

Postscripts to "A pragmatic approach to the problem of logical omniscience"

Paul Egré

► **To cite this version:**

Paul Egré. Postscripts to "A pragmatic approach to the problem of logical omniscience". 2002. ijn_00000285

HAL Id: ijn_00000285

https://jeannicod.ccsd.cnrs.fr/ijn_00000285

Submitted on 6 Nov 2002

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Postscripts to “A Pragmatic Approach to the Problem of Logical Omniscience”

Paul Égré

November 6, 2002

After my paper “A Pragmatic Approach to the Problem of Logical Omniscience” had been written and presented to a small audience of people, I received no feedback for a few months and turned to other projects. I only recently set out to work on logical omniscience again, at the time when two colleagues and friends of mine, namely M. Cozic and B. Spector, pointed out a few problems in my argument, some of which I had also become aware of in the meantime. Pending a full revision of the paper, I prefer to expose and respond to their criticisms in the form of a provisional postscript paper. I hope this will enhance both the good and the bad, and help for the discussion of these issues in seminar.

1 Representational versus metarepresentational

This distinction between a representational and a metarepresentational stance on belief sentences constitutes the core of my defense of the principle of substitution of logically equivalent sentences in belief contexts. I maintain that this distinction is essential to an understanding of the problem of logical omniscience. Yet it shouldn't be mixed up with neighbouring distinctions, although it may sometimes sound close to them.

The distinction is not a *normative* versus *descriptive* distinction. The latter distinction is more general. Yet one may think that a metarepresentational semantics of belief statements is descriptive, since it is relative to the utterance of belief reports in ordinary conversational contexts. And likewise that a representational semantics is normative, in particular because I describe Hintikka's semantics as a representational semantics, and moreover since the latter has a normative interpretation.

In the paper however, I purport to defend the principle of substitution of logical equivalents in belief contexts as a metarepresentational semantic principle. As such it is certainly a normative principle, since we find contexts where ascribers apparently fail to use it and since the main thrust of the paper remains to give an account (following Recanati's theory) of situations where the principle is not applied systematically. Likewise, it is correct to think that an

adequate representational semantics of belief should be as faithful as possible to the epistemic situation of real agents, whose knowledge is barely closed for any logical principle whatsoever. Such semantic frameworks have been defined and I discuss one of them, namely H. N Duc's system \mathbf{DEK}_N , in a sequel to this paper entirely devoted to the *representational* problem of logical omniscience ([2], [3]). Duc's logic is free of any principle of closure whatsoever (one may be happy to say that it is descriptive in this sense), but it is interesting precisely because it is still able to yield predictions (one may be happy to say that it is normative in this sense). It should be enough to suggest that the normative *vs* descriptive distinction often tends to get sloppy when we come to evaluate the adequacy of a semantic framework.

The distinction is not simply a *mental* versus *linguistic* distinction.

I suggest in the paper that the latter identification would be a "rough way" to understand what I mean by the representational-metarepresentational distinction, although I find it helpful to get a first grasp of it. So we need to be careful here. For sure the level of object-representations of an agent is mental, and a metarepresentational statement is a piece of language. But that is a trivial remark: what Recanati calls a metarepresentation is also a representation in the first place, and is at first distinct from the statement that expresses it. Hence my claim is different: the representational-metarepresentational distinction is about belief *sentences* in either case.

It is correct to say that a representational theory of belief focusses on the description of the content of belief of an agent, independantly of the beliefs are ascribed. Such a theory has to explain how belief contents are individuated and processed. For all that, a logical theory along these lines may perfectly well be a theory of belief sentences of the form $K\phi$. There are two objections here, inspired by questions that M. Cozic asked me:

- Here comes the first objection: *the fact that a theory deals with belief sentences necessarily implies the existence of an outside ascriber. Hence a so-called representational theory of belief sentences should be metarepresentational, which makes the distinction collapse.*

My answer is that we should carefully make explicit what the notion of *ascription* means in this case. There is a weak and default sense of the notion of *ascription* according to which any theory in which you find object-statements of the form: $K_i\phi$, where i denotes an agent and K refers to the knowledge of this agent, necessarily implies the existence of an outside modeller, namely the theorist who uses the metalanguage. This is the sense in which Fagin & alii write:

Thus the agents could be described as logically omniscient. This does not especially trouble us in the context of multi-agent systems, since in that context we view our notion of knowledge as *external*. That is, knowledge is *ascribed* by the system designer to the agents ([4]: 309).

But there is a stronger sense of the notion of ascription that is relative to the utterance of belief sentences by reporters in ordinary discourse situations. A theory of belief sentences can be ascriptional in the weak sense without aiming at the description of ascriptions in the strong sense: such a theory is likely to be a representational theory, and is certainly not a metarepresentational theory. Conversely, any metarepresentational theory of belief statements is ascriptional in both senses.

- Here comes the second objection: *you claim that Hintikka's theory of belief sentences is a representational theory because it is "thought-oriented". But it can very well be that such a theory is more well-suited for the purpose of providing the semantics of ordinary ascriptions than a theory like Recanati's whose ontology is supposed to be that of the ascriber only.*

This objection is more serious, and is actually twofold: first, the contrast between Hintikka's and Montague's respective theories of belief sentences is brought forward to enhance the distinction between a representational and a metarepresentational stance. It is true that Hintikka's and Montague's theories both make claims about belief reports as well as about the psychology of belief. But this shouldn't affect the well-foundedness of the distinction. It remains correct in my view to distinguish between a theory that focusses on what the ascriber endorses in a belief report, and a theory whose aim is to account for the way belief contents are individuated and processed independently of how attributions thereof are made in everyday discourse.

Now, it is indeed a real issue whether we take into account the belief worlds of the agents when making a belief attribution. As explained in the paper, Hintikka's theory straightforwardly accounts for the fact that someone can believe that Hesperus is a star while disbelieving that Phosphorus is. This kind of explanation may be extended to account for failures of substitution of synonymous predicates in belief reports, or even logically equivalent sentences, if one introduces the machinery of possible worlds in which standard logical equivalences break down (so-called *impossible possible worlds*). I agree that this way of accounting for discrepancies of identification on the part of an agent is *representationally* plausible and probably fundamental.

Whether this gives us a general principle concerning the way in which an ascriber is likely to make substitutions on the basis of a belief attribution is a distinct question. My claim in the paper is more modest and concerns specifically the problem of accounting for substitution of *logically equivalent* sentences. It may be that the situation is in fact more complex than Recanati claims in the case of proper names or synonymous predicates. But as regards mere logical equivalence, I tend to think like Recanati that substitution is *a priori* a correct semantic principle so far as only the perspective of the ascriber is involved. To quote from an example of Heim & Kratzer involving logically equivalent embedded sentences, I believe that a speaker who says:

- (1) Mary believes that if Robin competes, he will win.

is entitled to say:

(2) Mary believes that anyone who does not compete or loses will do something that Robin won't do

although I consider that in this case the second statement would not be uttered because it is less economical than the first in order to convey the same information.

I believe further that to make this transition, one needn't take into account the way Mary would assent to the embedded sentences nor care whether these would have the same meaning for her. So I tend to agree with Recanati that reluctance to make the transition here can occur only if the speaker, when uttering the first sentence, overtly suggests that Mary may be sensitive to something like the form of the words used (a *de re* vs *de dicto* distinction is also at stake here, cf. *infra*). In this situation, resort to alternative possible worlds serves as a way of representing what's going on in Mary's head, and this is consistent with what I then take to be a representational perspective on the semantics of belief sentences. It is not obvious whether this provides us with a semantic principle in the first place (such as: 'check whether the sentence is true in all of the belief-worlds of the agent before making the substitution').

The distinction is not a distinction between *beliefs as non-verbal dispositions toward action* versus *linguistic beliefs*. In his discussion of the logical omniscience problem, Dubucs (1991, [1]) has emphasized a distinction between N-beliefs and L-beliefs. N-beliefs or normalized beliefs are beliefs likely to be ascribed with respect to the non-verbal behavior of an agent, while L-beliefs or linguistics beliefs are beliefs ascribed to the agent "on grounds of his public assertions".

I agree with the importance of the distinction, yet a metarepresentational theory of belief statements must equally account for inferences concerning beliefs that are ascribed on the basis of overt behavior and beliefs that are ascribed on the basis of public assertions.

Is the distinction some sort of a generalization of the *de re* versus *de dicto* distinction? This suggestion was made in conversation by B. Spector. The fact that occurrences of proper names outside the scope of a belief operator are freely substitutable may be generalized to full sentences. Then, a translation of (1) above could be:

(3) Of the proposition that if Robin competes, he will win, Mary believes that it holds.

The proposition that if Robin competes, he will win, is the same as the proposition expressed by the embedded statement in (2).

Therefore it is equivalent to say:

(4) Of the proposition that anyone who does not compete or loses will do something that Robin won't do, Mary believes that it holds.

which may be a correct paraphrase of (2).

According to that suggestion, the distinction between *representational* and *metarepresentational* should rather be cashed out as a genuine semantic ambiguity of belief sentences. A *metarepresentational reading* could then be defined as a “generalized *de re* reading” of belief statements.

There is a correct intuition behind this suggestion: it suggests to analyze belief statements of the form ‘*X* believes that ϕ ’ as expressing a specific (representational) belief relation between an agent and a proposition. This corresponds to the reading advocated by Montague himself.

By contrast to the thesis of semantic ambiguity, my claim in the paper is that the distinction is *methodological* in the first place. I speak mostly of representational versus metarepresentational *perspectives, semantics, theories, principles* or *problems*. I also speak of representational *content*, when I want to contrast what the ascriber endorses, and what the ascriber is meant to endorse. In my discussion of visual ascriptions, I go so far as to speak of a metarepresentational versus a representational *reading* or interpretation of vision sentences. This is probably an unhappy extension, because I do not mean to say that the sentences themselves are semantically ambiguous, but again that they are subject to different semantic approaches.

Besides I am not sure whether we keep good track of the syntactic distinction between *de re* and *de dicto* if we talk of generalized *de re*. Is not the so-called generalized *de re* reading here just a way of describing the semantic interpretation we give in the metalanguage of a plain *de dicto* reading of the statement in the object-language? Consider the Kratzer-Heim example. A natural way to contest the equivalence between (1) and (2) above is to argue that (1) and (2) indeed have ambiguous readings. Under a standard *de re* reading, (1) becomes:

(5) Of Robin, Mary believes that if he competes, he will win.

and (2) becomes:

(6) Of anyone who either does not compete or loses, Mary believes they will do something Robin won’t do.

It is clear that under these readings, a different (representational) belief content is ascribed to Mary in each case. Clearly, this is not the same reading as the one corresponding to (3) and (4). I claim that the one corresponding to both (3) and (4) is just a standard *de dicto* reading.

2 Distribution of knowledge over conjunction

Most readers will probably have had the feeling that the section of the original paper where I discuss distribution of knowledge operators over conjunction is the weak part of my argument. I agree that the argument is not as clear as it should be, although I feel inclined to hold on to my conclusions. Here are a few specific objections to start with :

1. The parallel between knowing and seeing is puzzling : from the representational perspective taken by Dretske, seeing is a transparent attitude, whereas

from the same representational perspective, however we choose to individuate knowledge or belief contents, it seems that belief is not a transparent attitude. Therefore, how could such a parallel be illuminating?

2. The so-called metarepresentational analysis I give of seeing statements interprets “sees an A” as “sees an A as an A”: intuitively, this is not a transparent reading, and yet I claim in the paper that uniform substitution of logically equivalent predicates is acceptable. This seems to beg the question because it may be that to see an A as an A is not the same as to see an AB as an AB, even though A and AB do express the same property, so long as A and B do express distinct properties.

3. It may be that distribution of seeing over conjunction is not even a correct representational principle (a suggestion made to me by T. Williamson in conversation): for instance I may see John and Mary and Susan coming without seeing either John coming or Mary coming or Susan coming. This is perhaps what I had in mind in my discussion of *seeing as*. As Williamson pointed out to me, this is close to issues in the semantics of plurals for statements like: “John and Mary lifted the piano, yet neither did John lift the piano, nor did Mary lift the piano”.

I will not attempt to deal with these objections here and I leave them for discussion. I shall deal with a more fundamental objection that was made to me by B. Spector: *let us suppose that the argument from vision fails to establish that principle (DIST) is not a correct metarepresentational principle. Then, do we have a reply to the puzzle described in section 3.2?*

I recall that the puzzle is the fact that the combination of principle (RE) and (DIST) entails closure of knowledge under logical consequence.

Schematically: from rule (DIST): $K(\phi \wedge \psi) \therefore K\phi \wedge K\psi$, and (RE): $\phi \leftrightarrow \psi \therefore K\phi \leftrightarrow K\psi$, we get the following derived rule (CL):

$$\phi \rightarrow \psi, K\phi \therefore K\psi$$

According to B. Spector, the thesis defended in my paper is not as radical as it ought to be. In particular, I write:

the principle of closure of knowledge under logical consequence is unrealistic *whether we interpret it as a representational or as a metarepresentational principle* (p. 12).

But once one accepts the distinction between metarepresentational and representational, B. Spector asks: why should this principle be more problematic from a metarepresentational point of view than the principle of substitution of logically equivalent sentences? The dilemma is then recurrent: either to show without using the analogy with *seeing* that (DIST) is indeed fallible, or to accept (CL). Let us look again at each horn of the dilemma.

a) The first suggestion is to find another plausible way of reading belief statements, in agreement with Montague-Scott semantics. B. Spector here offers to analyse a statement of the form:

(7) X believes that ϕ

as

(8) X believes* a sentence that is logically equivalent to ϕ

Under this analysis, substitution of logically equivalent sentences is a correct principle, but distribution over conjunction is not. This analysis gives a suitable interpretation of **MS** and its possible world semantics, and B. Spector suggested that to assert (metarepresentationally) that an agent believes that ϕ would then amount to ascribing a specific (representational) belief* relation between an agent and a sentence.

I reply that such an analysis matches indeed the predictions of Montague-Scott semantics, but now has a very different flavour since it postulates a relation between an agent and a sentence, and no longer a relation between an agent and a proposition. One will have to specify what the belief* relation amounts to in order to avoid the traditional problems of sententialism.

b) Now for the second horn: I remark that the analysis suggested by B. Spector does not yield a direct natural language counterexample to the principle of distribution over conjunction. In this respect it is just another way of making sense of Montague-Scott semantics. Yet if the principle of distribution of knowledge over conjunction really is a basic principle governing knowledge ascriptions, then one will apparently be committed to the principle (CL) of closure under logical consequence so long as one accepts (RE). Couldn't the case I make for a pragmatic restriction of principle (RE) then be made against (CL)? Maybe, says B. Spector, but from a strict semantic point of view, an ascriber who utters $K\phi$ should anyway accept to say $K\psi$ as soon as ψ is a logical consequence of ϕ .

The latter move seems to me unsatisfactory and implausible however: my intuition is really that on the basis of (CL), belief contents will be ascribed to an agent by means of inferences that it is natural to reject, contrary to what is happening in the case of mere logical equivalence. The view I maintain is still the following: it is certainly right to infer from ' X knows *that* ϕ and *that* ψ ', ' X knows that ϕ and X knows that ψ '. But the statement $K(\phi \wedge \psi)$ of epistemic logic that is used to derive principle (CL) from (RE) and (DIST) has an ambiguous reading. It should then read: ' X knows *that* (ϕ and ψ)'.

I am aware that this is just another way of describing the fact that in Montague semantics, the semantic value of "that (ϕ and ψ)" need not be the same as that of "that ϕ and that ψ ". This still does not give us a natural language basis for the distinction.

References

- [1] Dubucs J. (1991), “On Logical Omniscience”, in *Logique & Analyse* 133-134, pp. 41-45.
- [2] Duc H.N. (1997), “Reasoning about Rational, but not Logically Omniscient, Agents”, *Journal of Logic and Computation*, vol. 7, v. 5, pp. 633-648.
- [3] Égré P. (2002), “Approches implicitement épistémiques et explicitement épistémiques du problème de l’omniscience logique”, manuscript, Université Paris 1. A first version is available on the web at : <http://www-ihpst.univ-paris1.fr/pics/papiers.htm>.
- [4] Fagin R., Halpern J, Moses Y. & Vardi M. (1995), *Reasoning about Knowledge*, MIT Press.