Ignorance under attitudes

Patrick D. Elliott
Heinrich-Heine University Düsseldorf
Düsseldorf, Germany
patrick.d.elliott@gmail.com

Filipe Hisao Kobayashi
Paris Lodron University of Salzburg
Salzburg, Austria
filipe.kobayashi@plus.ac.at

Abstract

Yablo (2014) observed that *Ann and Beth agree that p or q* conveys that p and q are both compatible with the beliefs of Ann and Beth. While Yablo argued for a radical departure from possible world semantics, we propose that this is due to local pragmatic reasoning. We frame our proposal within Update Semantics and develop a system where *p or q* requires the context to be compatible with both p and q. When disjunction is unembedded, our system yields ignorance inferences. Yablo's puzzle is accounted for because, under attitude verbs the requirement is imposed in the complement clause's local context.

1 Yablo's problem

Consider the following valid argument, based on a case discussed by Yablo (2014, p. 12). Even when x and y's individual beliefs differ, we can still truthfully say that x and y *agree* on certain propositions.

(1) a. Stephen believes [that snow is white and expensive].
b. Daniel believes [that snow is white, but not expensive].
c. (1-a) ∧ (1-b) ⇒ Stephen and Daniel agree [that snow is white].

A first stab at a semantics for *agree* that captures validity of (1) is given in (2) — it says that x and y *agree* on ϕ if each of x and y individually believe the proposition expressed by ϕ.[1] This is effective simply because a sentence of the form "ϕ and ψ" classically entails both ϕ, and ψ.

(2) Truth-condition for an *agree*-report (first attempt):
$[\![x \text{ and } y \text{ agree that } \phi]\!]^w = 1 \text{ if } \forall w' \in Dox^x_w, [\![\phi]\!]^{w'} = 1 \wedge \forall w' \in Dox^y_w, [\![\phi]\!]^{w'} = 1$

As pointed out by Yablo (2014), this simple-minded assumption about the semantics of *agree* immediately leads to strange results. For example, it predicts the validity of (3), which does not seem to be intuitively correct. The prediction straightforwardly follows from the classical validity of disjunction introduction: both "snow is white" and "snow is black" classically entail that "snow is white or black".

(3) a. Stephen believes [that snow is white].
b. Daniel believes [that snow is black].
c. Stephen and Daniel agree [that snow is white or black].

It is difficult to see how to block this prediction with classical resources while maintaining an account of (1), but it is very natural to suspect that the use of disjunction in (3-c) leads to additional non-semantic inferences that contribute to the sense that (3) is not valid. To spell this out, it is useful to consider a very simple case of disjunction embedded in a *belief* report.

(4) a. Daniel believes [that snow is white].

[1] We use Dox^x_w to talk about the set of worlds that x considers a candidate for the actual world in w, i.e., x's *doxastic alternatives*.

b. Daniel believes [that snow is white or black].

(5) $[\![x \text{ believes } \phi]\!]^w = 1 \text{ iff } \forall w' \in Dox^x_w, [\![\phi]\!]^{w'} = 1$

Intuitively, (4) is felt not to be valid, despite what is predicted by a relatively orthodox, Hintikkan semantics for *believe* Hintikka 1969, (5). There is however a pragmatic explanation for the feeling of non-validity, which follows from basic Gricean reasoning (Grice 1989). Let's make this concrete: the sentence in (4-b) evokes, among others, alternative *belief* reports involving the individual disjuncts, which are each presumably relevant to the current topic of conversation, (6-a). Given that the speaker has chosen to assert (4-b) rather than the logically stronger alternatives in (6-a), on the assumption that she is operating in accordance to the Maxim of Quantity, we can conclude that, for each $\psi \in$ (6-a), it's not the case that she believes ψ. Together with (4-b), the resulting implicature conveys speaker ignorance about the status of each alternative.

(6) a. $\{$ D. believes that snow is white, D. believes that snow is black $\} \subseteq$ *Alt*(4-b)

We can now consider the disjunctive *belief*-report together with its implicatures (7-a) and (7-b). Note that the resulting *enriched* meaning of the disjunctive *belief* report is incompatible with the each simpler alternative. This is plausibly responsible for the feeling that (4) is not valid.

(7) Daniel believes that snow is white or black.

a. $\rightsquigarrow$ *the speaker isn't certain that Daniel believes that snow is white.*
b. $\rightsquigarrow$ *the speaker isn't certain that Daniel believes that snow is black.*

This reasoning however does not extend to Yablo's problem (see Rothschild 2017 for discussion). To see this, consider the simpler alternatives evoked by a disjunctive *agree* report. In a context of (3), each of Alt_1 and Alt_2 are already known to be false, since Daniel and Stephen have incompatible beliefs about the color of snow.

(8) Daniel and Stephen agree [that snow is white or black].

a. Alt_1: *Daniel and Stephen agree that snow is white.*
b. Alt_2: *Daniel and Stephen agree that snow is black.*

Yablo uses the problem of agreement, among many other cases, to motivate a theory of partial content, couched in truthmaker semantics (see also Fine 2017). Within the framework of truthmaker semantics, a non-classical notion of entailment can be stated according to which "ϕ and ψ" entails "ϕ", and "ψ", but "ϕ" doesn't entail "ϕ or ψ". In this paper however we'll suggest that it would be too hasty to dismiss a pragmatic explanation for the invalidity of (3), in light of the possibility of local pragmatic enrichment. We turn to this notion in the next section.

2 Local pragmatic enrichment

2.1 Background: minimal worlds exhaustification

There is a significant body of evidence indicating that class of inferences traditionally derived as Gricean conversational implicatures are instead amenable to an analysis in terms of an embeddedable, covert operator *Exh* (see Chierchia, Fox, and Spector 2012 for an overview). One such inference is the so-called *scalar implicature* associated with indefinites

(9) Alex ate some of the cookies.

a. $\rightsquigarrow$ *Alex didn't eat all of the cookies*

There are various formulations of *Exh* (see especially Spector 2016 for discussion), and the one that we build upon in the present work is due to van Rooij and Schulz (2004): the so called *minimal worlds* exhaustivity operator (Exh_{mw}). As is standard, we relativize the interpretation of Exh_{mw} applied to a sentence ϕ to the set of *alternatives* to ϕ, $Alt(\phi)$. The definition of Exh_{mw} proceeds in two steps. First we define an ordering on worlds relative to $Alt(\phi)$: $w \leq_{Alt(\phi)} v$ when every element of $Alt(\phi)$ that is true in w is also true in v.

(10) $w \leq_{Alt(\phi)} v := \{\, \psi \in Alt(\phi) \mid [\![\psi]\!]^w = 1 \,\} \subseteq \{\, \psi \in Alt(\phi) \mid [\![\psi]\!]^v = 1 \,\}$

To see how this works, it will be useful to consider a schematic example. Consider a disjunctive sentence $a \vee b$. The standardly assumed alternatives (Sauerland 2004) are listed below:

(11) $Alt(a \vee b) = \{\, a \vee b, a \wedge b, a, b \,\}$

Assume the following worlds, where subscripts indicate which disjuncts are true: $w_{ab}, w_a, w_b, w_\emptyset$. First, let's determine which alternatives are true in which worlds.

- $\{\, \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_{ab}} = 1 \,\} = \{\, a \vee b, a \wedge b, a, b \,\}$
- $\{\, \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_a} = 1 \,\} = \{\, a \vee b, a \,\}$
- $\{\, \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_b} = 1 \,\} = \{\, a \vee b, b \,\}$
- $\{\, \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_\emptyset} = 1 \,\} = \emptyset$

The result is that the worlds which make the *fewest alternatives true* are the most minimal:

(12) $w_\emptyset < w_a, w_b < w_{ab}$

Now, minimal worlds exhaustification can be derivatively defined: it's true in any world w such that (i) the prejacent is true (i.e., the assertion) in w, and (ii) there is no world v that is more minimal than w and the prejacent is true in v.

(13) **Minimal worlds exhaustification:**
$[\![Exh_{mw}\, \phi]\!]^w = 1$ iff $\underbrace{[\![\phi]\!]^w = 1}_{\text{assertion}}$ and $\underbrace{\neg\exists v[[\![\phi]\!]^v = 1 \wedge v <_{Alt(\phi)} w]}_{\text{implicature}}$

Ignoring $w_\emptyset$ (since there, the prejacent is false), the minimal worlds in which the prejacent is true are w_a and w_b. It follows that the proposition expressed by $Exh_{mw}(a \vee b)$ is $[\lambda w \,.\, w \in \{\, w_a, w_b \,\}]$, i.e., either a or b is true, but not both — the scalar implicature is derived.

An interesting property of minimal worlds exhaustification is that it does not rely on the presence of the scalar alternative $a \wedge b$ — the so-called domain alternatives are sufficient in order to derive the scalar implicature. This is easy to see — even ignoring $a \wedge b$, both w_a and w_b are more minimal than w_{ab} with respect to alternatives $\{\, a, b \,\}$, since w_{ab} makes both alternatives true, and each of w_a and w_b make just one alternative true.

As it stands, note that minimal worlds exhaustification does not predict so-called ignorance inferences associated with disjunction. If a speaker asserts $a \vee b$, it typically implicates not just that she is certain that $a \wedge b$ is false (the scalar implicature), but also that she is not certain that a, and is not certain that b (the ignorance inferences). The classical strategy is to derive ignorance resources via Gricean reasoning (see, e.g., Fox 2007 for an explicitly hetereogeneous approach to scalar implicatures and ignorance inferences). Our main innovation will be to extend minimal worlds exhaustification in order to capture both of these inferences.

2.2 Exhaustification and local contexts

Since ignorance inferences amount to requirements on information states, it is not possible to develop an embeddable account of ignorance inferences without reference to a *local context*. We implement local contexts in a straightforward way by adopting *update semantics*; sentences describe ways of updating information states (Heim 1982; Dekker 1993). Certain operators, such as the logical connectives, modals/attitudes, and (crucially for us) *Exh*, are defined directly in terms of updates.

(14) For all sentences ϕ not defined directly as updates,
$c[\phi] := \{ w \in c \mid [\![\phi]\!]^w = 1 \}$

(15) For all sentences ϕ directly defined as updates,
$[\![\phi]\!]^w = 1$ iff $w \in W[\phi]$

(16) Updates for complex sentences:

a. $c[\phi \wedge \psi] := c[\phi][\psi]$
b. $c[\phi \vee \psi] := c[\phi] \cup c[\psi]$
c. $c[\neg\phi] := c - c[\phi]$

With the additional expressive power of update semantics, we will define a new formulation of minimal worlds exhaustification, crucially stated with respect to a local context c. We follow Bassi, Del Pinal, and Sauerland 2021 in that, like other presupposition triggers, *Exh* places constraints on c.[2] First we'll provide the definition of *Exh* we assume, and then we'll go through it step by step:

(17) $$c[Exh(\phi)] := \begin{cases} c[\phi] & \text{if } \forall\psi \in Alt(\phi)[\exists w[[\![Exh_{mw}\,\phi]\!]^w = 1 \wedge [\![\psi]\!]^w = 1] \leftrightarrow c[\psi] \neq \emptyset] \\ \text{undefined} & \text{otherwise} \end{cases}$$

$Exh(\phi)$ places a definedness condition on c. Informally, the definedness condition placed on c by $Exh(\phi)$ is that the alternatives in $Alt(\phi)$ that are compatible with c are exactly those alternatives in $Alt(\phi)$ that are compatible with $Exh_{mw}(\phi)$. If this definedness condition is met, then the context is simply updated with the prejacent ϕ.

Let's see how this works for disjunction. We already know that $Exh_{mw}(a \vee b)$ is the proposition is true in worlds where only one of a and b is true. We now must compute the definedness condition on c triggered by $Exh(a \vee b)$:

- $a \wedge b$ is logically incompatible with $Exh_{mw}(a \vee b)$, hence $c[a \wedge b] = \emptyset$.
- $a \vee b$ is logically compatible with $Exh_{mw}(a \vee b)$, hence $c[a \vee b] \neq \emptyset$.
- a is logically compatible with $Exh_{mw}(a \vee b)$, hence $c[a] \neq \emptyset$.
- b is logically compatible with $Exh_{mw}(a \vee b)$, hence $c[b] \neq \emptyset$

Thus:

(18) $$c[Exh\,(a \vee b)] = \begin{cases} c[a \vee b] & \text{if } c[a \wedge b] = \emptyset \wedge c[a] \neq \emptyset \wedge c[b] \neq \emptyset \wedge c[a \vee b] \neq \emptyset \\ \text{undefined} & \text{otherwise} \end{cases}$$

Updating with $Exh\,(a \vee b)$ therefore requires that $a \wedge b$ is known to be false (the scalar implicature), that the domain alternatives a, b are considered possible, and that the prejacent is considered

[2] See Bassi, Del Pinal, and Sauerland 2021 for discussion of this treatment of scalar implicatures as presuppositions.

possible. N.b., that *contextual ignorance* only follows in conjunction with the scalar implicature (see Degano et al. 2024 for experimental evidence that *possibility* should be derived separately from uncertainty).

3 Accounting for Yablo's problem

Back to Yablo's problem — recall that a simple-minded understanding of what it means for two individuals x and y to *agree* on a proposition ϕ is that both of x's and y's beliefs individually entail ϕ. This works for some simple cases involving disjunctive beliefs, but incorrectly predicts that x and y may agree on disjunctive propositions that are weaker than their respective beliefs, due to properties of classical entailment. Our central claim will be that Yablo's problem can be resolved without shifting to a non-classical notion of entailment by invoking Exh, in order to disrupt the predicted entailment. Concretely, we'll assume that the Logical Form of an *agree* report is as follows:

(19) Stephen and Daniel agree [Exh that snow is white or black].

In order to figure out what our analysis predicts, we need a principled account of the local context of *agree*, based on how presuppositions project from the complement of *agree*. Our contention is that "x and y agree that ϕ_π" (where π is the presupposition of ϕ) presupposes that each of x and y believe that π, mirroring the projection properties of other attitude verbs (Heim 1992). We demonstrate this via standard filtering tests:

(20) Context: *Sally doesn't smoke.*
 a. Both Stephen and Daniel believe that Sally used to smoke, and they agree that she stopped.
 b. Both Stephen and Daniel believe that Sally used to smoke, but they don't agree that she stopped.

Following Heim's entries for attitude verbs, we build this behavior into the update contributed by the *agree* report, as in (21). It presupposes that the context entails that x and y both individually believe the presuppositions of ϕ, and asserts that both x and y's beliefs entail that ϕ. In other words, we maintain a simple entailment-based semantics for *agree*, while being a bit more precise about how presuppositions project.[3]

(21) $c[x \text{ and } y \text{ agree } \phi] \neq \#$ when $\forall w \in c, Dox^x_w[\phi]$ is defined $\wedge$ $Dox^y_w[\phi]$ is defined.
When defined, $c[x \text{ and } y \text{ agree } \phi] = \{\ w \in C \mid Dox^x_w[\phi] = Dox^y_w \wedge Dox^y_w[\phi] = Dox^y_w\ \}$

We can now resolve Yablo's problem. Consider what is predicted as the update associated with "Exh that snow is white or black" ($Exh(Wh \vee Bl)$): it asserts that snow is white or black, presupposes that snow isn't both white and black, and that it's contextually possible that snow is white, and possible that snow is black.

(22) $c[Exh\,(Wh \vee Bl)] \neq \#$ only if $\neg\exists w \in c[Wh(w) \wedge Bl(w)], \exists w \in c[Wh(w)], \exists w \in c[Bl(w)]$

Now we can compute the result of updating the global context with "Stephen and Daniel agree that [Exh Snow is white or black]". What we predict is that this presupposes that Stephen and Daniel both consider it false that Snow is white and black, and that they each consider it possible that Snow is white, and possible that snow is black. This is demonstrated formally below:

[3] We set aside the presuppositions triggered by *agree* as they are not relevant for this particular set of data.

(23) $c[s \text{ and } d \text{ agree that } Exh\,(Wh \vee Bl)] \neq \#$ only if:
$\forall w \in c, \neg\exists w' \in Dox^s_w[Wh(w) \wedge Bl(w)], \exists w \in Dox^s_w[Wh(w)], \exists w \in Dox^s_w[Bl(w)],$
$\neg\exists w' \in Dox^d_w[Wh(w) \wedge Bl(w)], \exists w \in Dox^d_w[Wh(w)], \exists w \in Dox^d_w[Bl(w)]$
When defined, $c[s \text{ and } d \text{ agree that } Exh\,(Wh \vee Bl)]$
$= \{\, w \in c \mid \forall w \in Dox^s_w, Wh(w) \vee Bl(w) \text{ and } \forall w \in Dox^d_w, Wh(w) \vee Bl(w)\}$

Intuitively, the reasons that this doesn't jointly follow from "Stephen believes that snow is white", and "Daniel believes that snow is black", is because former doesn't entail that Stephen considers it possible that snow is black, and the latter doesn't entail that Daniel considers it possible that snow is white. To be concrete, we assume a simple Heimian semantics for *belief* reports as in (24).[4]

(24) $$c[x \text{ believes } \phi] = \begin{cases} \{\, w \in c \mid Dox^x_w[\phi] = Dox^x_w \,\} & \forall w \in c, Dox^x_w[\phi] \text{ is defined} \\ \text{undefined} & \text{otherwise} \end{cases}$$

4 Conclusion

In this paper, we have showed that Yablo's problem can be resolved once we take into consideration how implicatures are calculated within embedded clauses. We proposed a novel way of deriving the ignorance implicatures of disjunction via a silent *Exh* operator whose semantics imposes constraint on its local context.

A clear advantage of our proposal is that it can account for the observation of Rothschild (2017) that the following entailment goes through:

(25) Ann believes Beth ate the cookies
Cleo believes Deb ate the cookies ⇒ Ann and Cleo agree that a girl ate the cookies.

Our approach predicts that the validity of the inference in (25) should directly track the robustness of the corresponding *ignorance inference*. Here, we observe that there is a clear contrast between (26-a) and (26-b) — conveys that both *Cleo eating the cookies*, and *Beth eating the cookies* are open possibilities. (26-a) does not always lead to the same inference; for example, perhaps it is relevant to the topic of conversation that a *girl* rather than a *boy* ate the cookies.

(26) a. A girl ate the cookies.
b. Cleo or Beth ate the cookies.

Acknowledgements. We're grateful to Alexandros Kalomoiros, Jacopo Romoli, and Raven Zhang. This research was funded in whole or in part by the Austrian Science Fund (FWF) 10.55776/F1003.

References

Bassi, Itai, Guillermo Del Pinal, and Uli Sauerland (2021). "Presuppositional exhaustification". In: *Semantics and Pragmatics*. DOI: 10.3765/sp.14.11. URL: https://doi.org/10.3765/sp.14.11.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector (2012). "Scalar implicature as a grammatical phenomenon". In: *Semantics: an international handbook of natural language meaning*. Ed. by Claudia Maienborn, Klaus von Heusinger, and Paul Portner. Vol. 3. De Gruyter Mouton.

[4] There is however a sense in which entailment goes through. Specifically, "S. believes snow is white and D. believes snow is black" *Strawson entails* "S. and D. agree that *Exh* snow is white or black" (von Fintel 1999).

Degano, Marco et al. (2024). "The ups and downs of ignorance". Unpublished ms., lingbuzz/007389. URL: https://ling.auf.net/lingbuzz/007389.

Dekker, Paul (1993). "Transentential meditations: ups and downs in dynamic semantics". PhD thesis. University of Amsterdam.

Fine, Kit (2017). "Truthmaker Semantics". In: *A Companion to the Philosophy of Language*. John Wiley & Sons, Ltd, pp. 556–577. ISBN: 978-1-118-97209-0. DOI: 10.1002/9781118972090.ch22.

Fox, Danny (2007). "Free Choice and the Theory of Scalar Implicatures". In: *Presupposition and Implicature in Compositional Semantics*. Ed. by Uli Sauerland and Penka Stateva. London: Palgrave Macmillan UK, pp. 71–120.

Grice, Herbert Paul (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Heim, Irene (1982). "The semantics of definite and indefinite noun phrases". PhD thesis. University of Massachusetts, Amherst.

— (1992). "Presupposition Projection and the Semantics of Attitude Verbs". In: *Journal of Semantics* 9.3, pp. 183–221. ISSN: 1477-4593. DOI: 10.1093/jos/9.3.183. URL: http://dx.doi.org/10.1093/jos/9.3.183.

Hintikka, Jaakko (1969). "Semantics for Propositional Attitudes". In: *Philosophical Logic*. Ed. by J. W. Davis, D. J. Hockney, and W. K. Wilson. Synthese Library 20. Springer Netherlands, pp. 21–45.

Rothschild, Daniel (2017). "Yablo's semantic machinery". In: *Philosophical Studies* 174, pp. 787–796. DOI: 10.1007/s11098-016-0759-3. URL: https://doi.org/10.1007/s11098-016-0759-3.

Sauerland, Uli (2004). "Scalar implicatures in complex sentences". In: *Linguistics and Philosophy* 27, pp. 367–391. DOI: 10.1023/B:LING.0000023378.71748.db. URL: https://doi.org/10.1023/B:LING.0000023378.71748.db.

Spector, Benjamin (Apr. 27, 2016). "Comparing Exhaustivity Operators". In: *Semantics and Pragmatics* 9.0 (0), pp. 11-1–33. ISSN: 1937-8912. DOI: 10.3765/sp.9.11. URL: https://semprag.org/index.php/sp/article/view/sp.9.11 (visited on 01/15/2021).

van Rooij, Robert and Katrin Schulz (2004). "Exhaustive interpretation of complex sentences". In: *Journal of Logic, Language and Information* 13, pp. 491–519. DOI: 10.1007/s10849-004-2118-6. URL: https://doi.org/10.1007/s10849-004-2118-6.

Von Fintel, Kai (1999). "NPI Licensing, Strawson Entailment, and Context Dependency". In: *Journal of Semantics* 16, pp. 97–148.

Yablo, Stephen (2014). *Aboutness*. Princeton, New Jersey: Princeton University Press.