Logical Omniscience and Counterpart Semantics

Paul Égré*

Institut Jean-Nicod, CNRS, Paris

Abstract

This paper proposes a pragmatic approach to the phenomenon of hyperintensionality of belief reports in natural language, building on earlier work by J. Gerbrandy (2000) on counterpart semantics for first-order epistemic logic. Counterpart relations are used by Gerbrandy to model the notion of mode of presentation and to account for the context-dependency of de re beliefs. I propose to generalize Gerbrandy’s semantics to a second-order epistemic logic, in which not only individuals, but also properties and complex relations can have epistemic counterparts. The aim is to give a uniform treatment of cases of hyperintensionality involving expressions of distinct syntactic categories (coreferential proper names, cointensional predicates, logically equivalent sentences), by giving belief sentences a generalized de re logical form.

Key Words: de re Belief, Counterpart Semantics, Epistemic Logic, Hyperintensionality, Logical Omniscience, Quantified Modal Logic, Pragmatics

The aim of this essay is to propound a new approach to the problem of hyperintensionality of belief contexts, using the apparatus of quantified modal logic and the machinery of counterpart semantics. Consider the following sentence, in which we assume the conditional is a material conditional (but other examples would do):

(1) Peter believes that if door A is locked, then door B is not locked, but he does not believe that if door B is locked, then door A is not locked

In standard epistemic logic (Hintikka 1969) and in intensional logic (Montague 1970), an ascription of belief like this one, which involves two logically equivalent sentences, is predicted to be inconsistent. In Montague grammar, in particular, the proposition expressed by each embedded sentence is the same, and since beliefs are conceived as relations between individuals and propositions, the beliefs of the agents are predicted to be closed under logical equivalence, making them “logically omniscient” in that sense. In modal epistemic logic, the situation is potentially more problematic, since beliefs are predicted to be closed even under logical consequence. There clearly are, however, contexts in which a sentence like (1) above can be uttered consistently. On the other hand, there is also a sense in which the same content is ascribed (or denied) to Peter in either conjunct of (1). The aim of the present proposal is to reconcile these two intuitions, by offering a pragmatic account of hyperintensionality: the basic idea, which I try to articulate in this paper, is that Peter’s belief, even though opaque, can be analyzed as a de re belief about one and the same proposition (or about the same objects and relation each time), but under different counterpart relations, playing the role of modes of presentation, and acting as a pragmatic component in the evaluation of sentences.

The prime inspiration for this essay comes from the work of J. Gerbrandy (2000) on counterpart semantics for de re beliefs. In this paper, I propose to generalize Gerbrandy’s semantics to a second order modal logic, in order to account for cases of hyperintensionality involving expressions of distinct syntactic

*Version of the Proceedings of the ESSLLI 2006 Workshop on Logics for Resource-Bounded Agents, T. Agotnes and N. Alechina (eds.). This paper is based on chap. 1 of my PhD. dissertation (Égré 2004). Thanks to M. Aloni, D. Bonnay, E. Maier, O. Roy, F. Récanati, P. Schlenker, B. Spector and T. Williamson for helpful comments and criticisms on an earlier draft. Special thanks are due to D. Bonnay for several stimulating and helpful discussions. I also thank audiences in Paris, Amsterdam and Bordeaux, and an anonymous referee for detailed comments and suggestions.
categories (coreferential proper names, cointensional predicates, logically equivalent sentences). Thus the
idea is that cases of hyperintensionality should be analyzed on a par with other classic instances of opacity
for belief sentences, and the aim is to get a uniform treatment of all those cases. My proposal, however,
rests on the idea that belief sentences can be given a de re logical form, even in situations which would
standardly be analyzed as de dicto. This idea raises problems of its own, which will be discussed along
the way, but it also contains some potential benefits (like avoiding the resort to impossible worlds beyond
standard belief worlds).

The paper is structured as follows. In the first section, I briefly review the de dicto treatment of opacity
given in Hintikka’s modal framework. In section 2, I give some arguments in favor of a de re analysis
of opacity, and review the treatment of Gerbrandy in section 3. Section 4 presents the generalization of
Gerbrandy’s semantics to cases of hyperintensionality involving predicates and full sentences. I state a
compositional semantics for an appropriate system of second-order modal logic in section 5. The last
section offers an evaluation of the benefits and limitations of this proposal, and a comparison with other
semantic theories of hyperintensionality.

1 Opacity in epistemic logic

Ordinary belief sentences are notoriously opaque, in the sense that expressions which seem to convey the
same information outside the scope of the belief verb are no longer intersubstitutable salva veritate under
the scope of a verb like “believe”. Let us consider the following three examples:

(2) Peter believes that Cicero is a philosopher, but does not believe that Tully is a philosopher.
(3) Peter believes that John is an eye-doctor, but does not believe that John is an ophthalmologist.
(4) Peter believes that if door A is locked, then door B is not locked, but he does not believe that if door
B is locked, then door A is not locked.

There is a clear sense in which, in each of these examples, the embedded sentences say the same thing,
yet we do find contexts in which each of (2), (3) and (4) can be uttered consistently (which should not
be the case if the relevant expressions were always intersubstitutable salva veritate). Thus, the proper
names “Cicero” and “Tully” are synonymous in so far as they refer to the same individual. Likewise, the
predicates “eye-doctor” and “ophthalmologist” are synonymous in so far as they express the same concept.
Finally, under the assumption that the conditional is a material conditional in (4), the sentence “if door A
is locked, then door B is not locked” and its contrapositive are synonymous in so far as they are logically

1 Opacity in epistemic logic

salva veritate

salva veritate

equivalent. To clarify the problem raised by these sentences, let us represent them in a first-order modal
language, letting “□φ” stand for “Peter believes that φ”:

(5) □P(c) ∧ □¬P(t)
(6) □E(j) ∧ □¬O(j)
(7) □(L(a) → ¬L(b)) ∧ □¬□(L(b) → ¬L(a))

Relative to the standard semantics of modal logic (□φ is true at a world if φ is true at all accessible worlds),
a sentence like (5) is predicted to be inconsistent if one assumes Kripke’s thesis of the rigidity of proper
names, namely if c and t are taken to denote the same individual in every possible world. Likewise, (6) is
inconsistent if one assumes E and O to have the same intension, that is the same extension in each possible
world. Finally, (L(a) → ¬L(b)) and its contrapositive are logically equivalent, true together or false
together at every possible world, making (7) inconsistent as well. These examples all illustrate the problem
of hyperintensionality of belief contexts: given the standard semantics for epistemic logic, sentences which
express the same possible world proposition, that is sentences with the same intension, are predicted to be
substitutable salva veritate under the scope of a belief operator.¹

¹The concept of “hyperintensionality” was introduced by Cresswell 1973, to characterize contexts in which intensionally equiva-

lent sentences are not substitutable salva veritate. As a reviewer pointed out, one may distinguish two levels of hyperintensionality,
depending on how fine-grained one takes the notion of intensional equivalence to be. Strictly speaking, an operator o is hyperinten-

sional if there are two formulas φ and ψ, a model M and a world w such that M, w ⊨ (φ ↔ ψ) (that is φ and ψ are logically equivalent),

M, w ⊨ oφ and M, w ⊭ oψ. And following the reviewer’s suggestion, one may call an operator superintensional if there are two
formulas φ and ψ, a model M and a world w such that M, w ⊨ (φ → ψ) (that is φ and ψ are model-equivalent), M, w ⊨ oψ
and M, w ⊭ oφ. Hyperintensionality is stronger than superintensionality, since logical equivalence entails model-equivalence. What
The problem is in fact more dramatic in the case of epistemic logic (unlike in intensional logic), since beliefs are predicted to be closed even under logical consequence, due to the monotonicity of the $\square$ operator. Thus the following sentence is predicted to be inconsistent:

(8) Peter believes that John will come, but does not believe that John or Mary will come.

(9) $\square C(j) \land \neg \square (C(j) \lor C(m))$

The hyperintensionality of belief sentences therefore suggests that belief verbs shift the ordinary semantic value of sentences. For Hintikka (1969), for instance, two names like “Cicero” and “Tully” are not necessarily coreferential in a belief context simply because an agent can fail to know or realize that they denote the same individual. The same goes for predicates like “eye-doctor” and “ophthalmologist”, and also for logically equivalent sentences. Faced with the phenomenon of opacity, it is therefore natural to assume that words can take the meaning they have for the believer, rather than the meaning endorsed by the speaker.

Model-theoretically, thus, a sentence like (5) is satisfiable if $c$ and $t$ are allowed to take distinct values in at least one of Peter’s belief worlds. Likewise a sentence like (6) is satisfiable if the predicates $E$ and $O$, although coreferential in the actual world, do not overlap in at least one of Peter’s belief worlds. Sentences like (7) and (9) are prima facie more problematic, since the standard epistemic semantics gives logical constants like $\rightarrow$, $\neg$ and $\lor$ a uniform behavior throughout the models.

To go around this problem, a possibility (entertained successively by Montague, Cresswell, Rantala, and advocated by Hintikka 1975) is to enrich the space of worlds with logically impossible worlds, where logically equivalent sentences can take arbitrary values in a non-compositional way. In a way, made explicit by Muskens (1991) and anticipated by Thomason (1980), this solution can be seen as treating logical constants as part of the non-logical vocabulary, and to allow some variation in the denotation of the logical connectives at the belief worlds, in the same way in which constants and predicates are allowed to change their denotation at the belief worlds of the agent. In that manner, a compositional semantics can be given for beliefs, which accounts for opacity at all the relevant syntactic levels. Seen in this light, the problem of logical omniscience is ultimately a problem of semantic competence: by allowing proper names, predicates, and logical connectives to take on arbitrary values at special worlds, one accounts for the fact that agents can be confused about the objective synonymy of certain expressions.

## 2 De re beliefs and opacity

Despite the coherence of Hintikka’s treatment of opacity, there remain several reasons to look for a different solution. A first point of criticism, which has been recurrent on the side of the supporters of Kripke’s theory of proper names, concerns the leeway that allows proper names, for instance, to take distinct denotations at the belief worlds of the agent. For a strict Kripkean, proper names are rigid, which means that two names that are coreferential at the actual world should have the same denotation at all the worlds, including the epistemic worlds.\(^2\) A strict supporter of Kripke’s theory may allow the predicates “eye-doctor” and “ophthalmologist” to take different values at the belief worlds of Peter, if he grants that those two predicates have a descriptive content, about which Peter can make a mistake. But he will probably disagree with the case of proper names, on the ground that proper names have no descriptive content. This piece of criticism is not entirely convincing, however. For it is one thing to say that proper names are rigid and do not behave like hidden definite descriptions in general, and another thing to acknowledge that, as a matter of plain fact, one cannot fail to realize that two co-referential proper names are indeed coreferential.\(^3\)

A second, more compelling objection relates to the representation of de re beliefs. So far we have given each of the sentences (2)-(4) a de dicto logical form, giving the belief operator the largest possible example (3) suggests that belief operators are at least superintensional (since the predicates $E$ (“eye-doctor”) and $O$ (“ophthalmologist”) need not be equivalent at every world of every model – they are model-equivalent only in virtue of a meaning postulate). Example (4), on the other hand, suggests that belief operators are hyperintensional in the strict sense.

\(^2\)Such a view is expressed, in particular, by Recanati (2000: 395), who writes : “According to Hintikka (1962: 138-141), failures of substitutivity in belief contexts show that two co-referential singular terms, though they pick out the same individual in the actual world, may refer to different objects in the ascribee’s belief world. That option is ruled out in the present framework; for we want the ontology to be that of the ascriber all along: we want the singular terms to refer to the same objects, whether we are talking of the actual world, or about the ascribee’s belief world. That is the price to pay for semantic innocence.” Semantic innocence is the view according to which an attitude verb does not shift the semantic value of expressions occurring in its scope. This position is criticized in Égré (forthcoming), but I argue, as in what follows, that part of the intuition underlying this view can be reinterpreted.

\(^3\)This is argued quite persuasively by Gerbrandy (2000: 151), and by Aloni (2001: 44-45).
scope (over singular terms, in particular). Several authors, however, have defended the idea that a belief report can be opaque, and nevertheless be \textit{de re} at the same time (Kraut 1983, Heim 1992, Recanati 2000). Recanati, for instance, quoting Loar (1972), insists that “\textit{even on the opaque reading of a belief sentence in which a singular term occurs, reference is made to some particular individual}” (2000, italics his).

This claim reflects the following intuition: when one makes a belief ascription like (2), one does not simply mean, according to Recanati, that Peter makes a metalinguistic mistake on the meaning of the proper names “Cicero” and “Tully”. In Recanati’s theory, the use of two distinct, although coreferential proper names, is just a way of pragmatically indicating that Peter represents to himself one and the same individual under two distinct modes of presentation. When I say in different contexts:

10) Peter believes that Cicero is a philosopher

11) Peter does not believe that Tully is a philosopher

I am each time talking about Cicero-Tully, namely about one and the same individual. When I utter the conjunction of the two sentences, I’m still making reference to Cicero-Tully, but the names, being contrasted, pragmatically point to distinct modes of presentation.

Taking this intuition seriously, and considering that a belief can be opaque and nevertheless \textit{de re}, the paraphrase of (2) in modal epistemic logic might therefore be:

\begin{equation}
\exists x (x = c \land \Box P(x)) \land \exists y (y = t \land \neg \Box P(y))
\end{equation}

The problem here is that, if \(c\) and \(t\) denote the same individual \(d\) in the actual world (the world of evaluation), the standard semantics for first-order modal logic constrains the variables \(x\) and \(y\) to denote \(d\) across the belief worlds of Peter. Given a first-order model \(\langle W, R, D, I \rangle\), an assignment \(g\) and a world \(w\), \(M, w, g \models \exists x \Box \phi\) if and only if there is a \(d\) in \(D_w\) such that for every \(w'\) such that \(w Rw'\), \(M, w', g[d/x] \models \phi\). This, however, is problematic. First, it makes the ascription a plain contradiction, since the statement then is equivalent to \(\exists x (x = c \land \Box P(x)) \land \exists x (x = c \land \neg \Box P(x))\). However, it seems that a sentence like (2) can be uttered without contradiction.

Moreover, it makes a \textit{de re} belief ascription incompatible with situations of mistaken identity, which seems too strong. There are cases where a belief is clearly \textit{de re}, and yet does involve a failure to make a correct identification on the part of the ascribee. The paradigm case is Quine’s example of Ralph, who believes of Ortcutt, thought of as “the man seen at the beach”, that he is not a spy, and who also believes of Ortcutt, thought of as “the man in the brown hat”, that he is a spy. In this scenario, it seems one can make the following belief ascription:

\begin{equation}
\exists x (x = o \land \Box S(x)) \land \exists x (x = o \land \neg \Box S(x))
\end{equation}

This paraphrase, however, is a problem for the standard semantics of first-order modal logic, since it should then follow that Ralph believes of Ortcutt that he is and is not a spy (see Quine 1956, Aloni 2001), which seems too strong in this context:

\begin{equation}
\exists x (x = o \land \Box (S(x) \land \neg S(x)))
\end{equation}

Quine’s example, as argued quite convincingly by Gerbrandy and Aloni, provides a good indication that the standard semantics of first-order modal logic ought to be amended in order to account for \textit{de re} beliefs with mistaken identity. My suggestion here is that roughly the same semantic account which Gerbrandy gives to analyze Quine’s example can be extended to parse a sentence like (2) as \textit{de re} instead of \textit{de dicto}. I will first review Gerbrandy’s account, then I will attempt to show how to extend this analysis to cases of hyperintensionality involving predicates and full sentences.

\footnote{By singular term, Recanati means a proper name in the passage under discussion.}
3 Counterpart semantics

The two sentences “Ralph believes of Ortcutt that he is a spy”, and “Ralph believes of Ortcutt that he is not a spy” are intuitively compatible because the truth of each of them depends on a distinct underlying mode of presentation. This mode of presentation need not be explicitly expressed in the sentence, but can be salient to both the speaker and hearer in the discourse situation, so that the two sentences will be seen as mutually compatible. In Gerbrandy’s analysis, a mode of presentation is analyzed as a method of cross-identification, namely as a way of identifying an individual across epistemic alternatives.\(^5\) Thus one and the same actual individual can have distinct counterparts in the epistemic alternatives of an agent, corresponding to several identification methods. In Quine’s scenario, for example, there is one method of identification which “connects objects in Ralph’s epistemic alternatives just in case they are the man that Ralph saw...at the beach” (Gerbrandy 2000: 155), and another method which connects objects in Ralph’s belief worlds just in case they are the man Ralph saw in the brown hat. Both methods connect objects in Ralph’s epistemic alternatives to Ortcutt at the actual world.

In Gerbrandy’s semantics, epistemic sentences are evaluated relative to such identity relations, which play the role of a pragmatic parameter. Let us define an epistemic structure as a quadruple \((W,R,D,I)\), where \(W\) is a set of worlds, \(R\) is an epistemic accessibility relation between the worlds, \(D\) is a function which associates to each world \(w\) a domain of individuals \(D_w\), and \(I\) is an interpretation function for the non-logical vocabulary. Formally, a method of identification can be defined as a relation \(C\) between ordered pairs \((w,d)\) of worlds and individuals such that \(d \in D_w\). If \((w,d)C(w',d')\), \(d'\) will be called a counterpart to \(d\) in \(w'\). In Gerbrandy’s semantics, methods of identification are defined as equivalence relations. A further constraint, which is argued for quite persuasively in Aloni (2001), is to require those relations to be functional, namely such that one individual at a world has at most one counterpart at another world by the identification method, and also to be total, in the sense that if an individual has an epistemic counterpart at a belief world, it has a counterpart at every belief world.\(^6\) Intuitively, the reason why an agent can be mistaken about thinking there are two individuals whereas there is only one, is because this one individual is presented to him in two different ways (there are two different modes of presentation of the same individual). In other words, there should not be more counterparts of an actual individual at a belief world than there are modes of presentation of that individual. But moreover, once a counterpart of an actual individual “inhabits” a belief world, it should also persist at all the belief worlds.

Gerbrandy’s semantics for first-order modal logic is standard, except for the epistemic operators and the assignment of variables, which both have to be made sensitive to the identification relations. Thus, given a method of identification \(C\) and a pair of worlds \(w, w'\), two assignment functions \(g\) and \(h\) for the variables are in the counterpart relation induced by \(C\) if and only if \(h\) assigns to every \(x\) in \(w'\) the counterpart of the individual which \(g\) assigns to \(x\) in \(w\):

\[
\text{g} \mapsto_{\text{C}}^{w,w'} \text{ h iff for every variable } x, (w, g(x))C(w', h(x)).
\]

The specific satisfaction clauses are the following:

- \(M, w, g \models_C \exists x \phi\) iff there is a \(d \in D_w\) such that \(M, w, g[d/x] \models_C \phi\)
- \(M, w, g \models_C \Box \phi\) iff for every \(w'\) such that \(wRw'\), for every \(h\) such that \(g \mapsto_{\text{C}}^{w,w'} \text{ h} : M, w', h \models_C \phi\).

Given these definitions, Gerbrandy can account for the Ortcutt case. Take a model \(M\) with three worlds, where \(w\) is the actual world, in which \(o\) denotes \(d\), namely Ortcutt, and \(w'\) and \(w''\) are the two epistemic alternatives of Ralph. Let us call \(C_b\) the identification relation for the beach encounter, and \(C_h\) the identification relation for the hat encounter. In \(w'\) and \(w''\), \(d_b\) is Ortcutt as seen as the beach, namely the counterpart of \(d\) under \(C_b\), and in \(w'\) and \(w''\), \(d_h\) is Ortcutt as seen with the hat, namely the counterpart of \(d\) under \(C_h\). Supposing \(d_b\) is outside, and \(d_h\) is the hat, one then has:

\[
M, w \models_{C_h} \exists x (x = o \land \Box \neg S(x)) \land M, w \models_{C_b} \exists x (x = o \land \Box S(x))
\]


\(^6\)In what follows, I thus write \(C(w, d)(w')\) to denote the counterpart of \(d\) in \(w'\), relative to \(w\), or \(C(d)(w')\) when the world in which \(d\) lives is clear.
Using Gerbrandy’s counterpart semantics, it is therefore possible to account for Ralph’s beliefs, without ascribing a logical contradiction to Ralph.\textsuperscript{7}

Now, I would like to suggest that this machinery can be used to give a \textit{de re} analysis of the Tully-Cicero example by which we started. Remember sentence (5), here repeated as (17):

\begin{equation}
\text{(17)} \quad \text{Peter believes that Cicero is a philosopher, but he does not believe that Tully is a philosopher.}
\end{equation}

We can note that here, unlike in the Ortcutt case, the negation takes wide scope over the belief verb. Formally, this is not a problem, since in the model described for the Ortcutt case, it also holds that:

\begin{equation}
\begin{aligned}
M, w \models & C_o \exists x (x = o \land \neg \square S(x)) \\
& \\
& \\
&
\end{aligned}
\end{equation}

that is Ralph does not believe that Ortcutt is a spy under the beach identification relation. In the Tully-Cicero example, we can imagine in the same way that Peter’s modes of presentations are the names “Tully” and “Cicero”: Peter is simply not sure whether those two names of English denote the same individual or not. What this means is that there is at least one epistemic alternative where Cicero-thought-of-as-“Cicero” and Cicero-thought-of-as-“Tully” are distinct individuals. In that situation, the names play the role of methods of identification. It is therefore easy to define two identification relations, namely $C_c$ and $C_t$, such that, in a three-world model analogous to the previous one, it holds that:

\begin{equation}
\begin{aligned}
M, w \models & C_o \exists x (x = c \land \square P(x)) \land M, w \models C_c \exists x (x = t \land \neg \square P(x)) \\
& \\
& \\
&
\end{aligned}
\end{equation}

Using counterpart relations, the standard \textit{de dicto} analysis of a sentence like (17) can therefore be cast into a \textit{de re} analysis, in the spirit of Recanati’s suggestions concerning the relational character of opaque belief reports involving proper names. It has actually been suggested that proper names might systematically outscope attitude verbs. This was suggested even for definite descriptions, for totally different reasons, in Heim’s analysis of the presupposition projection of attitude verbs (Heim 1992).\textsuperscript{8} Other authors, like Kraut (1983), have been even more radical, by defending the idea that there are no \textit{de dicto} attitudes. In what follows, I shall use in a systematic way the idea that \textit{de dicto} belief reports can be restated as \textit{de re} belief reports involving specific modes of presentation or acquaintance relations on the part of the believer.\textsuperscript{9}

\section{Generalization}

In the previous section, we have seen that a certain \textit{de dicto} analysis of substitution failures of proper names can be expressed in terms of a \textit{de re} analysis of substitution failures, using the additional machinery of counterpart semantics. In this section my aim is to show that this analysis can be extended to handle cases of substitution failures involving predicates instead of proper names, and even full sentences, as in the examples by which we started, by allowing higher-order quantification over properties.

This generalization presupposes that it does make sense to talk about \textit{de re} belief about higher-order entities. Two independent arguments can be given for that, however. The first concerns the fact that one can devise scenarios analogous to Quine’s Ortcutt case of mistaken identity with properties. Imagine, for instance, a situation in which Peter has two friends, Jack and Jill, having exactly the same profession, namely being eye-doctor, but suppose that Peter is under the misconception that Jack’s job is scary (by coincidence, whenever he visits Jack, he sees him perform delicate eye-surgery) while he thinks Jill’s job is not (by coincidence, whenever he visits her, he sees her just testing people’s eyesight). Unbeknownst to Peter, Jack and Jill perform exactly the same tasks, but at different times. Peter thinks moreover that whoever does the same job has Jack does a scary job, and likewise whoever does the same job as Jill does not do a scary job. As in Quine’s example, it seems correct to say:

\textsuperscript{7}Gerbrandy gives a similar treatment to Kripke’s puzzle of belief (Kripke 1979), in which Pierre is confused about the meaning of the proper names “Londres” and “London”.

\textsuperscript{8}See Heim 1992: 210-211: “Another way of summarizing the suggestion I just made is this: there is not really just one \textit{de re} reading (for a given constituent), but there are many - one for each acquaintance relation that the context might supply. And some of those many, namely those where the acquaintance relation happens to include the subject’s awareness that the \textit{res} fits the same description used by the speaker, are very similar to the \textit{de dicto} reading: more precisely, they entail it. In a way, I am blurring the distinction between \textit{de re} and \textit{de dicto} readings. But that may not be such a bad thing. More often that not, the two are impossible to tell apart.”

\textsuperscript{9}While doing so, I do not mean to reject the well-foundedness of the \textit{de re-de dicto} distinction in general. An important issue, which I set aside here, concerns the syntactic restrictions that bear on the outscoping of constituents in a sentence, for instance in the case of indefinites.
The example invites to treat “being an eye-doctor” as a property about which Peter has two opposite beliefs. The two conjuncts of (20) may then be analyzed as $\exists x (X = E \land \Box \neg \text{Scary}(x))$ and $\exists x (X = E \land \Box \neg \text{Scary}(x))$ respectively.

A second reason to introduce higher-order quantification is that it is needed in order to account for the so-called non-specific de re readings of indefinite descriptions in attitude contexts, as in the sentence “Peter believes that some soccer player has a dog”, where “some soccer player” is taken de re by the speaker, but does not refer to any particular soccer player relative to the believer (see Bonomi 1995). This would happen in a context in which Peter sees a certain dog outside a restaurant, which he thinks belongs to one of the people he saw inside the restaurant, but such that only I, the speaker, know that these people are soccer players. In that case, both the first-order de dicto analysis, $\Box x(S(x) \land D(x))$, and the first-order de re analysis, $\exists x(S(x) \land \Box D(x))$ are false, and the correct analysis seems to be: $\exists x (X = S \land \Box \exists x(X(x) \land D(x)))$, namely “there are soccer players such that Peter believes that one of them has a dog”.

Granting the legitimacy of higher-order de re beliefs, we are in a position to give a sentence like “Peter believes that John is an eye-doctor” a generalized de re logical form, meaning that Peter believes of John and of the property of being an eye-doctor, that the latter applies to the former, namely:

$$\exists x \exists X (x = j \land X = E \land \Box X(x))$$

The same analysis can be given for “Peter believes that John is an ophthalmologist”, that is: $\exists x \exists X (x = j \land X = O \land \Box X(x))$. In many contexts, the two attributions will be used interchangeably: for instance, if someone tells me that “Peter believes that John is an eye-doctor”, without mentioning anything else about Peter, I may repeat this information to someone else by saying: “Peter believes that John is an ophthalmologist”. In such a case, the words are endorsed by the speaker only. In a situation in which I would utter (3), that is “Peter believes that John is an eye-doctor, but does not believe that John is an ophthalmologist”, the two de re logical forms remain compatible by assuming that each predicate is associated with a distinct counterpart relation, corresponding to a specific acquaintance relation on the part of the believer. Thus one can have:

$$\exists x \exists y \exists X (x = a \land y = b \land X = \lambda xy.(L(x) \rightarrow \neg L(y)) \land \Box X(xy))$$

This means that Peter believes of door A, of door B, and of the relation such that if one object is locked, then another is not locked, that this relation applies to those individuals. This generalized de re analysis allows to use the machinery of counterpart semantics in order to handle the substitution failure of logically equivalent sentences. All it takes is to suppose that the complex relation $\lambda xy.(L(x) \rightarrow \neg L(y))$ has different counterparts in the belief worlds of Peter, depending on the situation in which he is. Thus there might be two methods of identification (for relations), such that:

$$\exists x \exists y \exists X (x = a \land y = b \land X = \lambda xy.(L(x) \rightarrow \neg L(y)) \land \Box X(xy))$$

Again, one can imagine that these methods of identification are made salient to the hearer by the use of distinct syntactic expressions in each utterance. This reflects the intuition that Peter misperceives the identity of a logical relation between two objects and properties. At the same time, this allows to preserve the idea that Peter’s belief, although confused, can very well be a de re belief about the objects $a$ and $b$ (for instance in a situation in which he perceives the two doors A and B).

In the same way, closure under logical consequence can be blocked, despite the upward monotonicity of $\Box$, since if $P \subseteq Q$ holds at the actual world, the epistemic counterpart of $P$ need not be a subset of the

---

10I am indebted to M. Aloni for pointing out Bonomi’s example to me. Non-specific de re readings have been discussed independently by J. Fodor (1970) and R. Bauerle (1983). See Heim & von Fintel (2002).
epistemic counterpart of $Q$. For instance, the sentence “Peter believes John will come, but does not believe John or Mary will come”, can be paraphrased:

$$\exists x \exists y \exists X \exists Y (x = j \land y = m \land X = C \land Y = \lambda xy.(C(x) \lor C(y)) \land \Box X(x) \land \neg \Box Y(xy))$$

The sentence is consistent if one assumes that the counterpart of the relation denoted by $Y$ in the actual world does not hold of John and Mary (assuming those are correctly identified), even though in each of Peter’s belief worlds John will come.

5 A second-order epistemic logic

The present section gives the details of the generalization of Gerbrandy’s semantics to a second-order modal language. The language, which I call $L_2(\Box)$, is a second-order logic enriched with a unary modal operator (intended as an epistemic modality), in which it is possible to name complex predicates by the usual mechanisms of lambda-abstraction. The present treatment is inspired in part by the presentation of higher-order logic given in the first chapter of Fitting (2002). I only state the semantics here and do not investigate its possible axiomatizations: since the logic is higher-order, we can’t expect to have a full completeness result, but rather a Henkin-style completeness proof, but I shall leave this investigation for further work. Another central issue concerns the treatment of quantifiers, and in particular whether we should work with fixed or variable domains (see Fitting & Mendelsohn 1998). The use of counterpart relations makes it natural to let the domains vary, if we think of quantifiers as ranging over objects actually existing at a world, and of counterpart relations as establishing links between distinct ontologies (for instance that of the speaker, and that of the agent whose belief is reported). In what follows we thus place no restriction on the domains (except indirectly, by means of the conditions on counterpart relations), and likewise we give an actualist interpretation to the non-logical vocabulary.

The language $L_2(\Box)$

The definition of $L_2(\Box)$ is in two steps: first I introduce the language $L_1$ of first-order logic with predicates of arity 0 (propositional symbols), then the notion of a lambda-abstract, which is needed for the definition of the second-order part. The construction is in two steps in order to exclude lambda-abstracts of the form $\lambda x.\Box P(x)$, essentially for reasons of simplicity, so that only non-modal properties can be named. The language is built on an alphabet which includes:

(i) A denumerable set $\text{Var}$ of individual variables: $x, y, z, ...$

(ii) For all $n \geq 0$, a denumerable set $\text{VAR}_n$ of predicate variables of arity $n$: $X^n, Y^n, Z^n, ...$. By definition, $\text{VAR} = \bigcup_n \text{VAR}_n$

(iii) A denumerable set $\text{Cons}$ of individual constants: $c, c', c'', ...$

(iv) For all $n \geq 0$, a denumerable set $\text{CONS}_n$ of predicate constants of arity $n$: $P^n, Q^n, R^n, ...$. By definition, $\text{CONS} = \bigcup_n \text{CONS}_n$.

(v) Logical connectives: $\neg, \land, \exists$

(vi) Additional symbols: $\lambda, ,, (,)$

(vii) Modality: $\Box$

(viii) Equality symbol: $=$

Individual terms: every element of $\text{Cons}$ or $\text{Var}$ is an individual term.

$L_1$-formulas

If $t$ and $t'$ are individual terms, $t = t'$ is a $L_1$-formula

A predicate constant of arity 0 is a $L_1$-formula
If \( t_1, \ldots, t_n \) are individual terms, and \( P \) is a predicate constant from \( CONS_n \), then \( P(t_1, \ldots, t_n) \) is an \( L_1 \)-formula.

If \( \phi \) and \( \psi \) are \( L_1 \)-formulae, then so are \( \neg \phi \) and \( (\phi \land \psi) \).

If \( \phi \) is an \( L_1 \)-formula, then \( \exists x \phi \) is a \( L_1 \)-formula

Nothing else is an \( L_1 \)-formula.

**Lambda-abstracts:** if \( \phi \) is an \( L_1 \)-formula, and \( x_1, \ldots, x_n \) are distinct variables from \( \text{Var}, \lambda x_1, \ldots, x_n.\phi \) is a lambda-abstract of arity \( n \). We denote by \( ABS_n \) the set of lambda-abstracts of arity \( n \), and \( ABS \) the set of lambda-abstracts.

If \( \phi \) is an \( L_1 \)-formula, then \( \lambda.\phi \) is a lambda-abstract of arity 0.

**\( L_2(\square) \)-formulae**

Every \( L_1 \)-formula is a \( L_2(\square) \)-formula, and if \( X \in \text{VAR}_0 \), \( X \) is an \( L_2(\square) \)-formula.

If \( T \) and \( T' \) are respectively a variable in \( \text{VAR}_n \), a constant in \( CONS_n \), or a lambda-abstract in \( ABS_n \), then \( T = T' \) is an \( L_2(\square) \)-formula.

If \( T \) is a variable in \( \text{VAR}_n \), a constant in \( CONS_n \), or a lambda-abstract of \( ABS_n \), and \( t_1, \ldots, t_n \) are individual terms, then \( T(t_1, \ldots, t_n) \) is an \( L_2(\square) \)-formula.

If \( \phi \) is an \( L_2(\square) \)-formula, and \( X \) is variable of \( \text{VAR}_n(n \geq 0) \), then \( \exists X \phi \) is an \( L_2(\square) \)-formula

If \( \phi \) and \( \psi \) are \( L_2(\square) \)-formulae, then so are \( \neg \phi \) and \( (\phi \land \psi) \).

If \( \phi \) is an \( L_2(\square) \)-formula, then \( \square \phi \) is an \( L_2(\square) \)-formula.

Nothing else is an \( L_2(\square) \)-formula.

**Semantics for \( L_2(\square) \)**

**Model:** An \( L_2(\square) \)-model \( M \) is a quadruple \( \langle W, R, D, I \rangle \) where \( W \) is a non-empty set ; \( R \) is a relation on \( W \) ; \( D \) is a function which to each world \( w \) associates a domain of individuals \( D_w \); \( I \) is an interpretation function with domain \( W \times (\text{Cons} \cup CONS) \), such that: \( I_w(c) \in D_w \) if \( c \) is an individual constant and \( I_w(P) \subseteq (D_w)^n \) for \( P \) a predicate constant of arity \( n \).

If \( P \) is a predicate symbol of arity 0 (a propositional symbol), then one notes: \( I_w(P) = 0 \) instead of \( I_w(P) = \emptyset \) and \( I_w(P) = 1 \) instead of \( I_w(P) = \{ \emptyset \} \). More generally, if \( D \) is a set, one notes \( D^0 = 1 \).

**Assignment functions:** An assignment function \( g \) assigns to a variable \( x \) an element in \( \bigcup_{w \in W} D_w \), and to a variable \( X \) in \( \text{VAR}_n \) an \( n \)-ary relation over \( \bigcup_{w \in W} D_w \) (if \( X \) has arity 0, again one writes \( g(X) = 0 \) or \( g(X) = 1 \)).

An assignment \( g' \) is an \( x_1, \ldots, x_n \)-variant of an assignment \( g \) if \( g \) and \( g' \) give the same values to all variables except at most \( x_1, \ldots, x_n \).

**Method of identification:** A method of identification \( C \) is an equivalence relation between couples \( (w, d) \) such that \( w \in W \) and \( d \in D_w \), and between couples \( (w, R) \) such that \( w \in W \) and \( R \) is an \( n \)-ary relation over \( D_w \). Thus the counterpart of an \( n \)-ary relation is an \( n \)-ary relation. One supposes moreover that \( C \) is functional, that is, given \( (w, d) \) and \( w' \), there is at most one \( d' \in D_{w'} \) such that \((w, d)C(w', d')\), and similarly in the case of relations. A further natural constraint is to suppose that \( C \) is total in the following sense : if \( d \) has a counterpart at one world under \( C \), then it has a counterpart at every other world under \( C \). Thus one can designate by \( C(w, d)(w') \) and \( C(w, R)(w') \) the respective counterparts in \( w' \) of individual \( d \) and relation \( R \) of \( w \).

**Definition:** \( g \mapsto w, w' \) \( h \) iff for every variable \( x \in \text{Var} \) and \( X \in \text{VAR} \), \((w, g(x))C(w', h(x))\) and \((w, g(X))C(w', h(X))\).
Satisfaction of the formulae

In what follows, I note $T^{(n)}$ to mean that $T$ is a predicate constant, a predicate variable, or a lambda-abstract of arity $n$. Similarly, $P^{(n)}$ means that $P$ is a predicate constant of arity $n$, and $X^{(n)}$ means that $X$ is a predicate variable of arity $n$.

If $t$ is an individual term: one notes $(t)_{M,w,g,C} = I_w(t)$ if $t$ is a constant, and $(t)_{M,w,g,C} = g(t)$ if $t$ is a variable.

Likewise, one writes $⟨T⟩_{M,w,g,C} = I_w(T)$ if $T$ is a predicate constant, and $⟨T⟩_{M,w,g,C} = g(T)$ if $T$ is a predicate variable.

Finally, if $T = λx_1...x_n.φ$, then $⟨T⟩_{M,w,g,C} = \{ (g'(x_1),...,g'(x_n)) \}; g'$ is an $x_1,...,x_n$-variant of $g$ and $g'(x_i) ∈ D_w$ ($1 ≤ i ≤ n$), and $M, w, g' |=_C φ$.

If $T = λ.φ$, then $⟨T⟩_{M,w,g,C} = 1$ if $M, w, g |=_C φ$ and $⟨T⟩_{M,w,g,C} = 0$ if $M, w, g ∉_C φ$.

$M, w, g |=_C t = t′$ iff $⟨t⟩_{M,w,g,C} = ⟨t′⟩_{M,w,g,C}$.

$M, w, g |=_C T = T′$ iff $⟨T⟩_{M,w,g,C} = ⟨T′⟩_{M,w,g,C}$.

$M, w, g |=_C T^0$ iff $⟨T⟩_{M,w,g,C} = 1$, for $T ∈ VAR_0 ∪ CONS_0$.

$M, w, g |=_C T^{(n)}(t_1,...,t_n)$ iff $(⟨t_1⟩_{M,w,g,C},...,⟨t_n⟩_{M,w,g,C}) ∈ ⟨T⟩_{M,w,g,C}$.

$M, w, g |=_C ¬φ$ iff $M, w, g ∉_C φ$.

$M, w, g |=_C (φ ∧ ψ)$ iff $M, w, g |=_C φ$ and $M, w, g |=_C ψ$.

$M, w, g |=_C ∃xφ$ iff there exists $d ∈ D_w$ such that $M, w, g[d/x] |=_C φ$.

$M, w, g |=_C ∃X^{(n)}φ$ if there exists an $n$-ary relation $R ⊆ (D_w)^n$ such that $M, w, g[R/X] |=_C φ$.

$M, w, g |=_C □φ$ iff for all $w′$ such that $wRw′$, and for all $h$ such that $g R w'h : M, w', h |=_C φ$.

6 Evaluation

6.1 Comparisons

In principle, the present framework affords the same kind of fine-grainedness that is found in other approaches to hyperintensionality, as in Thomason’s treatment in terms of primitive propositions (Thomason 1980), and Muskens’ related treatment using impossible worlds (Muskens 1991). The reason is that counterpart relations can be determined by any syntactic component in the embedded sentence. For instance, a sentence like “John believes that if door A is locked then door B is not locked” can be given several logical forms, as we have seen, including one form in which all the embedded material is scoped out by means of a propositional variable (a variable of arity 0):

$$\exists X(X = λ.((L(a) → ¬L(b)) ∧ □X))$$

This means that of the proposition that if door A is locked then door B is not locked, Peter believes that it holds. Using appropriate counterpart relations, it is possible to say that Peter believes of that proposition, under one counterpart relation that it holds, and under a different counterpart relation that it does not hold.

In Thomason’s framework, this corresponds to the fact that the sentences “if door A is locked then door B is not locked” and its contrapositive can very well express different primitive propositions within a model, or be true and false at different non-standard worlds in the case of the impossible worlds approach. However, unlike Thomason or Muskens, our approach of hyperintensionality in terms of counterpart relations does not commit us to a domain of primitive propositions, or of impossible worlds. Ontologically, this is a gain, since all we need are the standard belief worlds of the agent, without having to treat logical constants as non-logical constants.

The closest antecedent to the present approach is Cresswell and von Stechow’s (1982) own generalization of the notion of de re belief, which they use in particular to represent mathematical beliefs. Thus they observe that the same sentence “Poirot believes that 2+2=4” can be given several distinct logical forms,
depending on the material that is taken de re or de dicto in the sentence. The present analysis, however, is
closer to Hintikka’s original motivation, since we see substitution failures first and foremost as failures of
the believer to grasp the identity of a property, or of a logical link, as represented by the use of epistemic
counterparts.

Also, we do not have to suppose that two syntactically distinct sentences necessarily express different
propositions: by default, two logically equivalent sentences will be substitutable in belief contexts, and
non-substitutability is seen as a context-dependent phenomenon. This is the sense in which this analysis
is fundamentally pragmatic: if φ and ψ are intensionally equivalent sentences (in the classic sense), then
in some contexts, when uttered after “Peter believes”, the clauses “that φ” and “that ψ” do express the
same information, whereas in other contexts they don’t. By giving belief sentences a generalized de re
logical form, one accounts for the intuition that the belief is about certain entities whose value is determined
first relative to the speaker. When I say: “Peter believes that John is an eye-doctor”, there is a sense in
which I say exactly the same thing as in: “Peter believes that John is an ophthalmologist”. In other words,
the same literal content is ascribed to Peter. As Recanati argues, situations in which it is appropriate to
utter “Peter believes that John is an eye-doctor but does not believe that John is an ophthalmologist” are
situations where this content is pragmatically enriched. For Recanati, this pragmatic enrichment does not
imply a modification of the basic semantic value of the predicates “ophthalmologist” and “eye-doctor”,
which should remain constant in the model. The nice feature of Gerbrandy’s semantics, in this respect, is
that this pragmatic enrichment is materialized by the additional parameter of cross-identification relations,
and that the value of lexical terms for which substitution fails needs only to be fixed relative to the speaker.
One can always assume, moreover, that the default parameter for contexts where substitution is not at stake
is the relation of plain identity. When I utter: “Peter believes that Cicero was poor”, without saying more
about Peter, the hearer should assume that Peter’s belief is about Cicero as commonly identified.

6.2 Refinements

Along with the benefits that we claim for this treatment of hyperintensionality, several imperfections may be
pointed out. One of them concerns the treatment we made of conjunction in all the examples we presented
so far. This limitation is already present in Gerbrandy’s treatment of Quine’s example, but is inherited in the
other examples we discussed. In order to get a consistent paraphrase of a sentence like (13) above, namely
“Peter believes that Ortcutt is a spy, and he believes that Ortcutt is not a spy”, we have used a metalinguistic conjunction in the form:

\[
\begin{align*}
M, w \models_{C_h} & \exists x (x = o \land \square S(x)) \land M, w \models_{C_h} \exists x (x = o \land \square \neg S(x)) \\
\end{align*}
\]

There is no way, in that system, to get a consistent reading of the conjunctive sentence \(\exists x (x = o \land \square S(x)) \land \exists x (x = o \land \square \neg S(x))\), since sentences are evaluated with respect to only one counterpart relation, which
we assumed to be functional. A possible way out would be to relax functionality, but this undermines the
intuitive one-to-one correspondence with modes of presentation. Furthermore, it will not be adequate to
handle non-conjunctive sentences of mistaken identity like “Peter does not believe that Cicero is Tully”, if
one analyzes the latter de re as \(\exists x \exists y(x = c \land y = t \land \neg \square (x = y))\).

To go around both problems, another possibility is to allow reference to modes of presentation directly
at the sentential level, by indexing variables, as done in Aloni (2001, 2005). In Aloni’s system, a sentence
like (13), that is “Peter believes that Ortcutt is a spy, and he believes that Ortcutt is not a spy”, is represented as:
\(\exists x_n(x_n = o \land \square S(x_n)) \land \exists y_m(y_m = o \land \square \neg S(y_m))\). Likewise, “Peter does not believe that Cicero
is Tully” can be paraphrased as “\(\exists x_n \exists y_m(x_n = c \land y_m = t \land \neg \square (x_n = y_m))\)”. In Aloni’s system, the
indices are indices of different conceptual covers. In the same way, we could let the indices denote different
counterpart relations supposed to be salient in the discourse context, and evaluate sentences with respect
to a family of counterpart relations. Given a family \(C\) of identification methods, we let \(C_i\) denote
the identification method indexed by \(i\). We write \(g^i \rightarrow_{C_i} w \leftrightarrow h\) if for every index \(i\) and every variables \(x\) and
\(X_i\), \((w, g(x))C_i(w', h(x))\) and \((w, g(X_i))C_i(w', h(X_i))\). Using this mechanism, one can account for the
consistency of mistaken beliefs about identity, and still maintain that the belief is about one and the same
actual entity, seen under different modes of presentation.

\[11\] I am indebted to M. Aloni for these suggestions.
7 Conclusion

In this paper I have defended the view that from a linguistic point of view, that is from the perspective of ordinary belief attributions, the problem of logical omniscience is to be treated on a par with the problems of opacity involving non-logical expressions of other syntactic categories, such as synonymous proper names or predicates, arguing that this problem calls for a pragmatic approach. From the point of view of epistemic logic, the present account can be seen as an extension of Hintikka’s original account, since the logic we used allows to give a de re representation of sentences that would standardly be analyzed as de dicto, and to get a consistent interpretation by means of counterpart relations instead of using impossible worlds. The inspiration remains fundamentally the same, however, since expressions which are synonymous for the speaker are allowed to take arbitrary denotations at the believer’s worlds, modulo the counterpart relations. While so doing, we have moved to a much more expressive logic, however. Further work needs to be done, in this respect, to investigate the proof-theoretic properties of the logic, and see what kinds of completeness results can be obtained. More fundamentally, it may be that the best reason we have to hold on to a plain de dicto analysis of the examples by which we started is that in all those situations, belief reports have a quotational component without which the ascription simply could not be made. In the present framework, however, this syntactic component is not eliminated, it is simply deferred to the level of counterpart relations.

References


