J. GERBRANDY

THE SURPRISE EXAMINATION IN DYNAMIC EPISTEMIC LOGIC

ABSTRACT. We examine the paradox of the surprise examination using dynamic epistemic logic. This logic contains means of expressing epistemic facts as well as the effects of learning new facts, and is therefore a natural framework for representing the puzzle. We discuss a number of different interpretations of the puzzle in this context, and show how the failure of principle of success, that states that sentences, when learned, remain to be true and come to be believed, plays a central role in understanding the puzzle.

1. THE PUZZLE

The large number of papers on the puzzle of the surprise exam is perhaps more due to the ambiguity of its formulation rather than its inherent difficulty. There is a lot that is imprecise in puzzle, and it has been used as an occasion for discussing topics in a wide range of fields, ranging from the problem of adding a knowledge predicate to a language with a certain expressive power (Montague and Kaplan 1960) to observations about specific types of probabilistic games (Sober 1998). Chow (1998) and Sorensen (1988) provide organized and relatively non-partisan overviews of representative selections from the literature.

The scenario is the following:

In the kind of school where you get exactly one exam every week, a teacher announces to his class: “This week, the exam will be a surprise.” It is commonly understood that an exam comes as a surprise if you do not know, the evening before, that it is given the next day.

A smart student, called Marilyn, reasons as follows.

“Suppose the exam is given on Friday. In that case, come Thursday evening, I will not have gotten an exam yet, and I will know that it must be on Friday, which means that it would not be a surprise. So, it is not on Friday. Suppose that it is on Thursday. Then, on Wednesday evening I will know that it must be

Synthese (2007) 155: 21–33
DOI 10.1007/s11229-005-2211-7

on Friday or on Thursday. I know it is not on Friday, so it must be on Thursday: again, it would not be a surprise. So Thursday is out as well. I can repeat this argument excluding all the other days of the week. So I will not get a surprise exam; in fact, I will not get an exam at all!"
The teacher gives the exam on Wednesday, surprising all students in the class. So, the teacher was right after all. What went wrong with Marilyn's reasoning?

In this paper, we will look at the puzzle in the context of dynamic epistemic logic. The point is not so much to present *the* solution to the puzzle; but rather to show how the some of the reasoning in it can be laid out in a precise formal logic, thereby explaining some of the aspects that might be confusing otherwise.

To be more specific: we first recast a classical analysis of the puzzle (worked out in detail by Sorensen (1988)) in a dynamic epistemic setting, and show how the assumption that sentences are successful, in the sense that they come to be believed after being learned, allows us to formally derive that Marilyn's knowledge is inconsistent from a fairly straightforward transcription of the teacher's announcement. We then show that the sentence in question is not successful in dynamic epistemic semantics, and how this blocks the reasoning of Marilyn at an early stage; basically, this is a simplified version of the analysis of Gerbrandy (1988), which is also presented in Van Ditmarsch and Kooi (2006). Because it may seem that by cutting Marilyn's argument short at such an early stage means that we have interpreted the teacher's statement by a sentence that is too weak, we discuss a number of ways that the statement can be strengthened: one contingent, one that is false, and one that is paradoxical.

2. DYNAMIC EPISTEMIC LOGIC

The language of Dynamic Epistemic Logic given by the following definition:

$$\varphi, \psi := p \mid \varphi \wedge \psi \mid \neg\varphi \mid K\varphi \mid [\varphi]\psi,$$

where p is an element of a given set of propositional variables P. A sentence of the form $K\varphi$ stands for the fact that our agent believes that φ.[1] The dynamic character of the language is obtained from the operators $[\varphi]$: a sentence of the form $[\varphi]\psi$ is to express that after our agent learns that φ, ψ is true.

A *possible world* s is a function that assigns to each propositional variable a truth-value (either 0 or 1). An *information state* is a subset of possible worlds. A *Kripke model* is a pair σ, s, where s is a world and σ is an information state. The idea is that s represents the state of affairs as it actually obtains, while σ represents the information of an agent as the set of possible worlds that are compatible with that information.[2]

The standard logical operators get their usual interpretation from epistemic logic:

$$\begin{aligned}
\sigma, s &\models p && \text{iff } s(p) = 1, \text{ i.e. if } p \text{ is true in } s,\\
\sigma, s &\models \neg\varphi && \text{iff } \sigma, s \not\models \varphi,\\
\sigma, s &\models \varphi \wedge \psi && \text{iff } \sigma, s \models \varphi \text{ and } \sigma, s \models \psi,\\
\sigma, s &\models K\varphi && \text{iff for all } t \in \sigma \text{ it holds that } \sigma, t \models \varphi.
\end{aligned}$$

The last clause says that a sentence φ is known to be true just in case it is true in all possible worlds in the information state σ.

The idea behind modeling the effect of getting new information is simple: we say that learning a sentence φ in an information state σ results in a new information state that contains exactly those possible worlds from σ where φ is true. Writing $\sigma[\![\varphi]\!]$ for the resulting state, we define:

$$\sigma[\![\varphi]\!] = \{s \in \sigma \mid \sigma, s \models \varphi\}.$$

With this function, we define that a sentence of the form $[\varphi]\psi$ is true just in case ψ is true in the model that results after learning φ:

$$\sigma, s \models [\varphi]\psi \quad \text{iff } \sigma[\![\varphi]\!], s \models \psi.$$

If φ contains no epistemic operators, $\sigma[\![\varphi]\!]$ is simply the intersection of σ with the set of worlds where φ is true. This way of modeling information change as the intersection of an information state with a proposition (both taken to be sets of possible worlds) goes back to at least Stalnaker (1978) and Heim (1982). Adding epistemic operators complicates the picture somewhat, but the idea remains basically the same.

There is a sound and complete axiomatization for this semantics, that contains the axioms and rules of $K45$ epistemic logic (or of $S5$ if we postulate that $s \in \sigma$ for each model σ, s) together with the following axiom schemes governing the behavior of the operators $[\varphi]$:

1. $[\varphi]p \leftrightarrow p$
2. $\neg[\varphi]\psi \leftrightarrow [\varphi]\neg\psi$
3. $[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$
4. $[\varphi]K\psi \leftrightarrow K(\varphi \rightarrow [\varphi]\psi)$.

The definitions of the semantics given here can be seen as a succinct version of what has by now become a standard definition of public announcement in multi-agent systems, in its different incarnations of Emde Boas et al. (1984), Landman (1986), Plaza (1989), Gerbrandy and Groeneveld (1997), Baltag and Moss (2004) or, from another angle, the definitions of Veltman (1996). Completeness of a similar logic was first proved by Plaza (1989).

Many of the articles quoted above revolve around a notion of getting information that is guaranteed to be true. This is also the case for Ditmarsch and Kooi (200x), who give an analysis of the puzzle similar to the one we give in Section 4. We think that assuming all information, whether existing or new, is true, is too strong an assumption in the present context. One of the points of the puzzle is that Marilyn draws a conclusion that is false, and the assumption that information is always true excludes this a priori.

3. BLINDSPOT

To simplify somewhat, let us suppose that the teacher makes his announcement on Tuesday. This leaves only three possible candidates for giving an exam: Wednesday, Thursday and Friday. Let the propositional variables we, th and fr stand, respectively, for the fact that exam is given on Wednesday, Thursday and Friday. For example, the proposition that Marilyn, after having observed that she did not get an exam on Wednesday, does not know if it is given on Thursday can be represented by the sentence $[\neg\mathsf{we}]\neg K\mathsf{th}$.

It seems now straightforward to transcribe the proposition that there will be a surprise exam.

$$
\begin{aligned}
(1) \qquad & \mathsf{we} \wedge \neg K\mathsf{we} \\
& \vee\, \mathsf{th} \wedge [\neg\mathsf{we}]\neg K\mathsf{th} \\
& \vee\, \mathsf{fr} \wedge [\neg\mathsf{we}][\neg\mathsf{th}]\neg K\mathsf{fr} \\
& \vee\, K\bot
\end{aligned}
$$

If the exam is on Friday, Marilyn does not know this after having learned that it is not on Wednesday and not on Thursday; if the exam is on Thursday, Marilyn does not know this after having learned that it is not on Wednesday; and Marilyn does not know it is on Wednesday. The last disjunct implies that when Marilyn has contradictory information the exam will be a surprise as well, which seems to be implied by the conclusion in the statement of the puzzle that the exam will be a surprise 'after all.'

Let us abbreviate this sentence as S. This sentence is a prediction about what Marilyn will know, solely on the basis of days passing by without an exam.

There is nothing inherently contradictory about this sentence, and we will see below that if the teacher plans to give his exam on Thursday, say, then this sentence is, indeed, true.

So how does Marilyn reach her false conclusion? The principle of *success* (Alchourrón et al. 1985) states that if you learn something, you come to believe it is true:

(2) $[\varphi]K\varphi.$

The property of success has long been considered uncontroversial for a certain notion of updating or learning. Most of the postulates for information change of Alchourrón et al. (1985) have been criticized at some point, but the postulate of success has rarely been challenged. Indeed, it seems evident to the point of triviality: what else could learning something mean if not coming to believe it is true?

With the assumption that success is valid we can reconstruct the entire informal argument in the puzzle as a formal proof in dynamic epistemic logic. We can prove that if the principle of success applies to S, then learning S provides Marilyn with inconsistent information; and if her information is inconsistent, then S is true.

(3) $[\mathsf{S}]K\mathsf{S} \vdash [\mathsf{S}]K\bot \quad \text{and} \quad [\mathsf{S}]K\bot \vdash [\mathsf{S}]\mathsf{S}.$

To see this, observe first that the axioms of dynamic epistemic logic guarantee that any sentence is equivalent to one without any dynamic operators. In particular, the sentence S is equivalent to:

(4) $$(\mathsf{we} \wedge \neg K\mathsf{we}) \vee (\mathsf{th} \wedge \neg K(\neg\mathsf{we} \rightarrow \mathsf{th})) \\ \vee (\mathsf{fr} \wedge \neg K((\neg\mathsf{we} \wedge \neg\mathsf{th}) \rightarrow \mathsf{fr})) \vee K\bot.$$

This means that S, in this reading, is equivalent to a sentence that only says something about what Marilyn knows at the initial situation. This equivalence fits nicely with observation of Sorensen (1988) that the temporal aspect of the situation does not seem to be essential to the puzzle, as is illustrated by the variation known as the ‘paradox of the designated student.’

It is now not very difficult to prove that believing that sentence (4) is true implies, in the logic $K45$ of belief, having inconsistent beliefs, by excluding all days of the week as possible days of the exam. Proofs to this effect can be found in Binkley (1969), Sorensen (1988) or Chow (1998), who all rewrite the teacher’s announcement into a statement of epistemic logic similar to that of sentence (4).

The basic point here is extensively discussed in Sorensen (1988): the teacher is taken to express a *blindspot* for Marilyn in the logic K45. The sentence S may be true, but it can never be consistently be believed by Marilyn.

4. TRUE BUT NOT SUCCESSFUL

There is something unsatisfactory about limiting the diagnosis to the observation that the announcement is a blindspot. Even if expressing a blindspot is a somewhat perverse way of packaging the information you want to convey, such a sentence *can* express information, and the hearer should be able to understand this and incorporate the new information into the information she already has.

Dynamic semantics explains how this is possible. To see how this works out for the sentence S, consider a model in which s_{we}, s_{th} and s_{fr} are possible worlds in which the exam is on Wednesday, Thursday, and Friday. At the outset Marilyn does not know which of these three worlds represents the actual situation. Her information state at this point can be represented by the set $\{s_{\mathsf{we}}, s_{\mathsf{th}}, s_{\mathsf{fr}}\}$; call this information state σ_{init}.

If the exam is given on Friday, it would not be a surprise, and indeed, $\sigma_{\mathsf{init}}, s_{\mathsf{fr}} \models \neg\mathsf{S}$. In the other two situations, S is true. So the teacher can only truthfully say that the exam will be a surprise if he plans to give it on Wednesday or on Thursday.

When Marilyn learns that S is true, she eliminates the world in which S is false from her information state. The state that results is $\{s_{\mathsf{we}}, s_{\mathsf{th}}\}$. Now, if the exam is actually on Thursday, it is not a surprise anymore: $\{s_{\mathsf{we}}, s_{\mathsf{th}}\}, s_{\mathsf{th}} \models \neg\mathsf{S}$. However, if the exam is given on Wednesday, it will remain to be a surprise. In either case, the sentence is not successful: Marilyn does not know whether the exam will be a surprise or not, even if she just learned that it would be.

If S correctly paraphrases the teacher's announcement, then Marilyn's reasoning is cut short after having excluded the last day as the day of the exam. She continues her argument by reasoning that the exam cannot be on Thursday either, because that would contradict the claim of the teacher that the exam comes as a surprise. To be sure, she is correct in concluding that, now, after the announcement, it will not be a surprise if the exam is on Thursday, and she is correct in that the teacher said that it would be, but she is not correct in seeing a contradiction between these two claims. If the exam is on Thursday, then S is true *before* the teacher makes his announcement, but it becomes false after she learns of its truth. This may be confusing, but it is not paradoxical.

We have a formal representation in dynamic epistemic semantics of the statement that the exam will be a surprise, for which it can be shown that the teacher expressed an 'obvious' truth and Marilyn can consistently learn the proposition expressed by the teacher. The analysis explains how the announcement exhibits unusual behavior only when made to the class, but is unremarkable when made to someone else (without the class hearing it). And we have an explanation of why the puzzle is puzzling: it is because it makes implicit use of the principle of success, a plausible principle that happens to fail in specific cases. It is a solution that satisfies the conditions for a good solution that were proposed by Wright and Sudbury (1977), but does without introducing any principles of reasoning that are introduced *ad hoc* only to interpret the paradox.

However, it may be objected that we missed an essential point in the puzzle. By representing the statement that the exam will be a surprise by the formula S, it became a prediction of what Marilyn can know solely on the basis of examless days passing by. But this seems to be too weak.

5. SENTENCES THAT STATE THEIR OWN SUCCESS

The teacher seems to be expressing something stronger than S: he predicts that the exam will be a surprise *anyway*, update and ratiocination included. This is the reason is why the proof seems convincing that the teacher's announcement is true *because* of the conclusions that Marilyn draws from it.

If the teacher makes a more general prediction, what could he have meant?

(5) "The exam will be a surprise, also after learning this"

Of course, if announcements were always successful, adding "also after learning this" would be superfluous. But as the previous analysis as shown, the statement that the sentence is a surprise is *not* always successful, so the addition may add non-trivial information.

There are a number of more or less natural interpretations of this statement, of which we will discuss three representative ones. The point here is not to argue that either reading is the right one, but rather to draw attention ways in which certain readings of the teacher's announcement can be represented in dynamic epistemic semantics.[3]

Before we start, however, observe that there is no proposition in dynamic epistemic semantics that will allow us to capture the whole informal argument. For suppose δ expresses such a proposition. Marilyn concludes that her information is inconsistent after learning δ; i.e., $[\delta]K\perp$ is true. This implies that $K\neg\delta$: Marilyn knew already, before the announcement, that δ was false. That leaves two possibilities: either δ was false, or her information was mistaken already before the teacher said anything. Since there is nothing in the puzzle that points to the latter, the former must have been the case. In short, if Marilyn's argument is valid in dynamic epistemic semantics, then teacher, whatever he meant precisely, must have been lying.

That said, let us look at some ways that sentence (5) could be interpreted.

One reading is to have 'this' refer simply to the statement that the exam is a surprise:

(6) $\mathsf{S} \wedge [\mathsf{S}]\mathsf{S}$.

This sentence expresses a straightforward proposition in dynamic epistemic semantics. In short: in the context in which Marilyn has not ruled out any days yet, the first conjunct provides Marilyn with the information that the exam is not on Friday, just as before. The second conjunct says that after the students learned that the exam is not on Friday, it will remain a surprise, which means that it will not be on Thursday either. So, the sentence (6) gives Marilyn exactly the information to exclude the last two days of the week.

Another reading of the teacher's announcement is to take the word 'this' to refer to the announcement itself. He would be expressing a sentence δ for which the following equation is valid:

(7) $\delta \leftrightarrow \mathsf{S} \wedge [\delta]\mathsf{S}$.

A sentence δ that satisfies this equivalence is interesting, because it allows us to mimic much of the reasoning in the puzzle in dynamic epistemic semantics. It is not difficult to show by an induction on days of the week, just like in the informal argument, that if Marilyn learns δ, then her information state becomes inconsistent: $[\delta]K\bot$ can be derived from (7) in dynamic epistemic logic. Moreover, the very fact that Marilyn's information is inconsistent allows us to conclude that the exam will be a surprise after the announcement: $[\delta]\mathsf{S}$ follows from $[\delta]K\bot$.

However, there is no sentence δ in dynamic epistemic logic that satisfies the equivalence of (7). Such a sentence would be 'contingently paradoxical:' in certain situations, it is both true and false.[4] To see this, take the case where the teacher plans to give the exam on Wednesday. Since $\sigma_{\mathsf{init}}, s_{\mathsf{we}} \models \mathsf{S}$ and $\sigma_{\mathsf{init}}, s_{\mathsf{we}} \models [\delta]\mathsf{S}$, it holds that $\sigma_{\mathsf{init}}, s_{\mathsf{we}} \models \delta$. But at the same time, since $\sigma_{\mathsf{init}}, s_{\mathsf{we}} \models [\delta]K\bot$ and $s_{\mathsf{we}} \in \sigma_{\mathsf{init}}$, it must hold that $\sigma_{\mathsf{init}}, s_{\mathsf{we}} \models \neg\delta$. Which is a contradiction.

One last reading that may be interesting is a sentence that states, as it were, its own success.

(8) $\delta \leftrightarrow \mathsf{S} \wedge [\delta]\delta$.

Any sentence δ that satisfies this equation must be a contradiction – the sentence can never be true. The falsity of any such δ can be proved by a very similar argument as the previous one; to see that it need not be paradoxical, it suffices to substitute $\bot$ for δ in the equation to obtain a valid statement.

6. CONCLUSIONS

We have examined the surprise exam paradox from the viewpoint of dynamic epistemic logic. Central in this analysis is the notion of success, or rather its failure: we argued that one possible way of explaining the reasoning in the puzzle is to assume that the teacher's announcement is successful, but that this assumption is not warranted in dynamic epistemic semantics.

Success can fail in many subtly different ways. Ditmarsch and Kooi (200x) introduce a useful distinction between three notions of success of formulas: those that when true, do not alter their truth value when learning them (sentences for which $\varphi \rightarrow [\varphi]\varphi$ is valid), those that are believed after learned ($[\varphi]K\varphi$ is valid), and, in the case where there are more agents involved, those that, when learned, become common knowledge. Whether a sentence is successful in any of these three senses clearly depends on the logic in which the definitions are framed: different logics of information (e.g., $K45$ or $S5$) will define different classes of successful sentences.[5]

These three classes of non-successful sentences can be divided into subclasses. Concentrating on the second notion of success, we find that some sentences, like Moore's paradox $p \wedge \neg Kp$, are 'anti-successful' in the sense that learning always leads to believing its negation ($[\varphi]K\neg\varphi$ is valid); some are successful only when inconsistent with previous information (such as the sentence S, for which $[\varphi]K\varphi$ is true only when $[\varphi]K\bot$), some are contingently successful ($[\varphi]K\varphi$ is not valid, but neither is its contrary). Many sentences are guaranteed to stabilize after a certain number of updates: if you repeat the sentence often enough, it will come to be believed or disbelieved, but there are, in the multi-agent case, also sentences that can be repeated infinitely often without coming to be believed.

This landscape of unsuccessful sentences has not been studied in any detail yet. One question that immediately arises is whether there exists a characterization of non-successful sentences that is stated in more direct terms than validity. Another point of interest is the precise relation between sentences that are blindspots and sentences that are not successful: clearly, the former is a subclass of the latter, but it seems that the relation must be deeper.

Even if blatant occurrences of unsuccessful sentences are rare in daily life, assertions of lack of knowledge of the hearer do appear, more or less hidden, in a range of puzzles, and it has been shown in several places that an analysis using dynamic epistemic semantics can give an independently motivated account of the reasoning involved in them.

One type of puzzle that has been analyzed using dynamic epistemic semantics involves a repetition of what seems to be the same piece of information. Examples are the puzzle of the muddy children (alternatively known as the wise men, the cheating husbands, etc) which are analyzed in Plaza (1989), Gerbrandy (1998) and in Ditmarsch and Kooi (200x), the Conway Paradox (van Emde Boas et al. 1984), and a puzzle known as 'Mr. Sum and Mr. Product' (Plaza 1989). Each of these puzzles can be analyzed as involving the repeated announcement with an unsuccessful sentence φ a number of times. The failure of success is crucial in explaining that such a sentence can also carry new information when it is announced a second or a third time.

A second type of puzzle revolves around the announcement of a true sentence φ in which it is stated that the hearer lacks information of a certain kind, together with an inductive argument in which the hearer reaches a false or contradictory conclusion. A typical example is, of course, the paradox of the surprise exam under scrutiny here (Sorensen 1988; Gerbrandy 1998; Van Ditmarsch and Kooi 2006). An analysis in terms of dynamic epistemic logic such as the one given here explains how the inductive argument revolves around the assumption that sentences are successful, and that the content of the announcement may not be.

Finally, dynamic semantics adds an extra dimension to truth-value based semantics, which provides us with some leeway in analyzing problems such as Moore's paradox (cf. Gillies 2001) and the Fitch Paradox (van Benthem 1997). The dynamic epistemic analysis has also been used by Veltman (1996) to model certain forms of non-monotonic reasoning, in which, typically, sentences that are first believed to be true can become false after adding more information.

To conclude, the strength of the analysis of such situations using dynamic epistemic semantics is that the sometimes complex behavior of learning certain sentences can be explained using a simple and straightforward model of learning. The subtle distinctions that one can make in dynamic epistemic semantics between the effect of learning different types of sentences throws light on a number of situations that are baffling otherwise.

ACKNOWLEDGEMENT

The author is supported by the Lagrange Project of the Fondazione CRT. He would like to thank Guido Boella, Paul Harrenstein, Barteld Kooi, Luigi Sauro and two anonymous referees for useful comments on earlier versions of this paper, and the University of Turin for its hospitality.

NOTES

[1] We will use the phrases like 'knowing,' 'believing' and 'having the information that' interchangeably. We do not think that the distinction between these concepts is directly relevant in the present context.
[2] More traditionally, a Kripke model is a tuple $(W, \rightarrow, V)$ that consists of a set of worlds, an accessibility relation over these worlds, and a valuation function. We get a Kripke model from a pair σ, s by setting $W = \sigma \cup \{s\}$, setting $s' \rightarrow s''$ iff $s'' \in \sigma$, and $V(s)(p) = s(p)$.
[3] This approach is similar in style to that of Halpern and Moses (1986).
[4] Baltag (2003) (who discusses the puzzle as well) defines a language for dynamic epistemic logic that allows for a certain limited form of self-reference, and makes a similar observation.
[5] Ditmarsch and Kooi (200x) show that the three notions of success are equivalent in $S5$. They are distinct in $K45$.

REFERENCES

Alchourrón, C. E., P. Gärdenfors, and D. Makinson: 1985, On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic* **50**, 510–530.

Baltag, A.: 2003, Logics for communication: reasoning about information flow in dialogue games. Lecture notes for NASSLLI 2003, http://www.indiana.edu/ nasslli/ program.html.

Baltag, A. and L. S. Moss: 2004, Logics for epistemic programs. *Synthese* **139**(2), 165–224.

Binkley, R.: 1969, The surprise examination in modal logic. *Journal of Philosphy* **65**, 127–136.

Chow, T. Y.: 1998, The surprise examination or unexpected hanging paradox. *American Mathematical Monthly* **105**(1), 41–51.

Gerbrandy, J.: 1998, Bisimulations on planet Kripke. Ph.D. thesis, Universiteit van Amsterdam. ILLC Dissertation Series DS-1999-01.

Gerbrandy, J. and W. Groeneveld: 1997, Reasoning about Information Change. *Journal of Logic, Language and Information* **6**(2), 147–169.

Gillies, A. S.: 2001, A new solution to Moore's paradox. *Philosophical Studies* **105**, 237–250.

Halpern, J. and Y. Moses: 1986, Taken by surprise: the paradox of the surprise test revisited. *Journal of Phosophical Logic* **15**(3), 281–304.

Heim, I.: 1982, On the semantics of definite and indefinite noun phrases. Ph.D. thesis, University of Amherst.

Landman, F.: 1986, Towards a theory of information. Ph.D. thesis, Universiteit van Amsterdam. Also appeared as GRASS 6 with Foris Publications, Dordrecht, Holland/Cinnaminson, U.S.A.

Montague, R. and R. Kaplan: 1960, A paradox regained. *Notre Dame Journal of Formal Logic* **3**.

Plaza, J.: 1989, Logics of public communications. In: M. Emrich, M. Pfeifer, M. Hadzikadic, and Z. Ras (eds.), *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*. New York: Academic Press, pp. 201–216.

Sober, E.: 1998, To give a surprise exam, use game theory. *Synthese* **115**, 355–373.

Sorensen, R.: 1988, *Blindspots*. Clarendon Press.

Stalnaker, R. C.: 1978, Assertion. In: P. Cole (ed.), *Pragmatics (Syntax and Semantics 9)*. New York: Academic Press, pp. 315–312.

van Benthem, J.: 1997, On what one may come to know. *Analysis* **64**(2), 95–105.

van Ditmarsch, H. and B. Kooi: 2006, The secret of my success. *Synthese* **151**(2), 201–232.

van Emde Boas, P., J. Groenendijk, and M. Stokhof: 1984, The Conway paradox: its solution in an epistemic framework. In: J. Groenendijk, T. M. V. Janssen, and M. Stokhof (eds.), *Truth, Interpretation and Information: Selected Papers from the Third Amsterdam Colloquium*. Dordrecht: Foris Publications, pp. 159–182.

Veltman, F.: 1996, Defaults in update semantics. *Journal of Philosophical Logic* **25**, 221–261.

Wright, C. and A. Sudbury: 1977, The paradox of the unexpected examination. *Australasian Journal of Philosophy* **55**, 41–58.

Dipartimento di Informatica
University of Turin
Corso Svizzera 185
10149 Torino
Italy
E-mail: jelle@gerbrandy.com