Natural Logic and Semantics

Lawrence S. Moss

Department of Mathematics, Indiana University, Bloomington, IN, USA 47405
lsm@cs.indiana.edu

Abstract. Two of the main motivations for logic and (model-theoretic) semantics overlap in the sense that both subjects are concerned with representing features of natural language meaning and inference. At the same time, the two subjects have other motivations and so are largely separate enterprises. This paper returns to the topic of language and logic, presenting to semanticists *natural logic*, the study of logics for reasoning with sentences close to their surface form. My goal is to show that the subject already has some results that natural language semanticists should find interesting. At the same time it leads to problems and perspectives that I hope will interest the community. One leading idea is that the target logics for translations should have a decidable validity problem, ruling out full rst-order logic. I also will present a fairly new result based on the transitivity of comparative adjective phrases that suggests that in addition to 'meaning postulates' in semantics, we will also need to posit 'proof principles'.

If we were to devise a logic of ordinary language
for direct use on sentences as they come,
we would have to complicate our rules of inference
in sundry unilluminating ways.

W. V. O. Quine, *Word and Object*

1 Natural Logic

By *natural logic*, I mean the study of inference in natural language, done as close as possible to the "surface forms". This work has various flavors and associated projects, and my goal in this talk is to present it to semanticists who know nothing about it. I would like to make the case that natural logic should be of interest in semantics, both the results that we have so far and the problems on the research agenda. I also want to comment at various points on the quote above from Quine, as just one example of an opinion that casts doubt on the whole enterprise of natural logic in first place.

My interest in the topic began in 2005 when I taught an introductory course in semantics for graduate students mainly from our linguistics department, with a few from philosophy and other subjects as well. One motivation for semantics found in textbooks is that it should be the study of *inference in language*: just as syntax has grammaticality judgments to account for, semantics has inference judgments. Now I happen to be mainly a logician, and this point resonates with me as a motivation for semantics. But from what I know about the semantics literature, it *almost never* gives a full account of *any* inferences whatsoever. It is seriously concerned with truth conditions and figuring out how semantics should work in a general way. But it rarely goes back and figures out, for various fragments, what the overall *complete stock of inferences* should be. I wanted to do just this, to introduce logic as another study of inference. In particular, I wanted to give examples of *completeness theorems* that were so elementary that they could be done *without the comparatively heavy syntax* of first-order logic.

Let me give an example of this, a real "toy." Consider sentences *All X are Y*, where X and Y are plural nouns. This is a very tiny fragment, but there certainly are inferences among sentences in it. For example,

$$\frac{\textit{All frogs are reptiles.} \quad \textit{All reptiles are animals.}}{\textit{All frogs are animals.}} \qquad \frac{\textit{All sagatricians are maltnomans.} \quad \textit{All sagatricians are aikims.}}{\textit{All maltnomans are aikims.}}$$

The inference on the left is valid, of course, and the one on the right is invalid. On the right, I have made up the nouns to hammer home the point that the validity or non-validity is not a matter of the nouns themselves, but rather comes from the form *All X are Y*. For sentences in this fragment, we can give an exact semantics: interpret each noun X as a subset $[\![X]\!]$ of an underlying universe M. This gives models. Given such a model, say $\mathcal{M}$, we say that *All X are Y* is *true in* $\mathcal{M}$ if $[\![X]\!] \subseteq [\![Y]\!]$. We can go on to define $\Gamma \models S$, for Γ a set of sentences and S a sentence, by saying that every model of all sentences in Γ is again a model of S. We can ask whether this semantics is adequate in the sense that intuitive judgments of valid inferences, presented in English, are matched by formal statements of the form $\Gamma \models S$. For this fragment, the semantics is basically adequate; the main issue with it is that sentences *All X are X* come out as valid even when the speaker knows or believes that there are no Xs. But putting this aside, the semantics is adequate[1]. Further, one can go on and ask for a *proof-theoretic characterization* of the relation $\Gamma \models S$. Here, it turns out that one can build *proof trees* using the following rules:

$$\frac{}{\textit{All } X \textit{ are } X} \qquad \frac{\textit{All } X \textit{ are } Z \quad \textit{All } Z \textit{ are } Y}{\textit{All } X \textit{ are } Y} \tag{1}$$

We write $\Gamma \vdash S$ if there is a tree all of whose nodes are either labeled from Γ or else match one of the two rules above, and whose root is labeled by S. Then one has the following *completeness theorem*:

Theorem 1 ([12]). *For all Γ and S, $\Gamma \models S$ if and only if $\Gamma \vdash S$*

[1] By the way, one can also change the semantics to require that $[\![X]\!] \neq \emptyset$ in order that *All X are Y* be true. One can make similar modifications to other semantics in the area. The point is that one can work with data provided by real people ignorant of logic and mathematics and then try to find logical systems for such data.

The completeness means that every valid semantic assertion is matched by a formal proof. Nothing is missing. This is not only the simplest completeness theorem in logic, but (returning to the motivations of semantics), it is a full account of the inferential behavior in a fragment. One would think that semanticists would have done this early on.

Then we can ask: given both a semantic account and a proof-theoretic account, why should we prefer the former? Why would we not say that the proof-theory *is* the semantics? After all, it covers the same facts as the semantic account, and it is an account of language use to boot. In addition, it is amenable to a computational treatment.

My suspicion is that inference as such is *not* what really drives semanticists. Just as getting the raw facts of grammaticality right is not the driving force for syntacticians, there are other matters at play. At the end of the day, one wants an explanation of how meaning works in language. And one wants a field that leads to interesting questions. Finally, there are all sorts of theory-internal questions that come up, and for semantics, these questions are not so close to the matter of inference.

In any case, I am interested in asking how far one can go with natural logic. A step up from the tiny fragment of *all* are the classical syllogisms. Here one can return to Aristotle, asking whether his system for *all*, *some*, and *no* (thought of as a formal system) is *complete*. The completeness of various formulations of syllogistic logic has already been shown, for example by in Łukasiewicz [9] (in work with Słupecki), and the basic completeness result was also rediscovered by Westerståhl [24]. There are also different formulations of what Aristotle was doing, and these lead to different completeness results: see Corcoran [4] and Martin [10].

In between the *all* fragment and full syllogistic logic, our paper [12] contains a series of completeness theorems: i) the fragment with *All* X *are* Y; (ii) the fragment with *Some* X *are* Y; (iii) = (i)+(ii); (iv) = (iii) + sentences involving proper names; (v) = (i) + *No* X *are* Y; (vi) *All* + *Some* + *No*; (vii) = (vi) + proper nouns; (viii) boolean combinations of (vii); (ix) = (i) + *There are at least as many* X *as* Y; (x) = boolean combinations of (ix) + *Some* + *No*. In addition, we have a completeness for a system off the main track: (xi) *All* X *which are* Y *are* Z; (xii) *Most* X *are* Y; and (xiii) = (ii) + (xii).

Note that the fragments with *Most* are not expressible in first-order logic. So in this sense, looking at weak fragments gives one more results.

We can go further, in a few ways. First, we can ask about *negation* on nouns, using set complement as the semantics. Then there is the matter of *verbs*, and as an initial step here we would look at *transitive verbs*, using arbitrary relations in the semantics. One could then mix the two enterprises by allowing negation on verbs alone, or on both nouns and verbs. The complete logic of *all*, transitive verbs, and negation on nouns may be found in Figure 1 below. Third, we could study *adjectives* in various ways, especially comparative phrases. We shall see some of this work later.

1.1 Objections to natural logic

I want to return to the quote from Quine at the beginning, and to put forth several reasons[2] why one might agree with it.

A. The logical systems that one would get from looking at inference involving surface sentences would contain many copies of similar-looking rules. Presenting things in this way would miss a lot of generalizations.
B. The systems would contain 'rules' that are not really rules at all, but instead are more like complex deduction patterns that need to be framed as rules only because one lacks the machinery to break them down into more manageable sub-deductions. Moreover, those complex rules would be unilluminating.
C. The systems would lack *variables*, and thus they would be tedious and inelegant.
D. Turning to the standard topic of quantifier-scope ambiguities, it would be impossible to handle inferences among sentences exhibiting this phenomenon in an elegant way.

My feeling is that all of these objections are to some extent apt, and to some extent miss the mark.

The first two points might be illustrated by the logic in Figure 1, a logic for sentences in the fragment shown. I have used *see* as a generic transitive verb just to simplify the presentation. I also have used the prime symbol $'$ for complement. But I intend this as a kind of variable over transitive verbs. We could as well write *All* X V *all* Z.

Here is an example of the kind of inference which could be captured in the system:

$$\begin{array}{l} \textit{All xenophobics hate all actors} \\ \textit{All yodelers hate all zookeepers} \\ \textit{All non-yodelers hate all non-actors} \\ \textit{All wardens are xenophobics} \\ \hline \textit{All wardens hate all zookeepers} \end{array} \tag{2}$$

Here is formal derivation corresponding to (2), using the rules in Figure 1:

$$\dfrac{\dfrac{\textit{All } X \textit{ hate all } A \quad \textit{All } Y \textit{ hate all } Z \quad \textit{All } Y' \textit{ hate all } A'}{\textit{All } X \textit{ hate all } Z}\,\textit{3pr} \quad \textit{All } W \textit{ are } X}{\textit{All } W \textit{ hate all } Z}$$

Figure 1 itself does not list all of the rules; the *monotonicity* rules are missing. For this fragment, there would be two of them: the first is the transitivity of *all* noted in (1) and also called Barbara in traditional syllogistics. The second is

$$\dfrac{\textit{All } X \textit{ are } U \quad \textit{All } U \textit{ see all } Z \quad \textit{All } Y \textit{ are } Z}{\textit{All } X \textit{ see all } Y} \tag{3}$$

[2] These objections are my formulations. I would not want to give the impression that Quine or anyone else agreed with them. In another direction, I do have to wonder how anyone could see what logic for sentences as they come would look like without actually doing it.

$$\frac{\textit{All } Y \textit{ are } Y'}{\textit{All } Y \ VP}\ \textit{Zero} \qquad \frac{\textit{All } Y' \textit{ are } Y}{\textit{All } X \textit{ are } Y}\ \textit{One}$$

$$\frac{\textit{All } Y \textit{ are } X'}{\textit{All } X \textit{ are } Y'}\ \textit{Antitone} \qquad \frac{\textit{All } Y \textit{ are } Y'}{\textit{All } X \textit{ see all } Y}\ \textit{Zero-VP}$$

$$\frac{\textit{All } X \textit{ see all } Y \quad \textit{All } X' \textit{ see all } Y}{\textit{All } Z \textit{ see all } Y}\ \textit{LEM}$$

$$\frac{\textit{All } X \textit{ see all } Y \quad \textit{All } X \textit{ see all } Y'}{\textit{All } X \textit{ see all } Z}\ \textit{LEM}'$$

$$\frac{\textit{All } X \textit{ see all } A \quad \textit{All } Y \textit{ see all } Z \quad \textit{All } Y' \textit{ see all } A'}{\textit{All } X \textit{ see all } Z}\ \textit{3pr}$$

Fig. 1. The *All* syllogistic logic with verbs and noun-level complements, leaving off the rules in (1) and the monotonicity rules *All* $X^{\downarrow}$ *are* $Y^{\uparrow}$ and *All* $X^{\downarrow}$ *see all* $Y^{\downarrow}$.

This is captured in the notation

$$\textit{All } X^{\downarrow} \textit{ see all } Y^{\downarrow}.$$

It was Johan van Benthem who first used this notation in [2]. (His work, and work influenced by it, is an important source of results and inspirations in the area, but I lack the space to discuss it.) Even more importantly, it was he who first recognized the importance of *monotonicity rules* for fragments of this form. The similarity of (3) and the Barbara rule from (1) illustrates objection (A): having both rules misses a generalization. At the same time, there is a rejoinder: using the arrow notation, or other "meta-rules", we can say what we want. Nevertheless, for some more complicated systems it is an open issue to present them in the "optimally informative" way. Indeed, it is not even clear what the criteria for such presentations should be.

Objections (A) and (B) are illustrated in the last rule in the figure, and to some extent in the two rules above it. The last rule says, informally, that if all X see all A, all Y see all Z, and all non-Y see all non-Z, then all X see all Z as well. Why is this rule sound? Well, take some $x \in X$. Then if this x is also a Y, then it sees all Z. Otherwise, x is a non-Y. But then x sees all non-A's. And since x was an X, it sees all A as well. Thus in this case x sees absolutely everything, a fortiori all Z.

This is not a familiar rule, and I could think of no better name for it than (3pr), since it has three premises. It is hard to take this to be a single rule, since it lacks the intuitively obvious status of some of the monotonicity rules. When presented to audiences or classes, hardly anyone believes that it is sound to begin with. Moreover, like the two *law of the excluded middle* rules, it really depends on the semantics of X' giving the *full* complement; so it might not even be the rule one always wants in the first place. But if one is committed to the semantics, one has to take it as a rule on its own because it cannot be simplified any further. In any case, I agree that the rule itself is probably not so illuminating.

Objection (C) is that systems for natural logic lack variables. I must again say that this is my point, and I make it to advance the discussion. It should be of interest in semantics to know exactly where variables really are needed, and to formulate logical systems that do involve variables.

It would take us too far afield in this short report to discuss objection (D). But one could see syllogistic logics capable of handling the different scope readings in ambiguous sentences, and yet do not have much in the syntax besides the disambiguation: Nishihara et al. [15], and Moss [13].

2 The Aristotle Boundary

Ian Pratt-Hartmann and I determined in [18] what I'll call the *Aristotle boundary.* This is the limit of how far one can go with purely syllogistic systems. We need some notation for logical systems taken from the paper.

$\mathcal{S}$	classical syllogistic: *all/some/no X are Y*
$\mathcal{S}^\dagger$	$\mathcal{S}$ with negation on nouns: *non-X*
$\mathcal{R}$	relational syllogistic: add transitive verbs to $\mathcal{S}$
$\mathcal{R}^\dagger$	relational syllogistic with noun-negations
$\mathcal{R}^*$	relational syllogistic, allowing subject NPs to be relative clauses
$\mathcal{R}^{*\dagger}$	relational syllogistic, again allowing subject NPs to be relative clauses and full noun-negation

In more detail, the syntax of $\mathcal{R}$ is

All X are Y	*All X aren't Y* $\equiv$ *No X are Y*
Some X are Y	*Some X aren't Y*
All X see all Y	*All X don't see all Y* $\equiv$ *No X sees any Y*
All X see some Y	*All X don't see some Y* $\equiv$ *No X sees all Y*
Some X see all Y	*Some X don't see any Y*
Some X see some Y	*Some X don't see some Y*

$\mathcal{R}^*$ allows the subject noun phrases to contain relative clauses of the form

who see all X	*who see some X*
who don't see all X	*who don't see some X*

Finally, $\mathcal{R}^{*\dagger}$ has full negation on nouns.

Theorem 2 ([18]). *There are complete syllogistic systems for $\mathcal{S}$ and $\mathcal{S}^\dagger$.*

There are no finite, complete syllogistic systems for $\mathcal{R}$. However, allowing reductio ad absurdum*, there is a syllogistic system for $\mathcal{R}$.*

Even allowing reductio ad absurdum*, there are no finite, complete systems for $\mathcal{R}^\dagger$ or for $\mathcal{R}^{*\dagger}$.*

These results begins to delimit the Aristotle boundary. It has much to do with negation, especially noun negation in connection with verbs.

Despite the negative results at the end of Theorem 2, the systems involved are *decidable*. This means that in principle one could write a computer program to decide whether a purported inference was valid or not. The complete story here is that the complexity of the validity problem for these logics is known.

Theorem 3 ([18]). *The validity problems for $\mathcal{S}$, $\mathcal{S}^\dagger$, and $\mathcal{R}$ are complete for nondeterministic logspace; for $\mathcal{R}^\dagger$, it is complete for deterministic exponential time; $\mathcal{R}^*$ for co-NPtime [11], and $\mathcal{R}^{*\dagger}$ for nondeterministic exponential time.*

Now, one can ask several questions. First, do the complexity results have any cognitive relevance? This seems to me to be a very good question, and it seems completely open. Second, one could ask for the *average-case* complexity results and to again ask for their cognitive relevance.

My feeling overall is that the Aristotle boundary should be of interest in semantics partly because of the prominence of *variables* in contemporary semantics. It would be good to pinpoint the features of language that necessitate going beyond a syllogistic presentation. This is what the results in [18] say. However, it should be noted that they do *not* say that one *must* use variables in the traditional way, only that one cannot do with a purely syllogistic presentation. In fact, one can also define logical systems for fragments like $\mathcal{R}^*$ and $\mathcal{R}^{*\dagger}$ which use something like variables, but with more restrictions. These restrictions correspond to the decidability of the system, a point which I return to in Section 3.

2.1 Fitch's "Natural Deduction Rules for English"

I would like to mention Fitch [6] as an early source on natural logic. This paper is not very well-known among people in the area, and I have seen few references to it by semanticists or anyone else for that matter. Frederic Fitch was one of the first people to present natural deduction proofs in what we call 'Fitch style'; Stanisław Jaśkowski also did this. For a good discussion of the history, see Pelletier [17]. Fitch's paper of 1973 presents a set of natural deduction rules for English. Figure 2 contains an example taken directly from his paper.

It should be noted that there is no formal syntax in the paper. His rules for *any* are thus ad hoc, and certainly there is more that one should say beyond his rules; they do show that he was aware of what we now call polarity phenomena. This lack of syntax is not terribly surprising, since he might not have known of Montague's work. But in addition there is no formal semantics either. From the point of view of natural logic, one can return to Fitch's paper and then ask whether his rules are complete. This question is open.

3 The Force of Decidability

I mentioned above that the Aristotle boundary should be of some interest in semantics. I want to end with discussion of the corresponding "Turing boundary". This would be the boundary between decidable and undecidable fragments.

1		John is a man	Hyp
2		Any woman is a mystery to any man	Hyp
3	Jane	Jane is a woman	Hyp
4		Any woman is a mystery to any man	R, 2
5		Jane is a mystery to any man	Any Elim, 4
6		John is a man	R, 1
7		Jane is a mystery to John	Any Elim, 6
8		Any woman is a mystery to John	Any intro, 3, 7

Fig. 2. An example from Fitch [6]

My feeling is that this boundary should be even more important to investigate. Formally-minded linguists should be more used to the rejection of undecidable frameworks following the Peters-Ritchie Theorem in formal language theory. There are certainly some who feel that semantics should make use of the strongest possible logical languages, presumably on the grounds that human beings can understand them anyways. But a wealth of experience stemming from computer science and cognitive science leads in the opposite direction. The feeling is that "everyday" deduction in language is not the same as mathematics, it might not call on the same mental faculty as deep reasoning in the first place. So one should investigate weak systems with an eye towards seeing what exactly can be said in them, before going on to more expressive but undecidable systems.

All of the logical systems mentioned so far in this paper have been decidable, including ones which need variables. (Incidentally, these fragments sometimes do not have the finite model property.) I am interested in finding yet stronger decidable fragments, and so this is how I end this paper. (For other work in the area, see Pratt-Hartmann [19, 19] and Pratt-Hartmann and Third [21].) One source of such stronger systems is *comparative adjective phrases*, such as *bigger than*, *smaller than*, and the like. These are always interpreted by *transitive relations* on a domain:

$$\text{If } a \text{ is bigger than } b\text{, and } b \text{ is bigger than } c\text{, then } a \text{ is bigger than } c. \qquad (4)$$

(The interpretations are also irreflexive: nobody is bigger than themselves. But this fact will not be relevant to our point in this section.) The transitivity corresponds to the validity of arguments like the following:

$$\frac{\textsf{Every sweet fruit is bigger than every kumquat}}{\textsf{Every fruit bigger than some sweet fruit is bigger than every kumquat}} \qquad (5)$$

That is, (5) is semantically valid, but only on the class of models which interpret *bigger than* by a transitive relation.

Now one might at first thing that what we need is a logical system which directly expresses transitivity using variables in some version of (4). We are already heading towards the use of variables, so what is the problem with (4)? The hitch is that (4) uses *three variables*, and it is known that a a logical system which can express all sentences in three variables is *undecidable*. Even more, a system which can express all of the *two-variable* sentences plus assertions of transitivity (as atomic sentences) is again undecidable, by a theorem of Grädel, Otto, and Rosen [8]. So if we believe that "simple" fragments of language should lead to decidable logics, then we cannot use a language which states (4) in a "first-class" way. Here is how this is done in [14]. The system uses variables, and also natural-deduction style rules. For transitivity, it uses

$$\frac{a(t_1,t_2) \quad a(t_2,t_3)}{a(t_1,t_3)}\ \text{trans}$$

Here a is an adjective phrase (it will be bigger below), and the t's are *terms* (variables, roughly). The derivation corresponding to (5) is

$$\cfrac{\cfrac{[\exists(\mathsf{sw},\mathsf{bigger})(x)]^3 \qquad \cfrac{\cfrac{[\mathsf{bigger}(x,y)]^2 \qquad \cfrac{[\mathsf{kq}(z)]^1 \qquad \cfrac{[\mathsf{sw}(y)]^2 \quad \forall(\mathsf{sw},\forall(\mathsf{kq},\mathsf{bigger}))}{\forall(\mathsf{kq},\mathsf{bigger})(y)}\ \forall E}{\mathsf{bigger}(y,z)}\ \forall E}{\mathsf{bigger}(x,z)}\ \text{trans}}{\forall(\mathsf{kq},\mathsf{bigger})(x)}\ \forall I^1}{\forall(\mathsf{kq},\mathsf{bigger})(x)}\ \exists E^2}{\forall(\exists(\mathsf{sw},\mathsf{bigger}),\forall(\mathsf{kq},\mathsf{bigger}))}\ \forall I^3$$

In the derivation, sentences like $\forall(\exists(\mathsf{sw},\mathsf{bigger}),\forall(\mathsf{kq},\mathsf{bigger}))$ stand for "everything which is bigger than some sweet fruit is bigger than every kumquat." The derivation also uses variables and temporary hypotheses. For example, $\exists(\mathsf{sw},\mathsf{bigger})(x)$ corresponds to a proof step like "let x be bigger than some sweet fruit." Even with all of this, the system is decidable. But again, the point I wish to make is that transitivity is a rule, not an axiom. This suggests an issue for semantics: what other constructions work this way?

Transitivity also plays a role in the recent literature on (of all things) avian cognition: see Guillermo Paz-y-Miño C et al [16]. For a different cognitive science connection of monotonicity rules to the modeling of inference (in humans), see Geurts [7].

Note on the references I have included in the references below many more papers than I actually reference in this, as a way of indicating much of what has actually been done in the area. Much of the history appears in van Benthem [3].

References

1. Gilad Ben Avi and Nissim Francez. Proof-theoretic semantics for a syllogistic fragment. In Paul Dekker and Michael Franke (eds.), *Proceedings of the Fifteenth Amsterdam Colloquium*, ILLC/Department of Philosophy, University of Amsterdam, 2005, 9–15.

2. Johan van Benthem. *Essays in Logical Semantics*. Reidel, Dordrecht, 1986,
3. Johan van Benthem. A brief history of natural logic. In M. Chakraborty, B. Löwe, M. Nath Mitra and S. Sarukkai, eds., *Logic, Navya-Nyaya and Applications, Homage to Bimal Krishna Matilal*. College Publications, London, 2008.
4. John Corcoran. Completeness of an ancient logic. *Journal of Symbolic Logic*, 37(4):696–702, 1972.
5. George Englebretsen. *Three Logicians*. Van Gorcum, Assen, 1981.
6. Frederic B. Fitch. Natural deduction rules for English. *Philosophical Studies*, 24:2 (1973), 89–104.
7. Bart Geurts. Reasoning with quantifiers. *Cognition* 86: 223-251, 2003.
8. Erich Grädel, Martin Otto, and Eric Rosen. Undecidability results on two-variable logics. Archive for Mathematical Logic, vol. 38, pp. 313–354, 1999.
9. J. Łukasiewicz. *Aristotle's Syllogistic*. Clarendon Press, Oxford, 2nd edition, 1957.
10. John N. Martin. Aristotle's natural deduction revisited. *History and Philosophy of Logic*, 18(1):1–15, 1997.
11. David A. McAllester and Robert Givan. Natural language syntax and first-order inference. *Artificial Intelligence*, 56:1–20, 1992.
12. Lawrence S. Moss. Completeness theorems for syllogistic fragments. In F. Hamm and S. Kepser, eds., *Logics for Linguistic Structures*, Mouton de Gruyter, 143–173, 2008.
13. Lawrence S. Moss. Syllogistic logics with verbs. *Journal of Logic and Computation, to appear*, 2010.
14. Lawrence S. Moss. Logics for two fragments beyond the syllogistic boundary To appear in A. Blass et al (eds.), Studies in Honor of Yuri Gurevich, Lecture Notes in Computer Science, Springer-Verlag, Berlin, 2010.
15. Noritaka Nishihara, Kenich Morita, and Shigenori Iwata. An extended syllogistic system with verbs and proper nouns, and its completeness proof. *Systems and Computers in Japan*, 21(1):760–771, 1990.
16. Guillermo Paz-y-Miño C, Alan B. Bond, Alan Kamil, and Russell P. Balda. Pinyon jays use transitive inference to predict social dominance. *Nature* 430, pp. 778–781, 2004.
17. F. J. Pelletier. A Brief History of Natural Deduction *History and Philosophy of Logic* 20, pp. 1–31, 1999.
18. Ian Pratt-Hartmann and Lawrence S. Moss Logics for the relational syllogistic. *Review of Symbolic Logic*, to appear 2009.
19. Ian Pratt-Hartmann. A two-variable fragment of English. *Journal of Logic, Language and Information*, 12(1), 2003, pp. 13–45.
20. Ian Pratt-Hartmann. Fragments of Language. *Journal of Logic, Language and Information*, 13:207–223, 2004.
21. Ian Pratt-Hartmann and Allan Third. More fragments of language. *Notre Dame Journal of Formal Logic* 47:2 (2006).
22. William C. Purdy. A logic for natural language. *Notre Dame Journal of Formal Logic* 32 (1991), no. 3, 409–425.
23. Fred Sommers. *The Logic of Natural Language*. Clarendon Press, Oxford, 1982.
24. Dag Westerståhl. Aristotelian syllogisms and generalized quantifiers. *Studia Logica*, XLVIII(4):577–585, 1989.
25. Anna Zamansky, Nissim Francez, Yoad Winter. A 'Natural Logic' inference system using the Lambek calculus. *Journal of Logic, Language and Information* 15(3): 273-295 (2006)