not allowed.

In Janssen (1999) an example is given where $\phi \vee \phi$ was not equivalent with ϕ. For conjunction the same phenomenon arises, due to related strategies of ∀belard. In the example below $\forall x/y$ says that ∀belard has to choose a value for x independent of the choice of y by ∃loise. First, notice that he has no strategy which guarantees him that (18) becomes false:

(18) $\exists u \forall x/u[x \neq u]$

One might expect that he has no a winning strategy either if he has the choice to falsify the formula on the left hand side of a conjunction or on the right hand side:

(19) $\exists u[\forall x/u[x = u] \wedge \forall x/u[x = u]]$

That is, however, not the case. ∀belard's strategy for this game could be to choose the left disjunct if $y = 3$, and the right one otherwise. On the left he plays $x := 4$, which makes the conjunct false, and on the right $x := 3$ with the same effect. So he has a strategy which makes the formula always false, hence (19) is false. This shows a remarkable property of GTS: *$\phi \wedge \phi$ is not in all contexts equivalent with ϕ.*

5 Discussion

The examples given in the previous sections show that strategies in GTS are on several points in conflict with intuitions on independence. Note that most of these conflicts also arise in case implicit slashing is not incorporated. Although the strategies for $\vee_{/x}$ and $\exists u_{/x}$ do not assume that the value of x is given, any information which can be deduced from other sources is used:

1. The information whether the slashed variable is universally quantified or existentially (examples (5) and (7)).
2. The context in which the formula is used (example (19)).
3. The strategies used elsewhere in the formula (examples (6), (8) and (12))
4. Even the value of the slashed variable is used if this can be deduced from other information (examples (10) and (16))

For these reasons, I conclude that game theoretical semantics is not a formalization of 'informational independence', but of 'imperfect information'. It is not impossible that Hodges would agree with this opinion, because the title of his paper Hodges (1997*a*) is 'Compositional semantics for a language of imperfect information', and *not* '... for a language of informational independence'.

6 Towards an alternative interpretation

The alternative which I propose is based upon the following three ideas:

1. There is no implicit slashing. If a choice has to be independent of a variable, than it has to be said so explicitly.
2. Independence is a local phenomenon. If one has to make a choice independent of x, then one can do so (or cannot do so) irrespective of the context.
3. Independence is a form of uniformity in other circumstances. If a choice promises something about other circumstances. if a choice can be made independent of x, it is a good choice for other values of x as well.

The first idea is implemented of course by not having implicit slashing, but also by enlarging the possibilities of explicit slashing. Not only variables bound by universal quantifiers, but also ones bound by existential quantifiers.

Following the second idea we considering subformulas (including formulas with free variables) as games on their own. This explains the name of the alternative: 'subgame semantics', henceforth SGS. ∃loise may have a plan which describes how she will react on moves of ∀belard and which guarantees her a win. This strategy may depend only on the values of variables occurring in the subgame, and not on the any information deduced from the context in which the subgame arises. So the strategy must be winning in that subgame for the current initial situation.

The third idea is realized by requirements on the admitted strategies for independent choices. Of course, a strategy for $\exists u_{/x}$ is a function which yields the same value for any x. The discussion in the previous sections show that this strategy should not be tailored on special (imperfect) information about x but should uniformly be good. I propose the following requirement: for other values of x the same choice wins if there is a winning choice at all.

One might describe the approach proposed in this paper by means of the following metaphor. Associated with a subgame there is a shelf of strategies, and when a game is played, a strategies is taken from the shelf for its subgames, and that strategy can be used in other situations as well.

References

Caicedo, X. & Krynicki, M. (1999), Quantifiers for reasoning ith imperfect information and σ_1^1 logic, *in* W. A. Carnielli & I. M.L d'Ottaviano, eds, 'Advances in Contemporary Logic and Computer Science. Proc. 11th Brazilian Conf. on Math.Logic 1996', Vol. 235 of *Contemporary Mathematics*, Americal mathematical Society, pp. 17–31.

Cook, R. & Shapiro, S. (1998), 'Hintikka's revolution "the principles of mathematics revisited"', *British Journal for the Philosophy of Science* **49**, 309–316.

Hintikka, J. (1974), 'Quantifiers vs. quantification theory', *Linguistic Inquiry* **5**, 153–177.

Hintikka, J. (1996), *The principles of mathematics revisited*, Cambridge University Press, Cambridge.

Hintikka, J. & Sandu, G. (1997), Game-theoretical semantics, *in* J. van Benthem & A. ter Meulen, eds, 'Handbook of logic and language', Elsevier, Amsterdam and The MIT Press, Cambridge, Mass., chapter 6, pp. 361–410.

Hodges, W. (1997*a*), 'Compositional semantics for a language of imperfect information', *Journal of the IGPL* **5**(4), 539–563.

Hodges, W. (1997*b*), Some strange quantifiers, *in* J. Mycielski et al, ed., 'Structures in Logic and Computer Science', number 1261 *in* 'Lecture notes in computer science', Springer, Berlin, pp. 51–65.

Janssen, T.M.V. (1999), IF logic and informational independence, *in* P. Dekker, ed., 'Proceedings of the Twelfth Amsterdam Colloquium, dec. 1999', ILLC/ Dept of Philosophy, University of Amsterdam, pp. 139–144.

Pietarinen, A. (2001), Games logic plays. Informational independence in game theoretical semantics, PhD thesis, School of cognitive and computing sciences. Preliminary version.

de Swart, H., Verhoeff, T. & Brands, R. (1997), 'Hintikka's "The principles of mathematics revisited"', *Logique et Analyse* **159**, 281–289.

Learning Word-to-Meaning Mappings in Logical Semantics

Makoto Kanazawa
University of Tokyo

Abstract. I generalize Siskind's problem of learning word-to-meaning mappings to logical semantics, and formulate the heart of the problem as a problem in typed lambda calculus. I show that every instance of this 'mapping problem' has infinitely many solutions, but many of them are equivalent in a certain sense. I show an algorithm for finding a reasonably small solution with respect to a certain 'definability' relation between terms.

Siskind (1996, 2000) discusses the problem of learning word-to-meaning mappings from sentences coupled with their meanings. In his setup, sentence meanings are represented by first-order terms built up from constants and function symbols (called *conceptual symbols*), and word meanings are represented by first-order terms with zero or more individual variables. For instance, the meaning of *John lifted Mary* may be represented by

$$\mathrm{CAUSE}(\mathbf{John}, \mathrm{GO}(\mathbf{Mary}, \mathrm{UP})),$$

where CAUSE, **John**, GO, **Mary**, and UP are conceptual symbols. The learner is fed with a stream of word strings coupled with expressions like the one above, and is expected to arrive at a word-to-meaning mapping like:

$$\begin{array}{lll} \textit{John} & \mapsto & \mathbf{John} \\ \textit{lifted} & \mapsto & \mathrm{CAUSE}(x, \mathrm{GO}(y, \mathrm{UP})) \\ \textit{Mary} & \mapsto & \mathbf{Mary} \end{array} \tag{1}$$

Semantic composition is supposed to be done by substitution. An equivalent way of looking at it is to view word meanings as λ-terms like $\lambda xy.\mathrm{CAUSE}(x, \mathrm{GO}(y, \mathrm{UP}))$, where all subterms are of *first-order* types, and semantic composition as application plus β-reduction.

Siskind shows that one can reduce the complexity of learning by breaking down the learning process into two phases. In the first phase, the learner forms, for each word, the set of conceptual symbols involved in the meaning of that word. This can be done efficiently by an online algorithm that maintains two sets of conceptual symbols: the set of symbols that are possibly in the correct word meaning, and the set of symbols that are necessarily in the correct word meaning. The second phase then builds the correct meaning expression out of the symbol set learned in the first phase. For instance, suppose that the learner has arrived at the following association of words with symbol sets:

$$\begin{array}{lll} \textit{John} & : & \{\mathbf{John}\} \\ \textit{lifted} & : & \{\mathrm{CAUSE}, \mathrm{GO}, \mathrm{UP}\} \\ \textit{Mary} & : & \{\mathbf{Mary}\} \end{array}$$

Then, given that the sentence *John lifted Mary* has meaning $\mathrm{CAUSE}(\mathbf{John}, \mathrm{GO}(\mathbf{Mary}, \mathrm{UP}))$, the learner can now uniquely determine the word-to-meaning mapping given in (1).

The aim of this paper is to study a generalization of Siskind's problem, where conceptual symbols, as well as variables, can be of arbitrary types, and meaning expressions are typed λ-terms built out of these, using λ-abstraction as well as application. This is the usual setup of logical semantics. For instance, the sentence *every man finds a unicorn* may have a meaning represented by $\forall x[\mathrm{MAN}(x) \to \exists y[\mathrm{UNICORN}(y) \wedge \mathrm{FIND}(x,y)]]$, which is just a shorthand for

$$\begin{array}{l} \forall^{(e\to t)\to t}(\lambda x^e.\to^{t\to t\to t}(\mathrm{MAN}^{e\to t}(x)) \\ \qquad (\exists^{(e\to t)\to t}(\lambda y^e.\wedge^{t\to t\to t}(\mathrm{UNICORN}^{e\to t}(y))(\mathrm{FIND}^{e\to e\to t}(y)(x))))). \end{array} \tag{2}$$

The meanings of the words in this sentence may be (suppressing the type when it is clear):

$$\begin{array}{lll} \textit{every} & \mapsto & \lambda u^{e\to t} v^{e\to t}.\forall(\lambda x^e.\to (ux)(vx)) \\ \textit{man} & \mapsto & \mathrm{MAN} \\ \textit{finds} & \mapsto & \mathrm{FIND} \\ \textit{a} & \mapsto & \lambda u^{e\to t} v^{e\to t}.\exists(\lambda y^e.\wedge(uy)(vy)) \\ \textit{unicorn} & \mapsto & \mathrm{UNICORN} \end{array} \tag{3}$$

(I mostly follow the notational convention of Hindley 1997, except that I may write the application of two terms M and N either MN or $M(N)$, depending on which is more readable.)

Just as in the simple first-order case of Siskind, the problem of learning word-to-meaning mappings can be broken down into two subproblems: the problem of finding for each word the set of constants that appear in the meaning of that word, and the problem of building the correct λ-term for each word using the symbol set associated with that word. The first problem can be solved efficiently in exactly the same way as in the case of Siskind. How to solve the second problem was rather obvious in the case of Siskind, but becomes far from so in our generalized setting. I formulate the heart of the problem as a problem in typed λ-calculus.

Mapping Problem. Given m sets of distinct variables

$$\begin{array}{c} \vec{x}_1 = x_{1,1}^{\sigma_{1,1}}, \ldots, x_{1,n_1}^{\sigma_{1,n_1}} \\ \vdots \\ \vec{x}_m = x_{m,1}^{\sigma_{m,1}}, \ldots, x_{m,n_m}^{\sigma_{m,n_m}} \end{array}$$

and a λI-term $T^\rho[\vec{x}_1, \ldots, \vec{x}_m]$ (with free variables $\vec{x}_1, \ldots, \vec{x}_m$) in $\beta\eta$-normal form, find m λI-terms $S_1^{\tau_1}[\vec{x}_1], \ldots, S_m^{\tau_m}[\vec{x}_m]$ and a $BCI\lambda$-term $D^\rho[y_1^{\tau_1}, \ldots, y_m^{\tau_m}]$, each in $\beta\eta$-normal form, such that

$$D[S_1[\vec{x}_1], \ldots, S_m[\vec{x}_m]] \rhd_{\beta\eta} T[\vec{x}_1, \ldots, \vec{x}_m].$$

I denote this problem by $T^\rho[\vec{x}_1; \ldots; \vec{x}_m]$, and I call $\langle S_1^{\tau_1}[\vec{x}_1], \ldots, S_m^{\tau_m}[\vec{x}_m]\rangle$ a *solution* to this problem.

In the solution to a mapping problem, $S_i^{\tau_i}[\vec{x}_i]$'s represent the 'meaning recipe' of individual words, where the free variables $\vec{x}_i$ are supposed to be filled by conceptual symbols. $D^\rho[y_1^{\tau_1},\ldots,y_m^{\tau_m}]$ then gives the meaning recipe for the sentence. Thus I assume that simultaneous binding of multiple occurrences of a variable is possible in word meanings, but not in meaning recipes for sentences.[1]

Note that in the above formulation, the free variables $\vec{x}_1,\ldots,\vec{x}_m$ in a mapping problem $T^\rho[\vec{x}_1;\ldots;\vec{x}_m]$ are all distinct; in actual learning, the conceptual symbol sets of some words in a sentence may overlap. So the problem for the learner facing a single sentence-meaning pair corresponds, in general, to (a disjunction of) multiple mapping problems. Also, the learner may have already formed a hypothesis meaning for some of the words in the sentence currently processed, in which case an additional issue of combining two constraints arises (more on this below). For these reasons, the mapping problem as formulated above models only the 'heart' of the problem for the learner. Nevertheless, it is a good place to start formal analysis.

A 'mapping problem' of a more restricted kind corresponding to Siskind's setup always has at most one solution, and the condition under which a solution exists is easy to state. As for the general version, we have

Theorem 1. *Every mapping problem has infinitely many solutions.*

Theorem 2. *For any sequence of terms* $\vec{M} = M_1^{\sigma_1},\ldots,M_n^{\sigma_n}$, *let* $U_\rho[\vec{M}] = \lambda y^{\sigma_1\to\cdots\to\sigma_n\to\rho}.yM_1\ldots M_n$. *Then, for every mapping problem* $T^\rho[\vec{x}_1;\ldots;\vec{x}_m]$ *such that* $T^\rho[\vec{x}_1,\ldots,\vec{x}_m]$ *is a BCIλ-term,* $\langle U_\rho[\vec{x}_1],\ldots,U_\rho[\vec{x}_m]\rangle$ *solves* $T^\rho[\vec{x}_1;\ldots;\vec{x}_m]$.

We can also prove a slightly more complex form of Theorem 2 for λI mapping problems in general.

'Universal' solutions like $U_\rho[\vec{x}]$ are not very interesting: they not only derive correct sentence meanings but also many unwanted ones. The following example illustrates the case of a λI mapping problem.

Example 3. Consider a mapping problem $T^t[\vec{x}_1;\vec{x}_2]$, where $\vec{x}_1 = x_{1,1}^{(e\to t)\to t}, x_{1,2}^{t\to t\to t}, x_{1,3}^{e\to t}$, $\vec{x}_2 = x_{2,1}^{e\to t}$, and $T^t[\vec{x}_1,\vec{x}_2] = x_{1,1}(\lambda z^e.x_{1,2}(x_{1,3}z)(x_{2,1}z))$. This corresponds to the situation where the learner is faced with a sentence *everyone walks* coupled with the meaning:

$$\forall z(\mathrm{PERSON}(z) \to \mathrm{WALK}(z)),$$

having arrived at the word-to-symbol-set association:

everyone	:	$\{\forall, \to, \mathrm{PERSON}\}$
walks	:	$\{\mathrm{WALK}\}$

A solution $\langle U_t[\vec{x}_1, \mathsf{W}^{(e\to e\to t)\to e\to t}],\ x_{2,1}\rangle$, where $\mathsf{W}^{(e\to e\to t)\to e\to t} = \lambda y^{e\to e\to t}x^e.yxx$, corresponds to the word-to-meaning mapping:

everyone	$\mapsto$	$\lambda y^{((e\to t)\to t)\to(t\to t\to t)\to(e\to t)\to((e\to e\to t)\to e\to t)\to t}.y(\forall)(\to)(\mathrm{PERSON})(\mathsf{W})$
walks	$\mapsto$	WALK

$\langle U_t[\vec{x}_1, \mathsf{W}^{(e\to e\to t)\to e\to t}],\ x_{2,1}\rangle$ solves not only $T^t[\vec{x}_1;\vec{x}_2]$, but also $V^t[\vec{x}_1;\vec{x}_2] = x_{1,1}(\lambda z^e.x_{1,2}(x_{2,1}z)(x_{1,3}z))$, corresponding to

$$\forall z(\mathrm{WALK}(z) \to \mathrm{PERSON}(z)).$$

[1]One might want to restrict $D^\rho[y_1^{\tau_1},\ldots,y_m^{\tau_m}]$ even further to the so-called Lambek-van Benthem fragment (van Benthem 1991); this will complicate some results slightly.

Of course, $\langle \lambda v^{e\to t}.x_{1,1}(\lambda z^e.x_{1,2}(x_{1,3}z)(vz)),\ x_{2,1}\rangle$, corresponding to the word-to-meaning mapping:

everyone	$\mapsto$	$\lambda v^{e\to t}.\forall z(\mathrm{PERSON}(z) \to vz)$
walks	$\mapsto$	WALK

only solves the former.

Certainly, the second solution in the above example is preferable to the first as a hypothesis for the learner. Perhaps what the learner should find is a 'minimal' solution, in the sense of solving as few unwanted problems as possible. To define a suitable sense of minimality, I introduce a 'definability relation' $\preceq$ between terms and a slightly weaker notion $\preceq_\rho$.

Definition 4. Let $M^\sigma[\vec{x}]$ and $N^\tau[\vec{x}]$ be two λ-terms in βη-normal form with the same set of free variables.

(i) We say that $N^\tau[\vec{x}]$ is *BCI-definable by* $M^\sigma[\vec{x}]$ and write $N^\tau[\vec{x}] \preceq M^\sigma[\vec{x}]$ if there is a closed *BCIλ*-term $P^{\sigma\to\tau}$ such that

$$P^{\sigma\to\tau}M^\sigma[\vec{x}] \rhd_{\beta\eta} N^\tau[\vec{x}].$$

We write $N^\tau[\vec{x}] \simeq M^\sigma[\vec{x}]$ if $N^\tau[\vec{x}] \preceq M^\sigma[\vec{x}]$ and $N^\tau[\vec{x}] \succeq M^\sigma[\vec{x}]$; and write $N^\tau[\vec{x}] \prec M^\sigma[\vec{x}]$ if $N^\tau[\vec{x}] \preceq M^\sigma[\vec{x}]$ but $N^\tau[\vec{x}] \not\succeq M^\sigma[\vec{x}]$.

(ii) We write $N^\tau[\vec{x}] \preceq_\rho M^\sigma[\vec{x}]$ if $\lambda y^{\tau\to\rho}.yN^\tau[\vec{x}] \preceq M^\sigma[\vec{x}]$. We write $N^\tau[\vec{x}] \simeq_\rho M^\sigma[\vec{x}]$ if $N^\tau[\vec{x}] \preceq_\rho M^\sigma[\vec{x}]$ and $N^\tau[\vec{x}] \succeq_\rho M^\sigma[\vec{x}]$; and write $N^\tau[\vec{x}] \prec_\rho M^\sigma[\vec{x}]$ if $N^\tau[\vec{x}] \preceq_\rho M^\sigma[\vec{x}]$ but $N^\tau[\vec{x}] \not\succeq_\rho M^\sigma[\vec{x}]$.

Clearly, $N^\tau[\vec{x}] \preceq M^\sigma[\vec{x}]$ implies $N^\tau[\vec{x}] \preceq_\rho M^\sigma[\vec{x}]$ for all ρ. Also, if τ 'ends in' ρ, $N^\tau[\vec{x}] \preceq_\rho M^\sigma[\vec{x}]$ implies $N^\tau[\vec{x}] \preceq M^\sigma[\vec{x}]$. The two relations do not coincide in general: $x^e \preceq_t \lambda y^{e\to t}.yx$, but $x^e \not\preceq \lambda y^{e\to t}.yx$. Both $\preceq$ and $\preceq_\rho$ are reflexive and transitive. Also, both relations are decidable. The latter follows from the known fact that every type is inhabited by finitely many *BCIλ*-terms in βη-normal form (van Benthem 1991).

Theorem 5. *If* $\langle S_1[\vec{x}_1],\ldots,S_m[\vec{x}_m]\rangle$ *solves* $T^\rho[\vec{x}_1;\ldots;\vec{x}_m]$, *and* $S_i[\vec{x}_i] \preceq_\rho P[\vec{x}_i]$, *then* $\langle S_1[\vec{x}_1],\ldots,S_{i-1}[\vec{x}_{i-1}],P[\vec{x}_i],S_{i+1}[\vec{x}_{i+1}],\ldots,S_m[\vec{x}_m]\rangle$ *solves* $T^\rho[\vec{x}_1;\ldots;\vec{x}_m]$.

Even though a mapping problem has infinitely many solutions, many of them turn out to be equivalent with respect to $\simeq$ and $\simeq_\rho$. A trivial example is $T^t[x^e;y^{e\to t}] = yx$, which has just two solutions modulo $\simeq$, namely $\langle x,\ y\rangle$ and $\langle \lambda y^{e\to t}.yx,\ y\rangle$ (which, in turn, are equivalent with respect to $\simeq_t$). I conjecture that every mapping problem has only finitely many solutions modulo $\simeq$.

The relations $\preceq$ and $\preceq_\rho$ can be naturally extended to solutions of mapping problems. In general, a mapping problem $T^\rho[\vec{x}_1;\ldots;\vec{x}_n]$ can have more than one $\preceq_\rho$-incompatible minimal solution.

Example 6. A problem $T^t[x^e;y^{e\to e\to t};z^e] = yzx$ has the following solutions, among others:

$$\langle \lambda y^{e\to e\to t}z^e.yzx,\ y,\ \lambda y^{e\to e\to t}v^{(e\to e\to t)\to e\to t}.v(\lambda xz.yzx)z\rangle,$$
$$\langle \lambda y^{e\to e\to t}v^{(e\to e\to t)\to e\to t}.vyx,\ y,\ \lambda y^{e\to e\to t}.yz\rangle.$$

Since $\lambda y^{e\to e\to t} z^e.yzx \succ_t \lambda y^{e\to e\to t} v^{(e\to e\to t)\to e\to t}.vyx$ and $\lambda y^{e\to e\to t} v^{(e\to e\to t)\to e\to t}.v(\lambda xz.yzx)z\rangle \prec_t \lambda y^{e\to e\to t}.yz$, the two solutions are incompatible with respect to $\preceq_t$. Their 'meet',

$$\langle \lambda y^{e\to e\to t} z^e.yzx,\ y,\ \lambda y^{e\to e\to t}.yz\rangle,$$

does not solve $T_t[x^e; y^{e\to e\to t}; z^e]$, and indeed the two solutions can be shown to be minimal with respect to $\preceq_t$. (Note that in this particular example, all solutions solve the same two problems, namely, $T^t[x^e; y^{e\to e\to t}; z^e]$ and $V^t[x^e; y^{e\to e\to t}; z^e] = yxz$. In the presence of an additional word-meaning recipe (e.g., $u^{e\to t}$), however, they generate different sets of sentence meanings, so they should not be treated as equivalent.)

The above example suggests that it may not be desirable for the purpose of learning to collect all minimal solutions to a given mapping problem. Below, I present a simple algorithm that finds a 'reasonably small' solution to a mapping problem.

In a term $\lambda x_1 \dots x_n.yN_1 \dots N_m$ in β-normal form, each N_i is called an *argument*. A *subargument* is either an argument or a subargument of an argument.

Algorithm A. Let a mapping problem $T^\rho[\vec{x}_1; \dots; \vec{x}_m]$ be given. Associate with each $\vec{x}_i$ the smallest subargument P_i that contains all occurrences of variables in $\vec{x}_i$. Reorder $\vec{x}_1, \dots, \vec{x}_m$ so that if P_i is a subargument of P_j, $i < j$. At the i-th cycle of the following iteration, a new variable u_i of a suitable type will be created. $V_i^\rho[\vec{x}_{i+1}, \dots, \vec{x}_m]$ will contain as its free variables some (but not necessarily all) of $u_1, \dots, u_i$, in addition to $\vec{x}_{i+1}, \dots, \vec{x}_m$.

Let $V_0^\rho[\vec{x}_1, \dots, \vec{x}_m] = T^\rho[\vec{x}_1, \dots, \vec{x}_m]$.

For $i = 1, \dots, m$, do the following:

- Find the smallest subargument $Q_i^\tau[\vec{x}_i, v_1, \dots, v_k, y_1^{\tau_1}, \dots, y_l^{\tau_l}]$ of $V_{i-1}^\rho[\vec{x}_i, \dots, \vec{x}_m]$ that contains all occurrences of variables in $\vec{x}_i$, where $v_1, \dots, v_k$ are among $u_1, \dots, u_{i-1}$, and $y_1^{\tau_1}, \dots, y_l^{\tau_l}$ are the free variables of Q_i^τ other than $\vec{x}_i, v_1, \dots, v_k$.
- Let $S_i[\vec{x}_i] = \lambda v_1 \dots v_k y_1^{\tau_1} \dots y_l^{\tau_l}.M^\tau[\vec{x}_i, v_1, \dots, v_k, y_1^{\tau_1}, \dots, y_l^{\tau_l}]$.
- Replace Q_i^τ by $u_i^{\tau_1\to\dots\to\tau_l\to\tau} y_1^{\tau_1} \dots y_l^{\tau_l}$ in $V_{i-1}[\vec{x}_i, \dots, \vec{x}_m]$, obtaining $V_i[\vec{x}_{i+1}, \dots, \vec{x}_m]$.

When the above procedure is over, $\langle S_1[\vec{x}_1], \dots, S_m[\vec{x}_m]\rangle$ is a solution to $T^\rho[\vec{x}_1; \dots; \vec{x}_m]$.

Algorithm A is nondeterministic and different executions can lead to different results when $P_i = P_j$ for some $i \neq j$. This can only be so if the head variable of P_i is not among $\vec{x}_i$ for some i; otherwise the output of the algorithm is unique.

Example 7. Let us apply Algorithm A to the mapping problem corresponding to the following situation. The learner has arrived at the word-to-symbol-set association:

every	:	$\{\forall, \to\}$
man	:	{MAN}
finds	:	{FIND}
a	:	$\{\exists, \wedge\}$
unicorn	:	{UNICORN}

and is faced with the sentence *every man finds a unicorn*, coupled with the meaning (2). Algorithm A produces just one result in this case, and the solution corresponds to the following word-to-meaning mapping:

$$\begin{array}{lll} every & \mapsto & \lambda u^{e\to t} v^{e\to t}.\forall(\lambda x^e. \to (ux)(vx)) \\ man & \mapsto & \mathrm{MAN} \\ finds & \mapsto & \mathrm{FIND} \\ a & \mapsto & \lambda u^{e\to t} w^{e\to e\to t} x^e.\exists(\lambda y^e.\wedge(uy)(wyx)) \\ unicorn & \mapsto & \mathrm{UNICORN} \end{array} \qquad (4)$$

Note that the solution corresponding to (4) in the above example is smaller ($\prec_t$) than the solution corresponding to (3). The relative merit of (4) vis-a-vis (3) is debatable. It is possible to modify Algorithm A, so that it produces (3) rather than (4) on this example, but the resulting variant will be a more complicated algorithm.[2]

A mapping problem is supposed to model an individual stage in the second phase of the learning process, when the learner is processing a single sentence coupled with its meaning. In actual learning, some words in the current sentence may have already appeared in the preceding sentences during the second phase, so that the learner has meanings already hypothesized for them. When an algorithm like Algorithm A is applied to the mapping problem corresponding to the current sentence-meaning pair, it may produce a meaning for some old word that is $\preceq_t$-incompatible with the meaning already hypothesized. In such a case, the learner presumably needs to find an upper bound with respect to $\preceq_t$ of the two word meanings, assuming that the old hypothesis is not to be completely abandoned. An upper bound always exists ($U_t[\vec{x}]$ is one), but to formulate an algorithm for finding a reasonably small upper bound is not trivial. A question like this motivates investigation of the structure of the partial orders on equivalence classes of terms that are induced by the two relations $\preceq$ and $\preceq_\rho$. I have to leave this and other issues for future work.

References

Benthem, Johan van. 1991. *Language in Action: Categories, Lambdas and Dynamic Logic*. Amsterdam: North-Holland. (Second edition, 1995, Cambridge, Mass.: MIT Press.)

Hindley, J. Roger. 1997. *Basic Simple Type Theory*. Cambridge: Cambridge University Press.

Siskind, J. Mark. 1996. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition* **61**(1–2), 39–91. (Reprinted in Michael Brent, ed., 1996, *Computational Approaches to Language Acquisition*, pp. 39–91, Amsterdam: Elsevier.)

Siskind, J. Mark. 2000. Learning word-to-meaning mappings. In Peter Broeder and Jaap Murre, eds., *Models of Language Acquisition: Inductive and Deductive Approaches*, pp. 121–153. Oxford: Oxford University Press.

[2]Algorithm A (and its variant) can produce a counter-intuitive result when a duplicating β-contraction is involved in the reduction $D[S_1[\vec{x}_1], \dots, S_n[\vec{x}_n]] \rhd_{\beta\eta} T[\vec{x}_1, \dots, \vec{x}_n]$, where $\langle S_1[\vec{x}_1], \dots, S_n[\vec{x}_n]\rangle$ is the desired solution.

Tense Probabilism Properly Conceived

Stefan Kaufmann
Kyoto University*

1 Introduction

"Tense Probabilism" (henceforth TP) is Barker's (1998) term for the doctrine that (i) the probabilities of conditionals are conditional probabilities and (ii) differences in probability between different "versions" of conditionals are due to differences in temporal reference. One of its tenets can be paraphrased as follows:

Thesis 1 (Tense Probabilism)
The value of a counterfactual at a given time is that of its indicative predictive counterpart at an earlier time.[1]

Thus assuming that Oswald killed Kennedy and that he acted alone, (1b) is unlikely now because (1a) *was* unlikely on the morning of November 22, 1963.

(1) a. If Oswald does not kill Kennedy, someone else will.
b. If Oswald had not killed Kennedy, someone else would have.
c. If Oswald did not kill Kennedy, someone else did.

Many authors have made this observation and incorporated it in their semantic accounts in various ways (e.g., Adams, 1975; Skyrms, 1981; Tedeschi, 1981; Ellis, 1984; Dudman, 1984, 1994; Edgington, 1995; Dahl, 1997; Dancygier, 1998).

Notice that the value of (1c) is not as clearly related to either one of the others. Here I will focus on the relationship between (1a) and (1b).

The problem

The above formulation of Thesis 1 leaves room for interpretation. The question of interest here is which "values" of sentences like (1a,b) are claimed to be correlated, and how.

Barker (1998) reads "value" as "probability" and shows that under this reading Thesis 1 is at odds with both naïve intuitions about certain examples and the commitment to non-determinism implicit in TP.

That TP implicitly assumes non-determinism is shown as follows: Despite the fact that Oswald killed Kennedy, (1b) may have high probability. Given Thesis 1 (under Barker's reading), this implies that (1a) had high probability at an earlier time, which by TP must be the conditional probability (at the time) that someone else will kill Kennedy, given that Oswald doesn't. But if Oswald's killing Kennedy had been brought about deterministically, the antecedent of (1a) would never have had non-zero probability, hence the required conditional probability would not be defined and Thesis 1 would be refuted.

*This paper owes much to comments and criticisms by Stanley Peters, David Beaver, Cleo Condoravdi, Peter Sells, and numerous others. The author gratefully acknowledges the generous support of the Japanese Society for the Promotion of Science (JSPS, Grant #P01063) and Satoshi Sato's hospitality during the completion of this manuscript.

[1]This statement is stronger than Barker's, which says that if the value of counterfactual is high now, then that of the indicative was high at an earlier time.

Barker argues that its alignment with non-determinism notwithstanding, TP makes the wrong predictions in situations involving genuine non-determinism. The argument involves sentences like (2a,b). Suppose the coin is fair and comes up heads.

(2) a. If I bet on tails, I will lose. [before the toss]
b. If I had bet on tails, I would have lost. [after the toss]

After the toss, since the coin came up heads, (2b) is highly probable. In contrast, before the toss, the probability that the coin would come up heads was .5, hence (2a) was not highly probable, barring the assumption (which is incompatible with TP) that the outcome was already determined.

Intuitively, the problem is that while the posterior fact that the coin came up heads is taken into account in interpreting (2b), it seems irrelevant to (2a); nor does there appear to be a coherent way within TP to bring it to bear on (2a). Barker concludes that TP has no promise of providing a unified account of counterfactuals and predictive indicatives. I am going to show that this conclusion is not inevitable.

2 Values and probabilities

Barker's is not the only reading of Thesis 1. I propose that a distinction be drawn between the *values* conditionals take at individual worlds on the one hand, and their *probabilities*, on the other. This distinction provides for an improved statement of Thesis 1 according to which conditionals like (2a,b) are *equivalent* though not *equiprobable*.

To prepare the ground for the argument, I will first pin down the relationship between truth values and probability. The basic framework is due to Jeffrey (1991); Stalnaker and Jeffrey (1994) and ultimately to De Finetti.

Let $\langle W, Pr\rangle$ be a set of worlds with a probability distribution $Pr : \wp(W) \mapsto [0,1]$, and let $X : W \mapsto \mathbb{R}$ be a *random variable* taking real numbers at the worlds in W. The *expectation* of X is as in (3).[2]

(3) $$E[X] = \sum_{x \in range(X)} x \cdot Pr(\{w \in W | X(w) = x\})$$

The denotations of atomic or truth-functional sentences, too, are random variables taking values in $\{0,1\}$ point-wise at worlds. The expectation of such a sentence is the probability that it is true. This connection is used to define a probability distribution P (distinct from Pr) over the sentences of the language. Let $V : \mathcal{L} \mapsto (W \mapsto \{0,1\})$ be a truth assignment to the language $\mathcal{L}$ of propositional logic. Then for each sentence $\varphi \in \mathcal{L}$, the probability $P(\varphi)$ is defined as the expectation of $V(\varphi)$; since the range of $V(\varphi)$ is $\{0,1\}$, this comes down to (4).

(4) $$P(\varphi) = \sum_{x \in range(V(\varphi))} x \cdot Pr(\{V(\varphi) = x\}) = Pr(\{V(\varphi) = 1\})$$

[2]Henceforth I abbreviate '$\{w \in W | X(w) = x\}$' as '$\{X = x\}$.' In the continuous case, the summation is replaced by an integral.

Extending the language to include the "natural" conditional '$\rightarrow$' (to be distinguished from the material conditional '$\supset$'), TP requires that the expectation of $V(A \rightarrow C)$ equal the conditional probability $P(C|A) = P(AC)/P(A)$ whenever the latter is defined. This has proven elusive. A long line of research, starting with the *triviality results* of Lewis (1976, 1986b) and summarized in Hájek and Hall (1994), has compiled overwhelming evidence that if the values of $V(A \rightarrow C)$ are to add up to the conditional probability, they cannot be truth values—assuming that the truth values of a sentence are, at each world, (i) either 0 or 1 and (ii) constant.

Rather than retracing these arguments here, I am going to build on an alternative proposal due to Jeffrey (1991) (cf. also van Fraassen, 1976; Stalnaker and Jeffrey, 1994): For (non-counterfactual) conditionals, the assignment is

(5)
$$V(A \rightarrow C)(w) = \begin{cases} V(C)(w) \text{ if } V(A)(w) = 1 \\ E[V(C)|V(A) = 1] \text{ if } V(A)(w) = 0 \end{cases}$$

An intuitive rationale for (5) may be gleaned from Stalnaker's (1968) theory of counterfactuals. If A is false at w, the value of the conditional depends on the "nearest" A-world, uniquely identified in Stalnaker's models by a selection function. (5) does not require that a single nearest A-world be found; instead, the selection function picks an A-world *at random* and assigns the conditional its value according to the upper line in (5). The value of the conditional at w is the expectation of this random trial.

Notice that according to (5), at worlds where A is false, the values of the conditional may be intermediate between 0 and 1, and furthermore, those values depend on Pr (via the expectation), hence are not constant. If Pr is interpreted as *subjective* probabilities representing the beliefs of an agent, mixing these intermediate values with (objective) truth values calls for some explanation (cf. Edgington, 1995). I will not discuss this issue here since does not arise with predictive conditionals. Here Pr is interpreted as objective chance, fully determined by the world and time of evaluation (cf. Lewis, 1980).[3]

Notice further that according to (5), the values of the conditional are uniformly distributed over all non-antecedent worlds, independently of any "third facts." Figuratively speaking, the set of "visible" A-worlds over which the expectation is taken is the same for all $\overline{A}$-worlds. This is the root of the problem pointed out by Barker.[4]

3 Counterfactuals

In the literature on counterfactuals in time, the lesson from examples like (2) has long been recognized to be that "Past Predominance" (Thomason and Gupta, 1981) must be balanced against "Overall Similarity" (Lewis, 1979). (2) has

[3]Formally, this can be made explicit by enriching the model with a temporal dimension to which the value assignment is made sensitive. A "$T \times W$-frame" as defined by Thomason (1984) is a suitable structure, but in the interest of simplicity I am not going to spell out the details here (see Kaufmann, 2002).

[4]There are a number of other reasons to look for improvements of (5), some of which (Lance, 1991; Edgington, 1991) involve compounds of conditionals and fall outside the scope of this paper. The solution proposed here addresses those problems as well, as I will show on some other occasion.

been discussed in this connection by Slote (1978); Bennett (1984); Mårtensson (1999), among others.

The underlying question is whether actual facts at times later than that of the (hypothesized) antecedent should be "carried over" in examining alternative worlds. (2) suggests that they should: Had I bet on tails, the coin would still have come up heads and I would have lost. (1) suggests that they should not: Had Oswald not killed Kennedy, Kennedy would not have been killed.

The consensus is that the relevant distinction between examples like (2) and those like (1) concerns the *causal dependencies* of the respective scenarios: The outcome of the coin toss is causally independent of my not betting (notice that judgments about (2) change if, e.g., a different fair coin is used when I bet on tails), whereas Kennedy's fate causally depends on Oswald's behavior. Barker's objection, too, is properly addressed by augmenting the model with causal information.

4 Causality

Causal relations are taken here to hold between event *tokens* such that the cause determines the probability of the effect (Hausman, 1998; Pearl, 2000). At each world, the presence or absence of cause and effect are indicated by random variables (not necessarily denoted by sentences of the language). Thus 'X *causally affects* Y' means 'the *value* of X determines the *expectation* of Y.'[5]

Given the set W of worlds, a set Φ of functions $X : W \mapsto \{0, 1\}$ is singled out as the collection of *causally relevant* variables. The '*causally affects*'-relation is encoded as a strict partial order '$\prec$' on Φ. All *descendants* of a variable X in $\langle \Phi, \prec \rangle$ (i.e., all Y such that $X \prec Y$) are causally affected by X.

Φ induces a partition on the set of worlds, each of whose cells comprises those worlds which agree on the values of all variables in Φ. Crucially, the assignment of values to conditionals, counterfactual or not, starts "locally" within those cells. Let $\mathcal{C}$ be such a cell and w a world in $\mathcal{C}$. Suppose $V(A)$ is false at w. Then for a conditional $A \rightarrow C$, if the conditional expectation of C given A is defined within $\mathcal{C}$, it is the value of $V(A \rightarrow C)(w)$. This is similar to (5) above, but now the conditional expectation is taken only over those antecedent-worlds which agree with w on the causally relevant variables.

If, on the other hand, the conditional expectation is not defined within $\mathcal{C}$ (i.e., if A has zero probability given $\mathcal{C}$), a larger cell $\mathcal{C}'$ is made available by "undoing" the falsehood of $V(A)$.[6] This is where the causal order on variables matters: The values of the variables that are causally affected by $V(A)$ (its descendants in the causal order) are "given up" as well, whereas the other variables retain the values they have at w. Thus in (2), where no causal influence of the bet upon the outcome is assumed, from a world at which I do not bet on tails and the coin comes up tails, only those "tail-betting" worlds are visible in which the coin comes up tails as well. Similarly for heads.

This is visualized in Figure 1, where the coloring symbolizes the values assigned to (2a) before the coin toss. The values according to 5 (on the left) are

[5]If Y causally depends on additional variables besides X, its expectation is determined by a joint setting of *all* those causal factors, including X.

[6]This resembles the "fattening" of Skyrms (1984, 1994), but there are differences when there are multiple mutually exclusive ways of making A true; I will not discuss this case here.

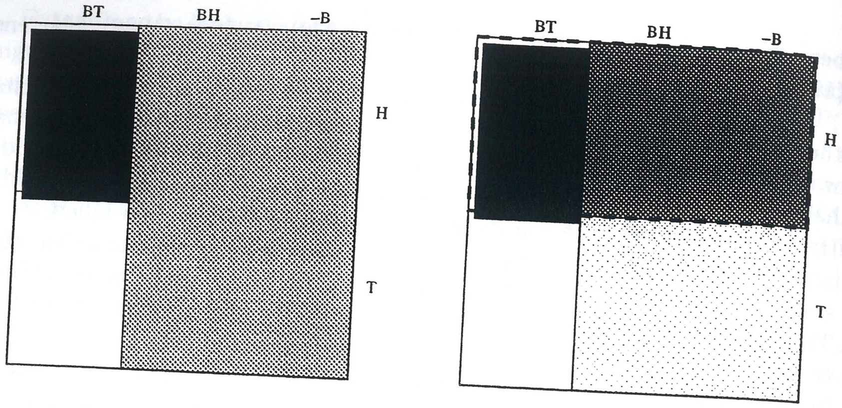


Figure 1: Distribution of values of (2); Black = 1; White = 0; Grey = intermediate, without (left) and with (right) causal information

distributed uniformly over all worlds at which I do not bet on tails, regardless of whether the coin comes up heads or tails. On the right, the distribution of values with reference to causality is given: At different non-antecedent worlds the conditional takes different values depending on the result of the toss.

To augment the interpretation function accordingly, the change to the clause for non-antecedent worlds in (5) consists in adding restrictions to the conditional expectation for worlds at which the antecedent is false: Besides '$V(A) = 1$', '$X = x$' is added for each *non-descendant* X of $V(A)$ in $\langle \Phi, \prec \rangle$ such that $X(w) = x$.

$$(6)\quad V(A \rightarrow C)(w) = \begin{cases} V(C)(w) \text{ if } V(A)(w) = 1 \\ E[V(C) | V(A) = 1, X_i = x_i] \\ \quad \text{for all } X_i \in \Phi \text{ s.t. } V(A) \prec X_i \text{ and } X_i(w) = x_i, \\ \quad \text{otherwise} \end{cases}$$

5 Conclusion

I have proposed a value assignment which makes reference to posterior facts (at times later than that of the antecedent) in the case of predictive conditionals. Is this legitimate? To see in what sense it is, we need to keep apart the notions of *truth* and *settledness*, familiar since Prior (1967) and Thomason (1970).

At the time the prediction in (2a) is made, there is no way of knowing which way the coin will land, thus both outcomes are open possibilities. Worlds of both types are indistinguishable with respect to past and present facts, so the uncertainty is objective and irreducible. It is not *settled* that the coin will come up heads; yet at each individual world it is already either *true* or *false*.

A similar statement can be made about conditionals: Although it is not *settled* what the actual value of the conditional is, at each individual world that value is already determined by the facts.

Thus as a predictive conditional "turns into" a counterfactual, it retains its values at individual worlds while its probability (i.e., the expectation of its values) may change simply due to the elimination of historical alternatives.

Before the toss, the value of (2a) is high at those non-antecedent worlds at which the coin comes up heads. Once the outcome (heads) is settled, only those worlds remain at which the value was high *already*. Hence the above restatement of Thesis 1 to the effect that the predictive indicative and its counterfactual counterpart are *equivalent* but not *equiprobable*.

Finally, an improved statement of the relationship between the probabilities of the two sentences becomes available as well: The probability of the indicative is the expectation of the probabilities that the corresponding counterfactual will have in the various cells of the partition induced by the causally related factors. Thus the probability of (2a) is the weighted sum of the probabilities of (2b), where the weights are the prior probabilities that (2b) will take those values.

References

Adams, E. 1975. *The Logic of Conditionals*. Reidel.

The Aristotelian Society. 1991. *The Symposia Read at the Joint Session of the Aristotelian Society and the Mind Association at the University of Durham*. Supplementary Volume 65.

Athanasiadou, A. and R. Dirven, editors. 1997. *On Conditionals Again*, volume 143 of *Amsterdam Studies in the Theory and History of Linguistic Science*. John Benjamins, Amsterdam / Philadelphia.

Barker, S. 1998. Predetermination and tense probabilism. *Analysis*, 58:290–296.

Bennett, J. 1984. Counterfactuals and temporal direction. *The Philosophical Review*, 93(1):57–91.

Dahl, O. 1997. The relation between past time reference and counterfactuality: A new look. In Athanasiadou and Dirven (1997), pages 97–112.

Dancygier, B. 1998. *Conditionals and Prediction*. Cambridge University Press.

Dudman, V. 1984. Parsing 'If'-sentences. *Analysis*, 44(4):145–153.

Dudman, V. 1994. On conditionals. *Journal of Philosophy*, 91:113–128.

Edgington, D. 1991. The mystery of the missing matter of fact. In Aristotelian Society Ari (1991), pages 185–209. Supplementary Volume 65.

Edgington, D. 1995. On conditionals. *Mind*, 104(414):235–329.

Eells, E. and B. Skyrms, editors. 1994. *Probabilities and Conditionals: Belief Revision and Rational Decision*. Cambridge University Press, Cambridge.

Ellis, B. 1984. Two theories of indicative conditionals. *Australasian Journal of Philosophy*, 62(1):50–66.

van Fraassen, B. C. 1976. Probabilities of conditionals. In Harper, W. L., R. Stalnaker, and G. Pearce, editors, *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, volume 1 of *The University of Western Ontario Series in Philosophy of Science*, pages 261–308. D. Reidel.

Gabbay, D. and F. Guenthner, editors. 1984. *Extensions of Classical Logic*, volume 2 of *Handbook of Philosophical Logic*. D. Reidel.

Grandy, R. and R. Warner, editors. 1986. *Philosophical Grounds of Rationality*. Clarendon Press, Oxford.

Hájek, A. and N. Hall. 1994. The hypothesis of the conditional construal of conditional probability. In Eells and Skyrms (1994), pages 75–110.

Harper, W., R. Stalnaker, and G. Pearce, editors. 1981. *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Reidel.

Hausman, D. M. 1998. *Causal Asymmetries*. Cambridge University Press, Cambridge, UK.

Jackson, F., editor. 1991. *Conditionals*. Oxford University Press.

Jeffrey, R. C., editor. 1980. *Studies in Inductive Logic and Probability*, volume 2. University of California Press.

Jeffrey, R. C. 1991. Matter-of-fact conditionals. In Aristotelian Society Ari (1991), pages 161–183. Supplementary Volume 65.

Kaufmann, S. 2002. *Aspects of the Meaning and Use of Conditionals*. PhD thesis, Stanford University.

Lance, M. 1991. Probabilistic dependence among conditionals. *Philosophical Review*, 100:269–276.

Lewis, D. 1976. Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 85:297–315. Reprinted with a Postscript in Jackson (1991), pages 76–101. Page numbers refer to Jackson (1991).

Lewis, D. 1979. Counterfactual dependence and time's arrow. *Noûs*, 13:455–476. Reprinted with Postscripts in Lewis (1986a), pages 32–66, and Jackson (1991); pages 46–75. Page numbers refer to Lewis (1986a).

Lewis, D. 1980. A subjectivist's guide to objective chance. In Jeffrey (1980), pages 263–293. Reprinted in Harper et al. (1981), pages 267–297; with a Postscript in Lewis (1986a), pages 83–132. Page numbers refer to Lewis (1986a).

Lewis, D. 1986a. *Philosophical Papers*, volume 2. Oxford University Press.

Lewis, D. 1986b. Probabilities of conditionals and conditional probabilities II. *Philosophical Review*, 95:581–589. Reprinted in Jackson (1991), pages 102–110; Lewis (1998), pages 57–65. Page numbers refer to Lewis (1998).

Lewis, D. 1998. *Papers in Philosophical Logic*. Cambridge University Press.

Mårtensson, J. 1999. *Subjunctive Conditionals and Time: A Defense of the Classical Approach*. Number 10 in Acta Philosophica Gothoburgensia. Department of Philosophy, University of Göteborg.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

Prior, A. 1967. *Past, Present and Future*. Oxford University Press.

Skyrms, B. 1981. The prior propensity account. In Harper et al. (1981), pages 259–265.

Skyrms, B. 1984. *Pragmatics and Empiricism*. Yale University Press, New Haven / London.

Skyrms, B. 1994. Adams conditionals. In Eells and Skyrms (1994), pages 13–26.

Slote, M. A. 1978. Time in counterfactuals. *The Philosophical Review*, 87(1):3–27.

Sosa, E., editor. 1975. *Causation and Conditionals*. Oxford University Press.

Stalnaker, R. 1968. A theory of conditionals. In *Studies in Logical Theory, American Philosophical Quarterly, Monograph: 2*, pages 98–112. Blackwell. Reprinted in Sosa (1975), pages 165–179; Harper et al. (1981), pages 41–55; Jackson (1991), pages 28–45. Page numbers refer to Harper et al. (1981).

Stalnaker, R. and R. Jeffrey. 1994. Conditionals as random variables. In Eells and Skyrms (1994), pages 31–46.

Tedeschi, P. 1981. Some evidence for a branching-futures semantic model. In Tedeschi, P. and A. Zaenen, editors, *Tense and Aspect*, volume 14 of *Syntax and Semantics*, pages 239–270. Academic Press, New York.

Thomason, R. H. 1970. Indeterminist time and truth value gaps. *Theoria*, 36:264–281.

Thomason, R. H. 1984. Combinations of tense and modality. In Gabbay and Guenthner (1984), pages 135–165.

Thomason, R. H. and A. Gupta. 1981. A theory of conditionals in the context of branching time. In Harper et al. (1981), pages 299–322.

Categorial-Hybrid Logical Grammar

Geert-Jan M. Kruijff

1. Introduction

In contemporary categorial grammar, the representation of an expression's meaning is the *only* level of representation. Categorial inference and the syntactic structure it creates are only interesting insofar as they assist in building a representation of meaning – they are, to all purposes and intents, mere vehicles. Naturally this raises the questions of *how* to represent a meaning a sentence realizes linguistically, and *what* to represent as part of such a meaning.

Traditionally, categorial grammar has captured meaning using a (simply typed) λ-calculus, building a representation in parallel to the categorial inference. Here, we propose an alternative way to representing linguistically realized meaning: Namely, as a term in a *hybrid modal logic* (Blackburn, 2000). There are two main linguistic motivations for doings so.

Firstly, we want to conceive of meaning as a *relational structure* in which the role of a node in a tree is indicated by the relation it bears to other nodes – unlike terms in a (typed) lambda calculus where the meaning of an argument is defined arbitrarily by its position in a predicate, cf. (Dowty, 1989). Not only does the relational view lead to a fine-grained description of meaning, but distinguishing argument roles is in fact necessary to give an adequate account of for example aspectual change (Dowty, 1979), or information structure and its realization (Kruijff, 2001). Modal logic is the most suitable logic for describing such relational structures.

Secondly, meaning is *ontologically rich* – as illustrated by for example theories of tense and aspect, but also of objects and properties. Hybrid modal logic enables us to shift semantic complexity to a sorting strategy, as for example suggested in (Van Benthem, 1996). This shift gives rise to the possibility of creating such ontologically rich logics of meaning. At the same time, hybrid modal logics usually do still enjoy decidability and computational tractability (Areces et al., 1999), despite their richness.

An overview of the extended abstract is as follows. In §2 we briefly explain hybrid logic, and how can we use it to model a relational perspective on meaning. In §3 we present a resource-sensitive proof calculus that enables us to construct such relational meaning representations in parallel to a categorial inference. We present a few examples of the approach in §4. The work presented in this paper is based on (Kruijff, 2001).

2. Hybrid logic and modelling meaning

Hybrid logic is a modal logic that provides us with means to logically capture two essential aspects of meaning in a clean and compact way, namely ontological richness and the possibility to refer. In this section we present the basic concepts of hybrid logic, based on (Blackburn, 2000): Sorts, modalities, and the @-operator.

Definition 1 (Basic hybrid multimodal language $\mathcal{H}(@)$). *Given a set of propositional symbols PROP=$\{p,q,r,...\}$, and a set of modality labels MOD=$\{\pi,\pi',\pi'',...\}$. Let NOM be a nonempty set of* nominals, *disjoint from PROP and MOD. Typically, elements of NOM are written as i, j, k. We define the basic hybrid multimodal language $\mathcal{H}(@)$ (over PROP, MOD, and NOM) to be the set of well-formed formulas such that:*

$$WFF\ \phi := i \mid p \mid \neg\varphi \mid \varphi\wedge\psi \mid \varphi\vee\psi \mid \varphi\rightarrow\psi \mid \langle\pi\rangle\varphi \mid [\pi]\varphi \mid @_i\varphi.$$

For any nominal i, we call the symbol sequence $@_i$ a satisfaction operator.

Definition 2 below gives the model theory for $\mathcal{H}(@)$. As with any sorted approach, to get the sorting strategy really working we need to ensure the semantics implements the strategy.

Definition 2 (Hybrid models, satisfaction, and validity). *A hybrid model $\mathfrak{M}$ is a triple $(\mathcal{W},\{R_\pi \mid \pi\in\mathcal{MD}\},V)$, where $(\mathcal{W},\{R_\pi \mid \pi\in\mathcal{MD}\})$ is a frame consisting of a set of states $\mathcal{W}$ and a set of relations R_π for each modal π in $\mathcal{MD}$, and V a* hybrid valuation. *A hybrid valuation is a function with domain $\mathcal{P}\cup\mathcal{N}$ and range $\wp(\mathcal{W})$ such that for all nominals i, V(i) is a singleton subset of $\mathcal{W}$. We call the unique state in V(i) the denotation of i. We interpret hybrid multimodal languages on hybrid models as follows:*

$$
\begin{array}{ll}
\mathfrak{M},w\models p & \text{iff } w\in V(p),\text{ where } p\in\mathcal{P}\\
\mathfrak{M},w\models\neg\phi & \text{iff } \mathfrak{M},w\not\models\phi\\
\mathfrak{M},w\models\phi\wedge\psi & \text{iff } \mathfrak{M},w\models\phi \text{ and } \mathfrak{M},w\models\psi\\
\mathfrak{M},w\models\phi\vee\psi & \text{iff } \mathfrak{M},w\models\phi \text{ or } \mathfrak{M},w\models\psi\\
\mathfrak{M},w\models\phi\rightarrow\psi & \text{iff } \mathfrak{M},w\not\models\phi \text{ or } \mathfrak{M},w\models\psi\\
\mathfrak{M},w\models\langle\pi\rangle\phi & \text{iff } \exists w'(wR_\pi w' \,\&\, \mathfrak{M},w'\models\phi)\\
\mathfrak{M},w\models[\pi]\phi & \text{iff } \forall w'(wR_\pi w' \Rightarrow \mathfrak{M},w'\models\phi)\\
\mathfrak{M},w\models i & \text{iff } w\in V(i),\text{ where } i\in\mathcal{N}\\
\mathfrak{M},w\models @_i\phi & \text{iff } \mathfrak{M},w'\models\phi \text{ where } w' \text{ is the denotation of } i.
\end{array}
$$

If ϕ is satisfied at all states in all hybrid models based on a frame $\mathfrak{F}$, then we say that ϕ is valid on $\mathfrak{F}$, which we can write as $\mathfrak{F}\models\phi$. If ϕ is valid on all frames, then we say that it is valid and write $\models\phi$.

The perspective on meaning we take here is *relational*: Propositions are related by named relations. Within the realm of grammar, we confine ourselves to *linguistic meaning*, i.e. meaning as far as linguistically realized. This perspective is reminiscent from approaches to dependency grammar like (Sgall et al., 1986) or thematic role-based semantics like (Dowty, 1989).

Logically, we can represent an expression's linguistic meaning as a conjunction of modalized terms, anchored by the nominal that identifies the head's proposition:

(1) $@_h(\mathbf{Proposition}\ \bigwedge\langle\delta_i\rangle(n_i\wedge\mathbf{dependent}_i))$

Dependency relations δ are modelled as modals $\langle\delta_i\rangle$, and with each dependent we associate a nominal n_i (its *discourse referent*). Technically, what (1) then states is that each nominal n_i, naming the state where a dependent expressed as a proposition $\mathbf{dependent}_i$ should be evaluated, is a δ_i successor of h, the nominal identifying the head.

Remark 1 (We obtain the intended interpretation). By inspecting how (1) gets interpreted on the model we can observe how it indeed yields the intended interpretation. Assume $h\wedge\langle\delta\rangle(n\wedge\phi)\equiv h\wedge\langle\delta\rangle\psi$. For $\mathcal{M},w\models h\wedge\langle\delta\rangle\psi$, we need $\mathcal{M},w\models h \wedge \exists w'(wR_\delta w' \,\&\, \mathcal{M},w'\models\psi)$. w is the state named by h. As $\mathcal{M},w'\models\psi\equiv\mathcal{M},w'\models n\wedge\phi$, we have that it must hold that $\mathcal{M},w'\models n$ and $\mathcal{M},w'\models\phi$. Thus, exactly at the state named by n, ϕ holds – or, put differently, n serves as the discourse referent for ϕ.

Observe how the perspective that thus arises differs from the more traditional picture emanating from type-logical approaches to natural language semantics, where the meaning of an argument is derived from its position in a predicate. For example, consider (2).

(2) $\lambda x\lambda y\lambda z.\mathbf{read}(x,y,z)$

This gives rise to an interpretation of a predicate and its arguments whereby the *order* in which the arguments appear (and are bound by the λ's) is essentially constitutive of their meaning. Thus, the first argument could be taken to hold what corresponds to the 'reader', whereas the second argument would then correspond to what is being read. This type of characterization is of course familiar from mathematical logic, and certainly gained prominence in formal semantics with the advent of Montague Grammar. But it is rather different from what we propose – here, the meaning of an argument is identified by the role it is assigned, not by its position.[1] A simple example is given in §4.

To round off the discussion, let us have a look at how we can model contextual reference. Principal to the approach we take here is that contextual reference is perceived of as the statement (and, eventually, the interpretation) of a *relation*– namely, the relation between the referent and its antecedent. A differentiation among types of anaphors can then be made by distinguishing different types of relations that model *accessibility*.

Abstractly, the specification of an anaphor's lexical meaning is given in (3) below.

(3) $(k \wedge @_k\langle\Xi\rangle(k' \bigwedge \mathbf{condition}_i))$

Here, Ξ models an accessibility relation. In words (3) thus states that there is a relation between the anaphoric expression (having a discourse referent) and a Ξ-accessible discourse referent for which particular **conditions**$_i$ hold. For example, consider a relation XS that makes entities in the discourse accessible. Then, a simplified specification of the lexical meaning of a pronoun "he" could look as in (4). From p, we try to relate to a discourse referent a for which it holds that it is male.[2]

(4) $(p \wedge @_p\langle XS\rangle(a \bigwedge \mathbf{male}))$

In (Kruijff, 2001) we present more details on how we can model different types of contextual reference in grammar, and how we can interpret them in a discourse theory.

3. Categorial-hybrid logical grammar

Traditionally, categorial type logics use a (typed) λ-calculus for specifying the meaning of a sentence. A convenient mathematical fact thereby is that there is a close correspondence between natural deduction and the λ-calculus - the *Curry-Howard correspondence.* An important result established by the Curry-Howard correspondence is that an elimination rule, eliminating an implication and thereby combining two elements, corresponds to *functional application* in the λ-calculus. Conversely, an introduction rule corresponds to *functional abstraction.* Thus, for example, when we apply an elimination rule to combine a function and an argument, we can *in parallel* apply the meaning of the argument to the meaning of the function (which is traditionally specified as a λ-term).

The issue now is, how can we establish a correspondence between natural deduction and operations in a hybrid logic, so as to compose a representation of a sentence's linguistic meaning in parallel to an analysis of the sentence's form? The answer is relatively simple, in fact.

[1] Combining proposals like (Hoffman, 1995) with (Dowty, 1989) we could of course create a lambda calculus-based representation where we are no longer dependent on the order in which arguments are filled in, and where arguments do bear roles. But this would not give us e.g. the advantages sorting gives us in hybrid logic.

[2] Or, from the viewpoint of linguistic meaning, the pronominal reference *states* that there is such an entity. Coherence then is the verification (in discourse interpretation) of that statement against the larger discourse context.

$$\frac{\alpha \vdash A : @_p\Psi \quad \beta \vdash (B\backslash_\mu A) : @_{p'}\Phi}{(\alpha \circ_\mu \beta) \vdash B : @_{p'}\Phi \wedge @_{p'}\Psi}\,E\backslash_\mu \qquad \frac{\beta \vdash (B/_\mu A) : @_{p'}\Phi \quad \alpha \vdash A : @_p\Psi}{(\beta \circ_\mu \alpha) \vdash B : @_{p'}\Phi \wedge @_{p'}\Psi}\,E/_\mu$$

$$\frac{\begin{array}{c}[\alpha \vdash A : @_h\top]\cdots \\ \vdots \\ (\alpha \circ_\mu \beta) \vdash B : @_{p'}\Phi \wedge @_{p'}\top\end{array}}{\beta \vdash (B\backslash_\mu A) : @_{p'}\Phi}\,I\backslash_\mu \qquad \frac{\begin{array}{c}\cdots[\alpha \vdash A : @_h\top] \\ \vdots \\ (\beta \circ_\mu \alpha) \vdash B : @_{p'}\Phi \wedge @_{p'}\top\end{array}}{\beta \vdash (B/_\mu A) : @_{p'}\Phi}\,I/_\mu$$

$$\frac{\begin{array}{c}[\gamma \vdash B : @_h\top]\ [\zeta \vdash C : @_{h'}\top] \\ \vdots \\ \alpha[(\gamma \bullet_\mu \zeta)] \vdash A : @_{p'}\Phi_\alpha \wedge @_h\top \wedge @_{h'}\top\end{array} \quad \beta \vdash (B \bullet_\mu C) : @_{p''}\Phi_\beta}{\alpha[\beta] \vdash A : @_{p'}\Phi_\alpha \wedge @_{p'}\Phi_\beta}\,E\bullet_\mu$$

$$\frac{\alpha \vdash A : \Phi_\alpha \quad \beta \vdash B : \Phi_\beta}{(\alpha \circ_\mu \beta) \vdash A \bullet_\mu B : @_{p'}\Phi_\alpha \wedge @_{p''}\Phi_\beta}\,I\bullet_\mu$$

FIGURE 1. Labelled natural deduction system for $\{\bullet, \backslash, /\}$

$$\frac{\alpha \vdash A : @_h\Psi}{\langle\alpha\rangle^i{}_\nu \vdash \Diamond_i A : @_h\langle\imath\rangle_\nu\Psi}\,I\Diamond \qquad \frac{\alpha \vdash \Diamond_i A : @_x\langle\imath\rangle\Psi \quad \begin{array}{c}[\gamma \vdash A : \phi] \\ \vdots \\ \beta[\langle\gamma\rangle^j{}_\nu] \vdash B : \Phi[\langle\jmath\rangle_\nu\phi]\end{array}}{\beta[\alpha] \vdash B : \Phi[\langle\jmath\rangle_\nu\Psi]}\,E\Diamond$$

$$\frac{\langle\alpha\rangle^i{}_\nu \vdash A : \Gamma[[\imath]_\nu\Psi]}{\alpha \vdash \Box^\downarrow{}_i A : \Gamma[[\imath]_\nu\Psi]}\,I\Box^\downarrow{}_i \qquad \frac{\alpha \vdash \Box^\downarrow{}_i A : @_x[\imath]_\nu\Psi}{\langle\alpha\rangle^i{}_\nu \vdash A : @_x[\imath]_\nu\Psi}\,E\Box^\downarrow{}_i$$

FIGURE 2. Labelled natural deduction system for $\{\Box^\downarrow, \Diamond\}$

First of all, recall that what we are building are *relational* structures. For a head h that means that it may be looking for an argument. That is, $@_h(\phi \wedge \langle\delta\rangle d)$ $(= @_h\phi \wedge @_h\langle\delta\rangle d)$, we have a nominal h that refers to some state where the head's proposition holds, and from where we should be able to link to some other (yet unspecified) state d along a δ (dependency) relation. Similarly, once we interpret a word group as a particular type of dependent, we specify that as saying that it is a dependent that is looking for a head. We have something like $@_h\langle\delta\rangle d$, but now δ and d are further specified and it is the h that we need to establish. In other words, to combine a head h with a dependent d, all we need to say is that d is what h is looking for, and vice versa.

Conversely, how do we model the analogon of functional abstraction? Functional abstraction corresponds to the application of an introduction rule, which discharges an assumption. For that discharge to work, the assumption must have been used earlier in the derivation. Given the above discussion, this must have lead to the introduction of a *link* (@) between the assumption's 'meaning' and an argument. Discharging the assumption then can be understood as simply severing that link: Formally, we replace the link $@_h x$ between the assumption's nominal h and the argument x by $\top$. Because $A \wedge \top \equiv A$, we thus *effectively* drop the assumption.[3]

Definition 3 (Base logic for DGL). *We define the base logic in terms of the proof calculus for pure residuation (Moortgat, 1997) to which we add operations acting on representations formulated in a hybrid logic. The part of the proof calculus dealing with the behavior of the product and its residuals* $\{\bullet, \backslash, /\}$ *is given in Figure 1. Note that we have defined the calculus using the Steedman-style notation of categories. The rules in Figure 2 define the behavior of unary modals* $\Box^\downarrow$ *and* $\Diamond$

[3] And, for that reason, we also drop the conjunct.

that are semantically relevant (Morrill, 1994). Unary modals that are semantically neutral leave the semantics untouched.

Remark 2 (Locality of unary modalities). We keep the relations between $\Box^{\downarrow}/\Diamond$ and $[\cdot]/\langle\cdot\rangle$ strictly local. We obtain this by labelling a structural modal with an index ν corresponding to the index given to the underlying modal relation. Observe that we allow for a more specific mode x to replace a less specific mode y in the representation of linguistic meaning in $E\Diamond$. In line with this possibility we drop the *ceteris paribus* condition usually assumed for structural rules: If a structural rule changes the mode of a structural modal, then the mode of the underlying modal relation changes accordingly. Finally, note that we do not have term constructors or deconstructors. They can be considered identity functions, by which we trivially obey the general residuation laws for unary modalities (Moortgat, 1997).

4. Examples

To give a simplified example of how the calculus works, consider first of all the lexical assignment for "sleeps", given below in (5).

(5) sleeps $\vdash \Box^{\downarrow}{}_{3rd}\Box^{\downarrow}{}_{sing}(s\backslash_{sc}\Diamond_{Actor}n) : @_h(\ sleep \wedge \langle\text{ACTOR}\rangle(x)\ )$

How does it get combined with its Actor? The steps are in given (6).

(6)

i. Axiom/lexicon:
$@_h(\mathcal{E} \wedge \textbf{sleep} \wedge \langle\text{ACTOR}\rangle(x)\ )$

ii. Axiom/lexicon:
$@_{h'}\langle\text{ACTOR}\rangle(\ e \wedge \textbf{Elijah}\ )$

iii. Elimination($\backslash$):
$@_h(\mathcal{E} \wedge \textbf{sleep} \wedge \langle\text{ACTOR}\rangle(x)\ ) \wedge @_{h'}\langle\text{ACTOR}\rangle(\ e \wedge \textbf{Elijah}\ ) \wedge @_h h'$

iv. Reduction:
$@_i\langle\text{ACTOR}\rangle j \wedge @_i\langle\text{ACTOR}\rangle k \rightarrow @_j k$

v. Result:
$@_h(\ \mathcal{E} \wedge \textbf{sleep}\ \langle\text{ACTOR}\rangle(\ e \wedge \textbf{Elijah}\ ))$

Because stating that $@_h h'$ means that h and h' refer to the same state, (6v) is model-theoretically equivalent to (6iii) together with the reduction in (6iv). (The reduction states that a head has only a *single* ACTOR dependent – i.e. an ACTOR is an inner participant, cf. (Sgall et al., 1986).)

In (Kruijff, 2001) we consider more elaborate examples, including illustrations of how we can use the multimodal, sortal setting to model phenomena like *information structure* and the *spatiotemporal-causal structure* ("aspectual category") of linguistic meaning.

Accordingly, when including these aspects, linguistic meaning takes the following more elaborate abstract form (with ι a specification of a node's informativity or *contextual boundness*, (Sgall et al., 1986; Kruijff, 2001)):

(7) $@_h([\iota_{\mathcal{E}}](\mathcal{E} \wedge \textbf{Proposition}) \bigwedge[\iota_i]\langle\delta_i\rangle(n_i \wedge \textbf{dependent}_i))$

Here, $\mathcal{E}$ is the *event nucleus* in the style of Moens & Steedman, cf. (Steedman, 2000). The specifications of informativity are (prototypically) underspecified lexically, and get specified in the process of grammatical analysis. Using the rules for unary modalities as given in Figure 2 establishing the interface between $\Box^{\downarrow}/\Diamond$ and there reflection in linguistic meaning, and structural rules that specify $\Box^{\downarrow}/\Diamond$ sensitive to for example word order or tune, we can use the calculus given here

to give a compositional, monotonic account of how information structure and its realization.

Finally, it should be remarked that the representations we obtain from the grammar specify *linguistic* meaning. We leave out the *interpretation* of such meaning with respect to a larger discourse context. In (Kruijff, 2001), we show how hybrid logic can be used to build a discourse representation theory following out the relational, information structure-sensitive perspective we adopt for linguistic meaning. The representations of linguistic meaning can be conceived of, in a straightforward way, as discourse representation structures.

5. Final remarks

An issue not addressed here is the relation between the model theory of the proof calculus (Moortgat, 1997), and the model theory of hybrid logic on the other hand. In categorial type logic, the model theory for the proof theory is the model theory for the semantic representation. Because of the *local* (rather than distributed/global) handling of unary modalities $\Box^{\downarrow}/\Diamond$ and their reflections $[\cdot]/\langle\cdot\rangle$, the situation is more complicated here. At the moment, soundness and completeness results for the calculus presented here remain therefore as open issues.

References

Carlos Areces, Patrick Blackburn, and Maarten Marx. 1999. A road-map on complexity for hybrid logics. In J. Flum and M. Rodríguez-Artalejo, editors, *Computer Science Logic*, number 1683 in Lecture Notes in Computer Science, pages 307–321. Springer-Verlag.

Patrick Blackburn. 2000. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Journal of the Interest Group in Pure Logic*, 8(3):339–365.

David R. Dowty. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Number 7 in Studies in Linguistics and Philosophy. Kluwer Academic Publishers, Dordrecht, Boston, London.

David R. Dowty. 1989. On the semantic content of the notion "thematic role". In Giennaro Chierchia, Barbara H. Partee, and Raymond Turner, editors, *Properties, Types, and Meaning: Volume II, Semantic Issues*, volume 39 of *Studies in Linguistics and Philosophy*, pages 69–129. Kluwer Academic Publishers, Dordrecht, Boston, London.

Beryl Hoffman. 1995. *Computational Analysis of the Syntax and Interpretation of "Free" Word-Order in Turkish*. Ph.d. thesis, ircs report 95-17, University of Pennsylvania.

Geert-Jan M. Kruijff. 2001. *A Categorial-Modal Logical Architecture of Informativity: Dependency Grammar Logic & Information Structure*. Ph.D. thesis, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic, April.

Michael Moortgat. 1997. Categorial type logics. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*. Elsevier Science B.V., Amsterdam New York etc.

Glyn V. Morrill. 1994. *Type Logical Grammar: Categorial Logic of Signs*. Kluwer Academic Publishers, Dordrecht, Boston, London.

Petr Sgall, Eva Hajičová, and Jarmila Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. D. Reidel Publishing Company, Dordrecht, Boston, London.

Mark Steedman. 2000. The productions of time. Draft 4.0, July 24.

Johan F.A.K. Van Benthem. 1996. Content versus wrapping: An essay in semantic complexity. In Maarten Marx, László Pólos, and Michael Masuch, editors, *Arrow Logic and Multi-Modal Logic*, Studies in Logic, Language and Information, pages 203–220. CSLI Publications, Stanford CA.

Modality in Comparative Constructions

Cécile Meier
Johann Wolfgang Goethe-Universität, Frankfurt/Main

1 Introduction

A maximality or minimality operator is an essential component of the meaning of comparative constructions in many recent analyses (von Stechow 1984, Rullmann 1995, Kennedy 1999, Heim 2001). The existence of such an operator, however, is primarily motivated in order to account correctly for the truth conditions of comparative constructions containing modals. In this paper, we investigate more closely the contribution of modal operators to the semantics of comparatives and we show that there is no need for a maximality or minimality operator. Following Kratzer's (1981, 1991) analysis of modal elements, we assume that the meaning of a modal sentence is dependent on a conversational background *and* an ordering source. For comparative environments, we demonstrate that the ordering source reduces a set of possible degrees to a single degree that is most (or least) wanted or expected, i.e., maximality and minimality readings of comparative constructions are an effect of the pragmatic meaning of the modal.

Comparative constructions with a modal in the subordinate clause may have a minimality or a maximality reading independent on the quantificational force of the modal. Consider the sentence in (1).

(1) The coach nominated more foreigners than he was allowed to nominate.

In order to interpret this sentence we may have a situation in mind where the rules of some national league determine the maximal number of foreigners that may enter a game. The sentence is true if the number of foreigners nominated exceeds this maximal number.

Imagine on the other hand the following scenario. Chuck is driving a truck full of eggs on a New Jersey highway. It is true that the slower Chuck drives the less eggs break. Moreover, Chuck wants to bring as many eggs as possible to their final destination and he does not want to risk a speeding ticket. The road traffic regulations limit the minimum speed on New Jersey highways to 45 mph and the maximum speed to 65 mph. All in all, it seems desirable that Chuck is driving as slow as possible within the regulations. In this scenario, the sentence in (2) might be true if Chuck is driving faster than 45 mph but slower than 65 mph.

(2) Chuck is driving faster than he is allowed to drive.

Modals denoting necessity show a similar behaviour. They can trigger a more-than-minimum reading or a more-than-maximum reading. The more-than-minimum reading is exemplified in (3).

(3) Mary was driving faster than she had to.

Assume that Mary has to get from A to B in at most an hour. In order to get to B in time she has to drive at least 140 km/h. If Mary, in fact, was driving 160 km/h she got to B early. In this case she exceeds the minimal speed.

A more-than-maximum reading is predominant if we assume a situation where John, a Porsche driver, is driving on a New Jersey highway. The highway speed regulations set the maximum speed to 65 mph. If John is, in fact, driving 95 mph the sentence in (4) is true.

(4) Jim is driving faster than he should be driving.

In general, it seems to be the case that constructions with *can*, *could* and *be allowed to* (possibility operators), but also constructions with *should*, *must* and *have to* (necessity operators) allow ambiguities with respect to the characteristics of the standard of comparison. How can we explain these facts?

In what follows, we discuss two proposals from the literature and show that they run into problems with respect to two critical examples: (a) comparatives with a minimality reading and a modal denoting possibility in the *than*-clause and (b) comparatives with a maximality reading and a modal denoting necessity in the *than*-clause.

2 Previous Analyses

Let us assume in the spirit of von Stechow (1984:35) that comparative constructions with *more* and *-er* are comparisons between two maximal degrees.[1] Furthermore, we assume that the modal element *be allowed* is represented by the embedded possibility operator POSS. This operator gets the usual semantics of an existential quantifier over worlds.[2] Under these assumptions, the example in (1) gets the representation in (5-a) and (2) the representation in (5-b).

[1] In particular, we are abstracting away from considerations of the internal structure and composition of these representations.

[2] POSS is true of a proposition p in a world w if there is an accessible world w^* such that p is true in w^*.

(5) a. MAX{d:the coach nominated d-many foreigners in w_0} >
MAX{d: POSS(λw.the coach nominated d-many frgnrs in w)}
b. MAX{d: Chuck is d-fast in w_0} >
MAX{d: POSS(λw.Chuck is d-fast in w)}

In (5-a), we assume that the accessible worlds are worlds where the coach observes the regulations of the league. In this sense, the representation seems to capture the meaning correctly. If the maximal number of foreigners permitted is seven, the sentence is true if the coach nominated more than seven and false otherwise. (5-b), however, seems not to be correct. In the accessible worlds no violations of the law in the actual world occur. Therefore everybody is driving with a speed between 45 and 65 mph. The maximal permitted speed is 65 mph. We predict (5-b) to be true if Chuck is driving faster than 65 mph. This reading is possible but not intended in the above mentioned scenario of Chuck the egg truck driver. Consequently, we predict (5-b) to be false if Chuck is driving 56 mph, contrary to the intuition: see the scenario above.

Modals denoting necessity reveal another problem: If we translate *had to* and *should* by a necessity operator[3], it turns out that the formulas in (6) are only defined if Mary and Jim, respectively, have only one speed in all accessible worlds. But in the given scenarios the admissible speed is 140 km/h or more in the case of (6-a) and the permitted speed varies between 45 mph and 65 mph in the case of (6-b).

(6) a. MAX{d:M. is d-fast in w_0} >
MAX{d: NEC(λw.M. is d-fast in w)}
b. MAX{d:Jim is d-fast in w_0} >
MAX{d: NEC(λw.Jim is d-fast in w)}

The deeper reason for this outcome is the fact that von Stechow's analysis presupposes a so-called "exactly" interpretation of the degree variable d.

Heim (2001:216) proposes an analysis for comparatives that has the effect of an "at least" interpretation for the degree variable. She assumes that gradable adjectives are subject to a monotonicity condition.

(7) *Monotonicity*:
A function f of type $< d, < e, t >>$ is monotone iff
$\forall x \forall d \forall d'[f(d)(x) = 1 \ \& \ d' < d \Rightarrow f(d')(x) = 1]$

Following Heim, we predict for both (6-a) and (6-b) a minimality reading. Why is this so? The maximal speed such that Mary (or Jim) has it in all

[3]NEC is true of a proposition p in a world w iff p is true in all accessible worlds w^*.

accessible worlds is the minimally permitted speed. In the case of (6-b), however, we attested a more-than-maximum reading.

In sum, von Stechow as well as Heim have difficulties to predict the available readings in some cases with embedded modals considered so far. Von Stechow's and Heim's analysis cannot predict the more-than-minimum reading in constructions with an embedded possibility operator. Constructions with an embedded necessity operator are undefined in von Stechow's account if the relevant degrees vary in the accessible worlds. Heim's analysis does better since it explains the more-than-minimum readings of constructions with a necessity operator. How could we fix these problems?

3 Doubly relative modality in comparatives

In the framework developed in Kratzer (1981), the interpretation of modal elements does not only depend on (a) the kind of modal relation (i.e. possibility and necessity) and (b) information that characterizes the accessibility relation but also (c) on an ordering source that induces an ordering on the accessible worlds: see also Kratzer (1991) for arguments for doubly relative modality. [4]

Consider the interpretation of the sentences that were problematic for von Stechow and Heim under the new perspective where we interpret the modals doubly relative.

Assume that modals like *be allowed to* in our example (2) are associated with two contextual parameters as follows: (a) the law, i.e., the road traffic regulations for New Jersey, in addition to some kind of complicated causal law that describes the proportional dependency of speed and damage to the truck load, (8-a), and (b) an ideal, i.e., what Chuck wants, in particular, that all 20000 eggs arrive undamaged at the point of final destination, (8-b).

[4]We adopt the definitions in (i) and (ii).

(i) POSS is true in a world w of a proposition p with respect to a modal base f and an ordering source o iff there is a world w' such that p is true in w' and (a) and (b) are satisfied:
(a) $w' \in \bigcap f(w)$ (b) $\neg\exists w^*[w^* \in \bigcap f(w) \ \& \ w^* <_{o(w)} w']$

(ii) NEC is true in a world w of a proposition p with respect to a modal base f and an ordering source g iff p is true in all worlds w' such that (a) and (b) are satisfied:
(a) $w' \in \bigcap f(w)$ (b) $\neg\exists w^*[w^* \in \bigcap f(w) \ \& \ w^* <_{o(w)} w']$

The versions differ from Kratzer's original version of human necessity and human possiblity in one respect: we assume simplifying that there *is* a world that comes maximally close to the world w.

(8) a. $\forall w : f(w) = \{p \mid \exists d[45 \text{ mph} \leq d \leq 65 \text{ mph} \,\&\, p = \lambda w.$Chuck is d-fast in $w\} \cap$ {the slower Chuck drives the more eggs arrive at their destination}

b. $\forall w : o(w) = \{p \mid \exists n[0 \leq n \leq 20000 \,\&\, p = \lambda w.n$-many eggs arrive undamaged in $w]\}$

The ordering source induces a ranking on the worlds that conform to the modal base.[5] The more eggs arrive undamaged the more propositions from the ordering source become true. The most eggs arrive undamaged in worlds where Chuck is driving slowest. Therefore, the set of degrees d such that there is a world that conforms to the wishes of Chuck as much as possible given the traffic regulations and Chuck is driving d-fast in that world is a singleton. And, we arrive at the following equivalencies.

(9) $\text{MAX}\{d : \exists w[w \in \bigcap f(w_0) \,\&\, \neg\exists w^*[w^* \in \bigcap f(w_0) \,\&\, w^* \geq_{o(w_0)} w] \,\&$ Chuck is d-fast in w]}
= 45 mph
= $\text{MIN}\{d : \exists w[w \in \bigcap f(w_0) \,\&\, \neg\exists w^*[w^* \in \bigcap f(w_0) \,\&\, w^* \geq_{o(w_0)} w] \,\&$ Chuck is d-fast in w]}

By similar reasoning, we may solve the problems in analyzing the more-than-maximum reading in constructions with a necessity modal. In the example with the Porsche driver Jim (4), we assume that *should* is associated with a modal base that conforms to the traffic regulations in New Jersey and relates Jims speed in a car to (abstract and subjective) degrees of his happiness. The faster he is driving the happier he is. In all those worlds where he is the happiest given the regulations, Jim is driving with 65 mph. And Jim's actual speed is compared to this value.

In sum, whenever an ordering source is involed in the interpretation of modals in comparatives this ordering source reduces a set of possible (i.e., accessible) degrees to a single degree that conforms to the ordering source as much as possible. The ambiguity of constructions with *less* might be explained in the same way: see Rullmann (1995) for an explanation of less-than-minimum readings in terms of a lexical ambiguity of *less*.

If we always interpreted the modals in comparative constructions doubly relative, there would be no need for a maximality operator or a minimality operator, in the first place. Therefore, it seems that Russell's classical analysis in terms of definite descriptions is good enough for the analysis of comparative constructions.

[5]Definition of ordering sources: For all $w, z \in W$, for any $A \subseteq \wp(W)$: $w \leq_A z$ iff $\{p : p \in A \text{ and } z \in p\} \subseteq \{p : p \in A \text{ and } w \in p\}$.

4 Concluding remarks

An analysis of comparatives in terms of definite descriptions is, however, difficult to defend. Ordering sources are contextual parameters and it is not clear, so far, how the accomodation of a suitable ordering source is restricted.

Furthermore, the acount presupposes that we analyze elements like *any*, *ever* and *or* as free choice items if they occur in the *than*-clause. An analysis in terms of negative polarity items or boolean *or* is no longer defined.

(10) a. 2 is greater than any rational lower approximation of $\sqrt{2}$. [Pinkal 1989]

b. In New Brunswick it was hotter than it ever was in L.A.

c. John is taller than Tim or Toby.

And it requires an explanation for the fact that minimality readings are indeed preferred over maximality readings with necessity operators and maximality readings are preferred over minimality readings in construction with possibility operators.

References

Heim, Irene (2001): Degree Operators and Scope. *In:* C. Féry & W. Sternefeld, eds, *Audiatur Vox Sapientiae: A Festschrift for Arnim von Stechow*. Berlin: Akademie Verlag.

Kennedy, Christopher (1999): *Projecting the Adjective. The Syntax and Semantics of Gradability and Comparison*. Outstanding Dissertations in Linguistics, New York: Garland Publishing, Inc.

Kratzer, Angelika (1981): The Notional Category of Modality. In: *Eikmeyer, H.J. ans H. Rieser (eds.) Words, Worlds and Contexts: New Approaches in Word Semantics*. Berlin: De Gruyter, pp. 38–74.

Kratzer, Angelika (1991): Modality. *In:* A. von Stechow & D. Wunderlich, eds, *Semantik: Ein internationales Handbuch zeitgenŽssischer Forschung*. Berlin: de Gruyter, pp. 639–650.

Pinkal, Manfred (1989): 'Die Semantik von Satzkomparativen', *Zeitschrift f§r Sprachwissenschaft* **8**(2), 206–256.

Rullmann, Hotze (1995): Maximality in the Semantics of WH-Constructions. PhD thesis, University of Massachussets, Amherst.

von Stechow, Arnim (1984): 'Comparing Semantic Theories of Comparison', *Journal of Semantics* **3**, 1–77.

λ-Grammars and the Syntax-Semantics Interface

Reinhard Muskens
Tilburg University

In this paper we discuss a new perspective on the syntax-semantics interface. Semantics, in this new set-up, is not 'read off' from Logical Forms as in mainstream approaches to generative grammar. Nor is it assigned to syntactic proofs using a Curry-Howard correspondence as in versions of the Lambek Calculus, or read off from f-structures using Linear Logic as in Lexical-Functional Grammar (LFG, Kaplan & Bresnan [9]). All such approaches are based on the idea that syntactic objects (trees, proofs, f-structures) are somehow prior and that semantics must be parasitic on those syntactic objects. We challenge this idea and develop a grammar in which syntax and semantics are treated in a strictly parallel fashion. The grammar will have many ideas in common with the (converging) frameworks of categorial grammar and LFG, but its treatment of the syntax-semantics interface is radically different. Also, although the meaning component of the grammar is a version of Montague semantics and although there are obvious affinities between Montague's conception of grammar and the work presented here, the grammar is not compositional, in the sense that composition of meaning need not follow surface structure.

λ-Grammars

We follow the tradition of Curry [6], Cresswell [4], and Oehrle [13, 14, 15] in representing syntactic information with the help of typed λ-terms.[1] For example, Oehrle [13, 14] considers multidimensional signs such as those in (1), consisting of a λ-term over strings, a semantic λ-term, and a type. (1a) can be combined with (1b) by applying (1a)'s first term to (1b)'s first term, applying (1a)'s second term to (1b)'s second term, and applying Modus Ponens to the types. The result is (1c).

(1) a. λxλy. y likes x : $\lambda x \lambda y.like(y, x) : np \to (np \to s)$

b. John : $j : np$

c. λy. y likes John : $\lambda y.like(y, j) : np \to s$

Oehrle's work will be our point of departure, but we deviate from it in two respects. First, while [13, 14] combine signs such as the ones in (1) with the help of the undirected Lambek Calculus, our signs will be combined using linear combinators. Mathematically this boils down to the same thing, as proofs in the undirected Lambek Calculus are in 1-1 correspondence with the latter (Van Benthem [1, 2, 3]). But a move from proofs to combinators spares the working linguist much technical overhead and it will serve our purpose to stress the point that semantics need not be dependent on any form of syntax. A second divergence from Oehrle's work is that we move from λ-terms over *structures* (strings in [13, 14], but also trees and f-structures in [15]) to λ-terms over *descriptions* of structures. This is in accordance with a general tendency in linguistics (starting with [9]) to replace structures with descriptions of structures as the prime vehicles of representation.

Consider the three 3-dimensional signs in (2).

(2) a. john : $\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3) : john$

[1] Curry considers expressions containing subscripted blanks, such as '—$_1$ is between —$_2$ and —$_3$' or '—$_1$ were eaten by the children'. Functors can apply to arguments and arguments are to be substituted for blanks in the order of the subscripts. Essentially then, although Curry does not explicitly mention this, these terms are lambda terms over syntactic objects.

syntax:	$k, f : \nu$	$t, F : \nu t$	$T, \mathcal{F} : (\nu t)(\nu t)$	
semantics:	$x, y : e$	$i, j : s$	$p : st$	$P : e(st)$

Table 1: Typographical conventions for variables used in this paper. *Var* : *Type* means that *Var* (with or without subscripts or superscripts) always has type *Type*.

b. mary : $\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3) : mary$

c. $\lambda t_1 \lambda t_2.[t_2$ [loves t_1]] :
$\lambda F_1 \lambda F_2 \lambda f \exists f_1 f_2[F_1(f_1) \wedge F_2(f_2) \wedge arc(f, cat, V) \wedge arc(f, tense, pres) \wedge$
$arc(f_1, cat, N) \wedge arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge$
$arc(f_2, pers, 3) \wedge arc(f, subj, f_2)]$:
$\lambda x \lambda y \lambda i.love(y, x, i)$

These signs each consist of a c-structure component, an f-structure component, and a semantic component. Expressions in sans serif in the c-structure terms are of type νt, and denote sets of nodes. For example, john can be thought of as the set of nodes that are labeled 'John', whereas an expression such as [loves mary] can be thought of as the set of nodes k directly dominating a node k_1 labeled 'loves' and a node k_2 labeled 'Mary', with k_1 preceding k_2.

The f-components of our signs consist of λ-terms over the first order feature language of Johnson [8] and the semantics in the third component is in accordance with a streamlined form of Montague's [10] theory. Constants *john* and *mary* are of type e and *love* is of type $e(e(st))$. Constants *cat*, *num*, *pers*, etc. are of a type a (attributes), while N, sg, 3, ... are of type ν (nodes). More typing information is given in Table 1. We consider a grammar with three dimensions here, but in general the number of dimensions of a grammar is arbitrary (though fixed). The terms that we are interested in are all closed and we require that lexical elements have closed terms in each dimension.

Signs can be combined by means of *pointwise application*. In general, if $M = \langle M_1, \ldots, M_n \rangle$ and $N = \langle N_1, \ldots, N_n \rangle$ are sequences of λ-terms such that $M_i(N_i)$ is well-typed for each i, the pointwise application of M to N is just

$$\langle M_1(N_1), \ldots, M_n(N_n) \rangle .$$

Generalizing the notation for application, we denote this as $M(N)$. It is easily seen that the result of pointwise application of (2c) to (2a) equals (3a) modulo standard equivalences and that (3a)((2b)) reduces to (3b).

(3) a. $\lambda t_2.[t_2$ [loves john]] :
$\lambda F_2 \lambda f \exists f_1 f_2[F_2(f_2) \wedge arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N)$
$\wedge\ arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge$
$arc(f, subj, f_2)]$:
$\lambda y \lambda i.love(y, john, i)$

b. [mary [loves john]] :
$\lambda f \exists f_1 f_2[arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N) \wedge$
$arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge$
$arc(f, subj, f_2)]$:
$\lambda i.love(mary, john, i)$

The three descriptions in sentential signs such as (3b) each denote a set in every possible model of the language; the first two denote sets of nodes (type νt), the third a set of possible worlds (a proposition, type st). The idea is that if the second set is non-empty in some model of the first four axioms in [8], then any node satisfying

abstract type	syntactic dimensions	semantic dimension
S	νt	st
NP	νt	e
N	νt	$e(st)$

Table 2: Concretizations of abstract types used in this paper.

the first description should be connected to the truth conditions expressed in the third element. The requirement that the second component should be satisfiable provides for a subcategorization mechanism. E.g., combining (3a) with a plural subject would have led to an f-description that can only denote the empty set.

In (4) and (5) some more lexical signs are given with two results of their possible combinations in (6).

(4) a. man : $\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3)$: man

b. woman : $\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3)$: $woman$

(5) a. $\lambda t \lambda T.T$([a t]) :
$\lambda F \lambda \mathcal{F}.\mathcal{F}(\lambda f.F(f) \wedge arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P' P \lambda i \exists x[P'(x)(i) \wedge P(x)(i)]$

b. $\lambda t \lambda T.T$([every t]) :
$\lambda F \lambda \mathcal{F}.\mathcal{F}(\lambda f.F(f) \wedge arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P' P \lambda i \forall x[P'(x)(i) \rightarrow P(x)(i)]$

(6) a. $\lambda T.T$([every man]) :
$\lambda \mathcal{F}.\mathcal{F}(\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P \lambda i \forall x[man(x, i) \rightarrow P(x)(i)]$

b. $\lambda T.T$([a woman]) :
$\lambda \mathcal{F}.\mathcal{F}(\lambda f.arc(f, cat, N) \wedge arc(f, num, sg) \wedge arc(f, pers, 3))$:
$\lambda P \lambda i \exists x[woman(x, i) \wedge P(x)(i)]$

The terms that our signs consist of are typed, but it is expedient to type the signs themselves as well. Types for signs will be called ***abstract types***. Abstract types in this paper are built up from ground types S, NP and N with the help of implication, and thus have forms such as NP S, N((NP S)S), etc. A restriction on signs is that a sign of abstract type A should have a term of type A^i in its i-th dimension. The values of the function $\cdot^i$ for ground types can be chosen on a per grammar basis and in this paper are as in Table 2. For complex types, the rule is that $(AB)^i = A^i B^i$. This means, for example, that $\text{NP(NP S)}^1 = \text{NP(NP S)}^2 = (\nu t)((\nu t)\nu t)$ and that $\text{NP(NP S)}^3 = e(e(st))$. As a consequence, (2c) should be of type NP(NP S). Similarly, (2a) and (2b) can be taken to be of type NP, (3a) and (3b) are of types NP S and S respectively, etc. In general, if M has abstract type AB and N abstract type A, then the pointwise application $M(N)$ is defined and has type B.

Abstraction can also be lifted to the level of signs. Supposing that the variables in our logic have some fixed ordering and that the number of dimensions of the grammar under consideration is n, we define the k-th n-dimensional variable ξ of abstract type A as the sequence of variables $\langle \xi_1, \ldots, \xi_n \rangle$, where each ξ_i is the k-th variable of type A^i. The ***pointwise abstraction*** $\lambda \xi M$ is then defined as $\langle \lambda \xi_1 M_1, \ldots, \lambda \xi_n M_n \rangle$. A definition of ***pointwise substitution*** is left to the reader.

With the definitions of pointwise application, pointwise abstraction, and n-dimensional variable in place, we can consider complex terms built up with these

constructions. (7a), for example, is the pointwise application of (6b) to the pointwise composition of (6a) and (2c). Here ζ is of type NP. (7a) can be expanded to (7b), where each dimension of a lexical sign is denoted with the help of an appropriate subscript (e.g. $(6b)_1$ is $\lambda T.T$([a woman])). The terms here can be reduced and the result is as in (7c), a sign coupling the c-description in its first dimension to one of its possible readings. The other reading is obtained from (7d), which reduces to (7e).

(7) a. $(6b)(\lambda \zeta.(6a)((2c)(\zeta)))$

b. $(6b)_1(\lambda \zeta_1.(6a)_1((2c)_1(\zeta_1)))$:
$(6b)_2(\lambda \zeta_2.(6a)_2((2c)_2(\zeta_2)))$:
$(6b)_3(\lambda \zeta_3.(6a)_3((2c)_3(\zeta_3)))$

c. [[every man] [loves [a woman]]] :
$\lambda f \exists f_1 f_2[arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N) \wedge$
$arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge$
$arc(f, subj, f_2)]$:
$\lambda i \exists y[woman(y, i) \wedge \forall x[man(x, i) \rightarrow love(x, y, i)]]$

d. $(6a)(\lambda \zeta_2.(6b)(\lambda \zeta_1.(2c)(\zeta_1)(\zeta_2)))$

e. [[every man] [loves [a woman]]] :
$\lambda f \exists f_1 f_2[arc(f, cat, V) \wedge arc(f, tense, pres) \wedge arc(f_1, cat, N) \wedge$
$arc(f, obj, f_1) \wedge arc(f_2, cat, N) \wedge arc(f_2, num, sg) \wedge arc(f_2, pers, 3) \wedge$
$arc(f, subj, f_2)]$:
$\lambda i \forall x[man(x, i) \rightarrow \exists y[woman(y, i) \wedge love(x, y, i)]]$

Let us call terms such as (7a) and (7d), which are built up from lexical signs with the help of n-dimensional variables, pointwise application and abstraction, *n-terms*. It is worth to note that n-terms are subject to the laws of α, β, and η-conversion, i.e. reasoning with them is as usual. But clearly, not every n-term makes for an acceptable coupling between syntax and semantics. We restrict ourselves to ***linear combinations*** of lexical elements. These are n-terms that are closed and conform to the condition that every abstractor $\lambda \zeta$, with ζ an n-dimensional variable, binds exactly one free ζ. n-terms conforming to this condition are called ***generated signs***.[2] Conditions such as the requirement that the third component of a generated sign must be satisfiable are ***admissibility*** conditions and a generated sign obeying them is called ***admissable***.

Multidimensional grammars that are set up in the way sketched here, with λ-terms in each dimension of the grammar and linear combination as a generative device, will be called *λ-grammars*.

Since any n-term M obeys the usual laws of λ-conversion, it can be written in the form $C(L_1) \cdots (L_m)$, where $L_1, \ldots, L_m$ are lexical signs and C is an n-term that does not contain any lexical material. If M is closed, C is a multi-dimensional (and typed) variant of a ***combinator*** in the sense of [5]. In case M is a generated sign, C will correspond to a ***linear*** (or **BCI**) combinator. For example, (7a) can be rewritten as (8), with $\lambda Q_1 \lambda R \lambda Q_2.Q_1(\lambda \zeta.Q_2(R(\zeta)))$ playing the role of the linear combinator combining (6b), (2c), and (6a).

(8) $\lambda Q_1 \lambda R \lambda Q_2.Q_1(\lambda \zeta.Q_2(R(\zeta)))((6b))((2c))((6a))$

From the fact that linear combinators play an important underlying role we see that λ-grammars have obvious affinities not only with LFG and Lambek Categorial

[2]Note that any linear combination of generated signs is itself a generated sign.

Grammar, but also with Combinatory Categorial Grammar (see e.g. [17, 18]). But λ-grammars should be distinguished from standard categorial grammars in that they are non-directional and do not use derivations.

Conclusion

It is a wideheld belief among linguists that semantics must in some sense be dependent upon syntax. Syntax first provides a scaffolding and semantics then follows the syntactic set-up, computing the meaning of a complex expression from the meanings of its syntactic parts. Such a compositional scheme works fairly well for English and related languages, although even in Montague's pivotal work it had to be assumed that a sentence can have as its parts (a) a noun phrase deep within that sentence and (b) the sentence lacking that noun phrase.

In Dalrymple et al. [7] it was observed that a problem arises when non-configurational languages such as Warlpiri are considered. In these languages combinations of words may form a semantic unit altough they are no syntactic unit. An example is (9) (see Simpson [16]), where the adjective *wita-jarra-rlu* is not adjacent to the noun (*kurdu-jarra-rlu*) that it modifies. The constituent structure of this sentence (and indeed of many Warlpiri sentences) is flat, with one S node dominating all preterminals.[3]

(9) Kurdu-jarra-rlu ka-pala maliki wajili-pi-nyi wita-jarra-rlu
child-dual-ergative pres-3ds dog chase-nonpast small-dual-ergative
'Two small children are chasing the dog', or
'Two children are chasing the dog and they are small'

Since *wita-jarra-rlu* and *kurdu-jarra-rlu* do not form a syntactic constituent but should still be considered a semantic whole, the idea of erecting a syntactic scaffold first and then using it in semantics breaks down. Constituent structure simply does not provide enough structure for interpretation. The case is typical for a wide class of languages.

In [7] this observation motivates a development in which semantics is read off from functional structure with the help of the $\{\multimap, \otimes\}$ fragment of intuitionistic Linear Logic (= the undirected Lambek Calculus). In this paper we have taken the more radical course of using linear combinators directly for combining syntactic / semantic signs. This simplifies the grammatical set-up, as in our approach there is no need for a phrase structure component as a separate generative engine. Linear combinations suffice. The toy grammar presented here shows that it is possible to do syntax and semantics really in tandem. Interpretation does not need any previous syntactic scaffolding, whether it be a constituent structure, a functional structure, a Lambek / Linear Logic proof, or any other syntactic structure. There is no space here to give a detailed analysis of (9), but an essential element should be that when the sign of an adjective is combined with the sign of a noun, the result in the c-dimension need not be a description of an adjective-noun constituent. A description requiring sisterhood suffices.

The grammar in this paper also shows that it is possible to import many ideas of Lexical-Functional Grammar into an essentially categorial framework. In this we build upon Oehrle [15]. That our move from structures to descriptions allows the incorporation of more ideas from LFG (constraining equations, path constraints for long distance dependencies) is shown in Muskens [12]. A development of the theory that is more geared towards categorial grammar and the multimodal enterprise can be found in Muskens [11].

[3] Any permutation of the words in (9) that leaves the 'auxiliary' element *ka-pala* in second position is also an acceptable Warlpiri sentence with the same two meanings.

References

[1] J.F.A.K. van Benthem. *Essays in Logical Semantics*. Reidel, Dordrecht, 1986.

[2] J.F.A.K. van Benthem. The Semantics of Variety in Categorial Grammar. In W. Buszkowski, W. Marciszewski, and J.F.A.K. van Benthem, editors, *Categorial Grammar*, pages 37–55. John Benjamins, Amsterdam, 1988.

[3] J.F.A.K. van Benthem. *Language in Action*. North-Holland, Amsterdam, 1991.

[4] M.J. Cresswell. *Logics and Languages*. Methuen, London, 1973.

[5] H. Curry and R. Feys. *Combinatory Logic*, volume I. North-Holland, Amsterdam, 1958.

[6] H.B. Curry. Some Logical Aspects of Grammatical Structure. In *Structure of Language and its Mathematical Aspects: Proceedings of the Twelfth Symposium in Applied Mathematics*, pages 56–68. AMS, 1961.

[7] M. Dalrymple, J. Lamping, and V. Saraswat. LFG Semantics via Constraints. In *Proceedings of the Sixth Meeting of the European ACL*. European Chapter of the Association for Computational Linguistics, 1993.

[8] M. Johnson. Logic and Feature Structures. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, Sydney, Australia, 1991.

[9] R. Kaplan and J. Bresnan. Lexical-Functional Grammar: a Formal System for Grammatical Representation. In J. Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 173–281. The MIT Press, Cambridge, MA, 1982.

[10] R. Montague. The Proper Treatment of Quantification in Ordinary English. In *Formal Philosophy*, pages 247–270. Yale University Press, New Haven, 1973.

[11] R.A. Muskens. Language, Lambda's, and Logic. Submitted for publication, 2001.

[12] Reinhard Muskens. Categorial Grammar and Lexical-Functional Grammar. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG01 Conference, University of Hong Kong*, Stanford CA, 2001. CSLI Publications. http://csli-publications.stanford.edu/hand/miscpubsonline.html.

[13] R.T. Oehrle. Term-Labeled Categorial Type Systems. *Linguistics and Philosophy*, 17:633–678, 1994.

[14] R.T. Oehrle. Some 3-Dimensional Systems of Labelled Deduction. *Bulletin of the IGPL*, 3:429–448, 1995.

[15] R.T. Oehrle. LFG as Labeled Deduction. In M. Dalrymple, editor, *Semantics and Syntax in Lexical Functional Grammar*, chapter 9, pages 319–357. MIT Press, Cambridge, MA, 1999.

[16] J. Simpson. *Warlpiri Morpho-Syntax: A Lexicalist Approach*, volume 23 of *Studies in Natural Language and Linguistic Theory*. Kluwer, Dordrecht, 1991.

[17] M. Steedman. *Surface Structure and Interpretation*. MIT Press, 1996.

[18] M. Steedman. *The Syntactic Process*. MIT Press, 2000.

Adverb order in Type-Logical Grammar

Øystein Nilsen, Utrecht University

Abstract

Empirical evidence is presented which 1) cannot be dealt with syntactically in terms of a sequence of functional projections (Cinque 1999, Alexiadou 1997) and 2) suggests that (some) adverb ordering phenomena should be treated on a par with polarity phenomena and quantifier scope rather than analyzed by enrichment of the semantic ontology (cf. a.o. Ernst 2001, Bartsch 1976). The proposed account relies on the derivability space generated by the residuated and Galois connected unary connectives in Type Logical Grammar, i.e. $\Diamond, \Box^{\downarrow}, {}^{0}\cdot, \cdot^{0}$ which has the crucial property of yielding a non-linear ordering of sentential categories. The proposed account treats adverb ordering in essentially the same way as quantifier scope (Bernardi and Moot 2000) and polarity phenomena (Bernardi 2001) have been treated in Type Logical Grammar and thus aims at a unification of the theoretical treatment of these phenomena.

The linearity assumption

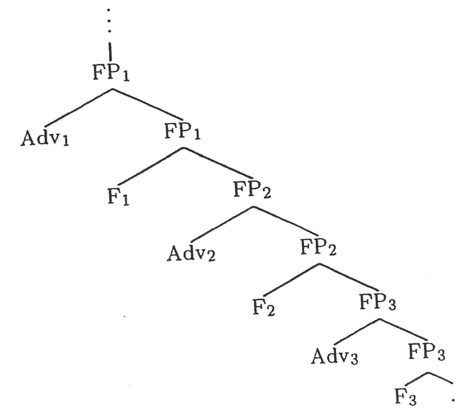
In their classical paper, Thomason and Stalnaker (1973) partitioned adverbial expressions into two classes, i.e. sentence modifiers (type $\langle t, t\rangle$) and predicate modifiers (type $\langle\langle e, t\rangle, \langle e, t\rangle\rangle$). Syntactically, this corresponds to the distinction between VP-adverbials and S-adverbials. However, this classification is uncontroversially much too coarse. Thus, Jackendoff (1972) distinguishes six distributinoal classes of adverbials, and Cinque (1999) as many as thirty two. Cinque observes that there are pairs of adverbs which would both be classified as S-adverbs, which cannot surface in certain orders.[1]

(1) a. Stanley **possibly never** ate his wheaties.
 b. *Stanley **never possibly** ate his wheaties.

By applying this test, one obtains a relation on classes of S-adverbs, call it RA, characterized by the predicate Adv_1 *can precede* Adv_2. At a purely descriptive level one may now ask what properties this relation has, i.e. whether it is transitive, antisymmetric, etc. As it turns out, existing analyses of adverb distribution seem to assume/presuppose that RA is antisymmetric, transitive and connected, i.e. that it is a *linear* ordering. For instance, Cinque (ibid.), Alexiadou (1997) and Nilsen (2000) analyze adverb ordering by assigning (classes of) adverbs to specifiers of minimalist functional projections which themselves are ordered by (syntactic) selection. This automatically renders the ordering of adverbs linear, since c-command restricted to specifiers (and heads) in an "extended projection line" (Grimshaw 1990) is linear. To see this, consider the tree in (2), where F_i are functional heads ordered by selection and Adv_i must occur in a spec-head relation with F_i. As the reader can easily verify, c-command restricted to adverbs in (2) gives rise to the linear sequence Adv_1, Adv_2, Adv_3.

[1] For the tuples of adverbs to be considered in the present paper, precedence maps directly to semantic scope, so "x precedes y" could just as well be rendered as "x outscopes y". More in particular, this holds of tuples of adverbs in the *mittelfeld* of the clause. When the adverbs occur finally, the mapping is sometimes the inverse and sometimes the same. See Cinque 1999, Nilsen 1997, 2000 for discussion of final adverbials.

(2)



Ernst (2001) gives a semantic account of adverb ordering by analyzing different classes of adverbs as modifiers of different kinds of semantic objects. By way of example, according to Ernst, an adverb like *completely* modifies EVENTS (i), whereas e.g. *not* modifies PROPOSITIONS (p), and *paradoxically* modifies FACTS (f). He then imposes the following system of type conversion on his ontological primitives:

$$i \Longrightarrow p \Longrightarrow f$$

This derives the fact that these three adverbs must occur in the following order when they cooccur: *paradoxically* > *not* > *completely*. Although Ernst's calculus is linear, he does not rule out non-linearity of adverb ordering. For instance, one could assign the type $\langle f, i\rangle$ to some adverbs, and these would then be able to precede or follow any other adverb. Furthermore, since Ernst's calculus is stipulated, he could in principle impose any other relation on his ontological primitives, and avoid linearity of the calculus in this way. Needless to say, the latter point may also be construed as a weakness of this kind of approach.

A counterexample to linearity

The Norwegian triple of adverbs *muligens* ('possibly'), *ikke* ('not') and *alltid* ('always') is not linearly ordered and, hence, cannot be accomodated into a sequence of functional heads. If we consider pairs of adverbs, we find that *muligens* must precede *ikke*; *ikke* must precede *alltid*; but *muligens* does not have to precede *alltid*. Thus, we have a counterexample to transitivity. This behavior is illustrated with the Norwegian examples below.[2]

(3) a. Ståle har **muligens ikke** spist hvetekakene sine.
 S has possibly not eaten the-wheaties his
 "Stanley possibly hasn't eaten his wheaties."

[2] In English the sequence *didn't possibly* may be slightly better than the sharply ungrammatical Norwegian example above, although most English speakers find it degraded. Certainly, English allows examples like (i), with two modals, whereas the corresponding Norwegian example (ii) is still sharply out.
(i) Stanley clouldn't possibly eat his wheaties.
(ii) *Ståle kunne ikke muligens spise hvetekakene. ('S could not possibly eat the-wheaties')
Similarly, in English the sequence *always not* might be marginally acceptable while the Norwegian counterpart (*alltid ikke*) is sharply ungrammatical.

b. *Ståle har **ikke muligens** spist hvetekakene sine.
S has not possibly eaten the-wheaties his

(4) a. Ståle hadde **ikke alltid** spist hvetekakene sine.
S had not always eaten the-wheaties his
"Stanley hadn't always eaten his wheaties."
b. *Ståle hadde **alltid ikke** spist hvetekakene sine.
S had always not eaten the-wheaties his

(5) a. Ståle hadde **muligens alltid** spist noen andres hvetekaker.
S had possibly always eaten somebody elses wheaties
"Stanley had possibly always eaten somebody elses wheaties.
b. Ståle hadde **alltid muligens** spist noen andres hvetekaker.
S had always possibly eaten somebody elses wheaties.
"Stanley always possibly ate somebody elses wheaties."

One could assume that some of these pairs of adverbs are constituents, and hence irrelevant for our purposes. However, this cannot be right, since the subject of the clause can appear anywhere in the sequence without altering the meaning (modulo information structure). This is illustrated below for the pair *alltid muligens* but it can be done for all the pairs above. Hence it really appears that the adverbs modify the sentence individually.

(6) a. Derfor spiste **Ståle** alltid muligens noen andres hvetekaker.
therefore ate S always usually somebody elses wheaties
b. Derfor spiste alltid **Ståle** muligens noen andres hvetekaker.
c. Derfor spiste alltid muligens **Ståle** noen andres hvetekaker.

Comparing now examples with two adverbs to examples with three of them, we find that orderings which were bad with pairs are good when a different adverb intervenes. In particular, *muligens* can follow *ikke* if *alltid* intervenes (7).

(7) Ståle har **ikke alltid muligens** spist noen andres hvetekaker.
S has not always possibly eaten somebody elses wheaties

Again the subject of the clause can appear anywhere in the sequence, so the adverbs are not constituents.

(8) a. Derfor har ikke **Ståle** alltid muligens spist noen andres hvetekaker.
therefore has not S always possibly eaten somebody elses wheaties.
b. Derfor har ikke alltid **Ståle** muligens spist noen andres hvetekaker.
c. Derfor har ikke alltid muligens **Ståle** spist noen andres hvetekaker.

Let us see how our ordering pattern can be accomodated in an Ernst-type calculus. We refer to the first element in the type $\langle x, y\rangle$ as its *argument* and write $_A(\text{adv})$ for the argument of an adverb. Similarly we write $_R(\text{adv})$ for the second element in the type, i.e. the *result* of the function. Let us write

$\text{Adv}_1 \prec \text{Adv}_2$ for "Adv_1 *can precede* Adv_2."[3] Finally, we operate with a calculus $a_1 \leq \ldots \leq a_n$ where a_i are types and $\leq$ is a linear ordering of the types. In other words, we abstract away from Ernst's particular interpretation of his types as "events", "propositions", etc. This is in order to be able to focus on the *structural* properties of his calculus. In general, the following biconditional can be seen to hold:

$$\text{Adv}_1 \prec \text{Adv}_2 \Leftrightarrow {_R}(\text{Adv}_2) \leq {_A}(\text{Adv}_1)$$

By this it follows that (contraposition of $[\Leftarrow]$ and $[\Rightarrow]$)

$$\text{Adv}_1 \not\prec \text{Adv}_2 \Leftrightarrow {_R}(\text{Adv}_2) \not\leq {_A}(\text{Adv}_1)$$

and, by linearity of $\leq$ (i.e. connectedness), that

$$\text{Adv}_1 \not\prec \text{Adv}_2 \Leftrightarrow {_A}(\text{Adv}_1) < {_R}(\text{Adv}_2).$$

We can now use the ordering facts above to reason with the types of our triple of adverbs. The ordering facts are given on the left and the conclusions to be drawn concerning their types are given on the right.

ikke $\prec$ alltid	${_R}(\text{alltid}) \leq {_A}(\text{ikke})$
alltid $\not\prec$ ikke	${_A}(\text{alltid}) < {_R}(\text{ikke})$
muligens $\prec$ ikke	${_R}(\text{ikke}) \leq {_A}(\text{muligens})$
ikke $\not\prec$ muligens	${_A}(\text{ikke}) < {_R}(\text{muligens})$
muligens $\prec$ alltid	${_R}(\text{alltid}) \leq {_A}(\text{muligens})$
alltid $\prec$ muligens	${_R}(\text{muligens}) \leq {_A}(\text{alltid})$

By compiling this into one sequence, we get the following:

$${_R}(\text{alltid}) \leq {_A}(\text{ikke}) < {_R}(\text{muligens}) \leq {_A}(\text{alltid}) < {_R}(\text{ikke}) \leq {_A}(\text{muligens})$$

If, in order to minimize the amount of types, we always read $\leq$ as $=$, we need three basic types a, b, c such that $a < b < c$ and the following types for the adverbs which we simply read off the ordering above:

ikke	$\langle a, c\rangle$
muligens	$\langle c, b\rangle$
alltid	$\langle b, a\rangle$

If we have a category S of type a, we can show that the following facts obtain which is exactly what we want; we give some derivations below in the Type-Logical presentation:

derivable	not derivable
ikke(alltid(S))	*alltid(ikke(S))
muligens(ikke(S))	*ikke(muligens(S))
muligens(alltid(S))	
alltid(muligens(S))	
ikke(alltid(muligens(S)))	

[3] Strictly speaking, $\prec$ should be read ‘[illegible]n *apply after*’, but since, in the relevant cases, precedence=scope, we ignore this.

A Type Logical account

We would like to know what our types a, b, c are. What is more, we would like the relation imposed on them to follow from some fundamental property of these types. Type-Logical Grammar (TLG, cf. Morill 1994, Moortgat 1996) provides us with an interesting alternative to Ernst's ontological interpretation of a, b, c. In addition to the residuated triple of binary connectives $\backslash, \bullet, /$, TLG has the residuated pair of unary connected $\Diamond, \Box^{\downarrow}$, characterized by the following law of residuation:

$$\Diamond A \vdash B \text{ iff } A \vdash \Box^{\downarrow} B$$

This translates into the following Gentzen-style sequent rules:

$$\frac{\Gamma[\langle A\rangle] \vdash C}{\Gamma[\Diamond A] \vdash C}\,\Diamond L \qquad \frac{\Gamma \vdash A}{\langle \Gamma\rangle \vdash \Diamond A}\,\Diamond R$$

$$\frac{\Delta[A] \vdash C}{\Delta[\langle \Box^{\downarrow} A\rangle] \vdash C}\,\Box^{\downarrow} L \qquad \frac{\langle \Gamma\rangle \vdash A}{\Gamma \vdash \Box^{\downarrow} A}\,\Box^{\downarrow} R$$

Given this, we can prove the following general theorems to hold for the unary connectives:

$$\underset{1.}{\dfrac{\dfrac{A \vdash A}{\langle \Box^{\downarrow} A\rangle \vdash A}\,\Box^{\downarrow} L}{\Diamond\Box^{\downarrow} A \vdash A}\,\Diamond L} \qquad \underset{2.}{\dfrac{\dfrac{A \vdash A}{\langle A\rangle \vdash \Diamond A}\,\Diamond R}{A \vdash \Box^{\downarrow}\Diamond A}\,\Box^{\downarrow} R}$$

Suppose we set $a = \Diamond\Box^{\downarrow} s$, $b = s$, $c = \Box^{\downarrow}\Diamond s$. This ensures that, indeed, $a < b < c$, i.e. $\Diamond\Box^{\downarrow} s \vdash s \vdash \Box^{\downarrow}\Diamond s$. Thus, we no longer have to stipulate a relation on our types. Furthermore, this account allows us to derive the ordering patterns observed without enriching our semantic ontology, as was the case with Ernst's account. Our lexicon looks as follows:

Lexicon	
Γ	$\in \Diamond\Box^{\downarrow} s$
alltid	$\in \Diamond\Box^{\downarrow} s/s$
muligens	$\in s/\Box^{\downarrow}\Diamond s$
ikke	$\in \Box^{\downarrow}\Diamond s/\Diamond\Box^{\downarrow} s$

This allows us to derive that $(ikke \circ (alltid \circ \Gamma)) \vdash \Box^{\downarrow}\Diamond s$, but not that $(alltid \circ (ikke \circ \Gamma)) \vdash \Box^{\downarrow}\Diamond s$. As shown in (10), the latter fails because $\Box^{\downarrow}\Diamond s \vdash s$ is *not* a theorem.

(9)
$$\dfrac{\Box^{\downarrow}\Diamond s \vdash \Box^{\downarrow}\Diamond s \qquad \dfrac{\Diamond\Box^{\downarrow} s \vdash \Diamond\Box^{\downarrow} s \qquad \overset{\vdots\,(1.)}{\Diamond\Box^{\downarrow} s \vdash s}}{((\Diamond\Box^{\downarrow} s/s) \circ \Diamond\Box^{\downarrow} s) \vdash \Diamond\Box^{\downarrow} s}\,/L}{((\underbrace{\Box^{\downarrow}\Diamond s/\Diamond\Box^{\downarrow} s}_{\text{ikke}}) \circ ((\underbrace{\Diamond\Box^{\downarrow} s/s}_{\text{alltid}}) \circ \underbrace{\Diamond\Box^{\downarrow} s}_{\Gamma}))}\,/L$$

(10)
$$\dfrac{\overset{\vdots}{\Diamond\Box^{\downarrow} s \vdash \Box^{\downarrow}\Diamond s} \qquad \dfrac{\boxed{\Box^{\downarrow}\Diamond s \nvdash s} \qquad \Diamond\Box^{\downarrow} s \vdash \Diamond\Box^{\downarrow} s}{(\Box^{\downarrow}\Diamond s/\Diamond\Box^{\downarrow} s) \circ \Diamond\Box^{\downarrow} s \vdash s}\,/L}{((\underbrace{\Diamond\Box^{\downarrow} s/s}_{\text{alltid}}) \circ ((\underbrace{\Box^{\downarrow}\Diamond s/\Diamond\Box^{\downarrow} s}_{\text{ikke}}) \circ \underbrace{\Diamond\Box^{\downarrow} s}_{\Gamma})) \vdash \Box^{\downarrow}\Diamond s}\,/L$$

Parallel facts obtain with respect to *muligens* and *ikke*, given the lexicon above. We give the proof that we can derive the order *ikke alltid muligens*:

(11)
$$\dfrac{\Box^{\downarrow}\Diamond s \vdash \Box^{\downarrow}\Diamond s \qquad \dfrac{\Diamond\Box^{\downarrow} s \vdash \Diamond\Box^{\downarrow} s \qquad \dfrac{\overset{\vdots}{s \vdash s} \qquad \Diamond\Box^{\downarrow} s \vdash \Box^{\downarrow}\Diamond s}{((s/\Box^{\downarrow}\Diamond s) \circ \Diamond\Box^{\downarrow} s) \vdash s}\,/L}{((\Diamond\Box^{\downarrow} s/s) \circ ((s/\Box^{\downarrow}\Diamond s) \circ \Diamond\Box^{\downarrow} s)) \vdash \Diamond\Box^{\downarrow} s}\,/L}{((\underbrace{\Box^{\downarrow}\Diamond s/\Diamond\Box^{\downarrow} s}_{\text{ikke}}) \circ ((\underbrace{\Diamond\Box^{\downarrow} s/s}_{\text{alltid}}) \circ ((\underbrace{s/\Box^{\downarrow}\Diamond s}_{\text{muligens}}) \circ \underbrace{\Diamond\Box^{\downarrow} s}_{\Gamma}))) \vdash \Box^{\downarrow}\Diamond s}\,/L$$

The behavior of *muligens* is stronly reminiscent of that of positive polarity items (PPI) and the analysis given here is similar to that given for PPIs like *something* by Bernardi and Moot (2000).

Consider now adverbs like *lenger* (any.longer) which behave as negative polarity items. In order to accomodate these, we follow Bernardi (2001) in assigning to them a category the result of which fails to derive $\Box^{\downarrow}\Diamond s$, the category of grammatical sentences. This cannot be done in our linear calculus, since all of $\Diamond\Box^{\downarrow} s$, s, $\Box^{\downarrow}\Diamond s$ derive $\Box^{\downarrow}\Diamond s$. However, by using the Galois connected unary connectives ${}^{0}\cdot$, $\cdot^{0}$ (see Areces et.al. 2001), we get the desired result. These connectives are governed by the following law:

$$A \vdash B^{0} \text{ iff } B \vdash {}^{0}A$$

Given this, one can prove that $A \vdash {}^{0}(A^{0})$, but not vice versa (see Areces et. al 2001). Thus, we have that $\Diamond\Box^{\downarrow} s \vdash {}^{0}(\Diamond\Box^{\downarrow} s^{0})$, etc. This allows us to assign the type ${}^{0}(\Diamond\Box^{\downarrow} s^{0})/{}^{0}(\Diamond\Box^{\downarrow} s^{0})$ to *lenger*. We now have that $(lenger \circ \Gamma) \nvdash \Box^{\downarrow}\Diamond s$, because ${}^{0}(\Diamond\Box^{\downarrow} s^{0}) \nvdash \Box^{\downarrow}\Diamond s$. By assigning the type $\Box^{\downarrow}\Diamond s/{}^{0}(\Diamond\Box^{\downarrow} s^{0})$ to *ikke*, we retain the results above, while incorporating the behavior of *lenger*. It also correctly derives that *alltid* cannot intervene between *ikke* and *lenger*, while it can follow them both:

(12) a. Jens liker *(ikke) lenger hvetekaker.
J likes not any.longer wheaties

b. Jens spiser ikke (*alltid) lenger hvetekaner.
J eats not always any.longer wheaties

c. Jens spiser ikke lenger alltid hvetekaker.
J eats not any.longer always wheaties

References

Alexiadou, A. *Adverb Placement: a case study in antisymmetric syntax* Amsterdam: Benjamins.

Areces,C., R. Bernardi, M. Moortgat, "Galois Connections in Categorial Type Logic" In *Proceeding of Mathematics of Language and Formal Grammar 2001* Helsinki.

Bernardi, R. (2001)"Types for linguistic typologies. A case study Polarity Items." This volume.

Bernardi, Raffaella and Richard Moot (2000) "Generalized quantifiers in declarative and interro gative sentences." in *Proceedings of ICoS-2*.

Cinque, G. (1999) *Adverbs and Functional Heads: A crosslinguistic perspective* Oxford: OUP.

Ernst, T. (2001) "The scopal basis of adverb licensing" ms. Amherst.

Jackendoff, R. (1972) Semantic Interpretation in Generative Grammar Cabridge Mass.: MIT Press.

Moortgat, Michael (1996), "Categorial type logics", in J. van Benthem and A. ter Meulen (eds.) *Handbook of Logic and Language* Cambridge, Mass.: MIT Press.

Morrill, G. (1994) *Type logical grammar: categorial logic of signs* Dordrecht: Kluwer, 1994.

Nilsen, Ø. (2000) *The Syntax of Circumstantial Adverbials* Oslo: Novus Press.

Thomasson and Stalnaker (1973) "A semantic theory of adverbs" *Linguistic Inquiry* 4, 195-220.

WHAT PSYCHOLINGUISTICS TELLS US ABOUT THE SEMANTICS / PRAGMATICS INTERFACE: THE CASE OF PRONOUNS

Massimo Poesio, University of Essex

Abstract

Whereas most theories of the meaning of pronouns view them as having a single, underspecified lexical meaning, psychological evidence about pronoun resolution and lexical access suggests that the lexicon does not assign pronouns an initial interpretation; rather, their interpretation is specified by competing defaults operating incrementally. We propose an incremental theory of the semantics and pragmatics of pronoun interpretation based on this evidence, that assumes that interpretation involves competing defaults activated in parallel.

1 Introduction

Montague's lexical rules for pronouns stipulating that pronouns are infinitely ambiguous at the lexical level (Montague, 1973) are one of the aspects of his theories most clearly at variance with a cognitively plausible view of the mental lexicon. Thus, the current consensus view about pronouns differs from Montagues's in that pronouns are assigned a preliminary underspecified interpretation, in the form of a free index / variable denoting a function from assignments to values, sometimes with additional presuppositional content specifying that the entity is female / male, and / or that the variable must denote 'the most salient individual matching the pronoun in gender' (see, e.g., Heim and Kratzer (1998); Roberts (2001); Groenendijk et al. (1996). Jacobson, as well (1999), assumes that pronouns have a single underspecified interpretation.

The psychological evidence about lexical access and pronominal interpretation, however, challenges this view, suggesting instead that lexical access for pronouns involves immediate access to all (contextually salient) interpretations. This might at first seem to indicate that Montague was right after all; but current theories of underspecification allow us to formulate an alternative explanation which does not rely on assuming an infinite number of interpretations at the lexical level. If we allow the semantic interpretation of certain utterances to be left empty by the lexicon, and instead to be supplied by contextual rules operating incrementally immediately after lexical access, we can hypothesize that what happens in the case of pronouns is that contextual resolution rules apply immediately after the lexicon has been accessed, and it is these rules that provide the interpretation of the pronoun. In this paper, we show how this hypothesis about the division of labor between semantics and pragmatics can be made more precise.

2 Psychological Evidence about Pronoun Interpretation

The fundamental experimental result about lexical access - reported by, e.g., Swinney (1979) and later confirmed by experiments by Seidenberg et al, Marslen-Wilson et al, and Tanenhaus, among others - is that all lexical entries of a homonym like *bank* are immediately retrieved. A crucial additional fact about lexical access was revealed by experiments by Frazier and Rayner (1990), showing that garden path diagnostics can be used to classify words in two classes. Words like *pitcher* or *bank* have multiple lexical entries, which are accessed in parallel and immediately

pruned, which causes garden path effects when subsequent context reveals that the preferred interpretation is not the right one). However, immediate garden-path effects are not observed with words like *newspaper*, suggesting that they have a single lexical entry assigning them an underspecified meaning covering all the possible interpretations. Frazier and Rayner thus proposed a 'weaker' form of the incrementality hypothesis, the Immediate Partial Interpretation Principle (IPIH), stating that while utterances have to be immediately interpreted, the initial interpretation may be underspecified and refined only later.

Experimental results on pronominal interpretation by, among others, Corbett and Chang (1983) and Gernsbacher (1989) show that lexical access for pronouns is more like lexical access for *pitcher* than lexical access for *newspaper*. In sentences like *Karen poured a drink for Emily and then she put the bottle down*, both *Karen* and *Emily* are activated after the pronoun is encountered (but not before it, or when the proper name *Karen* occurs in its place).[1]

3 Accounting for the psychological results

As said above, one way of explaining these results is to argue that Montague was right after all, and pronouns are infinitely ambiguous at the lexical level; but there are obvious problems with this idea from a cognitive perspective. Fortunately, current theories of underspecification allow us to formulate the alternative explanation mentioned above.

This hypothesis about the division of labor between semantics and pragmatics can be made more precise using the theory of incremental semantic interpretation and underspecification developed in (Poesio, 1994, 1995; Poesio and Traum, 1997; Poesio and Muskens, 1997; Poesio, 2002). According to this theory, semantic processing is best viewed as a process during which oncoming utterances are incrementally assigned an interpretation, in part by the lexicon, in part by syntactic and contextual rules. As in sign-based theories, this interpretation is characterized in terms of functions specifying the syntactic and semantic classification of utterances, among which a function **sem**(u) specifying an utterance's meaning. We briefly introduce the theory in this section; the interested reader is referred to the papers just cited for details.

The common ground as a record of the occurrence of utterances The theory we assume hypothesizes that the common ground includes a record of both the occurrence of utterances and their interpretation. This information is characterized using Compositional DRT (Muskens, 1995, 1994), modified as suggested in (Poesio and Muskens, 1997) to recover anaphoric accessibility. According to the theory, an utterance of the sentence *Kermit croaked* is incrementally interpreted as follows. First, the occurrence of an utterance of the word *Kermit* is recorded in the common ground: this amounts to updating the previous common ground K with the information that an event u_{pn} of uttering the word *Kermit* occurred, as follows:

(1) $K' = K; [u_{pn} | u_{pn}{:}\mathbf{utter}(a, \text{"}Kermit\text{"})]$

K' is then further updated as a result of lexical access and incremental parsing. As in sign-based theories, lexical access assigns a syntactic and semantic classification to u_{pn} in terms of the functions **cat** and **sem**; parsing hypothesizes that u_{pn} is a constituent of larger utterances–such an NP u_{np}.[2] The result of lexical access and parsing are the interpretations in (2) and (3), respectively.

[1] If only one antecedent matches the pronoun, the other interpretation is immediatly discarded; otherwise both competing interpretations remain active until the end of the sentence is reached.

[2] We assume here for simplicity a purely bottom-up parsing.

(2) $K'' = K'; [|\mathbf{sem}(u_{pn})\ \mathbf{is}\ \lambda P([x_n|x_n\ \mathbf{is}\ k]; P(x_n)), \mathbf{cat}(u_{pn})\ \mathbf{is}\ pn]$

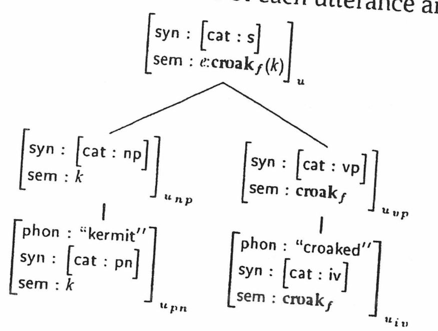
(3) $K''' = K''; [u_{np}|\mathbf{cat}(u_{np})\ \mathbf{is}\ np, u_{pn} \lhd u_{np}]$

The rules for semantic composition specify how the meaning of utterances such as u_{np} is compositionally derived from the meaning of their constituents, resulting in the further update:

(4) $K'''' = K'''; [|\mathbf{sem}(u_{np})\ \mathbf{is}\ \lambda P[x_{pn}|x_{pn} = k]; P(x_{pn})]$

The subsequent utterance of the word *croaked* results in further updates of the common ground, such as hypothesizing the occurrence of an utterance of a constituent u_{iv}, which is part of a larger VP constituent u_{vp} which, together with u_{np}, constitutes a sentence u_s. After the meaning of each of these constituents is derived, we obtain the interpretation in (5), in which we have used a attribute-value notation to indicate the classifications of each utterance and arcs to indicate domination.

(5)



Semantic interpretation as defeasible reasoning The fact that interpretations are generated incrementally and in parallel supports the view that interpretive processes such as lexical access, parsing, and pronoun resolution are defeasible inferences processes, advocated, e.g., by (Hobbs et al., 1993; Lascarides and Asher, 1993; Kameyama and Jaspars, 1996). Here these aspects of interpretation are uniformly modeled in terms of Prioritized Default Logic (PDL) (Brewka, 1991). In the somewhat simplified theory of the lexicon assumed here, the lexicon is a PDL theory with one default inference rule for each lexical entry; these rules are used to assign a syntactic category and a meaning to utterances during interpretation. For homonyms like *pitcher*, the lexicon contains one (normal) default inference rule for each interpretation (say, **pitcher**$_1$ and **pitcher**$_2$), possibly with different priorities (open variables are indicated by capitals, all but the essential aspects of lexical information are omitted):

```
U:utter(X,"pitcher") : cat(U) is n, sem(U) is PITCHER_1
----------------------------------------------------
             cat(U) is n, sem(U) is PITCHER_1

U:utter(X,"pitcher") : cat(U) is n, sem(U) is PITCHER_2
----------------------------------------------------
             cat(U) is n, sem(U) is PITCHER_2
```

These rules are immediately activated, and the interpretation with stronger priority chosen, pruning the rest; when the disambiguating context is encountered, reanalysis must take place. By contrast, the lexicon contains only one default inference rule for polysemous words like *newspaper*, assigning them an initial underspecified interpretation **newspaper**$_U$ (whose extension is the union of the extensions of its fully specified interpretations).

```
U:utterance(X,"newspaper") : cat(U) is n, sem(U) is NEWSPAPER_U
-----------------------------------------------------------
             cat(U) is n, sem(U) is NEWSPAPER_U
```


The rules that eliminate the underspecification are not part of the lexicon but operate at a later stage, after the sentence is completed, as suggested by Frazier and Rayner (see also (Lascarides and Copestake, 1998)). Thus, no reanalysis is required when the disambiguating context is encountered.

4 The semantics and pragmatics of pronouns

The lexical entry for *he*, on the other hand, only specifies its syntactic properties (again, we omit all but the essential information):

```
U:utterance(X,"he") : cat(U) = pro, gender(U) = male,
                      num(U) = sing
------------------------------------------------------
      cat(U) = pro, gender(U) = male, num(U) = sing
```

Because this lexical entry does not specify the meaning for the pronoun, this must be derived from contextual reasoning in order to satisfy Frazier and Rayner's IPIH. Our contextual interpretation rules for pronouns use a conceptual vocabulary derived from Centering Theory (Grosz et al., 1995). More specifically, we assume that the local focus is updated by C-UTTERANCES, each of which introduces into context a number of ranked FORWARD-LOOKING CENTERS (CFs). A complete formalization of Centering involves both 'weak' rules activated in the absence of a CB, and stronger rules that operate when a CB has been identified. We only discuss here the weak rule for pronoun resolution, that hypothesizes as a meaning of a pronoun utterance U_{pro} occurring in c-utterance U_{n+1} the discourse entity assigned as a meaning of utterance U_{np} occurring in the previous c-utterance U_n, provided that the two CF-introducing utterances match. We write **cf-utt**(U,U') for "U' is an utterance part of the c-utterance U that introduces a CF", and **prev-utt**(U,=) U' for "U' is the c-utterance preceding U".

$$\frac{[|\ \mathbf{cat}(U_{pro})\ \mathbf{is}\ pro,\ \mathbf{cf\text{-}utt}(U_{n+1},U_{pro}),\ \mathbf{cf\text{-}utt}(U_n,U_{np}),\ \mathbf{prev\text{-}utt}(U_{n+1})\ \mathbf{is}\ U_n,\ \mathbf{agr\text{-}match}(U_{np},U_{pro}),\ \mathbf{sem}(U_{np})\ \mathbf{is}\ x\ ] : [|\mathbf{sem}(U_{pro})\ \mathbf{is}\ x]}{[|\mathbf{sem}(U_{pro})\ \mathbf{is}\ x]}\ \text{PRO-MATCH}$$

Under the assumption that the example sentences in Corbett and Chang's experiments involve two c-utterances, when the pronoun *she* is encountered in the second c-utterance, in a context in which both *Karen* and *Emily* are introduced in the previous c-utterance, the rule above is activated twice, generating two conflicting extensions. We discuss in the longer version of the paper the implications of these findings concerning the form of the 'stronger' rules implementing centering preferences.

References

Brewka, G., 1991. *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge: Cambridge University Press.

Corbett, A. and F. Chang, 1983. Pronoun disambiguating: Accessing potential antecedents. *Memory and Cognition*, 11:283–294.

Frazier, L. and K. Rayner, 1990. Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, 29:181–200.

Gernsbacher, M. A., 1989. Mechanisms that improve referential access. *Cognition*, 32:99–156.

Groenendijk, J.A.G., M.J.B. Stokhof, and F. Veltman, 1996. Coreference and modality. In S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, chapter 7. Blackwell, pages 179–213.

Grosz, B. J., A. K. Joshi, and S. Weinstein, 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):202–225. (The paper originally appeared as an unpublished manuscript in 1986.).

Heim, I. and A. Kratzer, 1998. *Semantics in Generative Grammar*. Blackwell.

Hobbs, J. R., M. Stickel, P. Martin, and D. Edwards, 1993. Interpretation as abduction. *Artificial Intelligence Journal*, 63:69–142.

Jacobson, P., 1999. Towards a variable-free semantics. *Linguistics and Philosophy*, 22:117–184.

Kameyama, M. and J. Jaspars, 1996. Preferences in dynamic semantics. In M. Stokhof and P. Dekker (eds.), *Proc. of the 10th Amsterdam Colloquium*. ILLC, Amsterdam.

Lascarides, A. and N. Asher, 1993. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, 16(5):437–493.

Lascarides, A. and A. Copestake, 1998. Pragmatics and word meaning. *Journal of Linguistics*, 34(2):387–414.

Montague, R., 1973. The proper treatment of quantification in English. In K.J.J. et al. Hintikka (ed.), *Approaches to Natural Language*. Dordrecht: D. Reidel, pages 221–242. Reprinted in Thomason (1974).

Muskens, R., 1994. A compositional discourse representation theory. In P. Dekker and M. Stokhof (eds.), *Proceedings of the 9th Amsterdam Colloquium*.

Muskens, R., 1995. Tense and the logic of change. In U. Egli, P. Pause, C. Schwarze, A. von Stechow, and G. Wienold (eds.), *Lexical Knowledge in the Organization of Language*. Amsterdam / Philadelphia: John Benjamins, pages 147–183.

Poesio, M., 1994. *Discourse Interpretation and the Scope of Operators*. Ph.D. thesis, University of Rochester, Department of Computer Science, Rochester, NY.

Poesio, M., 1995. Disambiguation as (defeasible) reasoning about underspecified representations. In P. Dekker and M.Stokhof (eds.), *Proc. of the Tenth Amsterdam Colloquium*. ILLC.

Poesio, M., 2002. *Incrementality and Underspecification in Semantic Interpretation*. Lecture Notes. Stanford, CA: CSLI. To appear.

Poesio, M. and R. Muskens, 1997. The dynamics of discourse situations. In P. Dekker and M. Stokhof (eds.), *Proceedings of the 11th Amsterdam Colloquium*. University of Amsterdam, ILLC.

Poesio, M. and D. Traum, 1997. Conversational actions and discourse situations. *Computational Intelligence*, 13(3):309–347.

Roberts, C., 2001. Uniqueness presuppositions in english definite noun phrases. Available on the Semantics Archive.

Swinney, D. A., 1979. Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18:545–567.

Thomason, R. H. (ed.), 1974. *Formal Philosophy: Selected Papers of Richard Montague*. New York: Yale University Press.

On Certain Epistemic Implicatures in Yes/No Questions

Maribel Romero (University of Pennsylvania)
Chung-hye Han (Simon Fraser University)

1 Introduction

Inverted or preposed negation in *yes/no* (*yn*-)questions contributes the implicature that the speaker expects the answer to be in the affirmative (Ladd (1981), Han (1998), Gunlogson and Büring (2000)). For instance, the intuition is that (1) asks whether John drinks and implies that the speaker believes or at least expects that John drinks. Note also that *yn*-questions with non-preposed negation do not necessarily have this implicature. (2) can be another way of seeking information on whether John is a teetotaler (cf., Han (1999)). We will call this type of implicature *epistemic implicature*.

(1) Doesn't John drink?
Positive epistemic implicature: The speaker has the previous belief or expectation that John drinks.

(2) Does John not drink?
No epistemic implicature necessary.

The goal of this paper is to provide (tentative) answers to the following questions:

1. What property correlated with the existence of an implicature distinguishes preposed negation from non-preposed negation?
2. How does this property of preposed negation enforce an epistemic implicature?
3. Why is the implicature raised by preposed negation a *positive* epistemic implicature? That is, why are the polarity in the question and the polarity in the implicature opposite?

2 A first hypothesis

A first explanation of the contrast between (1) and (2) would maintain that: (i) preposed negation in *yn*-questions is sentential negation, whereas non-preposed negation is VP constituent negation; and (ii) sentential negation, when combined with the semantics of *yn*-questions, is responsible for the epistemic implicature.

Sentential vs. constituent negation will not do it. In (3a), negation is not just negating the event contributed by the VP and is more like a sentential negation negating the entire modal proposition. Still, (3a) does not give rise to a necessary epistemic implicature, in contrast with its preposed negation version in (3b):

(3) a. Does John not have to go to the meeting? ($\neg\Box$)
b. Doesn't John have to go to the meeting? ($\neg\Box$)

If we say that negation in (3a) is still constituent negation –negating a bigger constituent than VP–, the distinction between constituent and sentential negation becomes murky.

The semantics of *yn*-questions (Hamblin (1973)) According to Hamblin (1973), the denotation of a question is the set of its possible answers. A question operator -overt *whether* or the silent *Q*-morpheme- is in charge of taking the proposition expressed by the IP and turn it into the appropriate question denotation, as shown in (4).

(4) a. LF: [$_{CP}$ Whether/Q [it is raining]]
b. $[\![it\ is\ raining]\!] = \lambda w.\ it\ is\ raining\ in\ w$
c. $[\![whether]\!] = [\![Q]\!] = \lambda p_{<s,t>} \lambda w_s \lambda q_{<s,t>}\ [q = p\ \vee\ q = \neg p]$
d. $[\![whether/Q\ it\ is\ raining]\!](w_o)$
$= \lambda q\ [q = \lambda w.\ it\ is\ raining\ in\ w\ \vee\ q = \lambda w.\ \neg(it\ is\ raining\ in\ w)]$
= {"that it is raining", "that it is not raining"}

If we apply these semantics to *yn*-questions containing a (sentential) negation operator, we obtain exactly the same question meaning for (5a) as we did for (4a). No epistemic implicature follows from this semantic computation.

(5) a. LF: [$_{CP}$ Whether / Q [not [it is raining]]]
b. $[\![whether/Q\ it\ is\ not\ raining]\!](w_o)$
$= \lambda q\ [q = \lambda w.\neg(\ it\ is\ raining\ in\ w)\ \vee\ q = \lambda w.\ \neg\neg(it\ is\ raining\ in\ w)]$
= {"that it is not raining", "that it is raining"}

3 Focus is relevant

Interestingly, parallel effects to the ones associated with preposed negation can be reproduced in affirmative questions if we place Focus stress on the auxiliary (and on nothing else): (6) can be used to convey the negative implicature that the speaker believes that John does not drink. The non-stressed auxiliary version (7) is not biased in this way.

(6) DOES John drink?
Negative epistemic implicature: The speaker expects that John does not drinks.

(7) Does John drink?
No epistemic implicature.

Furthermore, if we take a *yn*-question with non-preposed negation and place focus stress on *not*, the epistemic implicature arises again:

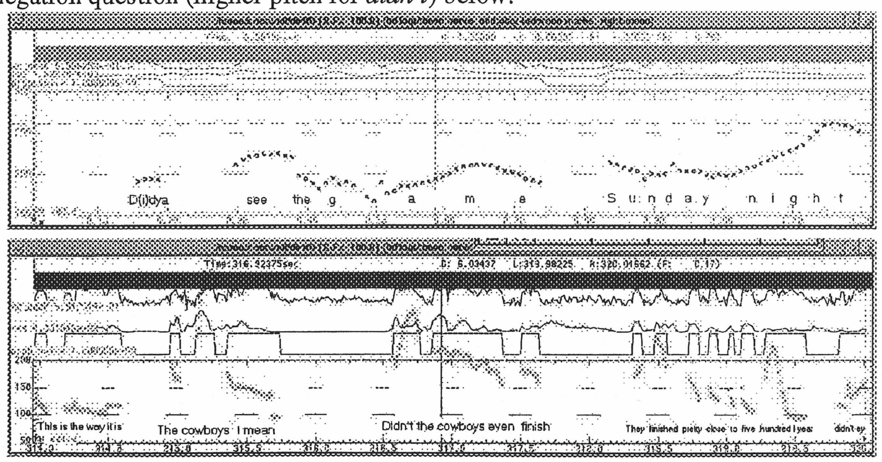
(8) Does John NOT drink?
Positive epistemic implicature: The speaker expects that John drinks.

Note that the polarity of a question carrying an implicature and the polarity of the implicature itself are opposite: i.e., *negative yn*-questions with preposed negation give rise to a *positive* epistemic implicature, and *positive yn*-questions with focus on the auxiliary give rise to a *negative* implicature. This crossed pattern of implicatures is the same as the distributional pattern of tag questions, which clearly bear focus stress on the auxiliary:

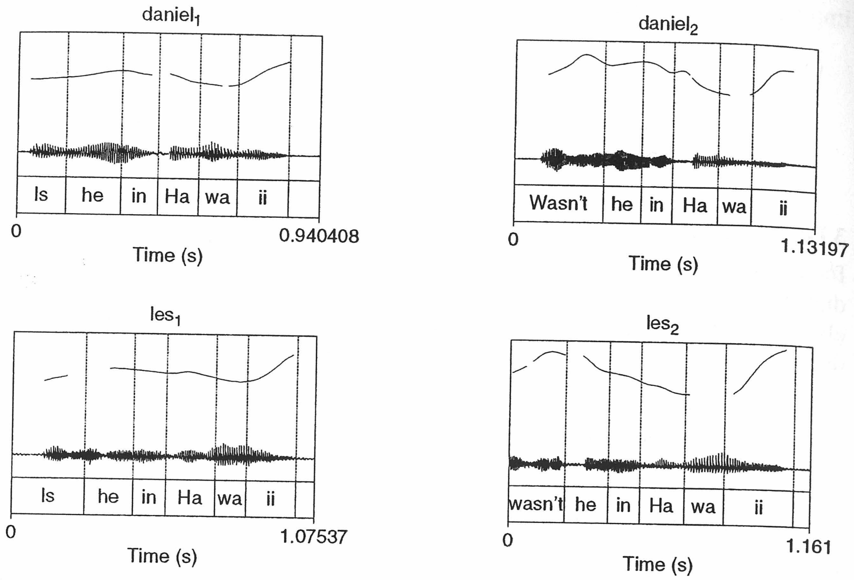
(9) a. John drinks, DOESN'T he? b. John doesn't drink, DOES he?

All this raises the question of whether the existence of epistemic implicatures and the crossed pattern of their polarities is related to focus. If so, we would expect our original sentences with preposed negation to involve focus-marking as well.

Preliminary evidence from naturally occurring data suggests that preposed negation does involve a special pitch curve different from non-focused auxiliaries. Compare the pitch track of the regular affirmative question (low pitch for *did*) with that of the preposed negation question (higher pitch for *didn't*) below:



We have also conducted a small experiment that elicits an (unfocused) affirmative *yn*-question and a negative *yn*-question with preposed negation in appropriate contexts. The results show that the negated auxiliary verb has relatively higher pitch than the auxiliary verb in affirmative questions.



In view of these data, we will assume that preposed negation bears Focus too, and it does so necessarily, whereas non-preposed negation, instead, can –but does not need to– be focused.[1] We will pursue a unified focus-based account of the positive and negative implicatures above: in all the examples with a necessary epistemic implicature, the negative/positive polarity inside the question is focus-marked (Verum Focus as in Höhle (1992)), lending (10a)-(11a) roughly equivalent to (10b)-(11b) respectively.

(10) a. Doesn't John drink? b. Is it FALSE that John drinks?

(11) a. DOES John drink? b. Is it TRUE that John drinks?

4 How polarity focus generates an epistemic implicature

4.1 A second hypothesis

Theories of Focus converge on the idea that non-focused material must be old, whereas focused material must be new (Rooth (1992), Schwarzschild (1999)). Applying that to (12a) (and obliterating the contribution of the question operator Q), it follows that the proposition in (12b) -with the positive polarity POS instead of *not*- must be old in the discourse.

(12) a. Wasn't$_F$ he in Hawaii?
b. The proposition $\lambda w. John\ was\ in\ Hawaii\ in\ w$ must be old.

One could argue that, since the proposition in (12b) is old but has not been explicitly expressed in the previous discourse, we infer that it is old in the *epistemic state* of the speaker. But, then, the proposition in (12b) certainly should count as old if explicitly expressed too, e.g. if the speaker A asserted it. This wrongly predicts (13) should be fine.

(13) A: John was in Hawaii last week. B: # Wasn't he (in Hawaii last week)?

In fact, the appropriate epistemically biased question to follow A's utterance does not contain a focused negative polarity, but a focused *positive* polarity, as in (14).

(14) A: John was in Hawaii last week. B: WAS he (in Hawaai last week)?

[1] It may be that preposed negation sometimes associates with other focused element instead of signaling focus-marking on itself. The present analysis can be extended to cover that case.

The generalizations we draw from this set of facts are: (i) B's question with focus on polarity α correlates with B's epistemic implicature of polarity β; and (ii) B's question with focus on polarity α correlates with A's utterance(/implication) of polarity α. The hypothesis presented in this subsection cannot explain these generalizations.

4.2 Our proposal

We assume that focus on polarity is evaluated with respect to a probabilistic epistemic model, where each proposition in the speaker's epistemic state is mapped to a probability value ranging from 1 ((TRUE-FOR-SURE, i.e, $[\![POS]\!]$) to 0 (FALSE-FOR-SURE, i.e., $[\![not]\!]$) (cf. Bayesian models in Gaerdenfors 1988). Other probability measures can be expressed by simple or complex expressions: e.g., *probably* (.9), *most likely* (.8), *likely* (.7), *possibly* (.5), etc. Each of these probability measures is an alternative to all the others. Focusing one of the expressions makes all the probability measures relevant in a way that will be important for the type of *yn*-questions at issue. (But see Höhle (1992)'s section 6.3.)

We further note that in a coherent discourse, we often find a hierarchy of superquestions and subquestions (Roberts (1996)). E.g., if we are searching for the answer to "Who is married to Bertha?", we may proceed by asking the subquestions "Is John married to Bertha?", "Is Paul married to Bertha?", etc. We propose that Focus can be used to mark this relation explicitly: focus in (15) presupposes that (15) is just a subquestion and that its superquestion "Who is married to Bertha?" is relevant and salient in the discourse.

(15) Is JOHN married to Bertha?
→It presupposes relevant superquestion: "Who is married to Bertha?"

Now, let us take the mini-discourse (16).

(16) A: I saw John at the movies last night (in Philadelphia).
B: Wasn't he in Hawaii? / Is it FALSE that he was in Hawaii?

As in (15), (16B) presupposes that there is a relevant, salient superquestion (SQ) "What is the probability assignment to the proposition "John was in Hawaii"?". In (16) (out of the blue), the only trigger for SQ is A's utterance. But A's utterance is not that question. How can it trigger or raise SQ? A's utterance triggers SQ if the acceptance of A's proposition into the epistemic state induces a revision of it. To see this, let us take an initial epistemic state S1, where the proposition P "John was in Hawaii" has probability 0.9 or 1. Then we get A's utterance, which implies that the proposition P "John was in Hawaii" is mapped to 0. This raises the question of what is the probability measure of P, after all. That is, this raises our SQ. In sum, the effect of Focus on the polarity –no matter whether it's on TRUE or on FALSE– is to presuppose the SQ and, from that, to imply that we had pre-existent beliefs about the answer to SQ and that they have been contradicted in the last update.

5 The crossed polarity pattern for epistemic implicatures

We now need to address the issue of why the focused polarity in the question and the polarity in the implicature are opposite. We will use the paradigm in (17) for illustration.

(17) Previous belief of B: The speaker B believed John was in Hawaii.
A: I saw John at the movies last night (in Philadelphia).
B: Wasn't he in Hawaii? B: Is it FALSE that he was in Hawaii?
B: # WAS he is Hawaii? B: # Is it TRUE that he was in Hawaii?

Let us first examine B's good response. We saw earlier that the Q morpheme in *yn*-questions takes the proposition P expressed by its sister node and makes a two-member set containing P and its negation $\neg P$. The resulting denotation is a set of propositions that divides the probability measure space for P in two *balanced* cells (cf. Groenendijk and Stokhof (1985)'s partition over the background set of possible worlds).

(18) $\boxed{\lambda w.\ it\ is\ raining\ in\ w\ \rightarrow\ 1}$ $\boxed{\lambda w.\ it\ is\ raining\ in\ w\ \rightarrow\ 0}$

Adding focus and the semantics of *yn*-questions, the denotation of a question with focus on epistemic polarity FALSE-FOR-SURE is as in (21).

(19) a. B: Wasn't he in Hawaii? b. B: Is it FALSE that he was in Hawaii?

(20) LF: $[_{CP}\ Q$ [FALSE-FOR-SURE$_F$ [he was in Hawaii]]]

(21) $\{ false - for - sure(\lambda w.\ he\ was\ in\ Hawaii\ in\ w),$
$\neg false - for - sure(\lambda w.\ he\ was\ in\ Hawaii\ in\ w) \}$

Assuming that a focused epistemic polarity makes all the gradient alternative probability measures salient, we obtain a partition of the probability continuum in two *unbalanced* cells:

(22)

$\lambda w.\ he\ was\ in\ Hawaii\ in\ w \ \rightarrow\ 0$	$\lambda w.\ he\ was\ in\ Hawaii\ in\ w\ \rightarrow\ n$, where $0 < n \leq 1$

Note that, given that accepting the proposition $\lambda w.\ you\ saw\ John\ at\ movies\ in\ Ph\ in\ w$ entails rejecting the proposition $\lambda w.\ he\ was\ in\ Hawaii\ in\ w$, the same partition would obtain if B had responded *DID you (see him at the movies last night)?*. That is, given the entailment between the two propositions, (22) can be further spelled out as in (23):

(23)

$\lambda w.\ he\ was\ in\ Hawaii\ in\ w\ \rightarrow\ 0$ $\lambda w.\ A\ saw\ John\ atm\ in\ w\ \rightarrow\ 1$	$\lambda w.\ he\ was\ in\ Hawaii\ in\ w\ \rightarrow\ n$, where $0 < n \leq 1$ $\lambda w.\ A\ saw\ John\ atm\ in\ w\ \rightarrow\ n$, where $0 \leq n < 1$

Let us now turn to B's bad response. The denotation of (24) is given in (26). In the resulting partition (27), the cells are unbalanced on the opposite extreme:

(24) a. B: # WAS he in Hawaii? b. B: # Is it TRUE that he was in Hawaii?

(25) LF: $[_{CP}\ Q$ [TRUE-FOR-SURE$_F$ [he was in Hawaii]]]

(26) $\{ true - for - sure(\lambda w.\ he\ was\ in\ Hawaii\ in\ w),$
$\neg true - for - sure(\lambda w.\ he\ was\ in\ Hawaii\ in\ w) \}$

(27)

$\lambda w.\ he\ was\ in\ Hawaii\ in\ w\ \rightarrow\ 1$	$\lambda w.\ he\ was\ in\ Hawaii\ in\ w\ \rightarrow\ n$, where $0 \leq n < 1$

The same partition would obtain if B responded *Didn't you (see him at the movies last night)?*, since accepting the proposition $\lambda w.\ he\ was\ in\ Hawaii\ in\ w$ entails rejecting the proposition $\lambda w.\ you\ saw\ John\ at\ movies\ in\ Ph\ in\ w$. That gives us the bad partition in (28):

(28)

$\lambda w.\ he\ was\ in\ Hawaii\ in\ w\ \rightarrow\ 1$ $\lambda w.\ A\ saw\ John\ atm\ in\ w\ \rightarrow\ 0$	$\lambda w.\ he\ was\ in\ Hawaii\ in\ w\ \rightarrow\ n$, where $0 \leq n < 1$ $\lambda w.\ A\ saw\ John\ atm\ in\ w\ \rightarrow\ n$, where $0 < n \leq 1$

The question is: what makes (23) acceptable and (28) unacceptable in the context illustrated in (17)? We propose that the contrast ultimately stems from whether or not the question is informative (Grice (1975)). We also assume, in the spirit of Gaerfenfors informational economy (p. 49), that a speaker wants to retain as much as possible from her old beliefs and that, hence, she only executes a revision of her previous epistemic state for a proposition P' if there is enough certainty about P'. In the examples at issue, A's utterance asserted or entailed any of the propositions P' in (29). Given the Gricean cooperation principles, this implies that A assigns P' a very high probability measure (.9 or 1). But, since accepting P' would suppose a revision of B's epistemic state, B will only execute such revision if P' is certain. That is, informational economy makes highly relevant a question that would distinguish between the measures .9 and 1 for P'.

(29) P' = $\lambda w.\ A\ saw\ John\ atm\ in\ w$, or
P' = $\lambda w. \neg(\ he\ was\ in\ Hawaii\ in\ w\ )$, or
P' = $\lambda w.\ (\ A\ saw\ John\ atm\ in\ w\ )\ \wedge\ \neg(\ he\ was\ in\ Hawaii\ in\ w\ )$

Let us see how these considerations impact our unbalanced partitions. The good partition carves the spectrum of probability measures so that the values .9 and 1 for the conflicting proposition P' are in two different cells. B asks A to choose one of the cells in (23). In this way, B is asking A to distinguish between a high probability belief for P' (.9) and certainty about P' (1). That is, B asks his question in a way coherent with the Gricean maxims and useful to the pursuit of informational economy. However, the bad partition draws the line between the probability measures 0 and .1 for the conflicting proposition P'. B asks A to choose the cell [0] or the cell [.1, .2, ..., .9, 1]. But, since A just asserted P', A must map P' to a very high probability measure, .9 or 1. But this means that the question B is asking with this partition has already been answered by the fact that A uttered P'. Hence, the partition induced by this question is bogus.

6 Conclusions and further issues

Preposed negation carries focus-marking on the negative polarity (Verum Focus in Höhle (1992)), and it does so necessarily. Non-preposed negation may or may not be focus-marked.

The felicity conditions of Polarity Focus in a question require that the corresponding superquestion (Roberts (1996)) be presupposed or salient in the previous discourse. Typically, a non-uttered question is salient if new information contradicts one speaker's previous beliefs, leading to a contradictory epistemic state that raises the superquestion.

Unfocused *y/n*-questions induce a balanced partition ([P'], [¬P']), whereas questions with Polarity Focus induce an unbalanced partition on the space of probability measures for P'. Only when the focused polarity in the question is opposite to that in the epistemic implicature (i.e., only when it its equal to that in A's utterance or its implication) is the unbalanced partitition informative and, hence, felicitous.

References

Grice, H. P. (1975) "Logic and conversation," in P. Cole and J. L. Morgan, eds., *Speech Acts*, volume III of *Syntax and Semantics*, Academic Press, NY, 41–58.

Groenendijk, J., and M. Stokhof (1985) *Studies in the Semantics of Questions and the Pragmatics of Answers*, Doctoral dissertation, University of Amsterdam.

Gunlogson, C., and D. Büring (2000) "Aren't positive and negative polar questions the same?" Presented at LSA 2000.

Hamblin, C. L. (1973) "Questions in Montague Grammar," *Foundations of Language* 10, 41–53.

Han, C.-H. (1998) "Deriving the interpretation of rhetorical questions," in E. Curis, J. Lyle, and G. Webster, eds., *Proceedings of West Coast Conference in Formal Linguistics*, volume 16, Stanford, CSLI, 237–253.

Han, C.-H. (1999) *The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar*, Doctoral dissertation, University of Pennsylvania.

Höhle, T. (1992) "Ueber Verum Fokus in Deutschen," *Linguistische Berichte*.

Ladd, R. D. (1981) "A First Look at the Semantics and Pragmatics of Negative Questions and Tag Questions," in *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society*, Chicago Linguistics Society, 164–171.

Roberts, C. (1996) "Information structure in discourse: Towards an Integrated Formal Theory of Pragmatics," in J. Yoon and A. Kathol, eds., *OSU Working Papers in Linguistics 49: Papers in SEmantics*, Ohio State University, Dept. of Linguistics, 91–136.

Rooth, M. (1992) "A theory of focus interpretation," *Natural Language Semantics* 1, 75–116.

Schwarzschild, R. (1999) "GIvenNess, AvoidF and Other Constraints on the Placement of Accent," *Natural Language Semantics* 7:2, 141–177.

Partial Adjectives vs. Total Adjectives: Scale Structure and Higher-Order Modification

Carmen Rotstein — Technion
Yoad Winter — Technion/UiL OTS

1 Introduction

Modifiers such as *slightly, nearly, completely, almost* and *very* apply to many adjectives and give rise to meanings that are hard to capture using first order logic. In this paper we show that the semantic behavior of such modifiers is sensitive to Cruse's (1980) and Yoon's (1996) typology of *total* and *partial* adjectives. These are pairs of adjectives such as *safe-dangerous, clean-dirty* or *healthy-sick*, where the first ("total") adjective in each pair describes lack of danger, dirt, malady etc., while the second ("partial") adjective describes the existence of such properties. We analyze the truth and acceptability conditions of constructions such as *almost safe/?dangerous* and *slightly dangerous/?safe* and show some systematic contrasts between total and partial adjectives. These contrasts are formally accounted for by giving total adjectives and partial adjectives a different *scale structure*. A total adjective is associated with a scale that is bounded on both ends, while a partial adjective is associated with a scale that is bounded by zero from below, and unbounded from above. Another assumption made is that the *standard value* of the total scale is constantly zero – the same as the lower bound of the partial scale.

2 Partial Adjectives vs. Total Adjectives

Cruse's and Yoon's distinction between partial adjectives and total adjectives can be exemplified by the following contrast in the intuitive meanings of the adjectives *dirty* and *clean*.

(1)
dirty	=	has some degree of dirtiness	≠	has no degree of cleanliness
clean	=	has no degree of dirtiness	≠	has some degree of cleanliness

Some more partial-total pairs are given below.

(2) *open-closed, sick-healthy, wet-dry, curved/crooked-straight, dangerous-safe, hungry-satiated, cracked-whole, incomplete-complete, rough-smooth, vague-clear, alive-dead*

Using this intuitive contrast, Yoon classifies partial adjectives like *dirty* as "existential" and total adjectives like *clean* as "universal". Yoon argues for some empirical differences between the two classes in their behavior with plurals and

donkey sentences. Whether these claims are warranted or not is not directly related to the focus of this paper, and we therefore do not discuss them here. Cruse notes a difference between "total" and "partial" adjectives in their compatibility with modifiers such as *almost, half* and *nearly*. This paper elaborates on some of Cruse's observations and offers a formal account of them.

We claim that the intuitive distinction between partial and total adjectives also manifests itself in truth-conditions. Consider for instance the following contrasts.

(3) a. The glass is dirty but it is almost clean.
b. #The glass is clean but it is almost dirty.

(4) a. The towel is wet but it is almost dry.
b. #The towel is dry but it is almost wet.

(5) a. The work is incomplete but it is almost complete.
b. #The work is complete but it is almost incomplete.

Under normal contexts, the *a* sentences are acceptable and contingent whereas the *b* sentences are unacceptable. We point out two factors that are responsible for this contrast:

(i) When *T* and *P* are a pair of antonymous total-partial adjectives, the phrase *almost T* is invariably acceptable whereas *almost P* is acceptable only under certain restrictions (see below);

(ii) Even when *almost P* is acceptable it stands in contradiction to *T*. By contrast, *almost T* does not contradict *P*.

For example, the phrases *almost dirty/wet* are often less acceptable than *almost clean/dry*. However, even when a statement *x is almost dirty/wet* is acceptable, it contradicts the statement *x is clean/dry*.

A related observation about the acceptability of *almost P* is the following:

(iii) The phrase *almost P* is acceptable if and only if *neither P nor T* is contingent.

For instance, if *x is almost wet* then it is neither wet (by co-restrictiveness of *almost*) nor dry (cf. (ii)). Conversely – with *T-P* pairs such as *dead-alive* or *complete-incomplete*, where an entity that is neither *T* nor *P* is hard to imagine, the phrases *almost alive* and *almost incomplete* are invariably odd. A similar distinction concerning the distribution of *almost* and related items is made in Morzycki ([4]).

3 Scale Structure and Higher-Order Modification

To capture these observations, we will make some assumptions regarding the scale structure of adjectives and modifiers. Some of the assumptions are standard and some are new and specific to T/P pairs.

The first assumption we make is that an adjective A denotes a subset of entities in a *scale* S_A (e.g. of real numbers), fully ordered by an asymmetric relation R_A. This assumption is similar to those made in previous works (e.g. [2,3,5]), and it is motivated by phrases such as *two meters tall(er)/shorter* and *two years old(er)/younger*, where a measure phrase modifies an absolute adjective or a comparative. The actual subinterval of a scale that an adjective A denotes is determined by a *standard value* d_A in the scale. The adjective denotes the subinterval $\{x \in S_A : R_A(x, d_A) \vee x = d_A\}$. The standard value is contextually determined. This is also assumed in previous works and is motivated by the vagueness of adjectives. For instance: a small building may be a lot bigger than a big elephant, so obviously, the standard value for *big* is different in the context of elephants than it is in the context of buildings.

If the adjectives A_1 and A_2 are antonyms, then R_{A_1} is $R_{A_2}^{-1}$, the inverse relation to R_{A_2}. This is motivated by equivalences such as the following.

(6) John is taller/less short than Mary $\Leftrightarrow$ Mary is shorter/less tall than John

In addition to those standard assumptions, we make some new assumptions specific to *total/partial* adjective pairs. When T/P is a pair of antonymous total/partial adjectives, we define their respective scales to be:

(7) $S_T = [-C, C]$ $\quad S_P = (0, \infty)$

The "boundedness" of total adjectives is motivated by their behavior with *completely* (cf. [3]). For instance, *"The red towel is completely dry and the green towel is completely dry"* entails *"Both towels are equally dry"*. Similar entailments do not appear with partial adjectives. For instance, "*The red towel is completely wet and the green towel is completely wet*" does not entail "*Both towels are equally wet*".

Therefore, the subinterval of the scale that is denoted by the adjective T, is bounded by the point $-C$ which is the denotation of the expression *completely T*. By contrast, there is no upper bound for the scale of partial adjectives. Note that the upper bound of S_T is set to C only for symmetry sake and is not essential for the theory.

Our second new assumption concerns the standard values of total and partial adjectives. We propose that while the d_P standard of partial adjectives is contextually determined as in other theories, the d_T standard of total adjectives is fixed to zero in any context. This is motivated by the intuitive "definition" in

(1): total adjectives but not partial adjectives entail zero amount of the relevant property.

We can now define the semantics of the modifier *almost* in the proposed scale structure of adjectives. We assume that the denotation of *almost A* is a small open interval which is adjacent but disjoint to the denotation of the adjective A. This assumption is standard and is compatible with the intuitive meaning of *almost* with other categories. For example, observe the following entailments:

(8) *every student came; almost every student came very early* $\Rightarrow$ *every student or almost every student came early.*

(9) *The white box is almost full; the black box is full* $\Rightarrow$ *every box that is fuller than the white box and less full than the black box is either full or almost full.*

The assumptions that were made above are summarized in figure 1.

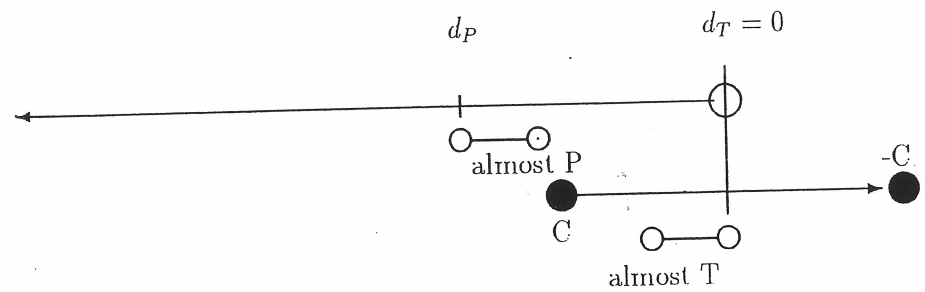


Figure 1: Total and Partial adjective scales

This scale structure directly accounts for observations (i)-(iii). For any total adjective T, the denotation of *almost T* is defined independently of contextual effects on the standard value d_T (which is invariably zero). Whether *almost T* contradicts P or not is contingent on the standard value d_P. By contrast, *almost P* is defined if and only if $d_P > 0$. This happens if and only if some object is neither P nor T. In such cases *almost P* contradicts T. When $d_P = 0$ (i.e. every object is either P or T), *almost P* is undefined, as there is no subinterval between d_P and 0.

4 Further Observations

In addition to the facts mentioned in the previous sections regarding *total/partial* pairs of adjectives and the modifier *almost*, there are some similar facts concerning those adjectives and related constructs. First, let us consider the modifier *slightly*.

The use of *slightly* with partial adjectives is always acceptable. However, its use with total adjectives is sometimes very odd. For example: *John is slightly dead*, or *the line is slightly straight* are rather strange sentences. In this case, as in the case of *almost*, we find that even when *slightly T* is acceptable, it is contingent with *P*, while *slightly P* contradicts *T*. Consider the following contrasts:

(10) a. I am still hungry but I am already slightly satiated.
b. #I am still satiated but I am already slightly hungry.

(11) a. The towel is still wet but it is already slightly dry.
b. #The towel is still dry but it is already slightly wet.

Another observation concerns the construction *except for*. Consider the following sentences:

(12) a. John is healthy except for an occasional flu.
b. #John is sick except for his healthy leg.

(13) a. The poem is complete except for the last stanza.
b. #The poem is incomplete except for the first three stanzas.

In this case as well, even if the sentences involving partial adjectives and *except for* are acceptable, the following contrast appears:

(14) a. The towel is wet except for its left corner $\Rightarrow$ The towel is not dry.
b. The towel is dry except for a few drops of water $\not\Rightarrow$ The towel is not wet.

The last observation to be discussed here is not about *total/partial* adjective pairs but about the comparative constructions *more ... than* and *as ... as*. These constructions behave in a way similar to this of *total/partial* pairs when used with *almost* and *slightly*.

The following sentences are acceptable:

(15) a. John is almost as old as Bill.
b. John is slightly older than Bill.

But the symmetric cases are unacceptable:

(16) a.# John is almost older than Bill.
b.# John is slightly as old as Bill.

Note that with other adjectives (e.g. *tall, short*), the contradiction might not be so obvious, but even then, "*John is almost taller than Bill*" has only a temporal interpretation, roughly meaning: "*John will soon be taller than Bill*". However, the sentence cannot be interpreted as meaning that John's height is almost greater than Bill's.

Acknowledgments

This research was partly supported by grant no. 1999210 ("Extensions and Implementations of Natural Logic") from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. The second author was also partly supported by the fund for the promotion of research at the Technion, research no. 120-042, and by an NWO grant for visiting the UiL OTS of the Utrecht University in summer 2001. We are grateful to Nissim Francez, Fred Landman, Tanya Reinhart and Susan Rothstein for discussion.

References

[1] Cruse, D.A. (1980). Antonyms and gradable complementaries. In Kastovsky D. ed., *Perspektiven der Lexikalischen Semantik*.
[2] Kennedy, C. (2000). Polar Opposition and the Ontology of 'Degrees'. *Linguistics and Philosophy* 24:33-70.
[3] Kennedy, C. and McNally, L. (1999). From event structure to scale structure. *Proceedings of SALT9*.
[4] Morzycki, M. (2001). *Almost* and Its Kin, Across Categories. *Proceedings of SALT11*.
[5] Seuren, P.A.M. (1978). The Structure and selection of Positive and Negative Gradable Adjectives. *Proceedings of CLS14*.
[6] Yoon, Y. (1996). Total and partial predicates and the weak and strong interpretations. *Natural Language Semantics* 4:217-236.

TOO

Rob van der Sandt & Bart Geurts
University of Nijmegen

Summary

We propose an account of the presuppositional contribution of *too* which deviates from the orthodox view in that we divide the standard presupposition in two: a descriptive part encoding information which must be matched by parallel information in the discourse context, and a hidden pronominal element. The two components are resolved independently and may link up to different levels of discourse structure. This yields a straightforward account of Kripke's observation that *too* seems to give rise to a more specific presupposition than is predicted by the standard view. The analysis also enables us explain why the presupposition of *too* is reluctant to accommodate, may access positions that are otherwise inaccessible, and gives rise to fully transparent readings in attitude contexts.

Three puzzles about 'too'

According to the standard view (e.g. Karttunen and Peters 1979) *too* contributes to (1a) by inducing the presupposition that there is someone other than Mary who lives in London. (1a) is thus said to presuppose (1b):

(1) a. $[\text{Mary}]_F$ lives in London too.

b. $\exists x\,[x \neq \text{Mary} \wedge \text{live\_in\_London}(x)]$

The standard accounts of projection moreover require that this presupposition should be given in the linguistic or extralinguistic context. If not, a cooperative speaker will revise the context of utterance, and accommodate the required information so as to license the presupposition. Kripke (as reported in Soames 1989 and Kripke ms.) noted that the standard presupposition is too weak for licensing *too*. Since we all know that many people live in London, the presupposition of (1a) is trivially fulfilled in any normal context. So we would expect (1a) to be fine. However, when uttered out of the blue this sentence sounds odd and calls for further identification of the person in question. Simply accommodating the information that there is someone other than Mary living in London will not do. Moreover, if the presupposition is satisfied by another clause, as in (2a) and should therefore be expected to be neutralized, we still conclude that the speaker holds (2b) to be true. This suggests that *too* does contribute to the semantic content, though not in the way predicted by the standard view.

(2) a. If Herb comes to the party, $[\text{the boss}]_F$ comes, too.

b. Herb is not the boss.

In view of these facts Kripke concluded that the standard view is wrong, that the presupposition of *too* contains an anaphoric element and that this presupposition arises from the anaphoric requirement that when someone uses

too he refers to some parallel information that is either in another clause or in the context. Furthermore, according to Kripke, the consequent of (2a) does *not* presuppose that someone other than Herb comes, but rather gives rise to the more substantial presupposition that Herb is not the boss.

The behaviour of *too* is remarkable in further respects, as well. Zeevat (1992, 2000) notes that this particle tends to access positions that are otherwise inaccessible:

(3) A. Harry may well have dinner in New York.

B. John is having dinner in New York, too.

Here the presupposition is licensed, apparently, by material in the subordinated, hence inaccessible, context.

Fauconnier (1985) and Heim (1992) draw attention to the remarkable behaviour of *too* and *also* in attitude contexts. When embedded in an attitude context the presupposition triggered by these particles may be interpreted as fully independent of the belief of the subject of the attitude. Heim imagines two children talking to each other on the phone:

(4) John: I am already in bed.

Mary: My parents think I am also in bed.

There is no suggestion here that Mary's parents share the information she just got. They may not even have any beliefs about John. Thus the presupposition is interpreted as fully independent of the parents' beliefs.

The anaphoric requirement

Though we agree with Kripke's observations we will not adopt his analysis. Following the mainstream of the presupposition literature we strictly distinguish between the presupposition induced and the result of the resolution process. We moreover adopt the anaphoric account of presupposition we have developed in earlier work (e.g. van der Sandt 1992, Geurts 1999). The main tenet of this theory is that so-called presupposition triggers are in fact anaphoric expressions that are just like pronouns in that they have to be bound to some accessible antecedent. If they cannot be bound, the descriptive content of the anaphoric expression will be accommodated at some suitable position which is accessible from the position where it is induced. The process of resolving presuppositional anaphors by binding or accommodation is subject to a number of constraints. Resolution of the presuppositional expression should respect the conditions on accessibility and not violate Gricean conditions on discourse acceptability. Ceteris paribus global accommodation, that is accommodation at the main level of discourse structure, is preferred to accommodation at subordinate levels.

In this framework the presupposition that is usually taken to be triggered by (2a) would come out as ∂[x: come(x), x $\neq$ y] (here we adopt Beaver's notation for marking presuppositional information). Resolution of this expression according to the standard algorithm yields wrong results for Kripke's examples. Assuming that the presupposition associated with the proper name has been dealt with, the initial representation of (2a) is as follows:

(5) [x, y: Herb(x), boss(y),
[: comes(x)] $\Rightarrow$ [: comes(y), ∂[z: comes(z), z $\neq$ y]]]

We might try to resolve the presuppositional expression at the main level, but this possibility, which would yield an incorrect reading, is excluded on the grounds of pragmatic infelicity. Thus the only alternative is to link the presuppositional expression to the parallel information in the antecedent, which yields after appropriate substitutions:

(6) [x, y: Herb(x), boss(y), [: comes(x), x $\neq$ y] $\Rightarrow$ [: comes(y)]]

This interpretation is too weak and does not account for the inference that Herb is not the boss.

The presupposition of 'too'

In order to account for Kripke's observations we propose an encoding of the presupposition of *too* which remains quite close to the orthodox view: in the presuppositional structure of TOO φ(a) we distinguish between a descriptive condition φ(x), and a pronominal part consisting of an anaphoric variable with the condition x $\neq$ a. The full presuppositional expression then comes out as ∂[: φ(x), ∂[x: x $\neq$ a]], where x is free variable which must be resolved in context. Note that the pronominal part is embedded in a larger presuppositional expression.

Assuming that the presuppositions of the proper name and the description have been taken care of the representation of (2a) is as follows:

(7) [x, y: Herb(x), boss(y),
[comes(x)] $\Rightarrow$ [: comes(y), ∂[: comes(z), ∂[z: z $\neq$ y]]]]

Resolution proceeds in the standard way. We first resolve the most deeply embedded anaphoric expression. In the present case this is the pronominal part of the presuppositional expression. After making the appropriate substitutions this yields:

(8) [x, y: Herb(x), boss(y), x $\neq$ y,
[: comes(x)] $\Rightarrow$ [: comes(y), ∂[: comes(z)]]]

Next the descriptive condition is resolved to parallel information in the antecedent, which gives (9) as the final representation:

(9) [x, y: Herb(x), boss(y), x $\neq$ y, [: comes(x)] $\Rightarrow$ [: comes(y)]]

This analysis deviates from Kripke's in one crucial respect but accounts for his observations. We deviate from Kripke's proposal in that we assign a presuppositional expression to the consequent of the conditional which is essentially the same as the standard presupposition, except that it is divided into two parts. As a result of splitting the presuppositional expression into a pronominal and a descriptive part, the presuppositional components may pick up their antecedents from different levels of representation. In the present case this results in a representation which entails what Kripke takes to be presupposed. The resolution of the pronominal part accounts for the inference that Herb is not the boss. The descriptive part of the presuppositional expression is absorbed in the antecedent of the conditional.

Whereas the pronominal element needs an antecedent, the descriptive condition can be resolved by way of accommodation. Therefore (10b), the representation of (10a), resolves to (10c):

(10) a. Either the boss will stay away from the party, or [John]$_F$ will come, too.
b. [x, y: boss(x), John(y),
[: stay_away(x)] $\vee$ [: comes(y), ∂[: comes(z), ∂[z: z $\neq$ y]]]]
c. [x, y: boss(x), John(y), x $\neq$ y,
[: stay_away(x)] $\vee$ [: comes(x), comes(y)]]

The pronominal expression is bound at top level. However, the descriptive condition cannot be resolved there in view of the infelicity of the resulting interpretation. This forces accommodation in the second disjunct, and yields (10c). This example also shows that, the received view notwithstanding, the presupposition of *too* contributes to the semantic content of the carrier sentence (cf. e.g. Stalnaker 1974, Karttunen and Peters 1979, Zeevat to appear). Not only does it enforce non-identity between the focused constituent and the antecedent of its pronominal component, but when accommodated it may also affect the semantic content in a more substantial way. Thus the second disjunct of (10) requires for its truth that both John and the boss will come.

Following Kripke, a number of authors (Asher and Lascarides 1998, Beaver 1997, Heim 1992, Van Rooy 1997, Zeevat 1992, to appear) have pointed out that the presuppositions of *too* are on a par with those of pronouns in that they demand an antecedent, and cannot be construed by way of accommodation. This peculiarity, which *too* has in common with *again* and other focus adverbs, proves difficult to explain on the standard view, which holds that resistance to accommodate is due to lack of semantic content. An alternative account is found in Zeevat (to appear) who claims that *too* is semantically redundant but obligatory in contexts that provide an antecedent for the presuppositional expression. Appealing to a result in Blutner's (2000) bi-directional optimality theory, he argues that if a triggering environment has a simple non-triggering expression alternative with the same meaning, it does not accommodate. As pointed out with respect to example (10), we reject the claim that *too* is semantically inane. On the account we endorse, the presupposition triggered by *too* contains a pronominal element, and this explains why *too* requires an explicit antecedent.

Inaccessible antecedents and transparent readings

When we inspect the presuppositional frame of *too* we observe a crucial difference between this presupposition and most other types of presupposition inducers. Consider the encoding of a definite description. The initial structure generated for 'The ψ is φ' is $[: \varphi(x), \partial[x: \psi(x)]]$. Here the anaphoric variable recurs in a condition in the matrix sentence. After projection of the presuppositional expression, the anaphor thus has to bind the variable in the non-presuppositional condition $\varphi(x)$. The presuppositional frame of *too* on the other hand does not share the anaphoric marker with any condition of its inducing matrix. The general format is: $[x: \varphi(x), \partial[: \psi(y), \partial[y: y \neq x]]$. Thus, after resolution of the presupposition the anaphoric variable will not enter in a binding relation with any condition in the matrix where it originated.

This gives an explanation for Zeevat's observation that the presupposition of *too* may access formally inaccessible antecedents. Clearly definite descriptions, pronouns and other presupposition inducers do not allow such antecedents. And the semantics of DRT provides a straightforward explanation, since the resulting DRS would not express a determinate proposition. In the case of descriptions or pronouns any attempt to access the modally embedded context in (3) results in an uninterpretable structure. This does not hold for the presupposition of *too,* however. Resolution of (11a), the representation of (3) after some preprocessing, results in (11b):

(11) a. [x, y: Harry(x), John(y), maybe[: have_dinner(x)],
have_dinner(y), ∂[: dinner(z), ∂[z: z $\neq$ y]
b. [x, y: Harry(x), John(y), maybe[: have_dinner(x)], have_dinner(y)]

We thus find that depending on their internal structure and the way they share their variables with their inducing matrix, different presupposition triggers may have different anaphoric properties. Which positions a presupposition can access depends on the interlinking between the anaphoric variable and the conditions it is intended to bind.

The same line of argument accounts for Heim's observation that the presupposition of (4) may be read as being fully independent of the beliefs of Mary's parents. In order to see this we should keep in mind that the binding theory of presupposition does not allow for copying routines. Presuppositional anaphors are created in situ, but will in the course of the

resolution process act as entities in their own right and be interpreted at the site where they are resolved. Descriptive information will land at the position where the anaphoric variable links up to or creates its antecedent. If a presupposition is induced in the scope of an attitude verb and does not find an antecedent in an accessible embedded context, it will project out to the main context. In the case of definite descriptions this yields a *de re* construal. The presuppositional material is interpreted externally and enables an anaphoric link to the content of the attitude (see Geurts 1999 for details). Thus (12b), which is the preliminary representation of (12a), resolves to (12c):

(12) a. Harry believes that the dog is hungry
b. [: Harry_believes: [∂[y: dog(y)], hungry(y)]]
c. [y: dog(y), Harry_believes : [hungry(y)]]

Here the marker correlated with *the dog* has a link into the content of the attitude, but it is the speaker, not the subject of the attitude, who is committed to the content of the description. When processing the presupposition of *too* we observe a different situation. Since the presupposition of *too* does not share its anaphoric variable with the inducing matrix, the content of the attitude is simply divorced from the presuppositional content, if the latter projects out.

References

Asher, N. and A. Lascarides, 1998: The semantics and pragmatics of presupposition. *Journal of Semantics* 15: 239-300.

Beaver, D. 1995: *Presupposition and Assertion in Dynamic Semantics*. PhD Thesis. University of Edinburgh, Edinburgh.

Beaver, D.I., 1997: Presupposition. In: J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*. Elsevier Science, 939-1008.

Blutner, R., 2000: Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17: 189-216.

Fauconnier, G. 1985: *Mental Spaces.* MIT Press, Cambridge, Mass.

Geurts, B. 1999: *Presuppositions and Pronouns*. Elsevier, Oxford.

Heim, I. 1992: Presupposition and the semantics of attitude verbs. *Journal of Semantics* 9: 183-221.

Karttunen, L. and S. Peters, 1979: Conventional implicature. In: C.-K. Oh and D. Dinneen (eds.): *Syntax and Semantics 11: Presupposition.* Academic Press, New York. 1-56.

Kripke, S. ms.: Remarks on the formulation of the projection problem. Manuscript, Princeton University.

Rooy, R. van 1997: *Attitudes and changing contexts*. PhD Thesis, Universität Stuttgart. Revised version to appear with Elsevier.

Sandt, R.A. van der 1992: Presupposition projection as anaphora resolution. *Journal of Semantics* 9: 333-377.

Stalnaker 1974: Pragmatic presuppositions. In: M. Munitz and P. Unger (eds.) *Semantics and Philosophy*, New York University Press, New York 197-214.

Soames, S. 1989: Presupposition. In: D. Gabbay and F. Guenthner (eds.): *Handbook of Philosophical Logic. Vol. IV. Topics in the Philosophy of Language*. Kluwer, Dordrecht, 553-616.

Zeevat, H. 1992: Presupposition and accommodation in update semantics. *Journal of Semantics* 9: 379-412.

Zeevat, H. to appear: Explaining presupposition triggers. In: K. van Deemter e.a. (eds.) *Information Sharing*. CSLI Publications, Stanford.

A Referential Analysis of Conditionals

Philippe Schlenker (USC & Institut Jean-Nicod)

0 Introduction

Lewis 1973 observed that his logic for counterfactuals could be applied to definite descriptions. His generalization was that the latter display the same non-monotonic behavior as the former, contrary to what Russellian or Strawsonian accounts would predict:

(1) a. If Otto had come, it would have been a lively party; but if both Otto and Anna had come it would have been a dreary party; but if Waldo had come as well, it would have been lively; but...
b. If O, L; but if O & A, ¬L; but if O & A & W, L; but ... [Lewis 1973]

(2) a. [Walking past a piggery] The pig is grunting; but the pig with floppy ears is not grunting; but the spotted pig with floppy ears is grunting; but ...
b. The P G; but the (P& F) ¬G; but the (P & F & S) G; but ... [Lewis 1973]

We take Lewis's observation literally and suggest that 'if' should be analyzed as the word 'the' applied to a description of possible worlds. We do not follow Lewis's own implementation, however, because it entails that definite descriptions do not refer. Rather, we use a combination of von Heusinger's Choice Function analysis of definiteness and of Stalnaker's Selection Function theory of *if*-clauses to analyze both 'if p' and 'the P' as referential terms: 'if p' denotes the most highly ranked p-world(s) under a measure of *similarity*, while 'the P' denotes the most highly ranked p-world(s) under a measure of *salience*. This analysis (i) reduces the non-monotonicity of conditionals to that of definite descriptions. In addition, it explains (ii) why *if*-clauses, like definites, may be topicalized; (iii) how a world pronoun ('then') can be coreferential with an *if*-clause; and it predicts (iv) that there should be syntactic constraints on coreference between world pronouns and their antecedents ('Condition C effects') which have not been hitherto described. Finally, (v) it allows us to analyze the indicative/subjunctive distinction in terms of a general system of referential classification, which appears to hold of reference to individuals, times and possible worlds alike.

1 Choice Functions and Selection Functions

Stalnaker's Conditions: Stalnaker 1968 analyzed 'if p, q' in terms of *Selection Functions* which were supposed to pick out the *closest* p-world from a given context. On his analysis, 'If p, q' is true in c iff $f(c, [[p]]) \in q$, where f is a function that satisfies the conditions in (3). We show that these conditions can be applied to von Heusinger's Choice Functions when these are used to model definiteness, and are thus taken to select the *most salient* element of a set:

(3) a. $\forall c\ \forall E\ f(c, E) \in E$
[Stalnaker's Condition 1]
b. $\forall c\ \forall E\ \forall E'\ \ [f(c, E') \in E\ \&\ f(c, E) \in E' \Rightarrow f(c, E)=f(c, E')]$
[Stalnaker's Condition 4]
b'. $\forall c\ \forall E\ \forall E'\ [E' \subseteq E\ \&\ f(c, E) \in E' \Rightarrow f(c, E')= f(c, E)]$
c. $\forall c\ \forall E\ [f(c, E)=\lambda \text{ iff } E=\emptyset]$
[Stalnaker's Condition 2]
d. $\forall c\ \forall E\ [c \in E \Rightarrow f(c, E)=c]$
[Stalnaker's Condition 3]

(i) (3)a is simply the definition of a Choice Function.
(ii) Under (3)a, (3)b is equivalent to (3)b', which can be interpreted in von Heusinger's terms as requiring that an element which is the most salient from a large set E should also count as the most salient from a subset E' if it belongs to E' (this condition is not discussed in the Choice Function literature but is clearly necessary to capture the notion of maximal salience).
(iii) In Stalnaker's system, λ is an impossible world in which every proposition is true. (3)c ensures that a conditional with an impossible antecedent is always true. But on the referential analysis we may re-interpret λ as denoting referential failure - not an unreasonable assumption given the deviance of conditionals with a *clearly* impossible antecedent (e.g. #If John were here and weren't here, Mary would be happy).
(iv) (3)d requires that the element selected out of a set E given a point of reference c be c itself if it belongs to E. When the point of reference is the world of utterance w*, this requires - correctly - that 'if p' denote w* if w* satisfies p. The condition is implausible for definite descriptions *if* the individual of reference is taken to be the speaker; for it would predict that 'the man in the white shirt' should of necessity denote *me* if I happen to be

wearing a white shirt - an overly strong prediction. However this problem arises only if the point of reference is taken to necessarily be the speaker. We suggest that this need not be the case. Centering is still useful, however, in cases such as the following (note that the ι-operator carries a subscript x which represents the point of reference):

(4) a. (In this school...) No class was so bad that the teacher abandoned the class.

b. [No x: class(x)] abandoned($[\iota_x y$ teacher(y)], ($[\iota_x z$ class(z)])

Here '$[\iota_x y$ teacher(y)]' selects the teacher that's most salient from the standpoint of a particular class x, and as desired this yields different teachers for different values of x. Furthermore, by Centering, '$[\iota_x z$ class(z)]' must have the same value as 'x' whenever 'x' denotes a class. This correctly derives the anaphoric reading of this definite description.

Adding Plurality: In order to address the arguments in Lewis 1973 against 'Stalnaker's Assumption' (=the assumption that a single world is selected), we treat *if*-clauses as *plural* descriptions (following a suggestion made in Schein 2001). This allows *if*-clauses to restrict generalized quantifiers over worlds (e.g. 'Necessarily/Probably, if p, q'), as in Kratzer 1991. But the construction is now analogous to partitives ('All *the students* came') rather than to simple quantifications ('Each student came'). The result, entirely derived from the semantics of definiteness and plurality, is a version of Lewis's system *with* the 'Limit Assumption' (=the assumption that a set of closest worlds can always be selected) but without 'Stalnaker's Assumption'.

2 Reference and Coreference

The referential analysis of *if*-clauses predicts that the latter should share other aspects of the behavior of definite descriptions. This leads to some welcome predictions:

Topicalization: As other definites (and unlike quantificational elements or mere restrictors), *if*-clauses can be topicalized and left-dislocated [Mary will die if you wait→If you wait, Mary will die;cf. The President, I like/I like him] (see Bhatt & Pancheva 2001)

'Then': When topicalized, *if*-clauses can be doubled by a world pronoun, the word 'then' (Iatridou 1994, Izvorski 1996). As noted in Iatridou 1994, this explains fine distinctions between sentences with and without 'then':

(5) a. If John is dead or alive, Bill will find him.

b. #If John is dead or alive, then Bill will find him.

(6) a. Les étudiants, eux ont compris [French]

The students, them-strong have understood

b. Assertion: f(c, [[students]])$\subseteq$[[understood]]

Implicature: $\exists$x (x$\notin$ f(c, [[students]]) & x$\notin$ [[understood]])

Iatridou suggests that 'if p, then q' should be analogized to (the German version of) (6)a, which carries an implicature that some non-students didn't understand. Similarly 'if p, then q' carries an implicature that some non-p worlds are non-q worlds, which accounts for the deviance of (5)b (since the antecedent is a tautology, there are no non-p worlds). Iatridou notes some counterexamples to her analysis, but these can be avoided by treating 'then' as a *strong* pronoun which, in this position, is necessarily focused. Furthermore, the implicature can then be derived from the semantics of focus (roughly, the implicature is that some propositions in the focus value of *[them]$_F$ have-understood*, i.e. {p: $\exists$X p=λw understood(w, X)}, should be false).

Condition C effects: Since *if*-clauses are now treated as referential expressions and 'then' is analyzed as a world pronoun, we expect syntactic constraints on coreference between a pronoun and a referential expression to apply in this case as well. One such constraint is the Condition C of Chomsky's Binding Theory, which prohibits a referential expression from being in the scope of ('c-commanded by') a coindexed pronoun (e.g. #He$_i$ likes Peter$_i$'s friends vs. ^{Ok}Peter$_i$ likes his$_i$ friends / ^{Ok}His$_i$ mother likes Peter$_i$'s friends). We display new examples that suggest that this somewhat surprising prediction is in fact borne out:

(7) a. [If it were sunny right now]$_i$ I would see [people who would then$_i$ be getting sunburned].

b. *I would then$_i$ see [people who would be getting sunburned [if it were sunny right now]$_i$].

c. Because I would then$_i$ hear lots of people playing on the beach, I would be unhappy [if it were sunny right now]$_i$

Although backwards anaphora with 'then' is possible, as shown by (7)c, (7)b is still deviant. This is presumably because 'then' does not merely precede the *if*-clause but also c-commands it. Thus *if*-clauses, as other referential expressions, must satisfy Condition C of the syntactican's

Binding Theory. This extends to Condition C the suggestion made in Percus 2000 that binding-theoretic conditions apply beyond individual-denoting expressions.

3 Referential Classification

We further suggest (tentatively) that the distinction between indicative and subjunctive conditionals should be seen as a special case of a general mechanism of referential classification, which applies to individuals (e.g. this/that), times and worlds (cf. Iatridou 1998 for the time/world analogy). **'This' vs. 'that':** Consider first the difference between 'this' and 'that'. 'This' incorporates a presupposition that its denotation should count as close to the speaker. We may be tempted to define the opposite presupposition for 'that', i.e. that 'that' may *not* denote an entity which is close to the speaker. This would be overly strong, however. Observing a scene in the mirror, I may find out that a table I was watching is in fact the table standing right next to me. I could then utter without presupposition failure: 'That is this!'. The presupposition isn't that 'that' *cannot* refer to an entity near the speaker, but rather that there is a *possibility* that it refers to an entity which is far away. A similar problem arises with pronouns: 'He is me!' uttered by a man who is observing someone in the mirror should not come out as a presupposition failure. The (weak) presupposition of 'he' the pronoun *might* refer to someone other than the speaker.

How can this notion of 'possibility' be analyzed? By evaluating sentences with respect to sets of assignments akin to Dekker 2000's information states We define for 'this' a predicate 'LOCAL' which requires that a term refer to an entity near the speaker; and for 'that' a predicate '<LOCAL' which requires that under *some* elements of an information state a term refer to an entity away from the speaker. We may thus define the following rules (note that an element s of an information state S has a distinguished position for the context of speech, so that the notion 'context of s' is well-defined).

(8) a. $[[\tau\{LOCAL\}]]^{s,S}\neq\#$ iff $\forall s'\in S$ $[[\tau]]^{s',S}\neq\#$ and $\forall s'\in S$ $[[\tau]]^{s',S}$ is close to the context of s'. If $\neq\#$, $[[\tau\{LOCAL\}]]^{s,S}=[[\tau]]^{s,S}$
b. $[[\tau\{<LOCAL\}]]^{s,S}\neq\#$ iff $\forall s'\in S$ $[[\tau]]^{s',S}\neq\#$ and $\exists s'\in S$ $[[\tau]]^{s',S}$ is not close to the context of s'. If $\neq\#$, $[[\tau\{LOCAL\}]]^{s,S}=[[\tau]]^{s,S}$

An update semantics may then be defined for 'That is this', as follows:

(9) a. That$_i$ is this$_k$
b. $x_i\{<LOCAL\}=x_k\{LOCAL\}$
c. $S[b]\neq\#$ iff $\exists s'\in S$ $s'(x_i)$ is not close to the speaker of s' & $\forall s'\in S$ $s'(x_k)$ is close to the speaker of s'. If $\neq\#$, $S[b]=\{s\in S: s(i)=s(k)\}$

Indicative vs. Subjunctive: With this background in mind, we can apply the same strategy to conditionals, claiming with Stalnaker 1975 that an indicative *if*-clause *must* select a world within the Common Ground, construed as the set of worlds that are 'close enough' to the actual world; while a subjunctive *if*-clause *might* (but need not) select a world outside the Common Ground. We may in fact use our predicates 'LOCAL' and '<LOCAL' as defined above, with the provision that a world is 'close to' a context iff it is in the Common Ground of that context. For the examples below, this results in potentially weak presuppositions for a subjunctive conditional *if p, q*, namely that for *some* element s of S, $f(s(w_0), [[p]])$ lies outside the Common Ground. Rules of interpretation are then as follows:

(10) a. If John is sick, Mary is unhappy
b. unhappy(Mary, $[\iota_{w_0} w_i\, sick(J,w_i)]\{LOCAL\}$)
c. $S[b]\neq\#$ iff $\forall s\in S$ $f(s(w_0), \{w\in W$: Mary is sick in $w\})$ is in the Common Ground of s. If $\neq\#$, $S[b]=\{s\in S$: Mary is unhappy in $f(s(w_0), \{w\in W$: Mary is sick in $w\})\}$

(11) a. If John were sick, Mary would be unhappy
b. unhappy(Mary, $[\iota_{w_0} w_i\, sick(J, w_i)]\{<LOCAL\}$)
c. $S[b]\neq\#$ iff $\exists s\in S$ $f(s(w_0), \{w\in W$: Mary is sick in $w\})$ is outside the Common Ground of s. If $\neq\#$, $S[b]=\{s\in S$: Mary is unhappy in $f(s(w_0), \{w\in W$: Mary is sick in $w\})\}$

References: Bhatt & Pancheva (2001) 'The Syntax of Conditionals'; von Heusinger (2000) 'Reference and Salience' ; Iatridou (1994) 'On the Contribution of Conditional *Then*'; Iatridou (1998) 'The Grammatical Ingredients of Counterfactuality'; Izvorski (1996) 'The Syntax and Semantics of Correlative Proforms'; Kratzer (1991) 'Conditionals'; Lewis (1973) *Counterfactuals;* Lewis (1979) 'Scorekeeping in a Language Game'; Percus (2000) 'Constraints on Some Other Variables in Syntax'; Schein, Barry (2001) 'Adverbial, Descriptive Reciprocals'; Stalnaker (1968) 'A Theory of Conditionals'; Stalnaker (1975) 'Indicative Conditionals'.

Two kinds of long-distance indefinites*

Bernhard Schwarz
The University of Texas at Austin

1. Introduction

Indefinites can often be interpreted as if they had scoped from a syntactic island. For example, (1) has a *long(-distance) intermediate scope* reading, expressed by the LF in (2), where *some teacher* escapes a relative clause.

(1) Every student read every book some teacher had praised.
(2) [every student] λ_1[[some teacher] λ_2[t_1 read every book t_2 praised]]

It has been proposed that such long scope shifts can be eliminated if it is assumed that *some* and *a* can be variables ranging over (Skolemized) choice functions (Reinhart 1997, Winter 1998, Kratzer 1998, Matthewson 1999, Chierchia 2001). In (3), *some* translates as a choice function variable which is existentially closed at an intermediate position (see Reinhart, Winter). In (4), *some* translates as a Skolemized choice function variable whose existential closure takes widest scope (see Matthewsson, Chierchia).

(3) [every student] λ_1[$\exists f$[t_1 read every book [f teacher] praised]]
(4) $\exists f$[[every student] λ_1[t_1 read every book [f_1 teacher] praised]]

(3) and (4), which are equivalent to each other, are also equivalent to (2).[1] This points to the two distinct choice function accounts of long scope described in (5) and (6) (cf. Chierchia).

(5) "multiple choice" analysis
    Indefinite articles can be choice function variables, bound by freely distributed existential closure (see (3)).
(6) "∃ sloppy choice" analysis
    Indefinite articles can be Skolemized choice function variables, bound by existential closure which takes widest scope (see (4)).

Building on observations due to Chierchia, this paper shows that (5) and (6) both undergenerate and overgenerate. Both analyses generate unattested readings for indefinites in non-upward monotone contexts. And

*For discussion and comments, I owe thanks to Fabian Heck, Peter Krause, Wolfgang Sternefeld, and especially, Hans Kamp and Agnes Bende-Farcas. Of course, none of these people should be blamed for any mistakes or shortcomings.

[1] Actually, (3) and (4) are equivalent to (2) only if we take the set of teachers to be non-empty. For the sake of the argument, let us assume that this assumption can be justified. The same sort of comment applies to all the other relevant examples below.

neither analysis accounts for the scope behavior of long-distance indefinites with *a certain*. These call for a sloppy choice function analysis where the function variable remains free (Kratzer).

2. Monotonicity and overgeneration

Chierchia observes that ∃ sloppy choice fails to derive long intermediate readings when the higher DP is not upward monotone. For example, (7) can be read as the negation of (1) in its long intermediate reading.

(7) Not every student read every book some teacher had praised.
(8) [not every student] λ_1[$\exists f$[t_1 read every book [f teacher] praised]]
(9) $\exists f$[[not every student] λ_1[t_1 read every book [f_1 teacher] praised]]

The LF (8) accounts for this reading, but (9) is too weak. For suppose Smith and Baker are the teachers, Mary and Sue are the students, both Sue and Mary read every book Smith praised, but only Sue read every book Baker praised. In this scenario, the relevant reading of (7) is judged false, yet (9) is true. For we can find a Skolemized choice function f such that f(Mary)(the teachers) = Baker, which verifies (9). With Chierchia, we conclude that the ∃ sloppy choice analysis undergenerates.

3. Monotonicity and overgeneration

To see the logic of ∃ sloppy choice more clearly, consider the simpler case in (10). In this case, ∃ sloppy choice yields a narrow scope existential reading. That is, (11a) is equivalent to (11b).

(10) Every student read a book I had praised.
(11) a. $\exists f$[[every student] λ_1[t_1 read f_1 [book I praised]]]
     b. [every student] λ_1[[some book I praised] λ_2[t_1 read t_2]]

Here is the sketch of an equivalence proof. We first consider the lambda abstracts in (12) and show that they relate as in (13). Then we exploit right upward monotonicity of *every* to prove the claim.

(12) a. λ_1[t_1 read f_1 [book I praised]]
     b. λ_1[[some book I praised] λ_2[t_1 read t_2]]

(13) a. For every g, if I praised any book, then $''(12a)\geq^{g} \subseteq ''(12b)\geq^{g}$
     b. For every g, if I praised any book, then for some Skolemized choice function f, $''(12b)\geq^{g} = ''(12a)\geq^{g^{f/f}}$.

This case study suggests that, more generally, (14) and (15) below are equivalent whenever δ is right upward monotone.[2]

(14) $\exists f[[\delta\ \alpha]\ \lambda_i[[f_i\ \beta]\ \gamma]]$

(15) $[\delta\ \alpha]\ \lambda_i[[\textit{some}\ \beta]\ \gamma]$

The ∃ sloppy choice analysis is in fact committed to this equivalence, as otherwise long intermediate readings could never be derived as intended. But, of course, we should now ask what happens if δ is not right upward monotone. Take (16), where *no* substitutes for *every* in (10) and (11).

(16) No student read a book I had praised.

(17) a. $\exists f$[[no student] λ_1[t_1 read f_1 [book I praised]]]

b. [no student] λ_1[[every book I praised] λ_2[t_1 read t_2]]

It turns out that in this case, ∃ sloppy choice interprets the indefinite as a narrow scope universal, that is, (17a) is equivalent to (17b)! The proof is analogous to the one on (11). First we show that (18a,b) relate as in (19). Then we exploit right downward monotonicity of *no* to prove the claim.

(18) a. λ_1[t_1 read f_1 [book I praised]]

b. λ_1[[every book I praised] λ_2[t_1 read t_2]]

(19) a. For every g, if I praised any book, then $''(18b)\geq^{g} \subseteq\ ''(18a)\geq^{g}$

b. For every g, if I praised at least one book, then for some Skolemized choice function f, $''(18b)\geq^{g} =\ ''(18a)\geq^{g^{f/f}}$.

More generally, if we insist on the equivalence of (14) and (15) for right upward monotone δ, we are also committed to the equivalence of (14) and (20) for right downward monotone δ.

(20) $[\delta\ \alpha]\ \lambda_i[[\textit{every}\ \beta]\ \gamma]$

Evidently, the latter equivalence is not at all welcome. Sentence (16) cannot mean that no student read every book I praised. So we need to somehow exclude (17a) as a LF for (16). Presumably, this means that LFs of the form (14) should be banned for any choice of δ.

4. Multiple choice and monotonicity

As it stands, ∃ sloppy choice both undergenerates (section 2) and overgenerates (section 3). What about the multiple choice analysis? We have not seen it undergenerate - it does derive the long intermediate reading of (7). Also, if we ban Skolemized choice function variables, it

2 Actually, (14) and (15) may certainly fail to be equivalent if there is a free occurrence of *f* in δ, α, β, or γ. But we can safely assume that there is no such occurrence.

does not derive any unattested reading for (16). Unfortunately, however, multiple choice overgenerates in other cases. First consider (21) below. If no two candidates wrote exactly the same papers (as seems plausible), the LF (22a) turns out equivalent to (22b), hence equivalent to (23).

(21) Every candidate submitted a paper he had written.

(22) a. $\exists f$[[every candidate] λ_1[t_1 submitted f [paper he_1 had written]]]

b. $\exists f$[[every candidate] λ_1[t_1 submitted f_1 [paper he_1 had written]]]

(23) [every candidate] λ_1[[some paper he_1 had written] λ_2[t_1 submitted t_2]]

This equivalence is welcome. However, turning to (24), it commits one to the unwelcome equivalence of (25a) and (26)!

(24) No candidate submitted a paper he had written.

(25) a. $\exists f$[[no candidate] λ_1[t_1 submitted f [paper he_1 had written]]]

b. $\exists f$[[no candidate] λ_1[t_1 submitted f_1 [paper he_1 had written]]]

(26) [no candidate] λ_1[[every paper he_1 had written] λ_2[t_1 submitted t_2]]

Thus ∃ closure must be restricted in its distribution. We need to stipulate that no operator can bind into a choice function indefinite from within the scope of its ∃ closure. Given the need for this constraint, call it *integrity condition*, one may conclude that long intermediate scope is after all better analyzed in terms of long distance scope shifts. Be this as it may, what we will see is that neither of these devices copes with *a certain* indefinites.

5. Functional indefinites

Winter (1998) observes that (27a) might be used to convey that every mother hating child will develop a complex. Standard scope shifting does not derive this *functional* reading. Winter credits it to the LF in (27b).

(27) a. Every child who hates a certain woman he knows will develop a serious complex.

b. $\exists f$[every [child λ_1[t_1 hates f_1 [woman he_1 knows]]] [... complex]]

Interestingly, (27b) is in conflict with our integrity condition. Does this condition have to be weakened? No. For (27b) is in fact weaker than any attested reading of (27a), including the functional one. *Every* being left downward monotone, (27b) says that every child who hates every woman he knows will develop a complex. While (27b) is thus not adequate, a minimal modification seems sufficient. Following Kratzer (1998), we omit ∃ closure and interpret *f* is a function the speaker has in mind, e.g. the mother function. This arguably accounts for the judgments on (27a).

6. Two kinds of long-distance indefinites

Kratzer proposed that functional construals are responsible for all long intermediate readings of indefinites (in extensional contexts). The data presented below challenge this view. More generally, they challenge the popular thesis that all long-distance indefinites (in extensional contexts) are subject to the same analysis (e.g. Reinhart, Kratzer, Winter). It seems true that functional indefinites can always give rise to something like long intermediate scope. Thus (28) can be judged true if every boy finished the cookies his mother brought, but none of those his sister brought.

(28) Every boy finished the cookies a certain woman he knows had brought.

But it seems that not all long intermediate indefinites can have functional readings. In its only sensible reading, (29a) says that for every boy, there is some person such that the boy ate all the cookies that person had brought. Thus, (29a) allows for intermediate scope of *someone*.

(29) a. Each boy ate all the cookies someone had brought.
b. Every boy who hates someone will develop a serious complex.

In (29b), *someone* might, perhaps, take widest scope, in which case it says that there is someone such that every boy who hates her will develop a complex. The indefinite may also take narrow scope, in which case (29b) says that every boy who hates anyone will develop a complex. But no third reading is available, in particular no functional reading of the sort found in (27a). Similar descriptions apply to the example pairs in (31) and (30).

(30) a. More than one boy devoured every cookie a girl from his class had brought.
b. Almost no boy invited a girl from his class.

(31) a. Most students have studied every article that some professor has published.
b. No student who some professor had invited showed up.

I conclude that not all long intermediate indefinites are functional. Moreover, so-called long intermediate readings with functional and non-functional indefinites can be shown to be rather different in nature. While both cases in (32) allow for a reading in which the indefinite has neither wide nor narrow ∃ scope, the two readings are not of the same kind.

(32) a. No boy finished the cookies someone had brought.
b. No boy finished the cookies a certain woman he knows had brought.

Suppose some boy finished the cookies a woman he knows had brought. This assumption makes (the relevant reading of) (32a) false whereas (the relevant reading of) (32b) may still be true. And the fact that no boy finished the cookies his mother had brought may be sufficient for the truth of (32b), but certainly not for the truth of (32a). Similar descriptions apply to the example pairs in (33) and (34).

(33) a. No student has studied every article that some professor has published.
b. No student has studied every article that a certain professor has published.

(34) a. At most one boy ate every cookie a girl from his class had brought.
b. At most one boy ate every cookie a certain girl from his class had brought.

Of course, these judgments are precisely what one should expect. Scope shifting derives the relevant reading of (32a) as shown in (35a). Free sloppy choice assigns (32b) the LF (35b).

(35) a. [no boy] λ_1[[someone] λ_2[t_1 finished the cookies t_2 had brought]]
b. [no boy] λ_1[[t_1 finished the cookies f_1 [woman he$_1$ knows] had brought]]

Given right downward monotonicity of *no*, (35b) is much weaker than (35a). This is why, in this case, the functional reading does not approximate a genuine intermediate scope reading as well as it does in (28).

References

Chierchia, Gennaro: 2001: 'A Puzzle about Indefinites', in Carlo Cecchetto, Gennaro Chierchia, and Maria Teresa Guasti (eds.), Semantic Interfaces: Reference, Anaphora, and Aspect, CSLI, Stanford.

Kratzer, Angelika: 1998, 'Scope or Pseudoscope? Are there Wide Scope Indefinites?', in S. Rothstein (ed.), Events in Grammar, Kluwer, Dordrecht, pp. 163-169.

Matthewson, Lisa: 1999: 'On the Interpretation of Wide-Scope Indefinites', Natural Language Semantics 7, 79-134.

Reinhart, Tanya: 1997: 'Quantifier scope: How labor is divided between QR and choice functions', Linguistics and Philosophy **20**, 335-397.

Winter, Yoad: 1998, Flexible Boolean Semantics, Doctoral Dissertation, Utrecht University

Existential Import

Pieter A.M. Seuren
Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

The present paper deals with the semantics of *all, some* and *not* in English and related languages. Since these semantics imply a logic, it also deals with the logic of quantification in natural language. To the extent that the linguistic findings are universally valid, the results are relevant to the study of human cognition in general.

Since standard quantification theory (SQT) replaced Aristotelian Predicate Calculus (APC), philosophers of language and semanticists have struggled with the fact that APC corresponds better with natural semantic intuitions than SQT. Pragmatics was called in to restore APC on non-logical, pragmatic grounds. The replacement of APC by SQT was motivated on the grounds that (a) APC suffers from improper existential import, and (b) SQT is a straightforward application of Boolean algebra, and as such preserves syllogistic reasoning on a mathematical basis.

Logically speaking, SQT differs from APC in that the Aristotelian subaltern entailments are abolished, which makes the Aristotelian Square collapse but saves the equivalences (conversions) between $\neg\forall\neg$ and $\exists$, and $\neg\exists\neg$ and $\forall$. In SQT, the quantifying predicate $\exists$ over pairs of sets <X, Y> yields truth iff $X \cap Y \neq \emptyset$ and $\forall$ does so iff $X \subseteq Y$. The author recently found that if the condition $X \neq \emptyset$ is added to the condition for $\forall$ and the conversions are changed into one-way entailments (i.e. $\forall x(Fx,Gx) \vdash \neg\exists x(Fx,\neg Gx)$ and $\forall x(Fx,\neg Gx) \vdash \neg\exists x(Fx,Gx)$, but not vice versa), APC is restored without improper existential import (only the subcontraries are lost). The Boolean basis is unaffected since whenever $\forall$ yields truth, it is still so that $X \subseteq Y$. The resulting revised Aristotelian predicate calculus (RAPC), is represented in the Hexagon of fig. 1 (arrows stand for entailments, 'C' for contraries, 'CD' for contradictories). Note that the traditional (Boethian) letter types **A**,**I**, **E** and **O** have been replaced with $\forall$, $\exists$, $\forall\neg$ and $\exists\neg$, respectively, since APC contains only the standard quantifiers $\forall$ and $\exists$, plus the standard negation $\neg$. (This answers the question, raised by Horn (1972, 1989:252-67) and Levinson (2000:69-71), of why the **O**-corner in APC is never lexicalised: there ís no **O**-corner!)

The revision of APC is easily shown by means of a valuation space interpretation (VSI) (Van Fraassen 1971). Let the valuation space (VS) of a sentence *A*, /*A*/, be the set of situations in a universe **U** of possible situations in which *A* is true. To say that *A* is true now amounts to saying that the actual situation $s_a \in /A/$. Clearly, $/\neg A/ = /A/$, $/A \wedge B/ = /A/ \cap /B/$, and $/A \vee B/ = /A/ \cup /B/$, and the whole of standard propositional calculus can be derived. APC can be rendered in VSI terms as in fig. 2, where each ring (circle) is marked for the VS of each Aristotelian sentence type. Given the fact that, as indicated in fig. 2, relations of entailment, contrariety,

subcontrariety and contradiction can be read from the VSI diagram, the whole of APC is represented in fig. 2.

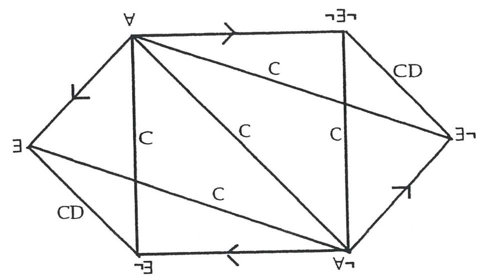
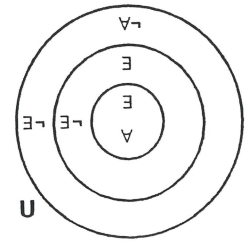


Figure 1



X entails Y iff /X/ is a subset of /Y/

X & Y are contraries iff the intersection of /X/ & /Y/ is empty

X & Y are subcontraries iff the union of /X/ & /Y/ equals **U**

X & Y are contradictories iff /X/ is the complement of /Y/ in **U**

Figure 2

APC, however, fails to take into account the set of situations where the F-class is empty. Therefore, **U** in fig. 2 is incomplete and must be extended with a further ring containing those situations in which there is no representative of the F-predicate, as shown in fig. 3a.

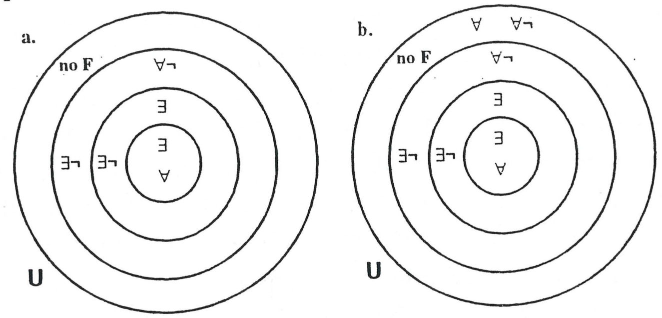


Figure 3

Now, however, the logical relations have changed. The subcontraries have gone and the conversions have been changed into one-way entailments as shown in the Hexagon of fig. 1. One notes that in fig. 3a, which represents RAPC, $\forall$, $\exists$, $\forall\neg$ and $\exists\neg$ are all false in cases where there is no F. SQT, on the contrary, declares both $\forall$ and $\forall\neg$ true in such cases, as shown in fig. 3b. This, in fact, is the only difference between RAPC and SQT.

From the point of view of natural intuitions, the loss of the subcontraries does not appear serious, as it is easy to see the simultaneous falsity of "Some F is G" and "Some F is not-G" for cases where the F-class is non-instantiated. The replacement of the equivalences by one-way entailments from the universal quantifier $\forall$ likewise looks empirically promising. The inference from "All students did not pass the exam" to "No student passed the exam" seems correct, whereas the inference from "No proof of the man's guilt has been found" to "All proofs of the man's guilt have not been found" does not. Similarly, "All students passed" seems to license

"No student did not pass", but not vice versa, since many speakers will judge the former false but the latter true in cases where there were no students. If these judgements are correct, the Hexagon, i.e. RAPC, corresponds even better to natural intuitions than the original APC.

Like SQT, however, RAPC still fails to account for intensional (i.e. imagined) entities and intensional predicates (i.e. predicates that do not require real existence of the argument term referent for truth). For example, a sentence like "Some gods are worshipped in that temple" may well be true without it being necessary to conclude that there exist real gods. For that reason, the quantificational calculus must be modified and generalised to account for intensional phenomena as well.

Since natural language refers to and quantifies over intensional objects in precisely the same way as it does with regard to extensional (really existing) objects, there appears to be a prima facie requirement that single theories of reference and of quantification should account for both the extensional and the intensional cases. This makes it mandatory to accept an ontology containing incompletely defined *intensional objects*, as proposed by the Austrian philosopher Alexius Meinong (1853-1920). In an intensional theory of quantification, the universe of individuals **I** must contain all really existing as well as all imagined entities (objects). The quantifiers are still higher order predicates over pairs of sets (generalized quantifiers). However, the restrictor set (R-set) is no longer the standard extension of the predicate Fx, [[Fx]], comprising the set of entities that satisfy Fx, but the *intensional extension* {Fx} or the set of entities that satisfy Fx plus those that satisfy Π(Fx) (where Π is an intensional predicate/operator). One notes that {Fx} cannot be empty, since whenever Fx is mentioned it has automatically been imagined. This makes it possible to remove the condition $X \neq \emptyset$ from the satisfaction conditions of ∀.

At this point the minimal, presupposition-preserving negation "~" must be defined. We take it that the satisfaction conditions of a predicate *P* are divided into two subsets, the *preconditions* and the *update conditions* (cp. Seuren at al. 2001). The former define the *presuppositions* of the propositional structure "Px", i.e. the conditions of contextual coherence ('discourse anchoring') for "Px"; the latter define the semantic contribution made by "Px" to the discourse at hand. Since an unanchored sentence lacks a truth value and does not express a proposition (as when I say to you now "The man was right after all", without any further explanation as to the identity of the man or the issue at hand), the preconditions are truth-conditional, not just pragmatic, as is widely held in pragmatic circles. For good functional reasons of coherent discourse, the normal default negation in natural language toggles between satisfaction and non-satisfaction of the update conditions, leaving the preconditions unaffected. This negation is called the *minimal negation*, represented by "~". In VSI terms, we say that for each sentence *A* there is a *subuniverse* of possible situations $\mathbf{U}_A$, where the presuppositions of *A* are true (cp. Seuren et al. 2001). If *A* has no

presuppositions (i.e. the predicate of *A* has no preconditions), $\mathbf{U}_A = \mathbf{U}$. We now say that for all sentences *A*, $/A/ \subseteq \mathbf{U}_A$ and $/{\sim}A/ \subseteq \mathbf{U}_A$, and $/{\sim}A/ = \mathbf{U}_A -$ $/A/$. There is also a *radical negation* ≃ , such that $/{\simeq}A/ = \mathbf{U} - \mathbf{U}_A$. This, however, is left out of consideration here. But note that $/{\sim}A/ \cup /{\simeq}A/ = /\neg A/$, as shown in fig. 4.

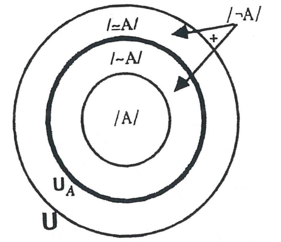


Figure 4

We now define [<Fx>], the *presuppositional extension* of Fx, as the set of entities that satisfy the preconditions of Fx, i.e. for which either Fx or ~Fx yields truth. Clearly, if Fx has no preconditions, [<Fx>] = **I** and {Fx} = [[Fx]]. We define the universal quantifier ∀ as taking the precondition that $\{Fx\} \cap [<Gx>] \neq \emptyset$, and the update condition that $\{Fx\} \cap [<Gx>] \subseteq [[Gx]]$. In other words, "All F is G" is true iff for all $e \in \{Fx\} \cap [<Gx>]$ (i.e. $e \in \{Fx\}$ qualifies for the predicate Gx), $e \in [[Gx]]$. For example, "All Englishmen are rich" is true iff all members of {Englishman(x)} that qualify for the predicate "rich" are indeed rich. Since the predicate "rich" has a precondition of existence for its subject term, the class of imaginary Englishmen is automatically excluded from consideration. On the other hand, "All unicorns are imaginary" is true in this world, since the predicate "imaginary" has no existential precondition, so that [<imaginary(x)>] = **I**. Since $\{unicorn(x)\} \cap \mathbf{I} = \{unicorn(x)\}$, it is sufficient that $\{unicorn(x)\} \subseteq [[imaginary(x)]]$, which is the case. Existential import is thus taken away from the existential quantifier and placed in the G-predicate. "Some Englishmen are rich" now entails the existence of Englishmen, but "Some gods are worshipped" does not entail the existence of gods, since "rich" is extensional but "be worshipped" intensional with regard to the subject term.

Whether the precondition $\{Fx\} \cap [<Gx>] \neq \emptyset$ specified for ∀ should also be specified for (the negation of) ∃, depends on natural intuitions. If a sentence like "No unicorn likes hay" is deemed true in this world, which is far from unlikely, then the answer is No, otherwise Yes. Pending the availability of reliable data, we leave both options open. In fact, French speakers report that "Aucun unicorne n'aime le foin" is more likely to be taken to be false, whereas "Il n'y a pas d'unicorne qui aime le foin" is clearly true. There may thus be different varieties of the (negated) existential quantifier.

This account restores the equivalence of "not-all F is G" and "some F is not G", which was lost in RAPC. In RAPC, $\neg\forall(Fx,Gx) \not\equiv \exists(Fx,\neg Gx)$, since when $[[Fx]] = \emptyset$, $\neg\forall(Fx,Gx)$ is true while $\exists(Fx,\neg Gx)$ is false. In the intensional calculus, however, when the matrix predicate Gx is extensional, ${\sim}\forall(Fx,Gx)$ is false in cases where $[[Fx]] = \emptyset$, owing to existential presupposition failure, while $\exists(Fx,{\sim}Gx)$ is likewise false. When, on the other hand, Gx is intensional, it is immaterial whether or not $[[Fx]] = \emptyset$, since with intensional matrix predicates the conditions for both $\forall$ and $\exists$ apply to $\{Fx\}$, which is automatically nonempty (see above). Fig. 5 shows this more clearly: the shaded area represents $/{\sim}\forall(Fx,Gx)/$, which coincides with $/\exists(Fx,{\sim}Gx)/$. If the G-predicate has no preconditions, $\mathbf{U}_\forall = \mathbf{U}$, which leaves the equivalence intact. This again improves the empirical status of the calculus.

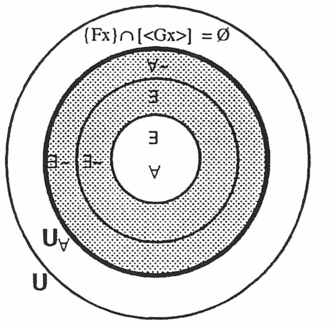


Figure 5

On the other hand, however, there is a problem (signalled by a number of authors, e.g. Zalta 1988, Castañeda in Haller 1985/6:58, Lejewski in Haller 1985/6:232), with the truth conditions of simple sentences of the form "Holmes is an Englishman". We want to say that "Some Englishmen are imaginary" is true, since $\{Englishman(x)\} \cap [[imaginary(x)]] \neq \emptyset$. If we are then asked to mention an instance of an imaginary Englishman, we want to be able to produce, for example, Sherlock Holmes as imagined by Conan Doyle. Yet under the terms specified so far "Holmes is an Englishman" is (radically) false owing to presupposition failure, since to be an Englishman one first has to exist, which Holmes does not do. We need, therefore, a second interpretation under which "Holmes is an Englishman" is true, which makes this sentence ambiguous. Our solution consists, in principle, in adding the hedge "who/which qualifies for the (main) predicate Gx" not only to all quantified terms but also to instantiations adduced in a chain of argument. Not only would "Some Englishmen are imaginary" then be read as "Some Englishmen who qualify for the predicate "imaginary" are imaginary", but "Holmes is an Englishman" would then likewise be read as "Holmes is an Englishman who qualifies for the predicate "imaginary", or "Holmes is a case in point". However, the mechanism for the distribution of such hedges has not been developed yet (see Castañeda in Haller 1985/6:58 for a similar view, but without formal elaboration).

One possible implication for the study of human cognition should be mentioned here. It appears that human cognition does not naturally develop the concept of null set until a very high degree of mathematical abstraction is achieved. The question is whether a satisfactory logic and semantics of quantification can be developed without the help of $\emptyset$ to account for the lower levels of abstraction where natural language operates. As has been shown above, intensional predicate calculus eliminates $\emptyset$ as an option for the R-set (F-predicate). It remains to eliminate $\emptyset$ for the matrix set (G-

predicate): a sentence like "Some logicians are 450 years old", where the predicate "450 years old" is uninstantiated, must be processable and result in the value False. This may be achieved by treating the quantifiers as binary predicates over pairs of R-sets and predicate intensions (satisfaction conditions of predicates). "All F is G" now means: "all members of $\{Fx\}$ that qualify for Gx satisfy the conditions of Gx". One notes that this would be a return to the Aristotelian notion of a proposition as the mental assignment of a property to an entity or set of entities. It would also place the medieval theory of distributive supposition in a new formal light.

There is, furthermore, the peculiar fact that, at a non-reflective default level of cognitive operation, *some* is naturally interpreted as "partial", and thus equivalent with "partial not". The corresponding extensional calculus is shown in fig. 6 (with "**P**" for 'partial' 'and "=" for 'equivalent'):

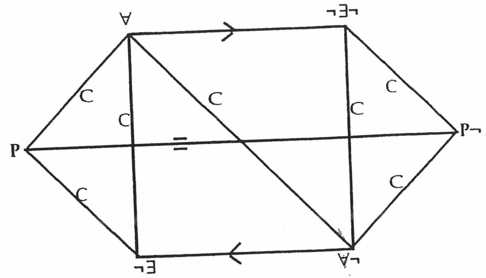


Figure 6

It may be assumed that at this default level of abstraction the notion 'subset' (denoted by "some F") is defaultwise interpreted as 'proper subset'. The combination of the 'no-null-set' hypothesis with the 'proper subset' hypothesis opens an interesting perspective on further research into the logical properties of cognition.

References:

Haller, Rudolf (ed.) (1985/86) *Non-Existence and Predication. Grazer Philosophische Studien* 25/6.

Horn, Laurence R. (1972) *On the Semantic Properties of Logical Operators in English.* PhD-dissertation, University of Southern Califiornia, Los Angeles CA. (Distr. Indiana University Linguistics Club 1976)

Horn, Laurence R. (1989) *A Natural History of Negation.* The University of Chicago Press, Chicago-London.

Levinson, Stephen C. (2000) *Presumptive Meanings. The Theory of Generalized Conversational Implicature.* MIT Press, Cambridge MA.

Seuren, Pieter A.M., Venanzio Capretta & Herman Geuvers (2001) The logic and mathematics of occasion sentences. *Linguistics & Philosophy* 24.5:531-95.

Van Fraassen, Bas C. (1971) *Formal Semantics and Logic,* Macmillan, New York-London.

Zalta, Edward N. (1988) *Intensional Logic and the Metaphysics of Intentionality.* Bradford book, MIT Press, Cambridge, Mass.

A Variable-Free Dynamic Semantics

Chung-chieh Shan
Harvard University

I propose a variable-free treatment of dynamic semantics. By "dynamic semantics" I mean analyses of donkey sentences (*Every farmer who owns a donkey beats it*) and other binding and anaphora phenomena in natural language where meanings of constituents are updates to information states, for instance as proposed by Groenendijk and Stokhof [3]. By "variable-free" I mean denotational semantics in which functional combinators replace variable indices and assignment functions, for instance as advocated by Jacobson [6, 7].

The new theory presented here achieves a compositional treatment of dynamic anaphora that does not involve assignment functions, and separates the combinatorics of variable-free semantics from the particular linguistic phenomena it treats. Integrating variable-free semantics and dynamic semantics gives rise to interactions that make new empirical predictions, for example "donkey weak crossover" effects.

1. Decomposing dynamism

Dynamic semantics combines *nondeterminism*, *input*, and *output* to interpret discourse fragments such as (1) in a process informally described in (2).

(1) A man walks in the park. He whistles.

(2) *Nondeterministically* select a man x.
Output x as a candidate antecedent for future anaphora.
Check to make sure that x walks in the park; if not, abort execution.
Input a previously encountered candidate antecedent y.
Check to make sure that y whistles; if not, abort execution.

In this section, I analyze each aspect in turn. In subsequent sections, I will then review the empirical and theoretical advantages gained in my variable-free treatment.

1.1. **Nondeterminism.** In the first half of (1), *a man* nondeterministically selects a man, who is then tested for the property *walks in the park*. If any choice of a man passes the test, the sentence is true; otherwise, it is false. Denotationally, I model nondeterminism by letting phrases denote sets of what they traditionally denote in Montague grammar. For example, I let noun phrases denote not individuals but sets of individuals: *John* denotes the singleton set containing John, and *a man* the set of all men. Formally, I assign the type $\mathrm{Set}(e)$ (or $e \to t$), rather than e, to noun phrases such as *John* and *a man*. Here e is the base type of individuals and t is the base type of truth values.

Now write 1 for the *unit type*, the identity for the binary type constructor $\times$ for product types. This type can be thought of as a singleton set, say $\{*\}$. Note that $\{*\}$ has two subsets, namely $\{*\}$ and $\{\}$. I treat these two subsets as true and false, respectively, thus establishing an isomorphism between the types t and $\mathrm{Set}(1)$. I then assign to *walks in the park* the type $\mathrm{Set}(e \to \mathrm{Set}(1))$ rather than $e \to t$. We can think of a property as a nondeterministic function that maps each e to a nondeterministic 1, returning either or nothing—that is, either $\{*\}$ or $\{\}$.

In general, nondeterminism can be added to any Montague grammar by replacing each semantic type τ with a transformed type $\mathrm{Set}(\lfloor\tau\rfloor)$, where $\lfloor - \rfloor$ is a map from

Thanks to Pauline Jacobson and Stuart Shieber for discussion and comments. This work is supported by the United States National Science Foundation under Grant IRI-9712068.

types to types recursively defined by

(3a) $\lfloor\tau\rfloor = \tau$ for any base type τ,

(3b) $\lfloor\tau_1 \to \tau_2\rfloor = \lfloor\tau_1\rfloor \to \mathrm{Set}(\lfloor\tau_2\rfloor)$ for any types τ_1 and τ_2.

My type for properties $\mathrm{Set}(e \to \mathrm{Set}(1))$, for example, is precisely $\mathrm{Set}(\lfloor e \to 1\rfloor)$. Once every type is transformed, it is straightforward to specify how semantic values compose, either by adding a composition method (as in Hamblin's interrogative semantics [4]) or by adding a type-shift operation. In programming language terms, I am using a lambda calculus that is *impure* because it incorporates call-by-value nondeterminism, as has been detailed by others [14, §8].

For brevity, I will henceforth write untransformed types in place of transformed ones. For example, I will write e rather than $\mathrm{Set}(e)$ for the type of an individual, $e \to 1$ rather than $\mathrm{Set}(e \to \mathrm{Set}(1))$ for the type of a property, and $e \to e \to 1$ rather than $\mathrm{Set}(e \to \mathrm{Set}(e \to \mathrm{Set}(1)))$ for the type of a two-place relation. Intuitively, a type $\tau_1 \to \tau_2$ is henceforth to be interpreted as a relation between τ_1 and τ_2, or equivalently, a function from τ_1 to the power set of τ_2. Accordingly, if f is of type $\tau_1 \to \tau_2$ and x is of type τ_1, then the term $f(x)$, of type τ_2, is to be interpreted as the image of the set x under the relation f.

We can now derive *A man walks in the park*:

$$\mathrm{A} : (e \to 1) \to e = \lambda p.\, \{\, v \mid * \in p(v)\,\},$$

$$\mathrm{MAN} : e \to 1, \qquad \mathrm{WITP} : e \to 1, \qquad \mathrm{WITP}(\mathrm{A}(\mathrm{MAN})) : 1.$$

1.2. **Input.** In the second half of (1), a man *he* is determined by the discourse context, who is then tested for the property *whistles*. The central idea of variable-free semantics is to model dependence on discourse context by letting phrases denote functions from *inputs* to what they traditionally denote in Montague grammar. For example, *he* will denote the identity function over men, and *he whistles* the function mapping each man to whether he whistles.

To restate this idea formally in terms of types, I introduce a new binary type constructor $\rhd$ ("in"). The type $\sigma \rhd \tau$ is like $\sigma \to \tau$ in that they may have the same models, namely functions from σ to τ. I use for both kinds of types the same λ-.- notation for abstraction and -(-) notation for application, but distinguish between them so that, for example, a value of type $(a \to b) \to c$ cannot apply directly to one of type $a \rhd b$. (This is equivalent to how, in Jacobson's formulation, syntactic categories regulate semantic combination to stop *loves him* from applying to *Mary* [6, §2.2.1.1].) By convention, all binary type constructors associate to the right.

I assume that *whistles* denotes some property $\mathrm{WHISTLE} : e \to 1$, and let *he* denote

$$\mathrm{HE} : e \rhd e = \lambda v.\, v.$$

Because HE does not have type e, the property WHISTLE cannot apply to it directly. I follow Jacobson in introducing a type-shift operation

(4) $$g^{\rhd} : (\alpha \to \beta) \to (\sigma \rhd \alpha) \to (\sigma \rhd \beta) = \lambda f.\, \lambda v.\, \lambda s.\, f(v(s)).$$

We can now derive *he whistles*:

$$g^{\rhd}(\mathrm{WHISTLE})(\mathrm{HE}) : e \rhd 1 = \lambda v.\, \mathrm{WHISTLE}(v).$$

For phrases containing more than one pronoun, for example *he loves her*, I generalize the type-shift operation $g^{\rhd}$ to a family of operations $g^{\rhd}_{i,j} = G^i(I^j(g^{\rhd}))$ for non-negative integers i and j, that is,

$$g^{\rhd}_{0,0} = g^{\rhd}, \qquad g^{\rhd}_{0,j+1} = I(g^{\rhd}_{0,j}), \qquad g^{\rhd}_{i+1,j} = G(g^{\rhd}_{i,j}),$$

where the "composition" function G and the "insertion" function I are defined by

(5) $$G : (\alpha \to \alpha') \to (\tau \to \alpha) \to (\tau \to \alpha') = \lambda g.\, \lambda f.\, \lambda v.\, g(f(v)),$$

$$(6) \qquad I : \big((\alpha \to \beta) \to (\alpha' \to \beta')\big) \to \big((\alpha \to \tau \to \beta) \to (\alpha' \to \tau \to \beta')\big)$$
$$= \lambda g.\, \lambda f.\, \lambda v'.\, \lambda x.\, g(\lambda v.\, f(v)(x))(v').$$

Assuming that *she* denotes SHE = HE and *loves* denotes LOVE : $e \to e \to 1$, we can now derive two denotations for *he loves her* with opposite scoping:

$$g^{\triangleright}_{1,0}(g^{\triangleright}_{0,1}(\text{LOVE}))(\text{SHE})(\text{HE}) : e \triangleright e \triangleright 1 = \lambda u.\, \lambda v.\, \text{LOVE}(v)(u),$$
$$g^{\triangleright}_{0,1}(g^{\triangleright}_{1,0}(\text{LOVE}))(\text{SHE})(\text{HE}) : e \triangleright e \triangleright 1 = \lambda v.\, \lambda u.\, \text{LOVE}(v)(u).$$

My generalization here of $g^{\triangleright}$ to handle multiple pronouns differs from Jacobson's, which does not posit $g^{\triangleright}_{i,j}$ for $i > 0$ and only generates the second scoping. The first scoping will be crucial as we consider output and binding below—typically, we need to use $g^{\triangleright}_{1,0}$ before the binding operation z, defined in (8), can apply.

1.3. **Output.** Informally speaking, *a man* can bind *he* in (1) by introducing a new discourse referent, i.e., a new candidate antecedent for future anaphora. I model this kind of addition to discourse context by letting phrases denote cartesian products between *outputs* and what they traditionally denote in Montague grammar. For example, *a man* will denote the set of all pairs $\langle v, v\rangle$ where v is a man, and *a man walks in the park* the set of all pairs $\langle v, *\rangle$ where v is a man who walks in the park.

Formally, I introduce a new binary type constructor $\ltimes$ ("out"). The type $\sigma \ltimes \tau$ is like $\sigma \times \tau$ in that they may have the same models, namely pairs between σ and τ. I will use for both kinds of types the same $\langle -, -\rangle$ notation for pairs, but distinguish between them so that, for example, a value of type $(a \times b) \to c$ cannot apply to another of type $a \ltimes b$. For simplicity, I treat as equivalent the isomorphic types

$$1 \ltimes \tau \quad \text{and} \quad \tau$$

for any type τ, and the isomorphic types

$$(\sigma_1 \times \sigma_2) \ltimes \tau \quad \text{and} \quad \sigma_1 \ltimes (\sigma_2 \ltimes \tau)$$

for any types τ, σ_1, and σ_2.

As with input, I introduce a type-shift operation

$$(7) \qquad g^{\ltimes} : (\alpha \to \beta) \to (\sigma \ltimes \alpha) \to (\sigma \ltimes \beta) = \lambda f.\, \lambda\langle s, v\rangle.\, \langle s, f(v)\rangle.$$

I then generalize $g^{\ltimes}$ to a family of type-shift operations $g^{\ltimes}_{i,j} = G^i(I^j(g^{\ltimes}))$ for non-negative integers i and j.

Recall from §1.1 that true and false are just nonempty and empty sets, respectively. Under this view, it is easy to define a concatenation function that conjoins the truth conditions of two discourse fragments:

$$; : 1 \to 1 \to 1 = \lambda *.\, \lambda *.\, *.$$

(Syntactically, I assume that the first argument to ; is the second of the two fragments to be conjoined, and vice versa.) Revising the denotation we specified earlier for *a*, we can now derive the reading of (1) where *a man* does not bind *he*:

$$\text{A} : (e \ltimes e \to \sigma \ltimes 1) \to \sigma \ltimes e = \lambda p.\, \{\, \langle s, v\rangle \mid \langle s, *\rangle \in p(\langle v, v\rangle) \,\},$$
$$g^{\ltimes}_{0,0}(\text{WITP})(\text{A}(g^{\ltimes}_{0,0}(\text{MAN}))) : e \ltimes 1,$$
$$g^{\ltimes}_{1,0}(g^{\triangleright}_{0,1}(;))(g^{\triangleright}_{0,0}(\text{WHISTLE})(\text{HE}))\big(g^{\ltimes}_{0,0}(\text{WITP})(\text{A}(g^{\ltimes}_{0,0}(\text{MAN})))\big) : e \ltimes e \triangleright 1.$$

For binding to take place, we need to feed outputs produced by semantically higher arguments into inputs solicited by semantically lower arguments. To implement this, I define one last type-shift operation

$$(8) \qquad z : \big(\alpha \to (\sigma \ltimes \beta) \to \gamma\big) \to \big((\sigma \triangleright \alpha) \to (\sigma \ltimes \beta) \to \gamma\big)$$
$$= \lambda f.\, \lambda v.\, \lambda\langle s, u\rangle.\, f(v(s))(\langle s, u\rangle)$$

and derive from it a family of type-shift operations $z_{i,j} = G^i(I^j(z))$ for non-negative integers i and j. We can now derive the reading of (1) where *a man* does bind *he*:

$$z(g^{\triangleright}_{1,0}(;))(g^{\triangleright}_{0,0}(\text{WHISTLE})(\text{HE}))\big(g^{\ltimes}_{0,0}(\text{WITP})(\text{A}(g^{\ltimes}_{0,0}(\text{MAN})))\big) : e \ltimes 1.$$

2. Empirical payoffs

Many variable-free analyses of empirical facts carry over in spirit to the variable-free dynamic semantics presented here, with extended coverage over dynamic phenomena. In this section, I give some simple examples that center around the classical donkey sentence (9).[1]

(9) Every farmer who owns a donkey beats it.

Before examining its variations, a derivation of (9) itself is in order. The critical lexical items are *every* and *who*. Given the denotation of *a* specified above, we expect *every* to have the semantic type

$$(e \ltimes e \to \sigma \ltimes 1) \to (\sigma \ltimes e \to \sigma' \ltimes 1) \to 1.$$

The same semantic type is also expected for other strongly quantificational elements, such as *most*. Following standard treatment in dynamic predicate logic, I let *every* denote[2]

$$\text{EVERY} = \lambda p.\, \lambda q.\, \{\, * \mid \forall s : \sigma.\, \forall v : e.\, \langle s, *\rangle \in p(\langle v, v\rangle) \Rightarrow \exists s' : \sigma'.\, \langle s', *\rangle \in q(\langle s, v\rangle) \,\}.$$

As for *who*, since I will only consider relative clauses with subject extraction in this paper, the following denotation is sufficient.[3]

$$\text{WHO} : (\sigma_2 \ltimes e \to \sigma_3 \ltimes 1) \to (\sigma_1 \ltimes e \to \sigma_2 \ltimes 1) \to (\sigma_1 \ltimes e \to \sigma_3 \ltimes 1)$$
$$= \lambda p.\, \lambda q.\, \lambda\langle s_1, v\rangle.\, \{\, \langle s_3, *\rangle \mid \langle s_2, *\rangle \in q(\langle s_1, v\rangle),\ \langle s_3, *\rangle \in p(\langle s_2, v\rangle) \,\}.$$

We are now ready to derive (9):

$$\text{FARMER, DONKEY} : e \to 1, \qquad \text{OWN, BEAT} : e \to e \to 1, \qquad \text{IT} : e \triangleright e,$$
$$x = \text{WHO}(g^{\ltimes}_{1,0}(g^{\ltimes}_{0,1}(\text{OWN}))(\text{A}(g^{\ltimes}_{0,0}(\text{DONKEY}))))\big(g^{\ltimes}_{0,0}(\text{FARMER})\big) : e \ltimes e \to e \ltimes e \ltimes 1,$$
$$y = g^{\ltimes}_{1,0}(z(g^{\ltimes}_{1,0}(\text{BEAT})))(\text{IT}) : e \ltimes e \ltimes e \to e \ltimes e \ltimes 1, \qquad \text{EVERY}(x)(y) : 1.$$

---

[1] My examples assume that all farmers are male.

[2] This meaning gives the donkey antecedent universal quantificational force; in other words, it makes (9) mean that every farmer who owns a donkey beats every donkey he owns. As Schubert and Pelletier [12] and others point out, sometimes the donkey antecedent seems to take existential quantificational force instead. For example, (ia) naturally means (ib).

(i) a. Every man who had a dime put it in the meter.
b. Every man who had a dime put a dime in the meter.

I leave it for future work to account for this variation within the present framework. One possible solution is to posit alternative denotations for *every*. Another is to treat *it* as a paycheck pronoun that repeats the existential force of *a dime*, effectively implementing the paraphrase in (ib).

Related is the *proportion problem*, noted by Kadmon [8] and others: Each sentence in (ii) has a different truth condition.

(ii) a. Most farmers who own a donkey beat it.
b. Most donkeys owned by a farmer are beaten by him.
c. Mostly, when a farmer owns a donkey, he beats it.

The type constructor $\ltimes$ is not symmetric; it distinguishes between the individual that participates immediately in predicate-argument combination and any additional output available as candidate antecedents for future anaphora. Thus the differences in (ii) can easily be modeled here by positing a natural denotation for *most*.

[3] It is no accident that the set comprehension notation used to specify this meaning is reminiscent of the list or monad comprehension notation used to express evaluation sequencing in programming languages [14].

2.1. **Weak crossover.** Variable-free dynamic semantics accounts for the "donkey weak crossover" contrasts in (10) and (11) in roughly the same way regular variable-free semantics accounts for the weak crossover contrast in (10) alone [6, §2.2.3].

(10) a. Every farmer_i who owns a donkey loves his_i mother.
b. *His_i mother loves every farmer_i who owns a donkey.

(11) a. Every farmer who owns a donkey_i loves the woman who beats it_i.
b. *The woman who beats it_i loves every farmer who owns a donkey_i.

More specifically, binding is disallowed in (10b) and (11b) because z forces the binder—or, in the case of donkey anaphora, the NP containing the binder—to c-command the bindee. For the disallowed binding configurations to be possible, the grammar would need an alternative binding operation

$$s : ((\sigma \ltimes \beta) \to \alpha \to \gamma) \to ((\sigma \ltimes \beta) \to (\sigma \rhd \alpha) \to \gamma) = \lambda f.\, \lambda \langle s, u \rangle.\, \lambda v.\, f(\langle s, u \rangle)(v(s)). \tag{12}$$

2.2. **Functional questions.** Regular variable-free semantics derives the functional question-answer pair in (13a) and predicts the weak crossover violation in (13b) under natural analyses of extraction [6, §3.1–2]. Variable-free dynamic semantics further derives the "donkey functional" question-answer pair in (14a) while predicting the donkey weak crossover violation in (14b). The answers are all of type $e \rhd e$.

(13) a. Who does every farmer_i who owns a donkey love? His_i mother.
b. Who loves every farmer_i who owns a donkey? ?His_i mother.

(14) a. Who does every farmer who owns a donkey_i love? The woman who beats it_i.
b. Who loves every farmer who owns a donkey_i? ?The woman who beats it_i.

2.3. **Across-the-board binding.** Typical compositional analyses of right node raising allow variable-free semantics to predict (15) [6, §3.4]; variable-free dynamic semantics further predicts (16). The conjuncts are all of type $(e \rhd e) \to 1$.

(15) Every farmer_i who owns a donkey loves—but no farmer_j who owns a donkey wants to marry—$\text{his}_{i/j/*k}$ mother.

(16) Every farmer who owns a donkey_i loves—but no farmer who owns a donkey_j wants to marry—the woman who beats $\text{it}_{i/j/*k}$.

2.4. **Sloppy readings.** Because VPs can take semantic type $e \ltimes e \ltimes e \to 1$ in variable-free dynamic semantics, it is straightforward to derive "sloppy" readings for sentences involving VP ellipsis (17) or association with focus (18).

(17) a. Every East Coast farmer who owns a donkey beats it, but no West Coast farmer who owns a donkey does.
b. Every farmer who owns a donkey showed it to his mother, but no farmer who owns a horse did.

(18) Only the farmer who owns THIS donkey beats it.

2.5. **Antecedent accessibility.** Muskens [11] notes the contrast in (19).

(19) a. No girl walks.
b. *No girl_i walks. If she_i talks, she_i talks.

The second sentence in (19b) accesses a discourse referent *she* but is a tautology. Muskens's semantic theory is one of many that cannot account for this contrast without introducing representational constraints.

In variable-free semantics, types keep track of the discourse referents free in each phrase. For example, the second sentence in (19b) has type $e \rhd 1$, not 1, even though it does denote a constant function. The theory here thus blocks (19b) and accounts for Musken's observation while remaining non-representational.

3. Discussion

From a theoretical perspective, variable-free dynamic semantics is appealing for the same reasons variable-free semantics and dynamic semantics are appealing—because it preserves direct compositionality, eliminates assignment functions, models updates to information states, and treats donkey anaphora.

There have been numerous proposals to combine Montague grammar with dynamic semantics, including ones by Groenendijk and Stokhof [2], Muskens [11], Kohlhase, Kuschert, and Müller [9], and van Eijck [13]. The variable-free approach here clearly shares the outlook of van Eijck's work, which uses de Bruijn indexing to eliminate variable indices from dynamic reasoning. By comparison, the theory here is more prominently guided by types, using them to record the *computational side effects* [10] of nondeterminism, input, and output that are incurred during dynamic interpretation. This type system seems related to but more simplistic than Fernando's proof-theoretic semantics [1]; for example, there are no dependent types.

I have tried to draw an analogy between incoming assignments ($\rhd$) and outgoing ones ($\ltimes$), characterizing with similar combinatorics the former with the exponential functor and the latter with the product functor. The success of the analogy—between, for example, the type-shift operations $g^{\rhd}$ and $g^{\ltimes}$—suggests that the same combinatorics may be applicable to an even wider range of linguistic phenomena. For instance, note that Hendriks's Argument Raising operation [5, §1.4.2] follows from the functoriality of $(- \to \sigma_1) \to \sigma_2$ for any σ_1 and σ_2, in the same way $g^{\rhd}$ and $g^{\ltimes}$ follow from the functoriality of $\sigma \rhd -$ and $\sigma \ltimes -$, respectively, for any σ.

References

[1] Tim Fernando. A type reduction from proof-conditional to dynamic semantics. *Journal of Philosophical Logic*, 30(2):121–153, 2001. Also as Technical Report TCD-CS-2000-59, Computer Science Department, Trinity College, Dublin (2000).

[2] Jeroen Groenendijk and Martin Stokhof. Dynamic Montague grammar. In László Kálmán and László Pólos, editors, *Papers from the 2nd Symposium on Logic and Language*, pages 3–48, Budapest, 1990. Akadémiai Kiadó.

[3] Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100, 1991.

[4] C. L. Hamblin. Questions in Montague English. *Foundations of Language*, 10:41–53, 1973.

[5] Herman Hendriks. *Studied Flexibility: Categories and Types in Syntax and Semantics*. PhD thesis, Institute for Logic, Language and Computation, Universiteit van Amsterdam, 1993.

[6] Pauline Jacobson. Towards a variable-free semantics. *Linguistics and Philosophy*, 22(2):117–184, 1999.

[7] Pauline Jacobson. Paycheck pronouns, Bach-Peters sentences, and variable-free semantics. *Natural Language Semantics*, 8(2):77–155, 2000.

[8] Nirit Kadmon. *On Unique and Non-Unique Reference and Asymmetric Quantification*. PhD thesis, Department of Linguistics, University of Massachusetts, 1987.

[9] Michael Kohlhase, Susanna Kuschert, and Martin Müller. Dynamic lambda calculus. Manuscript, Universität des Saarlandes, 28 October 1998.

[10] Eugenio Moggi. An abstract view of programming languages. Technical Report ECS-LFCS-90-113, Laboratory for Foundations of Computer Science, Department of Computer Science, University of Edinburgh, 1990.

[11] Reinhard Muskens. Combining Montague semantics and discourse representation. *Linguistics and Philosophy*, 19(2):143–186, 1996.

[12] Lenhart K. Schubert and Francis Jeffry Pelletier. Generically speaking, or, using discourse representation theory to interpret generics. In Gennaro Chierchia, Barbara H. Partee, and Raymond Turner, editors, *Properties, Types and Meaning. Volume 2: Semantic Issues*, pages 193–268. Kluwer, Dordrecht, 1989.

[13] Jan van Eijck. Incremental dynamics. *Journal of Logic, Language and Information*, 10(3):319–351, 2001.

[14] Philip Wadler. Comprehending monads. *Mathematical Structures in Computer Science*, 2(4):461–493, 1992.

Embedded Quantifiers and Quantificational Variability

Yael Sharvit, University of Connecticut

As observed in Szabolcsi (1997), many quantifiers in interrogative clauses embedded under verbs of the *know*-class may be interpreted as having scope over the embedding verb. This paper argues that this wide scope is only apparent, and is a Quantificational Variability (henceforth, QV) effect. Thus, (1a), paraphrased in (1b), receives in effect the interpretation in (1c):

(1)a. John knows which woman most men love.
b. "Most men are such that John knows which woman they love".
c. ∀Q[Q is a relevant question-part of [*which woman most men love*]][John knows Q]

This proposal predicts the (un)availability of apparent wide scope for quantifiers embedded under *know*-type verbs vs. *wonder*-type verbs to correlate with the (un)availability of QV:

(2) For the most part, Bill knows who cheated (cf. (1a))
(3)a. John wonders which woman most men love.
b. NOT: "Most men are such that John wonders which woman they love".
(4) ##For the most part, Bill wonders who cheated. (cf. (3a))

1. Szabolcsi's (1997) observations

Matrix pair-list (PL) questions are supported only by *every*. Many quantifiers support, what appear to be, PL questions embedded under *know*-type verbs:

(5) Q: Which woman does every man love?
A: John loves Mary and Bill loves Sally.
(6) John knows which woman every man loves.
"Every man is such that John knows which woman he loves".
(7) Q: Which woman do most men/(more than) five men love?
A: *John loves Mary, Bill loves Sally, Fred loves Kelly,...
(8) John knows which woman most men/(more than) five men love.
"Most men/(more than) five men are such that John knows which woman they love".
(9) Q: Which woman does no man love?
A: *John doesn't love Mary and Bill doesn't love Sally?
(10) John knows which woman no man loves.
NOT: "no man is such that John knows which woman he loves".

Note: the so-called 'choice'-question *who do two men love* is not, according to Szabolcsi (1997), a genuine pair-list question.

Wonder-type verbs do not allow "scoping" of any quantifier:
(11) John wonders which woman every man loves.

NOT: "John wonders about every man which woman he loves".
(12) John wonders which woman most men/(more than) five men love.
NOT: "most men/(more than) five men are such that John wonders which woman they love".

2. Szabolcsi's Analysis

PL questions embedded under *wonder*-type verbs are the same kind of "creatures" as matrix PL questions (i.e., "non-lifted" questions, interpreted via domain restriction). PL questions embedded under *know*-type verbs are "lifted-questions", which raise above the embedding verb and its subject, as schematized in (13). 'P' ranges over properties of questions:

(13)a. [[which-NP]$_1$ [Quantifier Embd-V e_1]]$_2$ [Subject Main-V e_2]
b. *[[which NP]$_1$ [Quantifier Embd-V e_1]] =>*
λP[Quantifier(λx[P(which-y∈NP[x Embd-V y])])]
(14)a. John knows which woman most men love.
b. [which woman most men love]$_1$ [John knows t_1]
c. λP[most(men, λx[P(which-y∈WOMAN[x loves y])])](λQ[John knows Q])
= Most(men, λx[John knows which woman x loves])

But this doesn't explain the correlation between the *know/wonder* contrast with respect to embedded quantifiers (i.e., the (8)/(12) contrast), and the *know/wonder* contrast with respect to QV (Berman (1991), Lahiri (1991)):

(15) For the most part, John knows who cheated.
"For most x that cheated, John knows that x cheated"
(16) *For the most part, John wonders who cheated.
"For most x, John wonders whether x cheated"

3. The QV Proposal

The wide scope of embedded quantifiers is a by-product of QV.

3.1. QV – Sharvit and Beck (to appear)

Since Berman (1991) and Lahiri (1991), it has been widely accepted that only verbs that have a proposition-taking meaning (the *know*-class) support QV. This is corroborated by the contrast between (15) and (16). However, as argued in Sharvit and Beck (to appear), some verbs that lack a proposition-taking meaning (e.g., *depend*) support QV and even *wonder* sometimes does, with the help of a presuppositional element such as *still*:

(17) Which candidates will be admitted, depends, for the most part, exclusively on this committee.
"Most candidates x, whether x will be admitted depends exclusively on this committee"

(18) Smith: Has John made up his mind about the cheating?
Jones: So far he has only made up his mind regarding Susie and Bill. For the most part, he is still wondering who cheated.
"For most relevant x, John is still wondering whether x cheated"

Sharvit and Beck: QV effects are the result of quantification over questions.
(19)a. For the most part, John knows who cheated.
b. MOST-Q[Q∈PART(⟦*who cheated*⟧)(w)][John knows Q in w]

PART(⟦*who cheated*⟧)(w) is the maximally salient division of ⟦*who cheated*⟧ into subquestions in w.

(20) A set of question-intensions S is a division of a question-intension Q into subquestions in a world w iff (i)-(ii) hold:
(i) Each member of S is a subquestion of Q;
(ii) Either a. $\cap\{$Ans-wk(Q')(w):Q'$\in$S$\} \subseteq$ Ans-wk(Q)(w), and $\neg\exists$S'$\subset$S such that $\cap\{$Ans-wk(Q')(w):Q'$\in$S'$\} \subseteq$ Ans-wk(Q)(w),
or b. $\cap\{$Ans-wk(Q')(w):Q'$\in$S$\} \subseteq$ Ans-strg(Q)(w), and $\neg\exists$S'$\subset$S such that $\cap\{$Ans-wk(Q')(w):Q'$\in$S'$\} \subseteq$ Ans-strg(Q)(w).
(21) A question-intension is a function from W to $D_{<<s,t>,t>}$ (Hamblin (1973)).
(22)a. Ans-wk(Q)(w) = $\cap\{$p:Q(w)(p) & w$\in$p$\}$
b. Ans-strg(Q)(w) = {w':Ans-wk(Q)(w')=Ans-wk(Q)(w)} (Heim (1994))
(23) A question-intension Q' is a subquestion of a question-intension Q iff: $\exists$w$\exists$p[Ans-strg(Q')(w) $\subseteq$ p & p is a partial answer to Q]
(24) A proposition p is a partial answer to a question Q iff: $\exists$w*[Ans-strg(Q)(w*) $\cap$ p = $\varnothing$] (Groenendijk & Stokhof (1984))

The idea is that the members of the division of Q together provide the complete answer to Q. Suppose John, Mary, and Bill cheated. Then{⟦*did John cheat*⟧, ⟦*did Bill cheat*⟧, ⟦*did Mary cheat*⟧} is a possible division of ⟦*who cheated*⟧.

(19b') MOST-Q[Q∈ {⟦*did John cheat*⟧, ⟦*did Bill cheat*⟧, ⟦*did Mary cheat*⟧}] [John knows Q in w]

3.2. Embedded quantifiers
Embedded quantifiers support QV, with a silent default adverb of quantification.

(25)a. John knows which woman most men love.
b. ∀Q∈PART(⟦*which woman most men love*⟧)(w)[John knows Q in w]

What can be a division of ⟦*which woman do most men love*⟧?

(26) Which woman do most men love?
No pair-list answer.
Functional answer: Most men love their mother.
(Engdahl (1986), Groenendijk & Stokhof (1984), Chierchia (1993))

The functional question is: "which function f is such that most men x love f(x)" (i.e., {most men love their mother, most men love their sister...}). Suppose John, Bill and Fred are our men, and John loves his mother and Bill loves his mother. (27) is a possible division of the functional ⟦*which woman do most men love*⟧:

(27) {⟦*does man John love his mother*⟧, ⟦*does man Bill love his mother*⟧}

(25b') ∀Q∈ {⟦*does man John love his mother*⟧, ⟦*does man Bill love his mother*⟧}[John knows Q in w]

But not all functions have obvious or salient names.

4. Predictions

4.1. The correlation with QV
Since apparent wide scope of embedded quantifiers is a QV effect, we predict "classical" QV constructions to be the ones where embedded quantifiers have apparent wide scope (AWS).

(28)a. For the most part, John found out who cheated. (QV)
b. John found out which woman most men love. (AWS)
(29)a. Which candidates will be admitted depends, for the most part, exclusively on this committee. (QV)
b. Which candidate most professors will interview, depends exclusively on this committee. (AWS)
(30)a. For the most part, John wonders who cheated. (no QV)
b. John wonders which woman most men love. (no AWS)
(31) a. Smith: Has John made up his mind about the cheating?
Jones: So far he has only made up his mind regarding Susie and Bill. For the most part, he is still wondering who cheated.
"For most relevant x, John is still wondering whether x cheated" (QV)
b. John is still wondering which woman most men love.
"For most men x such that John has been wondering which woman x loves, John is still wondering which woman x loves." (AWS)

(32) ∀Q[∃w'[w' is a belief-world of John in w and Q∈PART(⟦*which woman do most men love*⟧)(w')] and John has been wondering Q in w][John is still wondering Q in w]

4.2. The scope of 'no'

No cannot have wide scope at all – actual or apparent. It gets "trapped" in the scope of the silent universal adverb of quantification.

(33)a. John knows which woman no man loves
b. ∀Q∈ {⟦*which woman does no man love*⟧}[John knows Q in w]

4.3. 'De dicto' reading of the common noun part of the quantifier

⟦*did student Bill invite his mother*⟧ is a subquestion of ⟦*which woman did most of the students invite*⟧

(34) {<w, p>:p={w':student(w')-Bill invited his mother(w') in w'} or p={w':student(w')-Bill didn't invite his mother(w') in w'}}
(35) John and Mary mistakenly believe that the professors are the students. They agree on which woman most of the students invited to the party.
(36) ∀Q[John and Mary believe λw[Q∈PART(⟦*which woman most of the students invited to the party*⟧)(w)][John and Mary agree on Q]
=> John and Mary agree on the answer to ⟦*did student Bill invite his mother*⟧
=> John and Mary believe that Bill is a student.

According to Szabolcsi, in the 'every>>some' reading below, the common noun part of the quantifier is not read 'de dicto':

(37) Some librarian or other knows which book every student borrowed.
Some >> every, every>>some (Moltmann and Szabolcsi (1995))

This is supported by intuitions regarding the discourse in (38), with non-factive *certain*:

(38) There is chaos in the library. All ten librarians mistakenly believe that the 2^{nd} year students are 1^{st} year students, and they have lost all their records. But the crisis could be much worse. At least, some librarian or other is certain which book every 1^{st} year student borrowed.

5. Problems

(39) John knows which woman exactly five men/less than five men love.
(40) Szabolcsi's analysis: Exactly five men/less than five men are such that John knows which woman each of them loves.

(41) The QV analysis:
∀Q∈PART(⟦*which woman exactly 5 men...* ⟧)(w)[John knows Q in w]

The QV analysis enables John to have knowledge about more than five men.

Possible solution: suppose *exactly* is focused.
(42) John knows which woman $[\text{exactly}]_F$ five men love.

{John knows which woman exactly five men love,
John knows which woman more than five men love, ... }

Implicature: all alternatives to *John knows which woman exactly five men love* -- excluding it and excluding those that are entailed by it -- are false.

6. Summary

The apparent wide scope of embedded quantifiers is a QV effect. The QV analysis explains the correlation between "classical" QV effects and the apparent wide scope of embedded quantifiers.

References:

Berman, Steven (1991). *On the Semantics and Logical Form of Wh-Clauses*, Ph.D. dissertation, University of Massachusetts at Amherst.
Chierchia, Gennaro (1993). "Questions with Quantifiers", *Natural Language Semantics* 1/2:181-234.
Engdahl, Elisabet (1986). *Constituent Questions*, Kluwer, Dordrecht.
Groenendijk, Jeroen and Martin Stokhof (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*, Ph.D. dissertation, University of Amsterdam.
Hamblin, C.L. (1973). "Questions in Montague English", *Foundations of Language* 10:41-53.
Heim, Irene (1994). "Interrogative Complements of 'Know'". R. Buchalla and A. Mittwoch (eds.) *IATL 1*, Akademon, Jerusalem, pp. 128-144.
Lahiri, Utpal (1991). *Embedded Interrogatives and Predicates that Embed Them*, Ph.D. dissertation, MIT.
Moltmann, Friederike and Anna Szabolcsi (1994). "Scope Interactions with Pair-list Quantifiers", in M. Gonzalez (ed.), *Proceedings of NELS24*, 381-395. GLSA Publications, University of Massachusetts at Amherst.
Sharvit, Yael and Sigrid Beck (to appear). "Subquestions and Quantificational Variability Effects", in K. Megerdoomian and L.A. Bar-el (eds.), *Proceedings of WCCFL XX*, Cascadilla Press, 510-523.
Szabolcsi, Anna (1997). "Quantifiers in Pair-List Questions", in A. Szablcsi (ed.), *Ways of Scope Taking*, Kluwer, Dordrecht, 311-342.

The Semantic Role of Modal Particles

Henk Zeevat, ILLC/Computational Linguistics, Universiteit van Amsterdam

1 Introduction

In Zeevat (to appear), I defend the view that at least certain particles (*too*, *again*, *indeed*, German/Dutch *toch/doch*, Dutch *wel* and others) should be considered to be presupposition triggers of a rather special kind: they allow weak antecedents (antecedents that not are not common ground but of which it can only be common ground that they are believed by somebody, suggested by somebody or even only dreamt by somebody) as well as proper ground antecedents, do not allow accommodation and are obligatory in the weak sense that where they occur they can normally not be omitted.

I argued there that the first two special properties are derivable if we use Blutner(2000)'s reconstruction of Van der Sandt's theory of presupposition in bidirectional optimality theory and adopt the view that antecedents for presupposition triggers can generally be weak. Non-accommodation then follows by Blutner's theorem from the fact that sentences with particles have semantically equivalent (and simpler) non-presupposing expression alternatives in the sentence without the particle, because the particle does not have a proper semantic contribution itself[1] The semantic emptiness also explains why weak antecedents are allowed in this case and not for a range of other presuppposition triggers. For normal triggers like the definite article or the verb *to know*, a weak antecedent is just not enough. It is the function of a definite description (in a transparent position) to supply an existing object for a predication and weak antecedents just do not give existing objects. Similarly, knowledge can only be of facts and the content of a dream or a suggestion just does not meet the bill. For factive verbs expressing an emotional reaction however, weak antecedents are possible as long as a suitable bridging inference is possible, as in (1).

(1) John thinks that Mary left and he regrets that she did.
Bill said that Mary left and Tom regretted that she did. (bridging inference: Tom believed what Bill said.)

Particles without any semantic content are at the other end of the spectrum: the presupposition does have any contribution to make to the content of the sentence.

But the same does not hold for the third special property: the obligatory character of particle occurrences. For this, we require in the framework of bidirectional optimality theory special constraints requiring the marking of certain relationships between the current sentence. These closely match traditional classifications of the particles in question, that makes them additive (*too*, *also* and *again*), adversative (*doch* (German), *toch* (Dutch), *nevertheless*), contrastive (*aber* (German), *echter* (Dutch), *however*), or confirmative (*indeed*). An OT treatment of the particles must include a system of constraints that prefers realisations that contain the particle (or some other device that marks the same property) for a certain semantic input if the input has a certain relationship to the context.

In the next section I give a provisional sketch of the marking constraints for the particles above. In section 3, I show how the constraint by itself can account for the pragmatic interpretational effects of the particles, at least for some of their uses. In section 4, I will attempt a speculative historical account of why we have the constraints. The final 2 sections will try to deal with special

[1]To this date, this is the only serious explanat..n of why certain triggers do not accommodate. The suggestion that is repeated in Geurts et al. (this volume) that pronouns do not accommodate because they have too little descriptive content is not easy to defend. First of all, there are many triggers that do not accommodate which allow as much descriptive content in their presuppositions as anybody would care for. Take "indeed p" which presupposes "p". Further, the descriptive content of pronouns varies considerably with languages and it is hard to understand why "he" should lack sufficient descriptive content while at the same time "the man" or "the entity that could be referred to by a male pronoun" do not.

uses of the particles and will defend that it is quite normal for particles of this kind to be used for marking special speech acts.

2 Context Marking

We are used to think of the meaning of linguistic expressions as having some communicative intent: there is a message conveyed by the expression, or it contributes to defining a message. This seems the predominant view in the study of formal languages or in programming languages.

This section explores the view that the only reason why particles like *too* and *toch* are in the sentence is because a special relation of the message to the existing context needs to be marked. The message by itself is unchanged, but its integration in the context takes it into account that similar material is already sitting there.

A context here is a set of propositions and a topic, here again just a set of propositions[2]. We assume that the context is consistent. A sentence addresses the topic iff the proposition it expresses is an element of the topic. A topic is settled iff all its elements are decided by the set of propositions. We further assume that we have a history of the conversation. The extended context is the set of proposition that has been suggested and can be defined by letting EC be the smallest set such that $C \subseteq EC$ and whenever $q \in EC$ and $q = [O](p)$ with O one of *suggest*, *dream*, *think*, *maybe*, *believe* or *think*, $p \in EC$.

Topics are related to questions by the scheme (2).

(2) $\lambda p \exists x \Box (\varphi \leftrightarrow p)$

Here x translates one or more WH-phrases occurring in a question that is compositionally represented as φ. It follows that yes-no questions are singleton sets. Topics allow a simple inversion operator NEG that inverts the polarity of the elements occurring in them.

Additive Particles

If the topic is an old topic that was addressed by a proposition distinct from p in the extended context and that is now closed off, add an additive marker to p.

(3) John is also going to Spain.
Topic: Where is John going?/What is John doing?
John is having dinner in New York too.
Topic: Who is having dinner in New York?

The topic is partly determined by the placement of the focus accent in the sentence but —for the interpreter— largely by the sentence itself. For the speaker, this is different: she has decided what the topic is going to be as part of her determination what she is going to say.

A special case is where the topic has been addressed before but with a different focus value and the focus of the current sentence is meant to replace it. Compare (4).

(4) John is going to Spain instead.
John is having dinner in New York instead.

with the same topics as before. The *instead* is not additive since the current utterance takes the place of the of its antecedent. This is typically but not invariably a correction of earlier information, as can be seen from (5).

(5) John thinks that Bill is in Spain but Tom thinks he is in France instead.

[2]This view of topics is further elaborated in Zeevat (draft)

Adversative Particles

Markers of this kind indicate the presence of a suggestion that the current sentence is *not* the case. We can capture this partially by making adversative marking obligatory when the extended context contains a proposition q such that q is a reason for thinking that $\neg p$. A formalisation in terms of default logic is possible, but I will not pursue that here. A proper marker is the German/Dutch *doch/toch* which is rather imperfectly rendered by *nevertheless*, *nonetheless* or *all the same*.

(6) Peter ist doch hier gewesen.
Peter is toch hier geweest.
Peter has been here nevertheless/all the same.

Contrastive Particles

Contrastive marking is in order if the last element of the context set is addressing a topic T and the current sentences directly or indirectly addresses the topic $NEG(T)$. A sentence addresses T indirectly iff it addresses a different topic Q but entails or implicates elements of $NEG(T)$.

(7) John cannot make it tonight.
But Bill is coming.
Bill however is coming.
Bill is coming though.

Confirmative Particles

If the extended context contains p already, this must be indicated by a confirmation particle. But the nature of the speech act determines what particle. If we are repeating p because we want to remind the interlocutor of its truth (e.g. in an attempt to argue for something else, or in order to reopen an old topic) we use reminder particles such as German *ja* or *doch* or Dutch *immers* or *toch*. But if we reaffirm p because there is now new evidence, we use *indeed.*

(8) There is no more cake.
since John ate it all.
Maybe John has eaten the cake. ...
John has indeed eaten the cake.

3 Interpretational Effects

The last section sketched some constraints that could responsible for the occurrence of certain particles in sentences. The constraints in German and Dutch should be ranked high enough to make the particles obligatory in case they are in a certain context. Suppose they are.

The interpretation of a statement in bidirectional OT is finding the best possible input for the sentence such that the statement is a best way of realising that input. Finding an input is constrained by a number of constraints of which only one concerns us here: do not add new material to the context unless you are forced to.

Now sentences with a particle make it necessary for the interpreter to recognise the context as one that makes the particle obligatory. If not, the best realisation would be the one that omits the particle and not the actual input sentence. Sentences without a particle should not be used in a context in which the proposition is adversing, adding, confirming, contrasting or denying. That is why the bare sentence should be taken to imply that these special cases do not obtain. It follows from an economy principle that particled sentences are not preferred unless the triggering conditions for the markers occur.

4 Why Mark?

The marking that we considered can all be seen as special cases of marking the positive or negative presence in the context of the sentence expressed or of parts thereof.

We can try to reconstruct the genesis of marking devices in the following simple minded way. We assume a general principle of economy governing discourse. Probably the most uneconomical way to operate to generate confusion. And failure to mark is doing precisely that.

Absence of additive marking will leave open the possibility that the current information is not added to the topic or with respect to a different topic or that it is misrecognised as the same information. This will make it possible that the information provided in the current statement is not considered relevant for the old topic or inversely. (It is typically unnecessary in a list construction where the topic remains open.)

Absence of adversative marking will leave the possibility open that the grounds for disbelieving p will still be active when we try to remember whether p and will so lead to either the wrong answer or to unwarranted doubts.

Confirmation of old opinions without marking it can lead to double storage with loss of information at both spots. Likewise for unmarked reminders. Here we may also restart the discussion of the already established proposition. And finally unmarked correcting information may lead to a failure in recognition.

Contrast finally marks the reversal of the topic and possibly a topic change. Not marking it would be uneconomical and brings the danger of misrecognition.

If I am right in assuming that marking avoids confusion and so increases the efficiency of linguistic communication, in situations where a specific marking device is unavailable other devices will have to take over, in those situations where the marking is crucial. Adverbs often used in this role can then lose their descriptive meaning quickly and lose syntactic and phonological properties. But the emergence of such specialised material also makes the need of marking greater, since it becomes more and more reasonable to assume that in the absence of the marker the relations between the proposition and the context do not obtain. So avoiding confusion leads to marking devices under the assumption that non-marking creates confusion. The presence of light marking devices reinforces their use since the opposite assumptions will start attaching to the unmarked sentences.

This can be thought of as explaining both why for certain relations between the context and the proposition there are marking devices in many languages, but also why there is a good deal of variation: the presence of the marking device itself creates its need.

5 Speech Act Marking

There are basic unmarked speech acts with standard conditions and marked speech acts which violate some of the standard conditions. In addition there may be finer classifications of the basic speech acts which take into account the discourse function of the statement, i.e. the question by which discourse relation the current statemnt needs to be attached to the earlier discourse.

Many of the particles I discussed play a role in changing the standard conditions of the speech acts in which they occur. There is a natural relation between adversative markers and corrections. If reason for disbelief is either the opinion of the speaker that p is false, or the belief of the interlocutor that p is false or the context fact that p is false, the adversative marker in such a context is *ipso facto* a successful indicator that the utterance is a self-correction, a correction of the interlocutor or a correction of the common ground. All these corrections are special in that they break the standard conditions on assertions that the speaker nor the hearer is already committed to p or its negation. Common ground corrections and self-corrections force in addition to updates retractions in the context and are thereby more complicated than straightforward assertions.

Also *instead* which is related to the additive particles can give rise to corrections, again by the mechanism of finding an antecedent in the context.

Similarly confirmative particles have a relation with the speech act of reminding. Reminding is a break of the standard condition on assertions to contain new information. Again, if the confirmative particle is resolved to an element of the context or to a speaker belief in the context, this *ipso facto* makes it into a reminder.

The case of *indeed* is more subtle. Unlike the other confirming particles, it cannot be used with element of the context itself unless there are reasons to doubt their truth. *Indeed* suggests that the speaker has evidence for p which has not been used yet to endorse p in the common ground. When p is not in the context yet, this is invariably the case.

The obvious question is how it is possible that context markers can override standard conditions on speech acts. They have the capacity of being speech act markers, by forcing marked interpretations of the speech act. But they also play this role by convention.

6 *Toch* as a speech act marker

Toch is related to the English *though*, that marks concessions. It is not very controversial to assume that its origins are connected to adversativity (like the isolated *though* in English). In fact, outside dialogues, the most frequent use of *toch* in texts is as a pronominal concessive, referring to a salient reason why the sentence that it marks should not be true.

(9) (John has just had a positive experience).
Toch was hij niet tevreden.
He was not happy though.

If we put an accent on *toch* as in the following sentences it starts to have a special relation with the sentence with the opposite polarity. (10)is out of place if it is not common ground that he would not come.

(10) Hij komt TOCH.
He is coming after all.

The same we find in a confirmation question (a rising intonation declarative) as in (11). As a confirmation question it demands that that the other has said or implied that he would come. The accented *toch* makes a direct relation with the opposite opinion that was common ground. The implication is that the other has self-corrected when she said or implied that she was coming.

(11) Je komt TOCH?
You are coming after all, isn't it?

It is easiest to think of the accent on *toch* in these cases as a contrastive accent. The sentence contrasts with the context element that saying the opposite and that is identified both by the contrast intonation and by the adversative marker. If the context element is still salient, *wel* with the same intonation pattern can take over.

The accented *toch* is therefore a good marker for self-corrections and common ground corrections. And for asking the other to change plan. It asserts against the common ground or against the speaker, or it asks the interlocutor to confirm something that goes against common ground knowledge or to act agaisnt a common ground plan.

Deaccented *toch* on the other hand often marks reminders (see also Doherty for a discussion of this use), like the Dutch *immers* or the German *ja*. It is easy to see how that happens. We can object to what the other is saying or implying by reminding her of common ground opinions. *Toch* occurs because of the adversativity with what the other is saying or implying and the accent falls elsewhere (in (12) on Berlin).

(12) A: I saw John yesterday.
B: Hij is toch in Berlijn?
But he is in Berlin, isn't he?

But the use of deaccented *toch* can competely lack this adversative element [3].

(13) A walks into B's office.
Ik ben toch volgende week weg. (I told you I was away next week)
Can you take over my lecture?

The process underlying this is fairly transparent. At occasions like (12) the interpreter is forced to conclude not just that the other does not agree (adversativity) but also that the other thinks that the message is common ground between them. *Toch* can therefore mark common ground status of the sentence and be recategorised as a particle that marks reminders. Adversativity is however not ruled out. The context can lack an adversative element and this is fine, but it can have one and the link will be made.

German *ja* is probably best seen as a particle that marks reminders and not as a confirmative marker. (It is also the German *yes*.) For proper confirmative marker we would expect other people to be able to take over the role of the interlocutor as the bearer of the agreeing opinion. And this is not so. In (14), the speaker cannot agree with Peter's dream, but still agrees with the interlocutor.

(14) Peter dreamt that he would fail. Und er ist ja tatsächlich durchge-
fallen.
And as you know he failed indeed.

All the reminding uses of *toch* above can be replaced by German version using either *doch* (there is very little difference between *toch* and*doch*) or *ja*. This is illustrated in (15).

(15) Er kommt ja.
Komm ja.
Wenn du ja kommst, nimm das Buch mit.

Conclusion

I have tried to sketch a strategy for dealing with the complex modal particles that we find in German and Dutch. The strategy assumes that they start out as context markers in response to a functional need. The natural occurrence of context markers in marked speech acts leads to their recategorisation as speech act markers and causes the complexities of which many have complained. I have given the beginning of an account of why such a change should be expected. Context markers do not arise because of expressive needs but as warning signals. Bidirectional OT makes sure they have a bearing on the interpretation process, a bearing that can make them expressive of special speech acts again.

References R. Blutner: Some Aspects of Optimality Theory in Interpretation. *Journal of Semantics*,**17**,p. 189-216, 2000.

Doherty, M. *Epistemic meaning*, Springer-Verlag, 1987

Geurts & Van der Sandt. Too. This volume.

Elena Karagjosova, Towards a comprehensive meaning of German doch , Proceedings of the ESSLLI 20001 Student Session, Helsinki, August 2001;

Krifka, M. Additive particles under stress. Proceedings of SALT 8. Cornell, CLC Publications 1999, 111-128.

Zeevat (to appear) Explaining Presupposition Triggers. In Van Deemter & Kibble (eds.) Information Sharing. CSLI Volume.

Zeevat (draft) Splitting Topics.

[3] For a different view, see Karagjosova 2000

Addresses and Affiliations Authors Contributed Papers

Maria Aloni aloni@hum.uva.nl
University of Amsterdam, ILLC/Department of Philosophy
Nieuwe Doelenstraat 15
1012 CP Amsterdam
Nederland

Alon Altman alon@vipe.technion.ac.il
Technion - Israel Institute of Technology, Computer Science
Haifa 32000
Israel

Nicholas Asher nasher@bertie.la.utexas.edu
University of Texas at Austin, Department of Philosophy
Austin, TX, 78712
USA

Martin van den Berg vdberg@pal.xerox.com
FXPAL
3400 Hillview
Palo Alto, CA 94304
USA

Raffaella Bernardi Raffaella.Bernardi@let.uu.nl
University of Utrecht, UiL OTS
Trans 10
3512 JK Utrecht
Nederland

Patrick Blackburn patrick@aplog.org
INRIA
Lorraine
France

António Branco Antonio.Branco@di.fc.ul.pt
University of Lisbon, Departamento de Informática
Faculdade de Ciencias de Lisboa
Campo Grande, 1700 Lisboa
Portugal

Ronnie Cann ronie@ling.ed.ac.uk
University of Edinburgh, Dept. of Theoretical and Applied Linguistics
40 George Square
Edinburgh EH8 9LL
United Kingdom

Balder the Cate b.ten.cate@hum.uva.nl
University of Amsterdam, ILLC/Department of Philosophy
Nieuwe Doelenstraat 15
1012 CP Amsterdam
Nederland

Cleo Condoravdi condorav@parc.xerox.com
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304
USA

Dick Crouch crouch@parc.xerox.com
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304
USA

Paul Dekker dekker@hum.uva.nl
University of Amsterdam, ILLC/Department of Philosophy
Nieuwe Doelenstraat 15
1012 CP Amsterdam
Nederland

Daniela Dudau-Sofranie dudau@lifl.fr
LIFL-Grappa Team
Charles de Gaulle - Lille 3
59 653 Villeneuve d'Ascq Cedex
France

Jan van Eijck jve@cwi.nl
University of Amsterdam, CWI
PO Box 94079
1090 GB Amsterdam
Nederland

Katrin Erk erk@ps.uni-sb.de
Saarland University, Programming system Lab
Postfach 15 11 50
D-66041 Saarbrücken
Germany

Bart Geurts Bart.Geurts@phil.kun.nl
University of Nijmegen, Dept. of Philosophy
Postbus 9103
6500 HD Nijmegen
Nederland

Anthony S. Gillies gillies@mail.utexas.edu
University of Texas at Austin, Department of Philosophy
Austin, TX, 78712
USA

Philippe de Groote Philippe.de.Groote@loria.fr
INRIA-Lorraine
615 rue du Jardin Botanique - B.P. 101
54602 Villers-les-Nancy Cedex
France

Elena Guerzoni elenag@mit.edu
MIT, Department of Linguistics and Philosophy
22 Magazine Street 6
Cambridge MA, 02139
USA

Chung-hye Han chunghye@sfu.ca
Simon Fraser University, Linguistics Department
8888 University Drive
Burnaby BC V5A1S6
Canada

Juan Heguiabehere juanh@science.uva.nl
University of Amsterdam, ILLC
Nieuwe Achtergracht 166
1018 WV Amsterdam
Nederland

Helen de Hoop H.deHoop@let.kun.nl
Katholieke Universiteit Nijmegen, Algemene Taalwetenschap
Erasmusplein 1
6500 HD Nijmegen
Nederland

Gerhard Jäger jaeger@zas.gwz-berlin.de
Zentrum für Allgemeine Sprachwissenschaft
Jägerstrasse 10/11
D-10117 Berlin
Germany

Theo M.V. Janssen theo@science.uva.nl
University of Amsterdam, ILLC/Comp.Science
Plantage Muidergracht 24
1018 TV, Amsterdam
Holland

Makoto Kanazawa kanazawa@iii.u-tokyo.ac.jp
University of Tokyo, Interfaculty Initiative in Information Studies
7-3-1 Hongo, Bunkyo-ku
Tokyo, 113-0033
Japan

Stefan Kaufmann kaufmann@pine.kuee.kyoto-u.ac.jp
Kyoto University, Graduate School of Informatics
Yoshida-Honmachi, Sakyoku
Kyoto 606-8501
Japan

Ed Keenan ekeenan@ucla.edu
UCLA, Linguistics
UCLA 3125 Campbell Hall
Los Angeles, CA 90095-1543

USA

Alexander Koller koller@coli.uni-sb.de
Saarland University, Department of Computational Linguistics
Postfach 15 11 50
D-66041 Saarbrücken
Germany

Geert-Jan M. Kruijff gj@acm.org
University of the Saarland, Computational Linguistics
Saarbrücken
Germany

Maarten Marx marx@science.uva.nl
University of Amsterdam, ILLC
Plantage Muidergracht 24
1018 TV, Amsterdam
Nederland

Cecile Meier c.meier@lingua.uni-frankfurt.de
Johann Wolfgang Goethe-Universität Frankfurt/Main
Institut für Deutsche Sprache und Literatur II
Grüneborgplatz 1
60629 Frankfurt/Main
Germany

Reinhard Muskens R.A.Muskens@kub.nl
Tilburg University, Department of Linguistics
PO Box 90153
5000 LE Tilburg
Nederland

Østein Nilsen nilsen@let.uu.nl
Utrecht Institute of Linguistics OTS
Achter de Dom 24
3512 JK Utrecht
Nederland

Massimo Poesio poesio@essex.ac.uk
University of Essex, Department of Computer Science
Wivenhoe Park
Colchester CO4 3SQ
UK

Maribel Romero romero@ling.upenn.edu
University of Pennsylvania, Department Linguistics
610 Williams Hall
Philadelphia, PA 19104
USA

Carmen Rotstein carmen@cs.technion.ac.il
Technion - Israel Institute of Technology, Computer Science
Haifa 32000

Israel

Rob van der Sandt R.vanderSandt@phil.kun.nl
University of Nijmegen, Dept. of Philosophy
Postbus 9103
6500 HD Nijmegen
Nederland

Philippe Schlenker schlenke@usc.edu
University of Southern California, Department of Linguistics & Institut Jeon-Nicod
USC
Los Angeles, CA 90089-1693
USA

Bernard Schwarz schwarz@mail.utexas.edu
The University of Texas at Austin, Department of Linguistics
Austin, Texas 78712
USA

Pieter A.M. Seuren pieter.seuren@mpi.nl
Max Planck Institute for Psycholinguistics
PO Box 310
6500 AH Nijmegen
Nederland

Chung-chieh Shan ken@digitas.harvard.edu
Harvard University
33 Oxford Street
Cambridge, MA 02138
USA

Yael Sharvit yael.sharvit@uconn.edu
University of Connecticut, Dept. of Linguistics
Unit 1145
Storrs, CT 06269-1145
USA

Itai Sher
University of Texas at Austin, Department of Philosophy
Austin, TX, 78712
USA

Isabelle Tellier tellier@univ-lille3.fr
Université Charles de Gaulle
Charles de Gaulle - Lille 3
59 653 Villeneuve d'Ascq Cedex
France

Marc Tommasi tommasi@univ-lille3.fr
Université Charles de Gaulle
Charles de Gaulle - Lille 3
59 653 Villeneuve d'Ascq Cedex
France

Madison Williams mawilliams@mail.utexas.edu
University of Texas at Austin, Department of Philosophy
Austin, TX, 78712
USA

Yoad Winter Winter@cs.technion.ac.il
Technion - Israel Institute of Technology, Computer Science
Haifa 32000
Israel

Henk Zeevat zeevat@hum.uva.nl
University of Amsterdam, ILLC
Nieuwe Doelenstraat 15
1012 CP Amsterdam
Nederland