Now that you mention it: Awareness Dynamics in Discourse and Decisions

Michael Franke & Tikitu de Jager*

ILLC, Universiteit van Amsterdam
M.Franke@uva.nl, S.T.deJager@uva.nl

WORKING DRAFT, July 2008
Please do not cite or quote without permission.
All comments gratefully received.

Abstract

We model unawareness of possibilities in decision-making and (linguistic) pragmatic reasoning. A background model is filtered through a state of limited awareness to provide the epistemic state of an agent who is not attending to all possibilities. We extend the standard notion of awareness with ASSUMPTIONS (implicit beliefs about propositions the agent is unaware of) and define a dynamic update for 'becoming aware.' We give a propositional model and a decision-theoretic model, and suggest that decision problems should in general be seen as filtered models in this sense, describing only those features of the situation which the modeller considers relevant and the agent is aware of. We show how pragmatic relevance reasoning can be described in this framework, extending a standard definition to the case of awareness updates. An utterance can be relevant even if semantically uninformative, if it brings relevant alternatives to awareness. This gives an explanation for the use of possibility modals and questions as hedged suggestions, bringing possibilities to awareness but only implicating their degree of desirability or probability.

*The authors are listed in alphabetical order.

Contents

1 Introduction

EXAMPLE 1: Little Bo Peep. Little Bo Peep has lost her keys and doesn't know where to find them. She's checking her pockets, the nail behind the door, beside the telephone, and so on, but the keys are *nowhere* to be found. Frustrated and pouting, Bo slams onto the sofa. From his corner Little Jack Horner helps her out:

> JACK: Did you leave them in the car when you came in drunk last night?

Little Bo slaps her forehead (and his, impudent little scamp) and goes out to the car. At this point of the story, it is equally reasonable that the search ends successfully, or to imagine Bo's frustration continuing until she finds her keys several days later in the sugar jar.

Nobody should have trouble understanding at an intuitive level the kinds of changes Little Bo Peep's epistemic state undergoes in this example. Bo didn't think of the car as a possible place to look for the keys, but when Jack mentioned it to her, her oversight struck her as foolish and she promptly took the proper action. Should anybody be worried about a situation as commonplace as this?

Peculiarities show up, though, once we take a closer look. Most importantly, Jack manages to change Bo's epistemic state in quite a significant way, simply by asking her a question. Since under any standard semantic analysis questions are uninformative, there is some explaining to be done here. Before we get sidetracked by considering rhetorical questioning or possible pragmatic analyses, though, consider the following alternatives to Jack's helpful observation:

(1) JACK: Do you think it's possible the keys are in the car?

(2) JACK: The keys might be in the car.

(3) JACK: [*Not paying attention*] Hey, this TV show is really funny, this guy is looking everywhere for his keys and they were in his car the whole time!

(4) ADVERTISER ON TV: Do *you* forget your keys in the car? You need the ExtendaWristLock KeyChain! (Patent pending.) Order now and pay just $19.99!

(5) PASSING MOTORIST: *Honk honk!*

Bo's response to any of these might quite naturally be the same: she slaps her forehead at her foolishness and goes immediately to check the car. While the first two should be amenable to pragmatic explanation, this clearly won't do for the others.

Intuitively what's going on here is that Bo is failing to consider a possibility, which when brought to her attention she realises should not be ruled out. We will say that Bo is UNAWARE of the possibility that the keys might be in the car and we will investigate this kind of epistemic attitude both formally and in linguistic applications in this paper. Indeed, something like our notion of awareness has been approached sidelong in the linguistic literature, but until quite recently has received no thorough treatment. The recent exception is work by Eric Swanson on meaning and use of the epistemic modal 'might' ([Swa06b], extended in [Swa06a]) which is very close in spirit, but not in specifics, to our approach. Outside linguistics, however, a formal notion of unawareness of possibilities has been explored very fruitfully in theoretical computer science (the standard reference for the origin of the field is [FH88]) and has recently been applied variously in rational choice theory, i.e., game and decision theory ([Fei04; Fei05; HR06]; we describe other related work in Section 5). We aim to combine these insights with a linguistic treatment, and show the surprising range of outstanding linguistic puzzles that the notion is applicable to.

So what is unawareness of a contingency? Here are three basic interrelated properties of the notion in question.

Slogan 1: Unawareness is not uncertainty. Little Bo Peep's behaviour (pouting, sitting on the couch) does not indicate uncertainty about whether the keys are in the car. At the point of the story where Bo gives up the search, it is implausible to assume that she puts any credence, however small, on the possibility that the keys are in the car; if she did, she would surely have gone and checked. Judging from her behavior, if anything, it seems as if Bo believes that the keys are *not* in the car. We will say that Bo has an implicit belief that the keys are not in the car and argue that, firstly, implicit beliefs are different from the more familiar explicit beliefs (see the next two slogans) and that, secondly, unawareness of a possibility typically goes together with an implicit belief we call an assumption.[1] Implicit beliefs are typically assumptions of 'normality' (we'll discuss them in more detail, including the connection to closed-world reasoning, in Section 2.3).

Slogan 2: Unawareness is not introspective. Although Bo's behavior indicates an implicit belief, she does not *explicitly* (or consciously) believe that the keys are not in the car. In fact, she holds *no* explicit beliefs about the car, not even the tautological belief that the keys are either in the car or elsewhere. A self-referential way to get to grips with this fundamental intuition is to say that she is unaware of her own implicit beliefs. This failure of (negative) introspection[2] leads us to a definition of awareness in terms of the language Bo uses: if she were asked to describe all her beliefs about the keys she would not mention the car at all. (If prompted she might start enumerating tautologies: "The keys aren't on the table, so I suppose logically speaking they're either on the table or not on the table"; she would never, however, extend this sequence with "The keys are either in the car or not in the car".) The formal model we give in Section 2 will distinguish syntactically between the agent language Bo would use spontaneously to describe her explicit beliefs and the language we as modellers use to describe her implicit beliefs.

[1] Although the point will not be appreciated until we have presented our model, in the interests of fairness we should point out that the slogan "Unawareness is not uncertainty" has a slightly different connotation in the rational choice literature. We describe the differences between the two notions in Section 5.

[2] She does not know, but does not know *that* she does not know. [MR94] and [DLR98] show that unawareness must be more than simply failure of negative introspection, if we are to capture the properties we want. Still, starting with this notion is a good way to get the intuitive juices flowing in the right direction.

Slogan 3: Unawareness is easily overturned. Bo's implicit belief is very fragile, in a way that her explicit beliefs are generally not: it does not take any particularly convincing argument to overturn it. The example shares with other crucial examples in this paper what we might call the 'forehead-slap property': in becoming aware of her implicitly held belief, Bo realises the mistake she made in overlooking a certain (intuitively: relevant) possibility. We will concern ourselves exclusively with unawareness of this kind in this paper: unawareness through *inattentiveness* or *absent-mindedness*.[3] Indeed, as the alternative (5) in Example 1 shows, overturning this kind of unawareness need not even involve anything linguistic or intentional. But where language is concerned, we argue that the mere *mentioning* of some possibility that an agent is unaware of is sufficient for this unawareness to be overturned. This is why it is not possible to talk to Bo about her implicit beliefs: if Jack were to ask her the question in (1), we might imagine her answer to be something like: "Well, *now that you ask me* I do, but I wouldn't have if you hadn't."

1.1 Pragmatic considerations (Farmer Pickles bakes a cake)

Not being attentive to all possibly relevant factors in, say, a decision-making situation such as Bo's key-search is anything but unusual. Similarly, it is perfectly natural for conversationalists to attend to possible unawareness. The most interesting cases seem to occur when a speaker has enough information to motivate an awareness update, but not enough to say anything stronger (such as "I saw your keys in the car," "You always leave your keys in the car," or similar). Since questions usually don't carry assertive force, they are often a good way to produce *only* an awareness update, with no corresponding speaker commitments. (Other possibilities include possibility statements with "might" or "could", explicit epistemic hedging as in "They're not in the car, I presume", and so on.)

It is because Jack finds it likely that Bo is unaware of the car as a possible hide-out for the keys that he asks his question; we feel, even more strongly, that Jack first and foremost *intends* to make Bo aware with his question (he does not want an answer in the first place; he is not himself interested in the information he is 'officially' asking for). Moreover, it is natural for Jack to assume that thus making Bo aware will have a relevant impact on her decisions to act. And, Bo, in turn, might recognize that Jack is making her aware of something which he deems relevant to the case at hand. This recognition of the nature of Jack's conversational move might, of course, trigger (or at least license) further pragmatic reasoning on Bo's side: for instance, Bo might conclude that, since Jack is obviously a helpful (though impudent) little scamp, he himself must deem it sufficiently likely that the keys might be in the car.

In Bo's case the awareness update produces the desired effect without needing any pragmatic reasoning (this is why a passing motorist can unintentionally trigger exactly the same update for Bo). However this need not be the case, as shown in the following example.

Example 2: Bob the Baker. Bob (who is an expert baker) is visiting his friend Farmer Pickles (who isn't).

[3] Again, this distinguishes our approach from the rational choice literature; see Section 5.

PICKLES: I was going to bake a cake but I haven't got any eggs!
BOB: Did you think of making shortbread instead?
PICKLES: I didn't, in fact I didn't even know that you don't need eggs to make shortbread! Thanks, Bob!

Since Pickles isn't an experienced baker, his unawareness of shortbread as an option conceals a real uncertainty about whether the recipe requires eggs or not. Simply overturning this unawareness by accident (as the passing motorist might have done for Bo) would not produce the effect we see here: flipping through his cookbook he might see a photograph of shortbread, but he would have to check the recipe to see whether there are eggs in the ingredients list or not. Assuming that Bob is being helpful, though, Pickles can reason as follows: "Bob is deliberately bringing up a possibility because he thinks I'm overlooking it; that's only helpful if I should in fact consider it; then he should at least hold it possible that shortbread doesn't require eggs; but Bob is an expert baker, so he wouldn't be uncertain about such things; so he must *believe* that you don't need eggs to make shortbread." Pickles' response, showing that he makes this inference, seems perfectly natural; but it can only be justified (that crucial first step) by taking the sort of unawareness perspective we argue for.

1.2 Paper overview

We will be concerned with the dynamics of an agent's awareness and its role in conversation in this paper. Our main aim is to apply insights from the study of unawareness in rational choice theory to linguistics. In particular, we would like to show how awareness dynamics are applicable to a surprising range of outstanding linguistic puzzles. Towards this end, we seek to represent unawareness in (multi-speaker) discourse and investigate how and to what effect conversation changes the epistemic states of conversationalists.

In the next section we will introduce a formal representation of unawareness and implicit belief. Section 3 describes the dynamics of awareness updates and their interaction with factual information growth, while Section 4 works out an example of pragmatic reasoning based on awareness dynamics in detail. Finally, we site our work in the intersection between linguistics and rational choice in Section 5, before Section 6 concludes.

2 Formalising unawareness

We'll start by presenting a simple propositional model of awareness and its dynamics, just enough to model Little Bo Peep's predicament, in Section 2.1. We will then extend this basic propositional treatment to decisions under growing awareness in Section 2.2. In both sections, we will start by developing the basic formal notions alongside a few intuitive examples; in the decision-theoretic case in particular the definitions will undergo some revision as we introduce complications. For the reader's convenience, all final definitions are collected at the end of each section, in Definitions 1 and 2 on page 8 (for the propositional case) and Definitions 3–7 in Section 2.2.4 (for decision problems in full detail).

2.1 The propositional case

We start with a set $\mathcal{P}$ of PROPOSITION LETTERS, representing as usual statements about how the world might be. For Bo these express the location of the key: "(they're in her) pocket", "nail", "phone", "car", and "sugar-jar" and so on. A POSSIBLE WORLD w is associated with a valuation function $v_w : \mathcal{P} \to \{0,1\}$ as is usual. Since in fact (and according to Bo's most fundamental beliefs) the keys can only be in one place at a time, we only need to look at a few of the combinatorially possible worlds, and we can give them the same names as the propositions themselves: "pocket" is the world where the proposition "pocket" is true and none of the others is. We'll call this full set W: $W = \{\text{pocket}, \text{nail}, \text{phone}, \text{car}, \text{sugar-jar}\}$.

In a standard model, then, Bo's epistemic state would be the set of worlds she has not ruled out by observation or other trustworthy information sources. We'll call this her INFORMATION SET σ. We would standardly assume that Bo's initial information set σ_0, before her search starts, is simply W, i.e., her information rules out none of the worlds in question (this may of course change as she learns more). However according to our observation of Bo's behaviour, her epistemic state as our story opens seems instead to be $\{\text{pocket}, \text{nail}, \text{phone}\}$ (these are the places that she goes on to check before giving up in frustration). We'll capture this by filtering her information set through an AWARENESS STATE α, which models the proposition letters she is unaware of, and the assumptions she holds about their valuation. (Strictly this might be better named an "unawareness state"; we will use the terms interchangeably.) Formally, an awareness state α is a pair $\langle \mathfrak{U}, \mathfrak{v} \rangle$ where $\mathfrak{U} \subseteq \mathcal{P}$ is the set of UNMENTIONABLES (proposition letters the agent is unaware of; *we* will mention them frequently but the agent herself may not) and $\mathfrak{v} : \mathfrak{U} \to \{0,1\}$ is a valuation function giving the ASSUMPTIONS the agent holds. In Bo's case, we initially have $\mathfrak{U} = \{\text{car}, \text{sugar-jar}\}$ and $\mathfrak{v} = \{\text{car} \mapsto 0, \text{sugar-jar} \mapsto 0\}$, i.e., she assumes (in typical 'default' fashion) that the keys are not in the car and not in the sugar jar. Taking a different perspective, an awareness state α specifies a set of worlds $W_\alpha = \{w \in W \,;\, \mathfrak{v} \subseteq v_w\}$, those worlds in W which agree with the assumptions. This latter, equivalent view of awareness states facilitates the definition of filtering through awareness: we'll write $\sigma{\restriction}\alpha$ for Bo's information set FILTERED through her awareness, and define $\sigma{\restriction}\alpha = \sigma \cap W_\alpha$. Taken together, σ captures the complete factual information an agent like Bo has: σ would be her epistemic state if she was aware of all relevant contingencies; W_α is the set of worlds she ENTERTAINS given her (possibly limited) awareness; and $\sigma{\restriction}\alpha$ is the subset of these worlds that her information does not rule out, the ones which generate her beliefs.

As our story opens, Bo has no factual information, but she is unaware of some propositions: $\sigma_0{\restriction}\alpha_0 = W_{\alpha_0} = \{\text{pocket}, \text{nail}, \text{phone}\}$. But, if this is Bo's epistemic state, she should believe that the keys are not in the car. That's true from the modellers perspective (her *implicit* belief) but her explicit beliefs shouldn't mention the car at all. As we argued for in connection with the slogan "Awareness is not introspective", we rely on a *syntactic* notion to capture this: a belief formula ϕ can be EXPLICIT with respect to an epistemic state under unawareness $\sigma{\restriction}\alpha$ only if ϕ does not use any proposition letters in $\mathfrak{U}$. These are the *unmentionables* according to Bo's awareness, and her explicit beliefs must not mention them. (This story is explicitly intensional: what matters is not the extension of the proposition letters, but whether their names appear in $\mathfrak{U}$. This is what allows us to exclude a tautology such as "The keys are either in the car or not in the car" from Bo's explicit beliefs.)

Let's now try to model the different kinds of updates given in the story: factual and awareness. Given Bo's initial information set $\sigma_0 = W$ and awareness $W_{\alpha_0} = \sigma_0 \restriction \alpha_0 = \{\text{pocket}, \text{nail}, \text{phone}\}$, she begins to systematically investigate the three places she is aware of as possible hide-outs for the keys, and eliminates them one by one. Now $\sigma_1 = \{\text{car}, \text{sugar-jar}\}$ but she has not gained awareness of anything new: $\alpha_1 = \alpha_0$, so $\sigma_1 \restriction \alpha_1 = \varnothing$! This explains Bo's frustration: as far as she can see, she is in the inconsistent state. However inconsistency with unawareness is not as destructive as in the standard picture: it's quite natural for Bo to realise that there is (or that there must be) *some* possibility she has missed. Her frustration arises because nothing in the situation gives her any guidance as to what this might be, so there's no reasonable action she can take to get out of the trap she's entered.[4] But, then comes Jack's offhand question from his corner, and the scales fall from Bo's eyes! That is, on hearing an expression mentioning the proposition letter "car", Bo becomes aware of it: it disappears from her (un)awareness state. So $\alpha_2 = \langle \{\text{sugar-jar}\}, \{\text{sugar-jar} \mapsto 0\} \rangle$ and $W_{\alpha_2} = \{\text{pocket}, \text{nail}, \text{phone}, \text{car}\}$; $\sigma_1 \restriction \alpha_2 = \{\text{car}\}$, and it's easy to see why Bo immediately runs to check the car.

For convenience we collect here the formal features of this model.

Definition 1: Propositional unawareness. Let $\mathcal{P}$ be a set of proposition letters and W a set of worlds, with each world $w \in W$ associated with a valuation function $v_w \colon \mathcal{P} \to \{0, 1\}$. An epistemic state for an agent is a pair $\langle \sigma, \alpha \rangle$ with σ an information set (a subset of W representing the worlds that her information has not ruled out) and α is an awareness state. The awareness state $\alpha = \langle \mathfrak{U}, \mathfrak{v} \rangle$ specifies her unmentionables $\mathfrak{U} \subseteq \mathcal{P}$ and assumptions $\mathfrak{v} \colon \mathfrak{U} \to \{0, 1\}$, that is, the proposition letters she is unaware of and the truth-values she unconsciously assumes they hold. The state α gives rise to a set of worlds $W_\alpha = \{w \in W \,;\, \mathfrak{v} \subseteq v_w\}$, the worlds entertained by the agent. An information state under unawareness $\sigma \restriction \alpha$ (also to be read as σ filtered through α) is simply $\sigma \cap W_\alpha$.

For clarity we define here also the syntactic sublanguages we use, although these will feature only implicitly in the rest of the paper.

Definition 2: Syntax and belief statements. The language we define contains two belief operators for each agent: $\mathrm{B_i}$ (for implicit belief) and $\mathrm{B_e}$ (for explicit belief). An awareness state $\alpha = \langle \mathfrak{U}, \mathfrak{v} \rangle$ defines an agent language $\mathcal{L}_\alpha$, the language inductively defined using only the mentionable proposition letters $\mathcal{P} \setminus \mathfrak{U}$ and the explicit belief operator $\mathrm{B_e}$. Implicit belief corresponds to belief in a standard model: $\mathrm{B_i}(\phi)$ holds for an agent in epistemic state $\sigma \restriction \alpha$ iff $\sigma \restriction \alpha$ supports ϕ. However *explicit* belief has a stronger requirement: $\mathrm{B_e}(\phi)$ holds in $\sigma \restriction \alpha$ iff $\mathrm{B_i}(\phi)$ holds *and* $\phi \in \mathcal{L}_\alpha$. (Under this definition all explicit beliefs are implicit; we will often use "implicit belief" loosely where "*strictly* implicit belief" would be more correct.)

[4] Another natural reaction would be to search again the places she has already looked. This shows another way that inconsistency might not be fatal: if some of the information leading to it turns out to be incorrect. However this perspective requires belief *revision* as opposed to update, and has little to do with awareness.

2.2 Decision problems and awareness dynamics

Strictly speaking, the propositional treatment of Bo's growing awareness is a rather crude oversimplification: names such as "pocket" or "car" could at the same time represent states of the world ("(they're in her) pocket") or actions that Bo might wish to execute ("(search in her) pocket"). So, for example, when we concluded that in the epistemic state $\sigma_1 \upharpoonright \alpha_2 = \{\text{car}\}$, where Bo is aware of the car as the only open possibility, she would go check the car, we have silently succumbed to this equivocation between states and actions. But when the identification of propositions and actions is unwarranted, an extension of the analysis of awareness dynamics to decision problems is called for; not least because unawareness of propositions shows first and foremost in the agent's *behavior*. However, formalizing the dynamics of awareness of decision makers is not a trivial task and the final model is rather complex. In order to keep the exposition perspicuous we will have to work towards it in stages. We will start with a naïve approach, present a number of problems that arise, and thus hope to motivate the additional complexity of the solutions we've chosen to apply at each step. We end with final definitions in Section 2.2.4.

2.2.1 The basic picture

A decision problem is usually conceived as a tuple $\langle S, A, P, U \rangle$ where S is a set of relevantly distinct states of the world, A a set of possible actions, P a probability distribution over S, and $U : S \times A \rightarrow \mathbb{R}$ is a utility function giving a numerical desirability for each action in each state. There is a sense in which this definition already implicitly includes unawareness, in its limited set S and, more palpably even, in the limited actions A under consideration: common sense dictates that when modelling a particular decision problem we do not include in S every potentially relevant distinction of the state of the world that may affect the outcome of the agent's choice of action, but only certain distinctions that the agent can entertain herself (given her awareness-limited vocabulary); similarly, and more obviously even, we do not want to include in A all conceivably possible actions but only the ones that the agent is aware of as relevant to the task at hand. One of the main ideas we wish to stress in this paper is that a classical decision problem should be seen as an agent's limited subjective conceptualization of a decision making situation:

Slogan 4: Decision problems represent subjective awareness. A decision problem, which by definition includes only a small set of states and possible actions and thus restricts attention to only a small facet of reality, represents the agent's subjective assessment of the relevant factors of the situation, given her state of awareness.

Here is a simple example for the kind of subjective unawareness represented in decision problems. At the beginning of her search, Bo is aware of the nail, the phone and her pocket as places where her keys might be. Her decision problem $\delta = \langle S, A, P, U \rangle$ which comprises her limited awareness at this point of the story contains exactly these states:

$$S = \{\text{nail}, \text{phone}, \text{pocket}\} .$$

We assume that the actual state is "sugar-jar" but this is a state that Bo is neither entertaining, nor considering possible at the outset. Instead, Bo considers all and only the states in S possible. This is represented in the decision problem δ by assuming that the probability distribution P, which captures Bo's beliefs, has *full support*, i.e., assigns some non-zero probability to all states in S. (By definition it can assign no probability outside the states given by the decision problem.) The actions Bo can take in this key-search scenario correspond one-to-one with the possible states ("(the keys are in her) pocket" and "(search in her) pocket") and so we can, for modelling purposes, use the same names for states and actions: $A = S$. (Whence the constant equivocation in the exposition of the propositional case.) And, of course, since we assume that Bo wants to find the keys, her utility function should be something like (formally some positive linear transformation of):

$$U(s,a) = \begin{cases} 1 & \text{if } s = a \\ 0 & \text{otherwise.} \end{cases}$$

Taken as a whole, then, the decision problem δ represents Bo's own subjective assessment of the decision situation under her own limited awareness.

It is obvious how this model of Bo's epistemic state would treat factual information flow. If Bo learns (for instance, by checking) that the keys are not on the nail, she would revise her probabilistic beliefs (by a simple Bayesian update with the proposition "$\neg$nail"). But what about extending Bo's awareness? Suppose, whatever her probabilistic beliefs P might be, that she becomes aware of the car as a possible hide-out of the keys and of the corresponding action "car". Most straight-forwardly, we would like to update Bo's decision problem δ so as to include a state and action "car". This much is easy. But what should Bo's probabilistic beliefs be after she becomes aware of the new contingency? And what would her utilities be in the new updated decision problem?

Clearly, we would not want to specify these features by hand with every update. We would much prefer a model which fully determines the outcome of an awareness update. This is where the idea of *filtering* that we used in the propositional case applies: in order to model how a single agent's epistemic state changes under growing awareness we assume that there is a structure in the background, called a *background model*, which represents the agent's epistemic state under full awareness; unawareness is then modelled by an awareness state as a restriction, or filter, on the background model; the outcome of the filtering process is (or gives rise to) a decision problem, which is interpreted as the agent's assessment under limited awareness in line with the above slogan. Awareness updates are then fairly simple updates of the awareness state (basically: adding or removing elements from sets), which however may have rather far-reaching repercussions on the agent's decision problem via the background model and filtering.

Here is a first simplified attempt at implementing this architecture for Bo's decision problem. We assume in the background another decision problem $\delta^* = \langle S^*, A^*, P^*, U^* \rangle$ which represents Bo's decision problem under full awareness. According to our slogan this should also represent subjective awareness; indeed, it represents the features the *modeller* is aware of as possibly relevant. So, for this background model in Bo's case we have chosen

$$S^* = A^* = \{\text{nail}, \text{phone}, \text{pocket}, \text{car}, \text{sugar-jar}\}$$

(taking advantage again of the naming convention conflating states, propositions and actions) and appropriate beliefs P^* and utilities U^*. We should consider δ^* the equivalent of the information set σ in the propositional case: δ^* contains all the factual information that Bo would have under full awareness. This background structure δ^* is then filtered through an awareness state, as before in the propositional case. Of course, our propositional awareness states had no component to represent awareness of actions, while Bo's restricted awareness is both a restriction on the set of states S and on the set of possible actions A. Consequently, we need to enrich the notion of an awareness state to include a component $\mathfrak{A}$, analagous to $\mathfrak{U}$, which represents the *actions* the agent is unaware of: our new awareness states will be triples $\langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$ where $\mathfrak{A}$ is a subset of A^* giving the actions the agent does not consider.[5]

In Bo's case, it makes sense to assume that $\mathfrak{U}$ and $\mathfrak{v}$ are as before: Bo's initial awareness state α_0, before she starts her search, has her aware of "nail", "phone" and "pocket" as the possible states and possible actions, so that:

$$\begin{aligned} \mathfrak{U} &= \{\text{car}, \text{sugar-jar}\} \\ \mathfrak{v} &= \{\text{car} \mapsto 0, \text{sugar-jar} \mapsto 0\} \\ \mathfrak{A} &= \{\text{car}, \text{sugar-jar}\}\,. \end{aligned}$$

Just as in the propositional case we can define S_α as the set of states from S^* that are compatible with the assumptions of α.

Filtering δ^* through this awareness state gives us the restricted decision problem δ that we started with. In general, filtering in this case comes down to this: if $\delta^* = \langle S^*, A^*, P^*, U^* \rangle$ is a decision problem and $\alpha = \langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$ then the filtered decision problem $\delta{\restriction}\alpha$ is the decision problem $\langle S, A, P, U \rangle$ with[6]

$$\begin{aligned} S &= S^*_\alpha \\ A &= A^* \setminus \mathfrak{A} \\ P &= P^*(\cdot \mid S^*_\alpha) \\ U &= U^*{\restriction}(S \times A). \end{aligned}$$

The set S_α of states being entertained drives the agent's probabilistic beliefs under limited awareness by updating P, the agent's beliefs under full awareness, with all implicit assumptions the agent is making due to her unawareness. This is exactly what the beliefs in $P(\cdot \mid S_\alpha)$ represent.

Suppose that in Bo's δ^* the probabilities are 0.24 for each of "nail", "phone", "pocket" and "car", and 0.04 for "sugar-jar". In her initial state of unawareness she holds states "nail", "phone" and "pocket" possible, each with probability $\frac{1}{3}$ (because $P(\text{nail} \mid \{\text{nail}, \text{phone}, \text{pocket}\}) = \frac{0.24}{0.96} = \frac{1}{3}$). If she becomes aware of new possibilities without eliminating the existing ones by searching (if Jack helps her out before her search begins, for instance) these probabilities

[5] Some readers might wonder at this point whether it would pay to adopt a representation of decision problems in the style of [Jef65], where actions are treated as propositions. After all, this would allow us to treat unawareness of actions on a par with unawareness of propositions and therefore seems *prima facie* the simpler modelling solution. We reject this alternative for a number of reasons, the most interesting of which in the present context is that we would like to side-step the question which kinds of (implicit) beliefs an agent has about actions that she is unaware of.

[6] Strictly speaking, we'd have to define $P = P^*(\cdot \mid S^*_\alpha){\restriction}S^*_\alpha$. However, here and in the following we rule readability over formal precision.

decrease: to $\frac{1}{4}$ if she becomes aware of the car, and to the limit value of 0.24 under full awareness.

Awareness dynamics are now easy to define. If Bo becomes aware of the proposition and action "car" we simply remove this proposition from $\mathfrak{U}$ and the corresponding assumption from $\mathfrak{v}$ (thus enlarging the set of entertained states S_α), removing also the corresponding action from $\mathfrak{A}$. The background model makes sure that utilities and probabilities are defined in the updated decision problem which is retrieved from filtering through this updated awareness state. Effectively, this filtering process allows an easy implementation of deterministic awareness updates on decision problems: we, as modellers, specify the limit-stage of the agent's growing awareness to the extend that it is important for the modelling purposes. We also have a simple structure that captures which bits and pieces the agent is aware of. Simply adding the parameters that the agent becomes aware of to her awareness state in the most straight-forward fashion produces, via the background decision problem, the new and updated decision problem with all (numerical) information specified correctly.

2.2.2 Refinement 1: Individuating states, unawareness without assumptions

Bo's key-search example is fairly simple because propositions, states and actions correspond one-to-one and so the awareness update involved a nearly trivial extension of the idea of filtering from the propositional case to a richer decision-theoretic structure. Bob's shortbread suggestion in Example 2, on the other hand, requires further scrutiny of the notion of a state and another revision to the notion of an awareness state. Here is why.

Let's first consider the most intuitive background decision model for Farmer Pickles' epistemic state in Example 2. What would his decision problem look like if he were aware of all contingencies that we as modellers are interested in? First of all, Pickles considered baking a cake a possible action and he is made aware of a further possible action, namely baking shortbread. We should maybe allow Pickles to abstain from all baking, but further actions clearly do not play a role, so that for the background decision model $\delta^* = \langle S^*, A^*, P^*, U^* \rangle$ we should assume that

$$A^* = \{\text{cake}, \text{shortbread}, \text{abstain}\}\,.$$

But what should Pickles' assessment of the relevant states be (from the modeller's perspective)? Pickles knows that there are no eggs available, so this is not something that the model needs to distinguish. But there is a relevant piece of subjective uncertainty that we would like to model and that is whether the recipe for shortbread contains eggs or not. So, when fully aware, Pickles would make a distinction between two possible relevant states of affairs, one in which baking shortbread requires eggs and another one in which it does not:

$$S^* = \{\text{sb-req-eggs}, \text{sb-req-no-eggs}\}\,.$$

It is not significant at the moment whether Pickles has any beliefs as to which state is more likely, but we should assume that he does not rule out any state completely. So again we assume that P^* has full support on S^*. As for utilities, it is natural to assume that U^* is a function that orders state-action pairs as follows:

$$\langle \cdot, \text{cake} \rangle, \langle \text{sb-req-eggs, shortbread} \rangle \prec \langle \cdot, \text{abstain} \rangle \prec \langle \text{sb-req-no-eggs,shortbread} \rangle$$

In words: since there are no eggs baking a cake is as bad as baking shortbread if this does indeed require eggs; it's better not to bake anything; but most preferred, of course, is baking shortbread when it in fact does not require eggs. (We might call baking failed cakes simply as a waste of time, or go further and imagine the expressions of the two friends when they bite into a floury mess, as motivation for this utility ordering.)

So, how do we represent Pickles' epistemic state as a decision problem when he is unaware of shortbread as an option for baking? The obvious answer to simply leave out the action "shortbread" in his representation of the situation leaves us puzzling why, if Pickles is unaware of shortbread as an action alternative, his decision problem would nevertheless distinguish the state where the shortbread recipe specifies eggs and where it does not. Rather, an intuitively attractive representation of Pickles' decision situation *before* becoming aware of shortbread should only have one state: there is indeed no subjective epistemic uncertainty about whether baking shortbread requires eggs. But it is also not the case that we should simply *leave out* either one of the two states in S^* in the representation of Pickles' initial state. For, unlike in Bo's example, it does not seem defensible that Pickles holds any assumption about whether shortbread requires eggs. His unawareness of baking shortbread as an action shows in his behavior: he does not attempt to bake shortbread, does not mention it etc. But his behavior, in particular his answer to Bob's suggestion, also shows that he is uncertain about the ingredients of shortbread *after* becoming aware of this alternative action and *before* further pragmatic considerations. Of course, we could assume that Pickles indeed first has an implicit belief (say: shortbread requires eggs), which he then loses as soon as he becomes aware of shortbread baking (leaving him uncertain whether shortbread requires eggs). But this is at best a dirty hack. And it is also not necessary. In fact, we should improve the modelling attempt of the previous section based on this example in two respects: firstly, we should allow for unawareness of relevant propositions that *does not* go along with an assumption, and, secondly, we should consider the states in a decision problem as conglomerates of states (or possible worlds, as we will call them in the final model) that the agent does distinguish as potentially distinct.

To implement these amendments, we should firstly alter the definition of an awareness state $\langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$ to allow $\mathfrak{v}$ to be a *partial* function from $\mathfrak{U}$ to truth-values. This way we can represent which unmentionables an agent holds assumptions about, as well as what the assumptions are. We should also define a reasonable grouping mechanism that specifies which states (or worlds) of the background model together form a state in the decision problem under an agent's limited awareness. We will execute these ideas in section 2.2.4 after motivating a further slight but necessary extension to the model in the next section.

2.2.3 Refinement 2: Unawareness of outcomes

Example 3: Professor Branestawm. The scene is the banquet of this year's prestigious pepper conference.[7] First to the buffet is Professor Branestawm, with his friend Professor Calculus alongside. Branestawm is helping himself to a big bowl of fruit salad, when Calculus is suddenly taken by fear:

> calculus: Hey hold on a second! Won't fruit salad set off one of your allergies?
> branestawm: [*After some thought*] Ah, no, I don't think so. I haven't had an allergic reaction in months.
> calculus: [*Obviously still shaken*] Well, we wouldn't want to repeat the disaster of last year, would we?

Not only would we not want to repeat the disaster of last year, we would also not like to represent the possibility that Branestawm has an allergic reaction *as a state* that Branestawm is uncertain about, in the same way that Pickles in the previous example is uncertain about whether shortbread requires eggs or not. Of course, this is technically possible. We could represent Branestawm's fully aware decision situation with states and actions

$$S^* = \{\text{allergy}, \text{no-allergy}\}$$
$$A^* = \{\text{eat}, \text{abstain}\}\,.$$

But there is something decidedly odd about this representation: unlike in the shortbread-recipe example of the previous section, in which states were distinguished by some describable parameter (whether shortbread takes eggs or not), the only reason we can name for distinguishing states in the Branestawm example is the *outcome* of performing the action "eat". In other words, the only reason for wanting to distinguish states is in order to distinguish different possible outcomes of one of our actions. We prefer instead to add these outcomes explicitly, and to distinguish between two kinds of uncertainty: epistemic uncertainty (about what is currently true in the world) and metaphysical uncertainty about the outcome of inherently unpredictable events.[8]

Consequently, we take it to be much more natural to represent Branestawm's epistemic state under full awareness as a decision problem δ^* that distinguishes states before performance of actions from outcome states. In particular, δ^* has only one current state, two actions as before and three outcome states (of which one is the current state, because it is the outcome of "abstaining", i.e. an empty action that does not change any relevant parame-

[7] The initial P most likely stands for 'Pragmatics'.

[8] Of course, a determinist may wish to defend that the outcome of eating fruit salad is indeed fixed by the true state of affairs, just as a die-hard determinist might even argue that the outcome of a fair coin toss is fixed by the material facts just prior to the toss. Even so the point about naming remains: no determinist will be able find a name for the state leading to a coin-flip landing heads except for some variation on "pre-heads". A determinist not willing to give these arguments the right of way will probably also not be impressed by the argument from probabilistic independence, so we will omit it, except to remark that it is also a claim about relative ease of modelling rather than a statement of impossibility.

ter):

$$S^*_{\text{now}} = \{\text{current}\}$$
$$A^* = \{\text{eat}, \text{abstain}\}$$
$$S^*_{\text{fut}} = \{\text{allergy}, \text{no-allergy}, \text{current}\}$$

We will model the agent's metaphysical uncertainty about the outcomes of actions if performed in a given state as a function giving RESULT DISTRIBUTIONS $\Pi : S^*_{\text{now}} \times A^* \to \Delta(S^*_{\text{fut}})$ that maps each current state and action to a probability distribution on the set of outcome states. (As is standard we write $\Delta(X)$ for the set of all probability distributions on some set X.) In Branestawm's case, Π would be a function that might reasonably look like this:

$$\langle \text{current}, \text{abstain} \rangle \mapsto \begin{cases} \text{allergy} & \mapsto 0 \\ \text{no-allergy} & \mapsto 0 \\ \text{current} & \mapsto 1 \end{cases}$$

$$\langle \text{current}, \text{eat} \rangle \mapsto \begin{cases} \text{allergy} & \mapsto .001 \\ \text{no-allergy} & \mapsto .999 \\ \text{current} & \mapsto 0 \end{cases}$$

On this model utilities can be defined as a function from outcome states to reals: $u \colon S_{\text{fut}} \to \mathbb{R}$, and the utility of an action in a (current) state incorporates an expectation calculation over all possible outcomes.

After these conceptual considerations, it is high time to take stock. We still owe the reader a rigorous presentation of the final model, in particular (i) the background model, (ii) the awareness state and (iii) the filtering mechanisms which produces a decision problem as a representation of the agent's epistemic state given the first two components.

2.2.4 The final model

As in the propositional case of Section 2.1, we will model an agent's unawareness via possible restrictions in the language that she would use to describe her situation. Towards this end, we assume, as before, a set $\mathcal{P}$ of proposition letters that capture the model-relevant distinctions before and after the agent performs an action. Where before we had only one, we now consider two sets of model-relevant possible worlds: a set $\mathcal{W}$ of present worlds before the agent performs an action; and a set $\mathcal{O}$ (for 'outcome worlds') for the state of the world after the agent performed an action. Here, $\mathcal{W}$ and $\mathcal{O}$ need not have an empty intersection. Again, we associate with each world in $w \in \mathcal{W} \cup \mathcal{O}$ a valuation function $v_w : \mathcal{P} \to \{0, 1\}$.[9]

We define the background model in terms of these worlds. This will keep conceptually distinct *worlds* as the minimal modelling units in the background

[9]The division of worlds into current and future is natural in a decision-theoretic setting, where we only consider 'single-step' actions. In a full planning setting the underlying model should be some sort of extended temporal structure allowing for sequences of actions. This simplification will be harmless, so long as we are careful about distinguishing propositions that apply now or in the future 'by hand'.

model from *states* as they occur in decision problems, as representations of the agent's transient subjective epistemic state, even when worlds and states are to be identified, e.g. under an empty or otherwise trivial awareness state. A background model captures the agent's epistemic state under full awareness, just as an information state did in the propositional case.

Definition 3: Background Models. A background model is a structure with six components: $\langle \mathcal{W}, \mathcal{O}, \mathcal{A}, P, \mathcal{U}, \Pi \rangle$ where

- $\mathcal{W}$ and $\mathcal{O}$ are sets of current and outcome worlds;
- $\mathcal{A}$ is a set of actions;
- $P \in \Delta(\mathcal{W})$ a probability distribution on current worlds;
- $\mathcal{U} \colon \mathcal{O} \to \mathbb{R}$ is a utility function giving the desirability of future worlds;
- $\Pi : \mathcal{W} \times \mathcal{A} \to \Delta(\mathcal{O})$ is a function giving result distributions: for each $w \in \mathcal{W}$, $a \in \mathcal{A}$, and $o \in \mathcal{O}$ $\Pi_a^w(o)$ gives the probabilities of outcome o in state w by performing the action a.

Definition 4: Awareness States. An awareness state α is a triple $\langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$ such that $\mathfrak{U} \subseteq \mathcal{P}$ is a set of unmentionables, $\mathfrak{v} : \mathfrak{U} \to \{0, 1\}$ is a (possibly *partial*) valuation function on the set of unmentionables and $\mathfrak{A} \subseteq \mathcal{A}$ is a set of actions. The unmentionables $\mathfrak{U}$ are propositions that the agent is unaware of; the assumptions $\mathfrak{v}$ capture her implicit beliefs or assumptions (where an agent need not hold assumptions about all unmentionables); the actions $\mathfrak{A}$ are likewise those she is unaware of.

Based on an agent's awareness state we can define for future use the set of worlds and outcomes that the agent *entertains*, i.e., the set of worlds or outcomes not ruled out by her assumptions. Since according to our revised notion of awareness states the assumption function $\mathfrak{v}$ may be partial, the set of entertained worlds or outcomes is no longer necessarily the set of worlds the agent can distinguish given her awareness-language. In particular, she may entertain possibilities, because she does not hold any assumption that would rule them out, but still not be able to distinguish these possibilities in her limited vocabulary. (Think of Pickles in Example 2 who could not distinguish a state where shortbread requires eggs from one where it does not, because he is unaware of this distinction, but nevertheless held no assumptions about the recipe for shortbread.) We therefore also define how an agent's limited awareness *aggregates* worlds into states: here we should consider as a single state all those entertained worlds that agree on everything the agent can distinguish in her language. (The aggregation relation will define states in the agent's decision problem, see below.)

Definition 5: Entertaining and Aggregation. Let $\alpha = \langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$ be an awareness state. The worlds and outcomes that an agent in α entertains, i.e. the worlds the agent does not rule out by an assumption, are the sets

$$\mathcal{W}_\alpha = \{w \in \mathcal{W} \,;\, \mathfrak{v} \subseteq v_w\}$$
$$\mathcal{O}_\alpha = \{w \in \mathcal{O} \,;\, \mathfrak{v} \subseteq v_w\}.$$

Furthermore, the agent considers equivalent *by reason of unawareness* two worlds $w, w' \in \mathcal{W}_\alpha$, iff

$$v_w(p) = v_{w'}(p) \text{ for all } p \in \mathcal{P} \setminus \mathfrak{U}.$$

Obviously this is an equivalence relation on $\mathcal{W}_\alpha$, which we write $\equiv_\alpha$; the intuition is however very different from that of the epistemic accessibility relation. Two worlds are equivalent in this sense if the agent is not aware of anything that would distinguish them. We will define the states in a decision problem by AGGREGATION using this relation: a state is simply an equivalence class under $\equiv_\alpha$. (See below for the details.)

A background structure and an awareness state together give us the agent's subjective assessment of her situation. This includes limited awareness and possible implicit beliefs. We capture this in the notion of a filtered model.

DEFINITION 6: Filtered Models. Given a background model $\mathcal{M} = \langle \mathcal{W}, \mathcal{O}, \mathcal{A}, P, \mathcal{U}, \Pi \rangle$ and an awareness state $\alpha = \langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$, the FILTERED MODEL $\mathcal{M}\restriction\alpha$ is the structure (of the same type as the background model) $\langle \mathcal{W}', \mathcal{O}', \mathcal{A}', P', \mathcal{U}', \Pi' \rangle$ where:

$$\begin{aligned}
\mathcal{W}' &= \mathcal{W}_\alpha \\
\mathcal{O}' &= \mathcal{O}_\alpha \\
\mathcal{A}' &= \mathcal{A} \setminus \mathfrak{A} \\
P' &= P(\cdot \mid \mathcal{W}_\alpha) \\
\mathcal{U}' &= \mathcal{U}\restriction\mathcal{O}_\alpha \\
\Pi' &: \mathcal{W}' \times \mathcal{A}' \to \Delta(\mathcal{O}') \text{ is such that } \Pi'(w, a) = \Pi_a^w(\cdot \mid \mathcal{O}_\alpha)
\end{aligned}$$

A filtered model is the same kind of object as a background model; the only direct effect of filtering is to restrict attention to a sub-part of the background model. Both filtered and background models represent an agent's epistemic state (possibly given awareness restrictions) in a decision-making situation. These models *are*, in a sense, decision problems that just contain more information than the classical variety. We can obviously read off a decision problem in its classical guise from any such model, be that filtered or background. The only noteworthy elements in the following construction are the formation of states by aggregation, and the definition of the utilities: here we need to compute expected utilities where expectations are a mixture of epistemic uncertainty (which world in which state am I in?) and metaphysical uncertainty (what might happen if I do such and such?).

DEFINITION 7: Decision Problem. Let $\mathcal{M}$ be a background model $\langle \mathcal{W}, \mathcal{O}, \mathcal{A}, P, \mathcal{U}, \Pi \rangle$ and α an awareness state $\langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$. As above, call the elements of $\mathcal{M}\restriction\alpha$ (the filtered model) $\langle \mathcal{W}', \mathcal{O}', \mathcal{A}', P', \mathcal{U}', \Pi' \rangle$. The agent's DECISION PROBLEM $\delta(\mathcal{M}\restriction\alpha)$, defined on the filtered model, is of the classical form $\left\langle S, A, \hat{P}, U \right\rangle$ where:

$$\begin{aligned}
S &= \mathcal{W}' / {\equiv_\alpha} \\
A &= \mathcal{A}' = \mathcal{A} \setminus \mathfrak{A} \\
\hat{P}(s) &= \sum_{w \in s} P'(w \mid \mathcal{W}') \\
U(s, a) &= \sum_{w \in s} P'(w \mid s) \sum_{o \in \mathcal{O}'} \Pi'(w, a, o) \mathcal{U}'(o).
\end{aligned}$$

In words: S is the set of equivalence classes on $\mathcal{W}'$ given by the aggregation relation; A is simply the actions being entertained; $\hat{P}$ is in fact the same filtered probability distribution but interpreted on states (that is, on sets of worlds); and U gives the expected utility of a in s, under (epistemic) uncertainty about which world w from s obtains but also (metaphysical uncertainty) about which outcome o will result from doing a in w.

2.2.5 Example: Branestawm's allergies come to awareness

An example will help make clear the formal model given in the last section. Branestawm's allergy case in Example 3 makes for a fairly simple case of awareness dynamics involving decision problems. For the sake of a simple example, we merely want to model Branestawm's epistemic state before and after becoming aware of his allergies which might be set off by the fruit salad. In order to do so we will have to specify a reasonable language (for propositions and actions) restrictions of which will capture Branestawm's initial unawareness —indeed, the choice of a language thus becomes the initial modelling step— and subsequently define a background model based on that language. In Branestawm's case, this is rather easy. We get by perfectly with just two propositions, namely

"fruit": Branestawm has enjoyed some fruit salad;
"allergy": Branestawm's allergies go rock-a-doodle.

So, let's fix that $\mathcal{P} = \{\text{fruit}, \text{allergy}\}$. Similarly straight-forward is the choice of actions as $\mathcal{A} = \{\text{eat}, \text{abstain}\}$ containing actions representing Branestawm eating or not eating fruit salad. Next, consider the worlds and outcomes that should enter our background model. For the sake of simplicity, we identify worlds and outcomes with their valuation functions. So we will assume that $\mathcal{W} = \{w\}$ and $\mathcal{O} = \{w, o_1, o_2\}$ with the following valuation functions and utilities:

	fruit	allergy	$\mathcal{U}$
w	0	0	0
o_1	1	1	-10
o_2	1	0	1

Finally, we specify Branestawm's beliefs about the results of his actions if he were fully aware of all propositions in the result distribution Π as follows:

$$\langle w, \text{abstain} \rangle \mapsto w$$

$$\langle w, \text{eat} \rangle \mapsto \begin{cases} o_1 & \mapsto 0.001 \\ o_2 & \mapsto 0.999 \end{cases}$$

That is, all the probability mass of $\Pi^w(\text{abstain})$ is placed on w (abstaining does not change either of the proposition parameters we are concerned with); eating, on the other hand, leads to an allergic reaction (o_1) with very low probability, and otherwise to o_2.

This completely specifies Branestawm's background model $\mathcal{M}$. To fully specify his epistemic state *before* Calculus makes him aware of his allergies we need to specify in addition an appropriate awareness state α_0. In the case at hand this is the triple:

$$\begin{aligned} \mathfrak{U} &= \{\text{allergy}\} \\ \mathfrak{v} &= \{\text{allergy} \mapsto 0\} \\ \mathfrak{A} &= \varnothing \end{aligned}$$

This yields Branestawm's limited awareness of the decision situation as a filtered model $\mathcal{M}{\restriction}\alpha_0 = \langle \mathcal{W}', \mathcal{O}', \mathcal{A}', P', \mathcal{U}', \Pi' \rangle$ which comes out as:

$$\begin{aligned} &\mathcal{W}' = \{w\} \\ &\mathcal{O}' = \{w, o_2\} \\ &\mathcal{A}' = \mathfrak{A} \\ &P' = P \\ &\mathcal{U}' = \mathcal{U}{\restriction}\mathcal{O}' \\ &\Pi' : \mathcal{W}' \times \mathcal{A}' \to \Delta(\mathcal{O}') \text{ is such that} \\ &\qquad \langle w, \text{abstain} \rangle \mapsto w \\ &\qquad \langle w, \text{eat} \rangle \mapsto o_2 \end{aligned}$$

In words, in epistemic state $\mathcal{M}{\restriction}\alpha_0$ Branestawm believes that eating fruit salad will not trigger his allergies. This is modelled by the result distribution in his filtered model which assigns probability 1 to the outcome o_2, the outcome world where Branestawm does not have an allergic reaction.

The awareness dynamics in this example are fairly simple. Calculus's question simply has Branestawm become aware of the proposition "allergy". This is modelled by updating Branestawm's awareness state α_0 by removing this proposition from the set of unmentionables and assumptions. The resulting awareness state α_1 is trivial and the filtered model $\mathcal{M}{\restriction}\alpha_1$ in this simple example is the background model $\mathcal{M}$ itself.

One interesting point should be noted about this example: when Branestawm becomes aware of the possibility of an allergic reaction, his beliefs effectively do not change. That is, he gains a 0.1% uncertainty about the matter, but this is nowhere near enough to alter his choice of action. If Calculus knew this in advance he would have no reason to bring up the possibility; however because he himself is aware of the possibility of an allergy but uncertain about its relative probability, he feels compelled to mention the possibility. Branestawm, on the other hand, we take to be expert in the matter of his own allergies (at least when he is actively considering them). This is then an example of a pragmatically well-motivated 'awareness move' by Calculus, which nonetheless does not affect Branestawm's actions. If we were to describe his beliefs quantitatively, as the propositions his filtered model gives overwhelming probability mass to, we would say that the awareness update does not *overturn* his implicit belief due to assumption, but *ratifies* it.[10]

This is, perhaps, the simplest possible example of awareness dynamics, but, of course, there is much more to say about the changes in epistemic states

[10] Apart from this mention we stick to a purely qualitative view of beliefs, purely for convenience; that is, in terms of the worlds given non-zero probability mass. The extension may be straightforward but we have not looked at it in detail.

due to awareness and the interaction of awareness dynamics with factual information dynamics. Section 3 is devoted to these questions, but first we owe the reader a more explicit explanation of just where the convenient assumptions we have been making use of are supposed to come from.

2.3 Assumptions and associations

The first thing to notice about assumptions is that not just anything goes. A forgiving reader might not complain that we haven't sufficiently motivated Bo's assumption that the keys are not in her car as a cognitive reality, but should certainly object if in explaining a different scenario we had her unconsciously assuming that they were in fact hiding in the sugar-jar. We have appealed to intuitions of normality, without really making precise what we mean by this. Clearly 'normality' is sensitive to the details of the decision-making context; it is probably normal to assume the library is open when checking the remaining to-do list for an almost-finished essay, and equally reasonable to assume it is closed when plotting to break in at midnight for some clandestine reading.

The library example is not chosen at random: law-like conditionals such as "If she has an essay to write she studies late in the library" were used in a now classic experiment in psychology of reasoning, the 'suppression task' [Byr89], which shares many characteristics with the notions of awareness. The basic observation is that subjects asked to accept the truth of the conditional (as a premise in a logical argument) seem to implicitly hedge it with a normality assumption: "If she has an essay to write (and nothing unexpected happens) she studies late in the library". [SL08] gives an explanation of the data in terms of CLOSED-WORLD REASONING: Stenning and Van Lambalgen represent the implicit hedge as a 'dummy' proposition which is assumed false if there is no evidence that it is true. While the details do not concern us here (the parallel with unawareness is incomplete, although provocative), the closed-world reasoning is a perfect fit for our notion of assumptions.

That is, if our examples are to be intuitively satisfactory, assumptions should have a closed-world flavour: unusual events do not occur and the status quo is maintained, unless explicit reason is given to believe otherwise.

This formulation in turn suggests a loose probabilistic constraint on our assumptions due to unawareness. That is, it should generally be the case that the probability mass hidden by a particular assumption (an 'unusual event') is relatively small compared to the probability mass on the worlds being entertained (including, although not limited to, the 'status quo'). In other words, while becoming aware may qualitatively overturn an assumption, it should generally replace certainty that p only with uncertainty, not with near-certainty that $\neg p$.

We do not believe that this is a 'hard' semantic (or even pragmatic) constraint on acceptable states of awareness. However if we recall that our notion of unawareness is linked to absent-mindedness and cognitive limitations of attentiveness it seems that we should expect our cognitive apparatus (superbly evolved as it seems to be for problem-solving) to be reasonably good at prioritising attention, keeping focussed on the most probable and most utility-relevant contingencies and letting only the marginal ones slip beneath the surface.

Taking this cognitive perspective also solves a formal problem that we have so far managed to side-step by choice of easy examples. But consider again the case of Bob and Pickles. If Bob tells Pickles he could bake shortbread (making Pickles aware of a possible *action*), nothing in the formal setup we've given so far explains how Pickles gets to entertain new *outcomes* as well. Still, intuitively he should: when becoming aware of the action "baking shortbread" he should also become aware of certain natural outcomes of that action.[11] Although clearest in this case, the problem is not confined to actions and outcomes. The reality is that some possibilities are cognitively closely ASSOCIATED, so that becoming aware of one may bring on awareness of the other. However, very little formal or precise can be said in the present framework about this process of association in its full complexity. Hearing a possibility mentioned *at least* brings the possibility itself to awareness and mentioning a possible action certainly calls to mind stereotypical outcomes of the action. But beyond this we cannot say much more. That is why in this paper we've been careful not to make associations do any explanatory work. However, for the Pickles example discussed in detail in Section 4.1 we must at least rely on the association between the action "bake shortbread" and the shortbread-related propositions such as "shortbread has been baked", "the shortbread tastes awful (because of the lack of eggs)" and so on.

As in the case of assumptions, we may gesture at the adaptive nature of our cognitive capabilities in support of the idea that the right associations will spring to mind when they are needed. Apart from the formal definition of bringing to awareness propositions that are explicitly mentioned, however, the details of this association process must remain somewhat vague. In the following section we assume a mechanism giving associations and define the dynamic updates to awareness states it gives rise to, and the resulting updates to filtered information states and decision problems. As before, the strength of the account is that relatively simple changes in awareness can give rise to radical belief changes through the filtering process.

3 Information dynamics under awareness

The previous section laid out two closely related formal models for representing a single agent's unawareness, once for a mere propositional setting and once for a richer decision-theoretic structure. The models contained the basic ingredients for awareness dynamics: removing unmentionables from and adding actions to awareness states. We have not yet addressed the relation between awareness updates and the uptake of factual informative. This is what we will do presently. Again we proceed in stages from simple to complex, starting with the propositional case where we can focus on the main ideas that then carry over to the structurally more complex case of decision problems under growing awareness.

[11] The formal distinction between actions and propositions is of course a theoretical fiction which a shift to a first-order model (with possibilities of defining unawareness of terms such as "shortbread" be that in descriptions of actions or states of affairs) could alleviate. A first-order unawareness model has recently come on the market [BC07] however it's not yet clear how to combine this approach with implicit beliefs based on (possibly false) assumptions.

3.1 The propositional case

There are two fundamental ideas to the treatment of information dynamics under awareness. Firstly, we have already argued in the introduction, in particular in the slogan "Awareness is easily overturned", that unawareness from inattentiveness is lifted spontaneously whenever agents process linguistic information that contains mention of an unaware contingency. That is why we will assume that an agent who processes an utterance of some natural language sentence ϕ, be that for information uptake or anything else, will involuntarily become aware of all linguistic elements (proposition letters and actions) used in ϕ (or rather: a formal representation thereof in propositional logic) even *before* she can engage in any further processing.

A second key feature of information dynamics under unawareness is that information uptake can only take place, so to speak, within the window of awareness: more formally speaking, if, for the propositional case, an agent in the epistemic state $\langle \sigma, \alpha \rangle$ is aware of all proposition letters in (the formula) ϕ, an informative update with the (propositional) information in ϕ will be an update on the *filtered* state $\sigma \upharpoonright \alpha$ only (that is, only worlds being entertained are eliminated, not worlds from the background information set that are excluded by assumptions). This is fairly natural once appreciated: an agent who learns factual information can process this information only in the light of her (possibly limited) awareness. (Things become more complicated when the awareness state itself changes, a complication taken up in Section 3.3.)

These considerations lead to the following treatment of information updates for the propositional case. We will write $\langle \sigma, \alpha \rangle [\phi]$ for updating an epistemic state $\langle \sigma, \alpha \rangle$ with a propositional formula ϕ. This update can be considered a sequential update first of the awareness state, for which we will write $\alpha[\phi]$, and subsequently an update of σ with ϕ under the agent's updated awareness $\alpha[\phi]$. If ϕ is a propositional formula (representing an utterance), write $\mathcal{P}(\phi)$ for the proposition letters occurring in ϕ and $[\![\phi]\!]$ for the set of worlds where ϕ is true. Then we define propositional update with awareness as follows:

Definition 8: Epistemic update with (propositional) awareness. Let $\langle \sigma_0, \alpha_0 \rangle$ be an epistemic state. Then $\sigma_0 \subseteq W$ is an information set (the worlds not excluded by the agent's information) and α_0 is as always an awareness state. Let ϕ be an utterance. Then

$$\langle \sigma_0, \alpha_0 \rangle [\phi] \stackrel{\text{def}}{=} \langle \sigma_1, \alpha_1 \rangle$$

where $\alpha_1 = \alpha_0[\phi]$ is given by

$$\langle \mathfrak{U}, \mathfrak{v} \rangle [\phi] \stackrel{\text{def}}{=} \langle \mathfrak{U} \setminus \mathcal{P}(\phi), \mathfrak{v} \upharpoonright (\mathfrak{U} \setminus \mathcal{P}(\phi)) \rangle ,$$

and σ_1 is given by

$$\sigma_0 \setminus ((\sigma_0 \upharpoonright \alpha_1) \cap [\![\neg\phi]\!]).$$

For emphasis: updating σ_0 to σ_1 uses the *new* awareness state α_1, rather than the old one; first we make all proposition letters in ϕ mentionable and then we eliminate all entertainable worlds that are incompatible with ϕ.

3.2 Updates for decision problems

The main features of information dynamics under awareness carry over from the basic propositional case to the richer decision-theoretic models fairly straightforwardly. An epistemic state is now the pair $\langle \mathcal{M}, \alpha \rangle$ where $\mathcal{M}$ is a background model and α is an awareness state. Updating an epistemic state with (a formal representation of) an utterance ϕ proceeds analogously to the propositional case by first making the agent aware of all linguistic elements featured in ϕ, where this might now include actions as well, and subsequently updating the background model through 'the awareness window' of the filtered model $\mathcal{M}\restriction\alpha[\phi]$ with the information $[\![\phi]\!]$. This boils down to eliminating from the background model all worlds and outcomes where ϕ is not true that are visible in the awareness window *after* the agent became aware of all contingencies mentioned in ϕ. Let $\mathcal{A}(\phi)$ be all the actions mentioned in ϕ and define:

Definition 9: Epistemic update with (decision-theoretic) awareness. Let $\langle \mathcal{M}_0, \alpha_0 \rangle$ be the epistemic state of some agent, where now $\alpha_0 = \langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle$. Let ϕ be an utterance. Then

$$\langle \mathcal{M}_0, \alpha_0 \rangle\, [\phi] \stackrel{\text{def}}{=} \langle \mathcal{M}_1, \alpha_1 \rangle$$

where $\alpha_1 = \alpha_0[\phi]$ is given by

$$\langle \mathfrak{U}, \mathfrak{v}, \mathfrak{A} \rangle\, [\phi] \stackrel{\text{def}}{=} \langle \mathfrak{U} \setminus \mathcal{P}(\phi), \mathfrak{v}\restriction(\mathfrak{U} \setminus \mathcal{P}(\phi)), \mathfrak{A} \setminus \mathcal{A}(\phi) \rangle\,,$$

and $\mathcal{M}_1$ is derived from $\mathcal{M}_0 = \langle \mathcal{W}, \mathcal{O}, \mathcal{A}, P, \mathcal{U}, \Pi \rangle$ (indices omitted for readability) as follows:[12]

$$
\begin{aligned}
\mathcal{W}_1 &= \mathcal{W} \setminus (\mathcal{W}_{\alpha_1} \cap [\![\neg\phi]\!]) \qquad \text{(but see footnote 12)}\\
\mathcal{O}_1 &= \mathcal{O} \setminus (\mathcal{O}_{\alpha_1} \cap [\![\neg\phi]\!]) \qquad \text{(but see again footnote 12)}\\
\mathcal{A}_1 &= \mathcal{A}\\
P_1 &= P(\cdot \mid \mathcal{W}_1)\\
\mathcal{U}_1 &= \mathcal{U}\restriction\mathcal{O}_1\\
\Pi_1 &: \mathcal{W}_1 \times \mathcal{A}_1 \to \Delta(\mathcal{O}_1) \text{ is such that } \Pi_1(w,a) = \Pi^w_a(\cdot \mid \mathcal{O}_1)
\end{aligned}
$$

For clarity: the only non-trivial updates of the background model are the elimination of worlds and outcomes, which is but exactly the same procedure as in the propositional case. The restrictions to probabilities and utilities are simply required to keep the structure well-defined.[13]

[12] Note that we have to be a little careful interpreting the tense of the expression ϕ correctly: updating with the (true) information "Branestawm has not eaten any fruit salad" should not remove *outcomes* in which he does. On the other hand if his doctor declares "Branestawm will not suffer from an allergy", it is exactly the outcome worlds that should be removed. If the background model were a fully-fledged temporal model this difficulty would be avoided, but the construction of a decision problem would become much more complex. We prefer to stick to the simpler approximation, and apply common sense to the updates.

[13] A different route could also be taken: instead of removing worlds from the information set entirely, simply adjusting their degree of credence, assigned by P, to zero. Which is appropriate depends on whether you think possibilities ruled out by information are still entertained or not, which might even vary depending on the application under consideration.

3.3 Old information in the light of new awareness

The perhaps most fundamental idea behind out treatment of updates with factual information by agents with limited awareness is that factual information can only be evaluated (at the time it is observed) within the 'window of awareness' of the agent. But that may mean that assumptions can block the elimination of worlds which, when the implicit belief is given up by growing awareness, the agent might or might not want to rule out as well. Here is a simple example to illustrate the sequential interaction of awareness and information updates.

Suppose for simplicity that $\mathcal{P} = \{p, q\}$ and that our agent is unaware of p, assuming it to be true, and aware of q but uncertain about it; this gives us four possible worlds $W = \{pq, p\overline{q}, \overline{p}q, \overline{pq}\}$ (identifying them sloppily with their valuations). Then $\sigma_0 \upharpoonright \alpha_0 = \{pq, p\overline{q}\}$; if the agent now learns that q is true, she will erase the world $p\overline{q}$ and her information set will become (according to the definitions we've given) $\sigma_1 = \{pq, \overline{p}q, \overline{qp}\}$.

Now this means that *within her awareness window* she has come to believe q, because $\sigma_1 \upharpoonright \alpha_0 = \{pq\}$; this is an explicit belief by our definition, but, surprisingly, one that is *not* necessarily stable under awareness updates, because when the agent becomes aware of her implicit assumption about p, a mere awareness update that removes p from the set of unmentionables brings with it the world $\overline{pq}$ which has *not* been ruled out by the previous information update. So, taken together, when an agent processes factual information her implicit beliefs might in fact block correct information uptake. In order to rule out worlds that have not been ruled out by an informative update, because these worlds were hidden behind an implicit belief, the agent has to, in our system, *reprocess* or *reconsider* the previous factual information in the light of her extended awareness.

The reader's response at this point may be: "But then you have defined information updates in the wrong way." Indeed, it is tempting to give up the idea that information is processed only in the light of awareness and instead assume that information percolates, perhaps secretly, all the way up through to the background model. This would save us quite some trouble, not only in the definition of information uptake, but also in dispensing with the "reprocessing" of factual information.

However there is an important distinction between *observing* that q holds and merely *hearing reported* that q holds, and one that turns on unawareness. If our agent assumes p holds, she does not think to check whether a report of q is conditional on this assumption or not. The speaker, in turn, might hold the same assumption and might themselves *not* be willing to commit to the truth of q if they are made aware of p. The point is clearest in the case of lawlike conditionals discussed briefly in Section 2.3. If I hear "If she has an essay to write she will study late in the library" and I am assuming the library is open, it is simply unclear whether the speaker makes the same assumption or is trying to tell me something stronger (that the student is so fanatical she will find a way to sneak in anyway, for instance). Were this not a case of unawareness I could always ask the speaker for clarification, but the distinction hinges on possibilities I am *not yet entertaining*; it is only in retrospect, when they have been brought to my attention, that I realise the potential ambiguity of the speaker's intent.

This complicates the picture of epistemic update in conversation considerably. Rather than simply carrying around an epistemic state, agents must carry at least a rough memory of the updates that brought them to that state, in order to be able to reinterrogate that memory in the light of new possibilities. Of course this is a more realistic picture of real conversation, but it is a significantly less tractable one. However it raises one very interesting possibility: that a speaker might come to repudiate a statement she has previously accepted, or even made herself, without having in the strict sense learned anything new in the interim.

In general, these considerations play on the dynamics of awareness in conversation and open up the possibility for quite complicated pragmatic reasoning of various sorts. In the next section we give a concrete example, in (perhaps excruciating) detail, in order to show the power of the formal machinery we have defined.

4 Awareness dynamics, decisions, and pragmatics

So far we have seen how limited awareness and awareness growth can influence a single agent's decision to act. Our conjecture in this paper is that unawareness from inattentiveness is fairly natural and wide-spread. It is therefore not surprising to find that the notion of awareness also plays a role in a variety of pragmatic phenomena that arise in conversation. In this section we will revisit the introductory Example 2 where Bob, an expert baker, makes his friend Pickles *deliberately* aware of a contingency that he had overlooked. Adding to the brief informal discussion in Section 1, we will spell out formally the kind of pragmatic reasoning that revolves around the concept of awareness in this little dialogue. (We do not believe that this simple example covers all, or even necessarily the most important, aspects of pragmatic reasoning about awareness in decision-relevant conversations. We merely believe that this simple example is indicative enough of the kind of reasoning we have in mind, and its possible formalisation.)

4.1 Bob & Pickles revisited

In the Pickles dialogue (Example 2) we would like to model Pickles becoming aware of baking shortbread as a possible action. We will first spell out the background model in order to discuss a simple awareness update. Towards this end, let's first of all fix what Pickles' language should be able to distinguish when he is fully aware of all model-relevant contingencies. With quite some redundancy, we use the following set of proposition letters $\mathcal{P}$:

"eggs":	the recipe for shortbread requires eggs;
"yum-cake":	Pickles has baked a tasty cake;
"yuck-cake":	Pickles has baked a disgusting cake;
"yum-sb":	Pickles has baked tasty shortbread;
"yuck-sb":	Pickles has baked a disgusting cake.

Given these propositions we should distinguish in Pickles' background model certain worlds and outcomes, according to their associated valuation functions based on $\mathcal{P}$. But it should be clear that we do not have to consider all possible valuations, because two assumptions rule out quite a number of combinations: firstly, we assume that Pickles can only bake one item, so that there will be no world where Pickles has baked both a cake and shortbread; secondly, certain natural meaning postulates apply, (there can be no tasty cake if no cake has been baked, and so on).

Again for simplicity we identify worlds and outcomes with their associated valuation functions; the background model $\mathcal{M}$ for Pickles' case contains worlds $\mathcal{W} = \{w_1, w_2\}$ and outcomes $\mathcal{O} = \{w_1, w_2, o_1, \ldots, o_4\}$ with the following valuations and utilities (recall that "egg" means that the recipe for shortbread requires eggs; it is common knowledge that Pickles has no eggs so we don't bother including any uncertainty about that fact):

	egg	yum-cake	yuck-cake	yum-sb	yuck-sb	$\mathcal{U}$
w_1	0	0	0	0	0	0
w_2	1	0	0	0	0	0
o_1	0	0	1	0	0	-1
o_2	1	0	1	0	0	-1
o_3	0	0	0	1	0	1
o_4	1	0	0	0	1	-1

The actions Pickles is aware of in the limit are baking cake, baking shortbread and abstaining from all baking:

$$\mathcal{A} = \{\text{cake}, \text{sb}, \text{abstain}\}\,.$$

Pickles' beliefs about the true state of affairs are represented in his probability distribution P. We leave this parametrized for the sake of discussion below and set $P(\text{egg}) = p$. In turn, Pickles' beliefs about the outcomes of his actions are represented in the result distribution Π as follows (all probability distributions put probability mass 1 on exactly one outcome in this example, so we only need to specify which outcome that is):

$$\begin{aligned} \langle w_i, \text{abstain}\rangle &\mapsto w_i \\ \langle w_i, \text{cake}\rangle &\mapsto o_i \\ \langle w_1, \text{sb}\rangle &\mapsto o_3 \\ \langle w_2, \text{sb}\rangle &\mapsto o_4. \end{aligned}$$

So far for Pickles' epistemic state under full awareness. Let's now model his unawareness of the action "sb" and investigate in more detail the awareness update and its impact on Pickles' decision problem. Since Pickles is unaware of the action "sb" his initial unawareness state α_0 will have record this fact: $\mathfrak{A} = \{\text{sb}\}$. But also, Pickles is unaware of shortbread-related propositions, so his set of unmentionables $\mathfrak{U}$ is $\{\text{eggs}, \text{yum-sb}, \text{yuck-sb}\}$. Since, as we argued above, Pickles does not have any implicit beliefs about any of these, his assumptions $\mathfrak{v}$ are a trivial (empty) valuation function. For perspicuity: α_0 is given by

$$\begin{aligned} &\mathfrak{U} = \{\text{eggs}, \text{yum-sb}, \text{yuck-sb}\} \\ &\mathfrak{v} = \varnothing \\ &\mathfrak{A} = \{\text{sb}\}\,. \end{aligned}$$

When Bob mentions shortbread, we assume that Pickles becomes aware not only of the action "sb", but also of the naturally associated propositions, that is, the shortbread-related propositions "yum-sb" and "yuck-sb". The result of this awareness update is the trivial awareness state α_1 (full awareness and no assumptions): after Bob's remark about shortbread, Pickles represents his decision problem as $\mathcal{M}\restriction\alpha_1$ which is identical to $\mathcal{M}$.

But the decision problem represented in $\mathcal{M}$ has Pickles uncertain whether the recipe for shortbread contains eggs of not. This is, for the time being, how it should be. We have assumed that he himself does not know and that his initial unawareness of shortbread only concealed this true uncertainty. This uncertainty, however, can be overturned by a pragmatic inference that crucially relies on the idea that Bob, an expert baker who knows whether shortbread requires eggs, is cooperative much in the sense of [Gri89] and — enter awareness— has *ostensibly made Pickles aware* of shortbread. Here is a semi-formal account of this pragmatic reasoning.

Let's first of all ask ourselves what Pickles would do in the initial decision problem when he was still unaware of shortbread. Since eggs are unavailable baking a cake would seem stupid. In formal terms it has an expected utility of -1 in his decision problem $\delta_0 = \delta(\mathcal{M}\restriction\alpha_0)$, where expected utility of an action (in a classical decision problem δ) is defined as:

$$\mathrm{EU}_{\delta}(a) \stackrel{\text{def}}{=} \sum_{s \in S} P(s) \times U(s, a)$$

In contrast, abstaining from all baking has expected utility 0, so that in δ_0 this is clearly the preferred option. But now compare this with Pickles' decision problem under full awareness, δ_1. Clearly baking a cake and abstaining from baking altogether have the same expected utilities. But we have a new player in the race: baking shortbread. The expected utility of baking shortbread under full awareness is:

$$\mathrm{EU}_{\delta_1}(\mathrm{sb}) = -1 \times p + 1 \times (1 - p) = 1 - 2p$$

This means that under full awareness baking shortbread will be preferred to abstaining from baking (with expected utility 0) iff $p < 0.5$. In other words, barring pragmatic considerations Pickles will bake shortbread only if he thinks it is *more likely* that the recipe for shortbread does not require eggs. (Naturally the specific utilities chosen don't matter for the general point that there is *some* threshold beyond which p is high enough to justify baking.) But even if his subjective probability favored the possibility that shortbread *does* require eggs ($p > 0.5$) he could still revise these beliefs based on the following pragmatic reasoning: if Bob knows that Pickles faces the decision problem in question (including unawareness of shortbread) and if furthermore Bob is also helpful and cooperative, then his conversational move (deliberately bringing shortbread to awareness) can only be motivated if $p < 0.5$, for otherwise it would be futile, or lead him to choose an even worse action. If furthermore Bob is an expert baker who knows for sure whether w_1 or w_2 is the true state of affairs, Pickles is safe in concluding that the true state of affairs is w_1.

4.2 Decision-theoretic relevance

This intuitive formulation points towards the assumption of RELEVANCE as driving Pickles' inference. Pickles must be able to explain Bob's question as relevant to the purposes of the conversation, otherwise he would have to conclude that Bob was not being cooperative. If those purposes are, roughly, getting something tasty cooked, it's hard to imagine how Bob could have relevantly intended his question as a literal request for information. However if it is given the natural interpretation of deliberately and ostensibly bringing a possibility to awareness, the prospect looks much better.

If our promise of a formal solution is to be fulfilled, though, we need a formal notion of relevance that is appropriate for the decision-theoretic setting. As it happens one already exists for purely informational updates, which we can adapt with very minor changes to the current setting. The measure in question is called the VALUE OF SAMPLE INFORMATION (we use a variant of that defined in [RS61]).

The intuition is as follows: the relevance of (true) information in a decision problem can be measured as the change in the agent's expected utility with the information, compared to without it. If we assume the agent is a utility maximiser, they will choose one of the actions that appear best (in terms of expected utility) according to their information; comparing the actual value (given the new information) of these apparent best-action choices with the true best-action payoffs (again given the new information) gives the amount the agent believes her fortunes have improved in light of the information. Stated negatively, information that does not cause the agent to change her mind about her best action is irrelevant.[14]

We will first define the value of sample information for factual information uptake not involving expanding awareness. Towards this end, extend this definition of expected utility of an action in a (classical) decision problem to a set $B \subseteq A$ of actions by taking the average (we will use this for the agent's set of perceived best actions):

$$\mathrm{EU}_\delta(B) \overset{\text{def}}{=} \frac{1}{|B|} \sum_{a \in B} \mathrm{EU}_\delta(a).$$

We write $\mathrm{BA}(\delta)$ for the set of actions with maximal expected utility in δ:

$$\mathrm{BA}(\delta) \overset{\text{def}}{=} \{a \in A \,;\, \forall a' \in A \colon \mathrm{EU}_\delta(a') \leq \mathrm{EU}_\delta(a)\}.$$

Now suppose that δ is a decision problem representing the actual state of affairs (metaphysical uncertainty may still keep this nontrivial), while γ is the agent's conception of the decision problem she faces (limited by unawareness and complicated by epistemic uncertainty as usual). Then $\mathrm{EU}_\delta(\mathrm{BA}(\gamma))$ is perfectly well-defined (so long as δ and γ are 'compatible' in the obvious ways) and gives the actual expected utility of the actions the agent believes are best.

Now let δ represent some concrete decision problem, and write $\delta[\phi]$ for the same problem updated with some (true) factual information ϕ that does not

[14]This can easily be rejected as overly simplistic, since information making an agent more certain of a choice she has made can intuitively be highly relevant. We don't represent anywhere the higher-order notions of uncertainty that doing this intuition justice would require, however we also feel that this omission is harmless for the paradigm cases of unawareness that we treat.

involve awareness updates. The value of sample information ϕ in the original decision problem δ, written $\mathrm{VSI}_\delta(\phi)$, is given by

$$\mathrm{VSI}_\delta(\phi) \stackrel{\text{def}}{=} \mathrm{EU}_{\delta[\phi]}(\mathrm{BA}(\delta[\phi])) - \mathrm{EU}_{\delta[\phi]}(\mathrm{BA}(\delta)).$$

That is, we compare the actual expected utilities (given the information ϕ) of two sets of actions: those the agent considers best before she learns ϕ, and those she prefers after she learns ϕ.

One convenient feature of this definition is that information never has negative value. Another simple point to note is that information has strictly *positive* value only if it reveals some apparently optimal action to not in fact be so; that is, if it removes something from the set of best actions. (To see this, consider the alternatives: (i) the set of best actions stays the same, in which its value doesn't change, or (ii) something is *added* to the set, in which case the new action must have the same expected value as the previous elements —or they would be removed— so the average does not change.)

Turning now to awareness dynamics, we will take over this definition exactly as it stands, except for the update: we will consider also the value of changes to an epistemic state by updating awareness, rather than by incorporating information. The only difference, then, is in the definition of expected utility: rather than take this relative to a decision problem we define it relative to a filtered information state. If $\langle \mathcal{M}, \alpha \rangle$ is a decision-theoretic epistemic state (a background model and awareness state) then we define the expected utility $\mathrm{EU}_{\langle M,\alpha\rangle}(a)$ simply by taking $\mathrm{EU}_\delta(a)$ where δ is the classical decision problem 'read off' from $\mathcal{M}{\restriction}\alpha$ according to Definition 7; the definition of "$\mathrm{BA}(\cdot)$" is extended in the same way. Entirely analogous to the definition given above, we can write $\mathrm{EU}_{\langle \mathcal{M},\alpha\rangle[\phi]}(\mathrm{BA}(\langle \mathcal{M}, \alpha \rangle))$ for the consequences (judged in terms of awareness of ϕ) of the actions the agent considered best before ϕ was brought to her attention.

We might dub the new definition the "Value of Epistemic Change"; not that it is formally so different from the old, but the supporting intuitions certainly are. We are no longer dealing simply with sample information (intuitively obtained by direct observation) but with changes to the epistemic state of the agent herself (most likely obtained from conversation), albeit ones that we will assume bring her asymptotically closer to the truth. While her updated state may be closer to reality (if she learns true information *or* becomes aware of contingencies she was unwarrantedly excluding) it may still be far from the whole truth. These necessary caveats given, here is the definition:

$$\mathrm{VEC}_{\langle \mathcal{M},\alpha\rangle}(\phi) \stackrel{\text{def}}{=} \mathrm{EU}_{\langle \mathcal{M},\alpha\rangle[\phi]}(\mathrm{BA}(\langle \mathcal{M}, \alpha \rangle [\phi])) - \mathrm{EU}_{\langle \mathcal{M},\alpha\rangle[\phi]}(\mathrm{BA}(\langle \mathcal{M}, \alpha \rangle)).$$

4.3 Bob & Pickles made formal (at last)

Armed with this definition we can at last formalise the pragmatic reasoning we attribute to Pickles when he concludes that the recipe for shortbread should not require eggs. The formalisation rests on the assumption that although Pickles does not know which of the two possible worlds is actual (w_1 in which shortbread does not require eggs, or w_2 in which it does) he knows that Bob the Baker *does* know which world obtains.

Pickles began in state $\langle \sigma_0, \alpha_0 \rangle$, and Bob's suggestion has brought him (before any pragmatic reasoning takes place) to $\langle \sigma_0, \alpha_0 \rangle [\phi] = \langle \sigma_0, \alpha_1 \rangle$ (we'll assume that whatever logical form ϕ takes, (a) it confers no information directly —does not eliminate worlds or change probabilities— and (b) it mentions shortbread and thus induces the awareness updates we need).

Pickles can now imagine two possibilities for the decision problem as *Bob* sees it: if w_1 is the actual world, Bob sees $\langle \sigma_0[\neg\text{egg}], \alpha_1 \rangle$ (that is, the same decision problem Pickles does but updated with the information that the eggs aren't needed) or $\langle \sigma_0[\text{egg}], \alpha_1 \rangle$. Call the first possibility δ_1 and the second δ_2 (conflating the epistemic state with the corresponding decision problem, since these states occur with full awareness anyway).

Now Pickles can compute the value of his epistemic change, if each of these possibilities is the actual one:

$$\text{VEC}_{\delta_1}(\phi) = \text{EU}_{\delta_1}(\text{BA}(\langle \sigma_0, \alpha_0 \rangle [\phi])) - \text{EU}_{\delta_1}(\text{BA}(\langle \sigma_0, \alpha_0 \rangle))$$

and likewise for the alternative δ_2. This computation rests, of course, on the decision of best action Pickles would take. Let's assume the worst: he's uncertain enough about the recipe that he's unwilling to take the risk, and prefers to abstain from cooking. Then $\text{BA}(\langle \sigma_0, \alpha_0 \rangle) = \{\text{abs}\}$.

But then $\text{VEC}_{\delta_1}(\phi) = \text{VEC}_{\delta_2}(\phi) = 0$: for before Bob made his suggestion, Pickles had already decided not to cook anything.

If Pickles is to make sense of Bob's suggestion as relevant, he will have to conclude that it communicates something *beyond* its ostensive awareness update. It's easy to see that this extra content, if assumed to be true, could only be "¬eggs" if w_1 is the actual world, and "eggs" if w_2 is actual. So let's compare these possibilities:

$$\begin{aligned}
&\text{VEC}_{\delta_1}(\phi; \neg\text{eggs}) \\
&= \text{EU}_{\delta_1}(\text{BA}(\langle \sigma_0, \alpha_0 \rangle [\phi; \neg\text{eggs}])) \\
&\quad - \text{EU}_{\delta_1}(\text{BA}(\langle \sigma_0, \alpha_0 \rangle)) \\
&= \text{EU}_{\delta_1}(\text{sb}) - \text{EU}_{\delta_1}(\text{abstain}) \\
&= 1
\end{aligned}
\qquad
\begin{aligned}
&\text{VEC}_{\delta_2}(\phi; \text{eggs}) \\
&= \text{EU}_{\delta_2}(\text{BA}(\langle \sigma_0, \alpha_0 \rangle [\phi; \text{eggs}])) \\
&\quad - \text{EU}_{\delta_2}(\text{BA}(\langle \sigma_0, \alpha_0 \rangle)) \\
&= \text{EU}_{\delta_2}(\text{abstain}) - \text{EU}_{\delta_2}(\text{abstain}) \\
&= 0
\end{aligned}$$

That is, correctly assuming shortbread to not require eggs would lead Pickles to a positive expected utility gain, while correctly assuming that it *does* require eggs leads to no gain at all. Since Bob knows which of δ_1 and δ_2 actually obtains and since his update (by the relevance requirement) should lead to a positive expected value change, he must be in δ_1 and have intended to convey "¬eggs".[15]

The reader might be forgiven for thinking that this is making a very complicated mountain out of a very simple molehill. We give the derivation in all its considerable detail in order to make clear that this framework fulfills the promises we have made on its behalf: a fully formalised representation of awareness-related relevance reasoning can be represented. That is not to

[15] In fact something stronger can be said: if δ_0 is the actual state of affairs (and Bob knows this) then *no* action (that Pickles is aware of) does better than the one he has already chosen, so no reasoning he can conceive of would make Bob's suggestion relevant. This reasoning comes closer to the intuitively natural question "If you know shortbread needs eggs and we don't have any, why would you bring it up at all?"

say that the story we have given for Bob and Pickles necessarily covers all the possibilities, and it's also not to claim that the numerical details of this sort of reasoning are perspicuous for the modeller. We would be remiss in pushing the structural capabilities of the system, however, without substantiating our claims with at least one fully worked-out example.

5 Related work

We have tried to develop a formal notion of unawareness from inattentiveness in this paper and also to show that such a notion can be useful for accounts of natural language meaning. In other words, the purpose of this paper is explicitly a dual one: to propose a thoroughly grounded and explicit model of a sort of unawareness that has not yet been closely investigated in the rational choice literature, and also to make linguists aware of the possibility of awareness as a significant feature of semantics and pragmatics (to run the awareness flag up the mast, as it were). This section situates our contribution with respect to current models of unawareness and to linguistic theory; we show the contribution of our work and give some suggestions for further applications in linguistics.

5.1 Formal awareness models

The classical reference for the notion of unawareness is [FH88], whose original motivation was developing inference systems that did not suffer from the problem of logical omniscience (that an agent knows all logical consequences of the facts that she knows). Fagin and Halpern point out that there are several distinct reasons to want to do away with logical omniscience such as 'strict' unawareness of possibilities, computational limitations and resource bounds, lack of knowledge of rules of inference, or issues of attention and focus. Different modelling choices result from different conceptualizations of unawareness which in turn depend on the intended application of the unawareness model.

This is then also the primary difference between the models presented here and the majority of unawareness models presented in rational choice theory that have sprouted recently.[16] In rational choice theory, apart from a general interest in modelling reasoning about this notion (see for instance [DLR98; Hal01; HMS06; HMS08a; MR94; MR99]) and in including unawareness into game theoretic solution concepts (see for instance [Fei04; Fei05; HR06; HMS07; Ozb07]), most applications have focused on reanalyzing in the light of possible unawareness certain fairly strong game-theoretic predictions about rational behavior: [Fei04], for instance, shows how the possibility of unawareness helps establish cooperation as a rational solution in the prisoners dilemma; [HMS06] shows how possible unawareness has otherwise rational agents wholeheartedly engage in speculative trade despite the well-known class of "No-Trade Theorems" (for example [MS82]).[17]

[16] The online unawareness bibliography maintained by Burkhard Schipper (http://www.econ.ucdavis.edu/faculty/schipper/unaw.htm) is a good starting point for readers interested in exploring the rational choice literature further.

[17] Very roughly, a "No-Trade Theorem" shows that speculative trade should never take place. For an intuitive basis to this theorem, imagine you are bargaining at a street market for some item

The source of unawareness that we have been concerned with in this paper is inattentiveness. This is because we believe that it is this kind of unawareness that plays a key role in certain aspects of conversation (see also the next section). The crucial feature of unawareness from inattentiveness is the ease with which it is overturned. To appreciate the difference between unawareness from inattentiveness and that resulting from a lack of conceptual grasp, suppose a teenager is presenting a poorly-reasoned argument in favour of unprotected sex, and you mention the possibility of AIDS; the instant awareness update along the lines we have described is easy to imagine. Treating unawareness from lack of conceptual grasp is like imagining the same conversation as if it it were held in the '70s, when the disease was unidentified and the acronym not yet invented: the new possibility being brought to painful awareness was not forgotten but simply had not yet been imagined. It should be clear that where linguistic generalizations about extending awareness through dialogue are concerned it is the former, not the latter type of awareness dynamics that we should focus on. This is then the main difference in conceptual interpretation of unawareness between our models and the collection of models entertained in economic theory. The notional difference further cashes out in two major differences in the modeling.

The first difference between our linguistically-inspired models and the ones studied for economic applications is that the latter do not consider and spell out assumptions. Recall that in introducing the notion of assumptions as implicit beliefs, we referred to the intuition that in the initial example unaware Bo Peep *behaves as if* she believes the keys are not in the car. Interestingly, it seems to us that the motivation for explicit modelling of assumptions of agents is not exclusively linguistic. For instance, when Heifetz, Meier, and Schipper seek to explain how unawareness overturns the "No-Trade Theorem" [HMS07], they also need to assume (implicitly) a particular "as-if" behavior of unaware sellers and buyers, namely behavior as if certain favorable or unfavorable contingencies were believed to be true or false. We suggest that the notion of an assumption might be an interesting enrichment of existing unawareness models.

The second major difference between the two systems (or system types) also stems from our goal to apply an unawareness model to (generalizations about) cooperative conversation. For this end, we are interested in describing systematically the effects of awareness updates on decision problems, which requires specifying numerical probabilities and utilities for the newly-introduced possibilities. The main idea to achieve this end is filtering through an awareness state.[18] The problem of *changing* awareness has also been addressed, typically in game-theoretic settings where it is natural to assume that observing a player make a move you were unaware they could make overturns this unawareness. [HMS08b] gives a game-theoretic model and a vari-

of jewellery; you offer what you consider an outrageously low price, and the seller immediately spits in his palm, shakes your hand and shouts "Done!". Your immediate thought is likely to be "If he's so happy with the deal, I've been had", and if given the choice you would prefer to revoke the offer. The unawareness perspective suggests how a more careful buyer and seller can still both go away convinced that they have struck an advantageous deal: you are unaware that the necklace is stolen and thus a risky purchase, and the seller is unaware that you are leaving the country in the morning and thus have nothing to fear.

[18] We introduced this approach in [FJ07], in a preliminary and in many ways unsatisfactory model which nonetheless contains the seeds of the present account.

ant of rationalizability for games with possibly unaware players, and [Fei04; Fei05; HR06] have taken similar equilibrium-based approaches. However, the emphasis in these efforts is on non-cooperative game theory, whose solution concepts do not, strictly speaking, supply vanilla awareness updates irrespective of rationality considerations. The demands of linguistic pragmatics, based as it is on a fundamentally cooperative notion of interaction, are quite different: we would like to pin down pure awareness dynamics first and show how pragmatic reasoning con take place on top of it.

Seen in this light, the model of [Ozb07] deserves special mention. Ozbay gives in a non-cooperative setting a signalling games model with an equilibrium refinement somewhat similar to our notion of relevance as VSI or VEC. In the model an aware sender can make an unaware receiver aware of certain contingencies by her choice of signal, but the beliefs the receiver adopts when becoming aware are not determined, but subject to strategic considerations. Ozbay offers a refined equilibrium notion according to which the receiver should adopt beliefs under extended awareness that prompt him to choose a different action from the one that he had chosen under his initial unawareness. While this kind of constraint on belief formation seems to be what pragmatic reasoning based on a notion of relevance as VSI or VEC provides, it is unclear whether this should apply in all cases of (possibly) conflicting interests. It should, to our mind, apply for the cooperative case, and we have spelled out this kind of reasoning based on the example of Bob and Pickles.

In short, although our work is based on the standard models in the rational choice literature, our notion of unawareness is not quite the same. The linguistic application, and in particular the structural requirements imposed by decision-problem representations, have led us to develope a significantly different model based on similar, but subtly different, intuitions.

5.2 Unawareness in linguistics

Turning to the linguistic literature, the picture is quite different. The notions and intuitions have apparently been present since the work of Lewis, and probably before, but have never been treated as a distinct phenomenon amenable to a unified formal treatment.

In his seminal paper [Lew79], David Lewis gave a unifying account of a wide range of accommodation effects in terms of an evolving "conversational score". Awareness effects as we have described them make a somewhat uncomfortable fit with picture, since unawareness updates (if we are correct) proceed not by accommodation but by something akin to inherent salience or attention-focussing effects. However one class of observations given by Lewis fits the awareness story very comfortably: his Example 6, on *relative modality*.

Lewis is concerned here with modals such as "can" and "must", and their apparent restriction, in normal usage, to a subset of all 'metaphysical' possibilities. There is a large literature on this subject, of course, but certain features recur again and again: a restricted set of possibilities that are 'in play' at any given moment, against which modal statements should be evaluated, and the possibility to add hitherto unconsidered possibilities into this set as a conversation progresses.

The similarity to the unawareness picture is clear, so we should say something instead about the differences. It might be thought that our 'worlds being entertained' correspond directly to a Stalnakerian context set [Sta78]: the possibilities not ruled out by presuppositions in force. However there is a crucial difference between our assumptions and the presuppositions that this approach would conflate them with: assumption is typically something the agent would repudiate if she were made aware of it. This is implicit in our slogan "Unawareness is easily overturned": it is only when overturning unawareness also overturns an implicit belief (that is, when an assumption is given up as unfounded) that the epistemic update is as it were *visible* to the observer, since an awareness update that simply ratifies an implicit belief does not result in a change of behaviour.

Nevertheless the notions of assumption and presupposition are closely linked, and the exact relation between them remains a problem for further study. It seems, for instance, that assumptions can sometimes be 'converted into' presuppositions. Suppose you make a naïve statement due to unawareness of some contingency p. I am aware of p, and see that you seem to have neglected it, but even so I agree with your statement (suppose for example that I explicitly assign very low probability to p). If I choose not to object, it seems that all assertions in our further conversation are contingent on p being false, but in two quite different ways: we might say that you are assuming, while I am presupposing. Whether this is in fact the right distinction is unclear (the possibility of uncertainty about the awareness basis from which a speaker makes assertions complicates matters), but certainly the issue deserves further investigation.

Another difference to the standard approach is the inadvertency of an awareness update: the agent who undergoes such an update cannot choose rather to remain unaware, and no pragmatic reasoning can undo the *immediate* effects it produces. An interesting topic for further research is the interplay between such automatic updates and the explicit negotiation about which possibilities are 'on the table' displayed in sentences like "Let's leave that possibility out of the picture for the moment".

The closest account to ours that we are aware of is Eric Swanson's treatment of the language of subjective uncertainty [Swa06b], elaborated in [Swa06a]; his "coarse credal spaces" are very closely analogous to the aggregated states and outcomes in our decision-theoretic formulation. We agree wholeheartedly with his insight that "might"-statements can be appropriately used without any expectation that they will be informative in the usual propositional sense of removing possibilities from play, and that such use is in fact based on the hope that *adding* possibilities will be helpful to the addressee. What awareness brings to the party is a notion of excluded possibility that is both absolute and easily overturned.

Another area where notions of considering or ignoring possibilities appear very natural is in assessing the truth of conditionals in discourse. There has recently been some interest in the question how the acceptability of conditional sentences depends on their sequential presentation [F01; Gil07; Wil08]. Sensitivity to ordering sequence implies some sort of dynamic effect, and awareness dynamics indeed seem a fairly good intuition to explain the observed order sensitivity (for a proposal closely related to our views on the matter see [Mos07]).

A final, and much more speculative, area of potential linguistic application is to vagueness. We can relate this again to [Lew79]; there Lewis gave an account in terms of changing "standards of precision" that can make an utterance of "France is hexagonal" true (or acceptable) in one context and untrue (or unacceptable) in another. A hint in the direction of unawareness is given by Lewis' observation that, as in the case of possibility modals, 'accommodation' proceeds much more smoothly in the direction of *increasing* standards of precision than for decreasing them. If a standard is defined in terms of a set of alternatives (for instance "square", "boot-shaped", and "octagonal") and the best alternative from the set is considered "true enough" according to the standard, then introducing new alternatives via awareness can raise the standards of precision but never lower them. One can think of possibilities here as literally the measurement markers on some device: if we include only the 20-mile markers then it's 100 miles to Chicago, but add some more measurement possibilities and all of a sudden it's 106. As we've defined aggregation this won't work, since the equivalence relation gives rise to a sorites paradox; the open question is whether unawareness might add anything to your favourite sorites-proof account. At least this account gives an easy explanation for the asymmetry that Lewis has noted.

6 Conclusion

We have tried to cover a lot of ground in this paper; it's quite likely that we haven't succeeded in convincing the reader of everything. This is a good point to take stock, and state clearly which notions we think are central and which can be discarded while still agreeing with the endeavour in general.

We've used three slogans to give intuitions about unawareness:

1. Unawareness is not uncertainty (it cannot be represented formally by uncertainty; it typically takes the form of implicit beliefs).

2. Unawareness is not introspective (it must be represented intensionally; the modeller's language is not the agent's language).

3. Unawareness is easily overturned (it stems from absent-mindedness or inattentiveness; mere mention of possibilities, whatever the linguistic setting, suffices).

In particular the third slogan shows how our notion differs from the version common in the rational choice literature; as far as we can see, this characteristic is key for a *linguistic* application of the idea.

We've modelled unawareness in terms of FILTERING a background model through a set of UNMENTIONABLES (which define a limited agent language) and ASSUMPTIONS, and distinguished between IMPLICIT and EXPLICIT beliefs. These are key concepts we would like to see generally adopted, whatever the specific implementation.

In decision theory we've made a more specific suggestion: that decision problems be considered a subjective representation of the relevant features of the situation, and that unawareness models be used whenever that subjective notion of relevance may undergo revision over time. The technical details of our model produce numerically precise and potentially quite complex revisions of decision problems by way of simple updates to awareness structures.

In linguistics we've argued that awareness dynamics are a natural feature in conversation. We have offered a toy example of pragmatic reasoning centered on a conversational move intended first and foremost to bring a possibility to awareness. A further example is offered by the speculative suggestions of the previous section: they are intended not as assertions but simply to make the reader aware of the impressive range of possibilities this notion might fruitfully be applied to.

References

[BC07] Oliver Board and Kim-Sau Chung. "Object-Based Unawareness". In: *Logic and the Foundations of Game and Decision Theory, Proceedings of the Seventh Conference*. Ed. by G. Bonanno, W. van der Hoek, and M. Woolridge. 2007.

[Byr89] Ruth M. J. Byrne. "Suppressing Valid Inferences With Conditionals". In: *Cognition* 31.1 (1989), pp. 61–83.

[DLR98] Eddie Dekel, Barton L. Lipman, and Aldo Rustichini. "Standard State-Space Models Preclude Unawareness". In: *Econometrica* 66.1 (1998), pp. 159–173.

[Fei04] Yossi Feinberg. "Subjective Reasoning - Games with Unawareness". Research Paper No. 1875, Stanford University. Nov. 2004. URL: https://gsbapps.stanford.edu/researchpapers/library/RP1875.pdf.

[Fei05] Yossi Feinberg. "Games with Incomplete Awareness". Research Paper No. 1894, Stanford University. May 2005. URL: http://www.stanford.edu/%7Eyossi/Files/Games%20Incomplete%20DP.pdf.

[FH88] Ronald Fagin and Joseph Y. Halpern. "Belief, Awareness and Limited Reasoning". In: *Artificial Intelligence* 34 (1988), pp. 39–76.

[F01] Kai von Fintel. "Counterfactuals in a Dynamic Context". In: *Ken Hale: A Life in Language*. Ed. by Michael Kenstowicz. MIT Press, 2001, pp. 123–152.

[FJ07] Michael Franke and Tikitu de Jager. "The relevance of awareness". In: *Proceedings of the Sixteenth Amsterdam Colloquium*. Ed. by Maria Aloni, Paul Dekker, and Floris Roelofsen. 2007, pp. 91–96.

[Gil07] Anthony S. Gillies. "Counterfactual Scorekeeping". In: *Linguistics and Philosophy* 30 (2007), pp. 329–360.

[Gri89] Paul Herbert Grice. *Studies in the Ways of Words*. Harvard University Press, 1989.

[Hal01] Joseph Y. Halpern. "Alternative Semantics for Unawareness". In: *Games and Economic Behavior* 37 (2001), pp. 321–339.

[HMS06] Aviad Heifetz, Martin Meier, and Burkhard C. Schipper. "Interactive Unawareness". In: *Journal of Economic Theory* 130 (2006), pp. 78–94.

[HMS07] Aviad Heifetz, Martin Meier, and Burkhard C. Schipper. *Unawareness, Beliefs and Games*. Tech. rep. 6. Bonn Econ Discussion Papers, 2007.

[HMS08a] Aviad Heifetz, Martin Meier, and Burkhard C. Schipper. "A canonical model for interactive unawareness". In: *Games and Economic Behavior* 62 (2008), pp. 304–324.

[HMS08b] Aviad Heifetz, Martin Meier, and Burkhard C. Schipper. "Dynamic Unawareness and Rationalizable Behavior". Unpublished Manuscript. Apr. 2008.

[HR06] Joseph Y. Halpern and Leandro Chaves Rêgo. "Extensive Games with Possibly Unaware Players". Full paper unpublished; a preliminary version appeared in: Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems, 2006, pp. 744–751. 2006. URL: http://www.cs.cornell.edu/home/halpern/papers/aamas06.pdf.

[Jef65] Richard Jeffrey. *The Logic of Decision*. Chicago: Univeristy of Chicago Press, 1965.

[Lew79] David Lewis. "Scorekeeping in a language game". In: *Journal of Philosophical Logic* 8 (1979), pp. 339–359.

[Mos07] Sarah Moss. "On the Pragmatics of Counterfactuals". Unpublished Manuscript, MIT. 2007.

[MR94] Salvatore Modica and Aldo Rustichini. "Awareness and Partitional Information Structures". In: *Theory and Decision* 37 (1994), pp. 107–124.

[MR99] Salvatore Modica and Aldo Rustichini. "Unawareness and Partitional Information Structures". In: *Games and Economic Behavior* 27 (1999), pp. 265–298.

[MS82] P. Milgrom and N. Stokey. "Information, trade and common knowledge". In: *Journal of Economic Theory* 26 (1982), pp. 17–27.

[Ozb07] Erkut Y. Ozbay. "Unawareness and strategic announcements in games with uncertainty". In: *Proceedings of TARK XI*. Ed. by Dov Samet. 2007, pp. 231–238.

[RS61] H. Raiffa and R. Schlaifer. *Applied Statistical Decision Theory*. MIT Press, 1961.

[SL08] Keith Stenning and Michiel van Lambalgen. *Human reasoning and cognitive science*. MIT Press, 2008.

[Sta78] Robert Stalnaker. "Assertion". In: *Syntax and Semantics*. Ed. by Peter Cole. Vol. 9. New York: Academic Press, 1978, pp. 315–332.

[Swa06a] Eric Swanson. "Interactions with Context". PhD thesis. Massachusetts Institute of Technology, 2006.

[Swa06b] Eric Swanson. "Something 'Might' Might Mean". Unpublished Manuscript, University of Michigan. 2006.

[Wil08] J. Robert G. Williams. "Conversation and Conditionals". In: *Philosophical Studies* 138.2 (Mar. 2008), pp. 211–223.