UvA-DARE (Digital Academic Repository)

Knowledge Attributions in Context of Decision Problems

van Rooij, R.

DOI

[10.1007/978-3-319-02943-6_7](https://doi.org/10.1007/978-3-319-02943-6_7)

Publication date

2014

Document Version

Submitted manuscript

Published in

Epistemology, Context, and Formalism

[Link to publication](#)

Citation for published version (APA):

van Rooij, R. (2014). Knowledge Attributions in Context of Decision Problems. In F. Lihoreau, & M. Rebuschi (Eds.), *Epistemology, Context, and Formalism* (pp. 109-125). (Synthese Library; Vol. 369). Springer. https://doi.org/10.1007/978-3-319-02943-6_7

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Knowledge attributions in context of decision problems

Robert van Rooij

December 6, 2010

Abstract

1 Introduction

Ever since Dretske (1972) it has been argued that also standard knowledge attributions are context dependent. But in this, they are not alone. It is quite clear, for instance, that whether a particular attitude attribution is counted as true or not can vary from context to context. This is true, for instance, for *de re* belief attributions. Consider Quine's (1956) Ralph who, one evening, sees a man with a brown hat whose suspicious behaviour leads Ralph to believe that the man is a spy. This man happens to be Ortcutt. On another occasion, Ralph sees the same man at the beach, but he does not recognize him as the same man; and the thought that the man he sees at the beach is a spy does not even occur to him. Now, does Ralph believe that Ortcutt is a spy or not? That is, is the following sentence true or not?

- (1) Ralph believes of Ortcutt that he is a spy.

It is widely assumed (e.g. Stalnaker (1988)) that whether (1) is true or not depends crucially on the conversational context. If only the first half of the story is told, and nothing is presupposed, or salient, about Ralph's meeting of Ortcutt at the beach, (1) might intuitively be counted as true. Similarly, if only the second half of the story is told, (1) should be counted as being false. Only if both ways in which Ralph is acquainted with Ortcutt are equally salient in the conversational context, (1) doesn't seem to be (unambiguously) true nor (unambiguously) false. What is important about this example is that (1) can be true in one conversational context and false in another, although Ralph himself believes the same in both contexts.

As a second example, it is also uncontroversial that 'Knowing who' statements are context dependent. Intuitively, one knows who Pele is if one knows an appropriate complete answer to the question, *To what does Pele refer to?* The answer, *Except for Johan Cruyff and Maradona, the best soccer player of the world* seems clearly an adequate and appropriate answer (at least in one context). So if Mary can answer that question, then she knows who Pele is (at least in that context). But in that context, she may not know who Edison Arantes do Nascimento is, in the sense that she may not know who 'Edison Arantes do Nascimento' refers to, although this is the real name of Pele.

In this paper I will show that also knowledge attributions that involve *embedded questions* are context dependent, and that this context dependence involves *decision problems*, just as the interpretation of standard answers to questions. I will also indicate that *knowledge that* sentence are context dependent in a very same way. As a result, so I will argue, the analysis differs from the standard analyses by not just looking at relevant possible worlds. Instead, on this analysis the notion of fine-grainedness plays an important role. I will use the framework of Optimal Assertions, introduced by Benz (2007) and developed by Benz & van Rooij (2008) to account for optimal interpretations of assertions.

2 Optimal answers

2.1 Context dependence of questions and answers

There has been a controversial debate about whether or not strongly exhaustive answers have a prominent status among the set of all possible answers. Groenendijk & Stokhof (1997) are the most prominent defenders of the view that they constitute the basic answer, whereas other types of answers have to be accounted for pragmatically. For a constituent question like ‘*Who came to the party?*’ a complete answer has to tell us for each person whether he or she came to the party or not. This is important for the interpretation of embedded interrogatives: If Peter knows who came to the party, then Peter knows whether John came to the party, and whether Jeff came to the party, and whether Jane came to the party, etc. The set of all possible answers is then the set of all strongly exhaustive answers.¹ On the other hand, there are so-called ‘mention-some’ questions like ‘Where can I buy Italian wine?’ the ‘complete’ answers of which do not seem to imply exhaustivity. Similarly, the sentence ‘Peter knows where to buy an Italian wine in Amsterdam’ can be true, intuitively, without it being the case that for each x Peter knows whether he can buy Italian wine at x , where x ranges over all stores in Amsterdam. For reasons like this van Rooij (2003a,b,c) proposed that the conventional meaning of a question (and its answers) is underspecified, and that the exact meaning is determined by means of context, in particular, by means of the relevant *decision problem*. In Benz & van Rooij (2008) this analysis is embedded in the game theoretic setting of Optimal Assertions: what is expressed by an interrogative sentence, and what is implied by its answers, depends on the decision problem at stake in the conversation. It was argued that thus the truth value of the whole sentence ‘John knows where he can buy an Italian wine’ depends on Peter’s decision problem as well. In this paper we will argue that the meaning of the embedded question depends not only on Peter’s decision problem, but also on what is relevant to the participants of a conversation.

2.2 Decision problems

Let Ω be the set of all possible states of the worlds. For simplicity we restrict our considerations to situations with countable many possibilities, i.e. to countable Ω s.

¹If Ω is a set of possible worlds with the same domain D , and $\llbracket \phi \rrbracket^v$ denotes the extension of predicate ϕ in v , then a strongly exhaustive answer to question $?x.\phi(x)$ is a proposition of the form $\llbracket v \rrbracket_\phi := \{w \in \Omega \mid \llbracket \phi \rrbracket^w = \llbracket \phi \rrbracket^v\}$; i.e. it collects all worlds where predicate ϕ has the same extension. The set of all possible answers is then given by $\llbracket ?x.\phi(x) \rrbracket^{GS} := \{\llbracket v \rrbracket_\phi \mid v \in \Omega\}$. This poses a problem for mention-some answers as they are not elements of $\llbracket ?x.\phi(x) \rrbracket^{GS}$, hence not answers at all.

We represent an agent's expectations about the world by a probability distribution over Ω , i.e. a real valued function $P : \Omega \rightarrow \mathbf{R}$ with the following properties: (1) $P(v) \geq 0$ for all $v \in \Omega$ and (2) the sum of all $P(v)$ equals 1. For sets $A \subseteq \Omega$ we set $P(A) = \sum_{v \in A} P(v)$. Hence $P(\Omega) = 1$. We represent the agent's preferences over outcomes of actions by a real valued utility function over action–world pairs. We collect these elements in the following structure:

Definition 2.1 *A decision problem is a triple $\langle (\Omega, P), \mathcal{A}, U \rangle$ such that (Ω, P) is a countable probability space, $\mathcal{A}$ a finite, non-empty set and $U : \mathcal{A} \times \Omega \rightarrow \mathbf{R}$ a function. $\mathcal{A}$ is called the action set, and its elements actions. U is called a payoff or utility function.*

Let us now assume that our agent, Ann, faces a *decision problem*, i.e. she wonders which of the alternative actions in $\mathcal{A}$ she should choose. It is standard to assume that rational agents try to maximise their expected utilities. The *expected utility* of an action a is defined by:

$$EU(a) = \sum_{v \in \Omega} P(v) \times U(a, v). \quad (2.1)$$

Suppose that Ann receives some information, e.g. proposition C . In probability theory the effect of learning a proposition C is modelled by *conditional probabilities*. Let H be any proposition, e.g. the proposition that one sells Italian wine at the station. H collects all possible worlds in Ω where this sentence is true. Let C be some other proposition, e.g. the answer given by Bob. Then, the probability of H *given* C , written $P(H|C)$, is defined by:

$$P(H|C) := P(H \cap C) / P(C). \quad (2.2)$$

This is only well-defined if $P(C) \neq 0$. In terms of this conditional probability function, we can now define the *expected utility of a after learning C* by:

$$EU(a|C) = \sum_{v \in \Omega} P(v|C) \times U(a, v). \quad (2.3)$$

Assuming that Ann is a utility maximizer, it can be expected that after Ann learned C , Ann will choose that action which maximizes the expected utility after learning C , i.e. $\max_a EU(a|C)$.

2.3 Support problems and optimal assertions

Ann, the hearer, learned C from somebody else, a speaker. Assuming that a speaker is rational as well, he chose to inform Ann of that proposition which he thinks is most useful to give. Instead of hearer and speaker, we will from now on talk about the *decision maker*, D , and the *expert*, E . The decision maker has a decision problem, $\langle (\Omega, P_D), \mathcal{A}, U \rangle$ and we will assume for simplicity that this problem is common knowledge (perhaps after she asked a question). In order to get a model for the questioning and answering situation we have to add a representation for the answering expert's situation. In principle this involves a whole decision problem for the expert. For reasons of simplicity, however, we will assume that the expert's utility function coincides with that of the decision maker, and only add a probability distribution P_E that represents his expectations about the world:

Definition 2.2 A support problem is a five-tuple $\langle \Omega, P_E, P_D, \mathcal{A}, U \rangle$ such that (Ω, P_E) and (Ω, P_D) are countable probability spaces, and $\langle (\Omega, P_D), \mathcal{A}, U \rangle$ is a decision problem. We call a support problem well-behaved if (1) for all $A \subseteq \Omega$: $P_D(A) = 1 \Rightarrow P_E(A) = 1$ and (2) for $x = D, E$ and all $a \in \mathcal{A}$: $\sum_{v \in \Omega} P_x(v) \times U(a, v) < \infty$.

The first condition for well-behavedness is included in order to make sure that E 's answers cannot contradict D 's beliefs. It implies that for sets $A, B \subseteq \Omega$: $P_E(A) = 1 \Rightarrow P_D(A) > 0$ and $P_D(A|B) = 1 \ \& \ P_E(B) = 1 \Rightarrow P_E(A) = 1$. The second condition in the definition is there in order to keep the mathematics simple.

A support problem represents just the fixed static parameters of the answering situation. We assume that D 's decision does not depend on what she believes that E believes. Hence her epistemic state (Ω, P) represents just her expectations about the actual world. E 's task is to provide information that is optimally suited to support D in her decision problem. Hence, E faces himself a decision problem, where his actions are the possible answers. The utilities of the answers depend on the way how they influence D 's final choice. We look at the dependencies in more detail. We find two successive decision problems:

We assume that the answering expert E is fully cooperative and wants to maximise D 's final success. Hence, E 's payoff is identical with D 's (our representation of the *Cooperative Principle*). E has to choose his answer in such a way that it optimally contributes towards D 's decision. Due to our assumption that D 's information is mutually known, E is able to calculate how D will decide. Hence, we represent the decision process as a sequential two-person game with complete coordination of preferences. We find a solution, i.e. optimal answers and choices of actions by calculating backward from the final outcomes. The following model will be worked out using standard techniques of game and decision theory. We concentrate on *ideal* dialogue.

2.4 Calculating optimal answers by backward induction

D 's decision situation First we have to consider the final decision problem of D . In the previous section we have determined the *expected utility after learning A* by:

$$EU(a, A) = \sum_{v \in \Omega} P(v|A) \times U(a, v).$$

If the decision maker D tries to maximise expected utilities by her choice, it follows that she will only choose actions that belong to $\{a \in \mathcal{A} \mid EU_{(\Omega, P_D)}(a, A) \text{ is maximal}\}$. In addition we assume that D has always a preference for one action over the other, or that there is a mutually known rule that tells D which action to choose if this set has more than one element. In this case we can write a_A for this unique element. In short, we assume that the function $A \mapsto a_A$, for $P_D(A) > 0$, is known to E .

***E*'s decision situation** According to our assumption, questioning and answering is a game of complete coordination (Principle of Cooperation). We have implemented this assumption by taking *E*'s payoff function to be identical with *D*'s payoff function *U*. In order to maximise his own payoff, *E* has to choose an answer such that it induces *D* to take an action that maximises their common payoff. We use definition (1.1) for calculating the expected utility of an answer $A \subseteq \Omega$. With a_A as defined above we get:

$$EU_E(A) := \sum_{v \in \Omega} P_E(v) \times U(a_A, v). \quad (2.4)$$

We add here a further Gricean maxim, the *Maxim of Quality*. We call an answer *admissible* if $P_E(A) = 1$. The Maxim of Quality is represented by the assumption that the expert *E* does only give admissible answers. This means that he believes them to be *true*. For a support problem $s = \langle \Omega, P_E, P_D, \mathcal{A}, U \rangle$ we set:

$$Adm_s := \{A \subseteq \Omega \mid P_E(A) = 1\}. \quad (2.5)$$

Hence, the set of optimal answers for *s* is given by:

$$Op_s = \{A \in Adm_s \mid \forall B \in Adm_s : EU_E(B) \leq EU_E(A)\}. \quad (2.6)$$

Assuming that the expert is making an optimal assertion, the inquirer can conclude from *E*'s assertion that *A* that she is in a support problem *s* where it holds that $A \in Op_s$. Because, by assumption, she knows already *E*'s utility function, *D* can learn a lot about what kind of worlds *E* takes to be probable.

We have argued that an informed speaker can and should use backward induction to determine which answer he should give. Notice that by our use of backward induction, the informed speaker assumes that the hearer will perform that act which has the highest expected utility after she has updated her belief by standard Bayesian conditionalization with the *semantic* meaning of the answer. This doesn't mean, however, that the hearer just interprets the answer simply at face value: On the assumption that the speaker is informed of her own decision problem and chooses his answer by making use of backward induction, the hearer can conclude more from the answer than just its standard semantic meaning.

We saw that an answer must be an element of Op_s for any support problem $s = \langle \Omega, P_E, P_D, \mathcal{A}, U \rangle$. If we assume that *E* has complete knowledge of all relevant facts, then it follows that for the actual world *w* it holds that $P_E(w) = 1$. But let's consider only the weaker assumption that *E* knows which action $a \in \mathcal{A}$ is optimal:

$$\exists a \in \mathcal{A} \forall v \in \Omega (P_E(v) > 0 \rightarrow \forall b \in \mathcal{A} : U(a, v) \geq U(b, v)). \quad (2.7)$$

Let's assume that *E* answered *A*, then the decision maker *D* knows that *E*'s answer is optimal, i.e. that $A \in Op_s$, hence that:

$$P_E(A) = 1 \wedge \forall B : (P_E(B) = 1 \rightarrow EU_E(B) \leq EU_E(A)). \quad (2.8)$$

As $EU_E(B) = \sum_{v \in \Omega} P_E(v) \times U(a_B, v)$, it follows with (2.7) and (2.8) that

$$P_E(\{v \in \Omega \mid \forall b \in \mathcal{A} : U(a_A, v) \geq U(b, v)\}) = 1. \quad (2.9)$$

From this *I* can infer that:

$$w \in \{v \in \Omega \mid \forall b \in \mathcal{A} : U(a_A, v) \geq U(b, v)\}. \quad (2.10)$$

Interesting about this inference is that although *E* determines via backwards induction what he should assert by making use of naive Bayesian updating on the hearer's side, the speaker in fact realizes that on the basis of this assumption *I* will update her beliefs via a more sophisticated method than conditionalization.

3 Decision problems and embedded questions

3.1 Mentions some questions

In the last section we showed that the actual interpretation of an answer can be calculated from the fact that this answer is optimal. Before we consider the interpretation of embedded questions, let us first look at what our analysis predicts for standard mention-some questions.

Let us consider the mention-some question in (2) in the natural situation.

(2) I: Where can I buy Italian wine?

E: At the station and at the Bijenkorf but nowhere else. (*SE*)

E: At the station. (*A*) / At the Bijenkorf. (*B*)

The answer (*SE*) is called *strongly exhaustive*; it tells us for every location whether they sell Italian wine or not. The answers (*A*) and (*B*) are called *mention-some* answers.

We denote by a, b the actions of going to the station and going to the Bijenkorf. There may be other actions too. Let $A \subseteq \Omega$ be the set of worlds where one sells Italian wine at the station, and $B \subseteq \Omega$ where one sells Italian wine at the Bijenkorf. We represent the payoffs just as in the previous section: For every possible action $c \in \mathcal{A}$ the utility value is either 1 (success) or 0 (failure); especially we assume that $U(a, v) = 1$ iff $v \in A$, else $U(a, v) = 0$; $U(b, v) = 1$ iff $v \in B$, else $U(b, v) = 0$.

It is easy to see that $EU_D(A) = EU_D(SE)$ and it similarly holds that $EU_D(B) = EU_D(SE)$. This shows that for the inquirer it doesn't matter which information – exhaustive or not – she receives, as long as it is true. Thus, all the answers are equally useful with respect to conveyed information and the inquirer's goals. What we have to show now, however, is that all answers are equally optimal for the answering *expert*. We will show that $EU_E(A) = EU_E(B) = EU_E(SE) = 1$ if A, B and SE are admissible answers, and thus known to be true by the expert.

We start with answer *A*: If E knows that A , then A is an optimal answer. If learning A induces I to choose action a , i.e. if $a_A = a$, then the proof is very simple:

$$EU_E(A) = \sum_{v \in \Omega} P_E(v) \times U(a_A, v) = \sum_{v \in A} P_E(v) \times U(a, v) = 1.$$

Clearly, no other answer could yield a higher payoff. Obviously, also B is optimal if E knows that B . The same result follows for any stronger answer, including the strongly exhaustive answer SE , $A \wedge B$ or $A \wedge \neg B$. This shows that their expected utilities are all equal as long as they are admissible answers. Hence, all these answers are equally good and E can freely choose between them.

3.2 Context-dependence of embedded questions

Let us now see whether our analysis of optimal answers is relevant for the interpretation of embedded questions as well. At first it might seem that this cannot be the case. Consider

(3) a. Peter knows where he can buy Italian wine.

b. Peter knows where he can *best* buy Italian wine.

As argued for in the previous sections, the first sentence is true if Peter knows any place where he can buy Italian wine. If the set of optimal answers is identical with the meaning of the embedded sentence, however, then we should expect that (3a) has the same meaning as (3b). Intuitively, however, this is not the case: if the price of the wine counts as well, although is only of marginal importance, for (3b) to be true, Peter has to know not only that he can buy Italian wine in a certain store, but also that he can't buy this wine cheaper in any other place.

Still, we will argue that we can account for this distinction by making the meaning of (3a), but not that of (3b), depend on what is relevant in the conversational situation. In the following we present a number of examples that strongly suggest that this is the correct way to proceed. If the only relevant property of stores is the fact that they sell Italian wine, then any mention–some answer is an optimal answer, and the only thing that can be inferred from the fact that Peter knows an optimal answer is the fact that he knows some place where it is possible to buy Italian wine. But if there are more relevant properties, we get different interpretations, including interpretations where it is implied that Peter knows where he can best buy Italian wine. The following examples should all be read in the context where Peter, the office assistant, was sent to buy Italian wine for an evening dinner.

- (4) In the afternoon Ann tells Bob that Peter went shopping but that he returned without wine. Bob gets very angry about it.

Ann: Maybe, it was not his fault.

Bob: Oh, Peter knows where he can buy Italian wine.

- (5) In the afternoon Ann tells Bob that Peter bought some Italian wine but it was obviously completely overpriced. Bob gets very angry about it.

Ann: Maybe, it was not his fault.

Bob: Oh, Peter knows where he can buy Italian wine.

- (6) In the afternoon Ann tells Bob that Peter bought some Italian wine but it took a long time because he went to one of the wine shops in the centre and he was caught in the city traffic. Bob gets very angry about it.

Ann: Maybe, it was not his fault.

Bob: Oh, Peter knows where he can buy Italian wine.

In (4) we get the same interpretation as in (3a). In (5) and (6), however, there are other things that count as well. In (5) “Peter, knows where he can buy Italian wine” must intuitively be interpreted as meaning that he knows a place where he can buy it cheaply, while from (6) we can conclude that Peter knows a place close by where he can avoid the traffic in the city.

3.3 Accounting for the examples

As illustrated with examples (4), (5), and (6), knowledge attributions involving embedded questions depend crucially on context. In particular, it depends on which attributes are relevant in the conversational situation.

If we want to apply the game theoretic model of optimal answers to the semantics of embedded interrogatives, we have first to decide whose role, i.e. the expert's or the

interrogator's, Peter takes in a sentence like: *Peter knows where to buy Italian wine*. The following examples indicate that he takes the expert's role:

- (7) a. Ann: Where can I buy Italian wine?
 Bob: Ask Peter, he knows where to buy Italian wine.
- b. Ann: Does this train stop in Flensburg?
 Bob: I am the first time on this train but Peter knows whether the train stops there.

As before we assume here that the answering expert (Peter) knows all relevant facts, i.e. we assume that (2.7) holds. This implies that:

$$\exists a \in \mathcal{A} : P_E(\{v \in \Omega \mid \forall b \in \mathcal{A} : U(a, v) \geq U(b, v)\}) = 1.$$

Let's set $O(a) := \{v \in \Omega \mid \forall b \in \mathcal{A} : U(a, v) \geq U(b, v)\}$, the set of worlds where action a is optimal. Then we say that E knows *what to do* in a given support problem $s = \langle \Omega, P_E, P_D, \mathcal{A}, U \rangle$ if $\exists a \in \mathcal{A} : P_E(O^s(a)) = 1$. Notice that this means that if we look only at the embedded question 'What should John do?', we might give it the following representation: $\{O^s(a) \mid a \in \mathcal{A}\}$. In the case of our example (3a) in the situation as described before, for example, it results in $\{\{w_1, w_3\}, \{w_2, w_3\}\}$. In w_1 and w_3 , one of the optimal actions is going to the station, while in w_2 and w_3 one of the best actions is going to the Bijenkorf.

In this model of (3a) we assumed that Peter doesn't care where he should go to (in Amsterdam), as long he goes to a place where he can buy Italian wine at that place. In that case, (3a) and (3b) are actually equivalent. But now suppose that in w_3 , it is closer to walk to the station than to walk to the Bijenkorf: $U(a, w_3) > U(b, w_3)$. We may then assume our tourist to prefer to walk to the station in w_3 . So, in that case, the Bijenkorf is not one of the best places to walk to in w_3 , and the question 'What is the best action to do?' is not represented by $\{\{w_1, w_3\}, \{w_2, w_3\}\}$, but rather by $\{\{w_1, w_3\}, \{w_2\}\}$. In this case we know that Peter's actual utility function depends on (i) whether one sells Italian wine at that place, and (ii) how far it is to walk to that place. However, in a particular conversational context, it might be common ground among the conversational partners that only one of those – the first one, for instance – really counts. In such a situation (3a) can still be counted as true in w_3 , although Peter takes w_2 to be a possible alternative. Notice that in this case (3b) is intuitively counted as false, because by mentioning 'best' the utility function looks at all attributes that Peter considers relevant himself.

If two conversational situations differ in what is considered to be relevant by the interlocutors, then this defines two different support problems. Hence, if we ask whether it is true in a world w that *Peter knows where he can buy Italian wine*, then we have to consider w together with a support problem. Let $s = \langle \Omega, P_E, P_D, \mathcal{A}, U \rangle$ be a given support problem and $w \in \Omega$. We can then say that *E knows where he can buy Italian wine* is true in (w, s) iff $\exists a (P_E(O^s(a)) = 1 \wedge P_E(w) > 0)$. But, as we assume that the expert knows all relevant facts, the last condition is always true. The difference between the truth values of (3a) and (3b) is then due to the difference of the associated support problems. The effect of *best* in *E knows where he can best buy Italian wine* is to accommodate additional attributes which we would not consider if we read *E knows where he can buy Italian wine*.

4 Knowledge that

4.1 Context dependence and decision problems

Until now we considered knowledge where attributions. It seems intuitively clear that knowledge who, and knowledge what behave in very much the same way: also here the interpretation seems to depend on what is at stake. More interesting is the questions whether our analysis of optimal assertions can account for the context dependence of knowledge that attributions as well, and in what sense the resulting analysis differs from the standard analyses.

Suppose Peter is at the zoo next to the zebra cage with his son. The zebras are in plain view and when his son asks him what they are, Peter tells him. It make all sense for us to say that Peter *knows* that they are zebras. Still, Peter can't really rule out completely that they are not mules cleverly disguised by the zoo authorities to look like zebras. Dretske (1970) suggests that although Peter can't he still *knows* that the animals he saw were zebras. It is perhaps not true anymore, however, when the possibility that they are cleverly disguised mules is brought up. Thus, the meaning of knowledge that attributions are context dependent.

But not just any context dependence will do for 'know'. For instance, although there is some similarity with gradable adjectives like 'tall', Stanley (2004) quite clearly showed that there are important disanalogies as well. Whereas 'tall' allows for modifiers (*very*), fits well in comparatives, and is sensitive to comparison classes ('the fly is tall, but the elephant is not'), 'know' does not. This is all just to show that the context dependence won't consist in relating the meaning of the word *with respect to a particular scale*. But that leaves open many other alternatives.

According to the standard context-dependent analysis (e.g. Lewis, 1996), Peter knows A is true iff Peter can rule out all (relevant) $\neg A$ worlds. It depends on context what the relevant $\neg A$ worlds are. Thus, $know(p, A)$ is true in w for modal model M iff for all of Peter's epistemically accessible worlds v it holds that, if v is among the (most) relevant worlds, A must be true in v (and M). It depends on the conversational context what are the (most) relevant alternative worlds. The sentence 'Peter knows that he looks at zebras' is true because worlds where he looks at cleverly disguised mules are not considered.

If one determines relevance with respect to decision problems, a slightly different picture arises. Suppose that the relevant decision problem is which action in $\mathcal{A}$ should be chosen. In that case it seems most natural to say that John knows A is true in w (and M) iff this entails that he knows which action in $\mathcal{A}$ he should choose. Thus, just like for knowledge-where attributions, it is natural to assume that in terms of the Optimal Assertion framework, Peter plays the expert role.

Let $s = \langle \Omega, P_E, P_D, \mathcal{A}, U \rangle$ be a given support problem and $w \in \Omega$, where Peter is the expert. The above discussion suggests that $Know(p, A)$ is true in (w, s) iff $\exists a (P_E(O^s(a)) = 1 \wedge P_E(w) > 0)$. But this seems much too weak: the truth of the knowledge attributions seems to be *independent* of (the semantic value of) A . There is an obvious way, however, to make the embedded sentence relevant: it should be the case that A 'corresponds' somehow to the optimal action in the support problem. The natural idea would be to say that Peter knows that there is water in the bathtub really means (in this context) that Peter knows that he looks at zebras rather than *lions*, because in the relevant support problem two actions are at stake: a and b , and a is best iff it are zebras (or cleverly disguised mules) that he looks at, and b is best iff it are lions. Thus, the actions correspond with alternative sentences, and $Know(p, A)$

is true iff the knowledge attributions with the alternatives are false. In this sense the analysis is close to Schaffer's (2004) contrastive analysis of knowledge attributions.

Notice that context-change might influence the truth of knowledge attributions. If the support problem were more fine-grained, involving also an action c that is better than a in mules-worlds, but worse than a in zebra-worlds, Peter doesn't know anymore that he is looking at zebras, meaning that he wouldn't know anymore which action is preferred in all worlds of the relevant decision problem. Thus, even if nothing changed about Peter's knowledge, the truth of the knowledge attribution might still change. Thus, if the stakes higher, more actions become relevant, and it is more difficult for the knowledge attribution to be true.

4.2 Granularity

In the previous section I suggested that indistinguishability, and thus, coarse-grainedness, plays a role in knowledge attributions. But we haven't been very clear yet about the semantics. If we want to account for knowledge, and knowledge attributions, we have to consider modal models with Peter's epistemic accessibility relation R_p . But some modal models make more fine-grained distinctions than others. In one modal model, M' , we don't distinguish situations, or worlds, where the animals Peter sees are zebras or mules cleverly disguised in that way, while in another more fine-grained model, M , we do. Let M' be a model in which we don't distinguish zebra-worlds from mules-worlds. Then we would like to say that Peter knows that he looks at zebras is true with respect to M' iff in all 'worlds' v in M' that are epistemically accessible to Peter it is true that he looks at zebras. But recall that v is now a coarse-grain world, i.e. not really a single world, but stands proxy for a *set* of worlds. How do we determine these coarse-grained worlds, and under what circumstances is a sentence true in such a world?

We will first consider the last question, but ignore for the moment the possible worlds. Rather than looking at when a sentence is true in a coarse-grained world according to a modal model, we will first look at the parallel question when a predicate is true of a coarse-grained 'individual' in a model.

In natural language we conceptualize and describe the world at different levels of granularity. Hobbs (1985) argues that to represent or conceptualize the world at a coarser-grained level, we can just restrict ourselves by looking only at the *relevant* predicates of our original language. Consider a model $M = \langle D, I \rangle$ for the first-order language $\mathcal{L}$, and take $\mathcal{L}'$ to be a sublanguage of $\mathcal{L}$ containing only its 'relevant' predicates. In terms of the monadic predicates of $\mathcal{L}'$ we can now define an equivalence relation ' $\sim_{\mathcal{L}'}$ ' with respect to language $\mathcal{L}'$: $a \sim_{\mathcal{L}'} b$ iff $a, b \in D_M$ and for all monadic predicates P of $\mathcal{L}'$: $M \models P(a) \Leftrightarrow M \models P(b)$.² In terms of this equivalence relation, Hobbs (1985) proposed to construct a coarse-grained model M' as follows: (i) the domain $D_{M'}$ is just the set of equivalence classes $D_{M'} = \{\{y \in D : y \sim_{\mathcal{L}'} x\} : x \in D_M\}$, and (ii) the interpretation function is such that for all monadic predicates $P \in \mathcal{L}'$, $M' \models P([a])$ iff $M \models P(a)$, where $[a]$ denotes the equivalence class containing $a \in D_M$.

Hobbs (1985) appealingly suggests to account for *type*-identity as identity at a more coarse-grained level of description. He so explains why we cannot say 'A Ford Falcon was heading south on U.S. 101, went out of control, and crashed into the

²In general this obviously won't do: one has to look at relations as well. Let us ignore this, however.

same car' to mean that it hit another Ford Falcon. The reason is that *type*-level identity is just indistinguishability, but only restricted to distinguishable predicates that are *relevant*. Unfortunately, Lasersohn (2000) showed that the truth definition at the coarse-grained level proposed by Hobbs (1985) does not capture the intuitive motivation. It is clear that (i) 'I own a Ford Falcon. The same car is owned by Enzo.' should be interpreted with respect to a coarse-grained model. According to Hobbs' construction, $M' \leq M$ just in case if for every monadic predicate $P \in \mathcal{L}'$, if $P([a])$ is true in coarse-grained model M' , it has to be the case that $P(b)$ is true in fine-grained model M , for *every* $b \in [a]$. However, it is clear that in (i) the predicates 'Owned by me' and 'Owned by Enzo' are relevant, and thus part of $\mathcal{L}'$. Because in M' it is the same car that has both of these properties, Hobbs' construction falsely predicts that every token of this car should have both properties in M as well.

Instead of making use of *universal* quantification as proposed by Hobbs (1985), why not make use of *existential* quantification? We assume a surjective function f from the domain of M , D_M , to the domain of coarser-grained model M' , $D_{M'}$,³ that preserves each relevant predicate P : if $x \in P_M$, then $f(x) \in P_{M'}$. The other direction follows by contraposition: if $x \notin P_{M'}$, then there is no $y \in f^{-1}(x)$ such that $y \in P_M$. To capture the idea of simplification, or coarsening, it is natural to assume that f is *not injective*: it might be that $f(x) = f(y)$, although $x \neq y$. Of course, we want refinements to preserve all the predicates and relations of the restricted language $\mathcal{L}'$, but this preservation is now stated as follows: $M' \leq_{\mathcal{L}'} M$ just in case if $x \in P_{M'}$, then $\exists y \in f^{-1}(x) \in P_M$, for each $P \in \mathcal{L}'$. But is it not problematic that the predicates 'Owned by me' and 'Owned by Enzo' are both relevant, and thus part of $\mathcal{L}'$? Yes, and no! Yes, if one wants the logic of M' to be the same as the logic of the finest-grained M . But no, if one is more flexible here. In particular, if we don't mind that at coarse grained M' for some sentences A it might hold that A is both true *and* false. To account for this feature, I will define two notions of truth simultaneously: a notion of *weak truth* and a notion of *strong truth*.⁴ Intuitively, it is the notion of weak truth that we are after, but it is convenient to state this notion partly in terms of the strong notion. In general, the weak and strong truth conditions of sentences in course-grained model M' are simultaneously defined in terms of their truth conditions in fine-grained model M and function f as follows (where we make use of the substitution analysis for simplicity and where $\underline{a}$ is the unique name of a):

$$\begin{aligned} M' \models_w P(\underline{a}) & \text{ iff } \exists d \in f^{-1}(a) : M \models P(\underline{d}) \\ M' \models_w \neg A & \text{ iff } M' \not\models_s A \\ M' \models_w A \wedge B & \text{ iff } M' \models_w A \text{ and } M' \models_w B \\ M' \models_w \forall x A & \text{ iff for all } d \in I_{M'} : M' \models_w A[\underline{d}/x]. \end{aligned}$$

$$\begin{aligned} M' \models_s P(\underline{a}) & \text{ iff } \forall d \in f^{-1}(a) : M \models P(\underline{d}) \\ M' \models_s \neg A & \text{ iff } M' \not\models_w A \\ M' \models_s A \wedge B & \text{ iff } M' \models_s A \text{ and } M', g \models_s B \\ M' \models_s \forall x A & \text{ iff for all } d \in I_{M'} : M' \models_s A[\underline{d}/x]. \end{aligned}$$

In particular, it holds that $M' \models_w \neg P(\underline{a})$ iff $\exists d \in f^{-1}(a) : M \not\models P(\underline{d})$. Notice that for all A it holds that if $M' \models_s A$ it follows that $M' \models_w A^f$ (where A^f is the translation

³A function f from D to D' is surjective iff the range of f is D' .

⁴These notions are intimately related with the *tolerant* and *strict* notions of truth defined in van Rooij (2011) and Cobreros et al. (2011). For formal developments of these notions see especially the latter paper.

of A corresponding to the surjective function f) but not necessarily the other way around. As it turns out, the weak logic at M' is rather like Priest's (1979) Logic of Paradox.

4.3 Knowledge attributions again

Back to our modal models. The indistinguishability relation should now be one between worlds. Thus, a modal model consists now of coarse-grained worlds and a sentence is strong or weakly true with respect to such a coarse-grained world (given a more fine-grained modal model and the surjective function f). The parallelism can be highlighted by making use of two-sorted logic: worlds are just seen as arguments of predicates, and we say that atomic proposition q is (weakly) true in coarse-grained world v' in model M' , $M' \models_w q(v')$, just in case $\exists v \in f^{-1}(v') : M \models q(v)$.

Now we have to decide how the indistinguishability relation between worlds is determined. The idea is that this indistinguishability relation is determined by the support problem s . We don't distinguish worlds where the utility ordering of actions in $\mathcal{A}_s$ (the actions in s) is the same with respect to U_s (D 's utility function in s). Thus, $u \approx_s v$ iff_{def} $\forall a, b \in \mathcal{A}_s : U_s(a, u) \leq U_s(b, u) \text{ iff } U_s(a, v) \leq U_s(b, v)$. As for the above example, two worlds u and v that are exactly alike except that Peter looks at zebras in v and at cleverly disguised mules in u are considered to be irrelevantly different with respect to $\mathcal{A}$ and U , i.e. $u \approx_s v$, iff the utility orderings of the actions in $\mathcal{A}$ are the same.

Now we say that $Know(p, A)$ is (weakly) true in v' of coarse-grained modal model M' iff $\forall u' \in R_p(v') : M' \models_w A(u')$. If A is atomic, this, in turn, is true iff $\forall u' \in R_p(v') : \exists u \in f^{-1}(u') : M \models_w A(u)$. This is a very weak notion: if Peter cannot distinguish zebras from mules cleverly disguised to look like zebras, it predicts that Peter also knows that the animals he sees are cleverly disguised mules. But, then, it is not true at coarse-grained world w' that we were after; we wanted to determine when the knowledge attribution is true with respect to our fine-grained world w . The idea is that the truth or falsity of this attribution is determined *via* the truth or falsity of the attribution of the coarse-grained model, but that doesn't mean that it is equated with it. In particular, if in w it are zebras Peter is looking at, we should predict that the attribution that "Peter knows that the animals he sees are cleverly disguised mules" is false. This can be done by assuming not only that the knowledge attribution is true in coarse-grained world w' as seen above, but also that the embedded sentence is true in fine-grained world w (one might call this an implementation of Lewis's (1996) rule of actuality).⁵

In what sense is this different from just restricting the worlds as Lewis proposed? Indeed, it is closely related with it. The Lewisian strategy would be to stay with one pointed modal model $\langle \Omega, w, R, D, I \rangle$ (where $w \in \Omega$ represents the actual world and where R_p is Peter's epistemic accessibility relation), but when interpreting a knowledge claim, one only looks at those epistemically accessible worlds that are most *relevant*. The actual world w is obviously among the relevant worlds, but which

⁵Bach (2005) argued against context dependent analyses of knowledge attributions, because even with respect to any such model it can still be asked whether Peter *really* knows A , and thus that the sceptical charge is not really met. Peter *really* knows A , if A is true in all (relevant *and irrelevant*) epistemically accessible worlds. I believe that this criticism is unacceptable: the worlds in Ω are just *representations*, and one of them is the representation of the actual world. But representations can be finer and finer grained, and there need not be a *finest grained* representation. It is very natural to relate this impossibility with the power of the sceptic: you will never quite satisfy her.

other worlds should be in there as well? In terms of our framework, I would like to make the following proposal (which we already argued for implicitly): just as with knowledge-where attributions, also for knowledge-that attributions it is Peter that functions as the expert (at least with respect to the issue at stake). Now look again at our definition of a support problem in section XXX:

Definition 4.1 A support problem is a five-tuple $\langle \Omega, P_E, P_D, \mathcal{A}, U \rangle$ such that (Ω, P_E) and (Ω, P_D) are countable probability spaces, and $\langle (\Omega, P_D), \mathcal{A}, U \rangle$ is a decision problem. We call a support problem well-behaved if (1) for all $A \subseteq \Omega : P_D(A) = 1 \Rightarrow P_E(A) = 1$ and (2) for $x = D, E$ and all $a \in \mathcal{A} : \sum_{v \in \Omega} P_x(v) \times U(a, v) < \infty$.

What is important for now is the first constraint for support problems to be well-behaved: for all $A \subseteq \Omega : P_D(A) = 1 \Rightarrow P_E(A) = 1$. By our assumption that E is an expert (with respect to the relevant issue), it seems only natural to say that the expert *knows* A if $P_E(A) = 1$. But this means that if E considers a world epistemically possible, this also holds for D . It follows that the set of epistemically possible worlds for the expert (at least as far what is relevant for D 's decision problem) is a subset of the set of epistemically possible worlds for the decision maker. For knowledge attributions I think it is natural to assume that D 's decision problem represents the decision problem of at least one of the participant of the conversation. But this means that all the worlds that D takes to be possible are also worlds that are compatible with what is presupposed by the participants of the conversation. Thus, if E considers a world epistemically possible, this world is also compatible with what is presupposed by the participants of the conversation. It follows that if a world is incompatible with what is presupposed by the participants of the conversation, this world is also not among E 's epistemically possible worlds.

Of course, this need not really be the case in general: Peter does not know everything *we* presuppose. But my claim is that this is the case as long as we limit ourselves to what is relevant in the conversational situation. Putting this all together, this comes down to the proposal that in the context of support problem s , the set of relevant possible worlds in a conversational situation is just the set of worlds that are compatible with what we presuppose, which might be thought of as the following set: $\{w \in W : P_D(w) > 0\}$. And this helps: even if Peter cannot distinguish zebras from cleverly disguised mules, if *we* presuppose that the animals Peter looks at are zebras, the knowledge attributions that Peter knows that it are zebras that he looks at is counted as true as well.

What is the relation between Lewis's analysis (with the assumption that world v is taken to be irrelevant iff $P_D(v) = 0$) and ours? Obviously, the two come down to the same iff there is a one to one correspondence between the worlds that have a non-zero probability w.r.t. P_D and the coarse-grained worlds determined by the support problem s . I take this to be the typical case, which means that the two indeed give rise to the same predictions in these case. But perhaps our analysis does something extra: it looks in addition to what the agent (Peter) thinks from his own perspective.

References

- [1] Benz, A. (2006), 'Utility and Relevance of Answers', in: A. Benz, G. Jäger, R. v. Rooij (eds.): *Game Theory and Pragmatics*, Palgrave, Mcmillan.
- [2] Benz, A. and R. van Rooij, (2007), 'Optimal Assertions and what they implicate, A uniform game theoretic approach', *Topoi*, **26**: 63-78.

- [3] Cobreros, P., P. Egge, D. Ripley, and R. van Rooij, (2011), ‘Tolerant, classical, strict’, *Journal of Philosophical Logic*, to appear.
- [4] Does, J. van der & M. van Lambalgen, (2000), ‘A logic of vision’, L&P;
- [5] Dretske, (1970), ‘Epistemic operators’, *Journal of Philosophy*;
- [6] Grice, P. (1989), *Studies in the Way of Words* Harvard University Press, Cambridge Massachusetts.
- [7] Groenendijk, J.. and M. Stokhof, (1984), *Studies in the Semantics of Questions and the Pragmatics of Answers*, Ph.D. thesis, University of Amsterdam.
- [8] Hobbs, J. (1985), ‘Granularity’, *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-85)*.
- [9] Lasersohn (2000), ‘Same, Models and representations’, *Proceedings SALT X*.
- [10] Levinson, S. *Presumptive Meanings. The Theory of Generalized Conversational Implicatures*, MIT Press, Cambridge, Mass.
- [11] Lewis, D. (1996), ‘Elusive knowledge’, *Australian Journal of Philosophy*;
- [12] Nunberg, (1985), ‘Individuation in Context’, *Proceedings of WCCFL*;
- [13] Priest, G. (1979), ‘The logic of paradox’, *Journal of Philosophical Logic*, **8**: 219-241.
- [14] Rooij, R. van (2003a), ‘Questioning to Resolve Decision Problems’, *Linguistics and Philosophy*, **26**: 727–76.
- [15] Rooij, R. van (2003b), ‘Utility of Mention some Questions’, *Research on Language and Computation*, **2**: 401-416.
- [16] Rooij, R. van (to appear), ‘Vagueness, tolerance and non-transitive entailment’, to appear.
- [17] Schaffer, J. (2004), ‘From contextualism to contrastivism’, *Philosophical Studies*, **119**: 73-104.
- [18] Schulz, K. and R. van Rooij, (2006), ‘Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation’, *Linguistics and Philosophy*, 2006, **29**: 205-250.
- [19] Stalnaker, R. (1988), ‘Belief attribution and context’, In: R. Grimm, D. Merrill (eds.): *Contents of Thought*, U. of Arizona Press, Tuscon.
- [20] Stalnaker, R. (1993), ‘Twin Earth revisited’, *Proc. Aristotelian Society*;
- [21] Stalnaker, R. (1996), ‘Knowledge, belief and counterfactual reasoning in games’, *Economics and Philosophy*
- [22] Stanley, J. (2004), ‘On the linguistic basis of contextualism’, *Philosophical Studies*.