Presuppositions, provisos, and probability*

Daniel Lassiter
Stanford University

Abstract Theories of presupposition in the tradition associated with Karttunen, Stalnaker and Heim relate presupposition satisfaction to the content of conversational participants' epistemic states, usually modeled as sets of worlds. However, converging evidence from recent work on modality and from other areas of cognitive science suggests that epistemic states are better thought of as having the richer structure of probability distributions. I describe an account of semantic and pragmatic presupposition which combines core ideas from dynamic semantic treatments with a probabilistic model of information states and their dynamics in conversation, and argue that it predicts the core data of the proviso problem (Geurts 1996) without invoking *ad hoc* mechanisms as conditional strengthening accounts typically do. The frequently cited intuition that (ir)relevance is crucial follows without stipulation, and I present new cases which suggest that irrelevance is too weak to predict all cases of unconditional presuppositions, problematizing strengthening accounts which rely on it. The proposed theory is able to account for this new data and also for semi-conditional presuppositions, a sticking point for previous theories of presupposition projection. I argue that this perspective also gives us a reasonable line on several related issues, including the divergence between presupposed conditionals and conditional presuppositions, instances of the proviso problem in counterfactuals, and the contextual variation in the difficulty of accommodation.

---

1 The proviso problem: Core issues

Theories of presupposition projection following Heim (1983) are built around the notion of **satisfaction in a local context**. Geurts (1996) points out that these theories predict weak conditional presuppositions like (1a) in many cases in which the actual inferences are unconditional, as in (1b). Geurts dubs this the **proviso problem**.

(1)   If Theo's wife hates sonnets then his manager does too. (Geurts 1996: 268-9)

     a.   ↝̸ If Theo's wife hates sonnets then he has a manager.
     b.   ↝ Theo has a manager.

However, as Geurts (1996) and Beaver (2001) point out, genuinely conditional presuppositions do arise in certain cases.

(2)   If John is a diver, he'll bring his wetsuit on vacation.

     a.   ↝ If John is a diver, he has a wetsuit.
     b.   ↝̸ John has a wetsuit.

The basic puzzle is to explain why (1) has the unconditional presupposition (1b), while (2) has the conditional presupposition (2a).

Satisfaction theorists have generally opted for some kind of strengthening account: effectively, the idea is that the presupposition of (1) really is (1a), but some secondary mechanism strengthens (1a) to (1b). In addition to being somewhat *ad hoc*, this approach suffers from a number of empirical problems noted by Geurts (1996), of which two of the most important are discussed here. First, such a theory must explain how the mechanism which strengthens the conditional presupposition in (1) can avoid doing the same in (3):

(3)   Sam knows that if Theo's wife hates sonnets then he has a manager.

     a.   ↝ If Theo's wife hates sonnets then he has a manager.
     b.   ↝̸ Theo has a manager.

Satisfaction theories predict that the presuppositions of (1) and (3) should be the same, and yet the strengthening mechanism must be able to distinguish them. Second, the theory must explain "semi-conditional" presuppositions like (4b). These are problematic for all major theories of presupposition projection (cf. Schlenker 2011).

(4)   If John is a diver and wants to impress his girlfriend, he'll bring his wetsuit.

      a. ⇴ If John is a diver and wants to impress his girlfriend, he has a wetsuit.

      b. ⇝ If John is a diver, he has a wetsuit.

      c. ⇴ John has a wetsuit.

2 Overview

I present a modified satisfaction theory built around the assumption that information states are probabilistic in form. On this account, presuppositions are not propositions *per se*, but conditions on conversational participants' probabilistic epistemic states. Very roughly, the idea is that presupposed information is taken for granted, and it is inappropriate to take a proposition for granted unless it is highly probable (subject to the usual caveats about accommodation, to be discussed below).

In many cases the account proposed makes the same predictions as standard satisfaction theories which treat information states as sets of worlds. However, the move to probabilities brings along a number of benefits. Most importantly for our purposes, the contrast in (1-2) is predicted as a consequence of the fact that, depending on speakers' and listeners' qualitative knowledge about which probabilistic dependencies hold among relevant propositions, the constraints on conditional probabilities that the theory assigns as pragmatic presuppositions to conditional sentences may actually entail the constraints assigned to the corresponding unconditionals. In combination with some pragmatic considerations about the best English paraphrases of probability statements, the theory predicts that an unconditional sentence should appear as the best rendition of a presupposition in the consequent of a conditional when it is probabilistically independent of the antecedent. Conditional sentences, on the other hand, are predicted to be the best paraphrase when the conditional probability is known to be greater than the unconditional probability. I suggest that these predictions match well our intuitions about dependencies in core cases of the proviso problem.

The present theory improves in several ways on previous accounts of the proviso problem, including those which invoke relevance and/or probabilistic independence. Many previous approaches invoke strengthening mechanisms which can reasonably be criticized as *ad hoc* and excessively powerful (e.g., Singh 2007; see Schlenker 2011 for discussion). The proposal of Schlenker (2011), while more constrained, does not include a clear theoretical rationale for its occasional invocation of probabilistic independence. In contrast,

my probabilistic solution to the proviso problem follows directly from the architecture of the theory, rather than being invoked as a *deus ex machina* strengthening mechanism.[1]

§6 presents new data which suggests that unconditional inferences arise not only in cases of independence between antecedent p and consequent presupposition q, but also when the conditional probability of q given p is known to be *less* than the unconditional probability of q. This is unexpected if independence is the crucial factor, but the present theory is able to account for it. In the final sections I argue that the theory gives us a simple solution to the problem of semi-conditional presuppositions, defuses Geurts' objection based on (3), yields a reasonable measure of the difficulty of accommodation, and can be extended to account for cases of the proviso problem in counterfactuals.

3 Motivation and formal details

Satisfaction theories generally assume that presuppositions place conditions on the information states of conversational participants in some way. Indeed, as Beaver (2001: 79) points out, the key feature uniting dynamic theories — including both dynamic semantic theories like Beaver 2001, Heim 1983 and dynamic pragmatic accounts such as Gazdar 1979, Schlenker 2009, Stalnaker 1973 — is that presupposition satisfaction and projection are tied in one way or another to the epistemic states of conversational participants and the way that these change in the course of a conversation. (Theories built around DRT such as Geurts 1999, van der Sandt 1992 rely on rather different assumptions, of course. I will not try to discuss the many differences between my account and DRT in this paper, though the topic is well worth considering.)

The default assumption for such approaches, I take it, is that the information states relevant to presupposition are the same ones that are used to evaluate epistemic modals, as claimed explicitly by e.g. Beaver (2001),

1 Other than the works of Singh and Schlenker already cited, the closest predecessor to my knowledge is Beaver (1999), who considers a Bayesian account as one among several ways of clarifying his notion of a plausibility ordering on epistemic states. An account built around a different non-probabilistic concept of "independence" which makes reference to a notion of orthogonality of questions is proposed by van Rooij (2007). This account and the present one are close in spirit, if not in implementation; a detailed comparison of the predictions of the two theories would be interesting, but will not be pursued here.

Klinedinst & Rothschild (2012), Veltman (1996), Yalcin (2007).[2] In standard accounts of both epistemic modality and presupposition, information states are assumed to be sets of worlds. However, if this assumption were problematized on the basis of evidence from the semantics of epistemic modals or from other fields concerned with the way that people represent information, it would be reasonable to revise our assumptions about the structure of the information states relevant to presupposition and consider whether this has ramifications for other aspects of the theory. In fact, the sets-of-worlds conception of epistemic states has been called into question in recent work from a variety of perspectives, including formal semantics, psychology, and artificial intelligence. In each case, the best available formal tools represent information using structures at least as rich as **probability distributions**.

This section sets up my account by outlining a satisfaction theory due to Klinedinst & Rothschild (2012) which is close to standard dynamic semantics, and then describing briefly some motivations external to the theory of presupposition for upgrading this account to a probabilistic one. §4 then gives some formal details of a probabilistic theory of presupposition along with general considerations about the type of pragmatics appropriate to the account. Although the technical changes that I will propose to Klinedinst & Rothschild's (2012) account are relatively small and the pragmatics is (I hope) fairly intuitive, these modifications have important consequences for presupposition projection, which I explore in the following sections.

3.1 A satisfaction theory based on information states

Yalcin (2007) uses facts about the sentence-internal dynamics of epistemic modals to argue that interpretation is relativized to an **information state parameter** s which can be manipulated by various operators including attitude verbs and *if*. Klinedinst & Rothschild (2012) adopt Yalcin's proposal and use it to construct a static variant of dynamic approaches to presupposition projection (e.g., Beaver 2001, Heim 1983; see also Cresswell 2002 for a related static semantic treatment of dynamic phenomena). I will quickly present a summary/interpretation of Klinedinst & Rothschild's variant of dynamic

---

2 This equation might appear to be undesirable if we think that presuppositions are constraints on a body of knowledge shared by speaker and listener (common ground), since this would seem to render epistemic modals generally uninformative; see however Dekker 2010, Roussarie 2009 for discussion and proposals that avoid this problem by enriching the notion of common ground.

semantics to use as a jumping-off point for my own account.[3] I will focus only on their account of presupposition, but note that this is not Klinedinst & Rothschild's (2012) main concern: they give a number of arguments for an interaction of connectives and information states along these lines which are independent of issues around presupposition projection.

Klinedinst & Rothschild make the customary assumption that information states are sets of worlds. They treat semantic presuppositions as lexically triggered, and define conditions for the use of presuppositional expressions in terms of the form of the relevant information state. Although their account is framed in terms of definedness conditions on propositions, I will state it in terms of a pragmatic usage constraint (this does not affect their predictions, and it makes the relation to my own theory clearer). In the simplest case of atomic sentences, we require that an expression p with semantic presupposition $\underline{p}$ should not be used unless $\underline{p}$ is entailed by the information state. More explicitly:

(5)   **Usage constraint** (atomic sentences): If p is an atomic sentence with semantic presupposition $\underline{p}$, then p can be uttered felicitously only if all worlds in the contextually relevant information state s satisfy $\underline{p}$, i.e. if $s \subseteq \underline{p}$.

(Note that I adopt the convention of underlining presuppositions: for any q, if q is associated with a semantic presupposition it will be written $\underline{q}$.)[4]

The next step is to define connectives which manipulate s. The key to presupposition projection in this system is that clauses which occur in non-initial clauses of a complex sentence are evaluated with respect to a **local information state** which may take into account the information contained in earlier clauses. This may lead to different restrictions on s than the same semantic presupposition would have if it were to occurred in a sentence with only one clause, where it would be evaluated with respect to the **global**

---

3 I suspect that better-known varieties of dynamic semantics would also suffice for this purpose, though I will not explore this possibility here. The choice of Klinedinst & Rothschild 2012 as a starting point is not totally innocent, though: their theory invites an interpretation in terms of domain-general principles of information flow which is somewhat closer to the Bayesian perspective that I will argue for than typical dynamic semantic theories are.

4 Some complex theoretical issues are hidden in the description "the contextually relevant information state". We might construe this state as a representation of common ground (Clark & Marshall 1981, Stalnaker 1974, 1978), or take (5) to constrain each conversational participant's personal epistemic state separately, as I will in my proposal below. I don't know if these choices make empirically different predictions, either for Klinedinst & Rothschild's theory or for the modification I suggest below (cf. fn.10).

information state s.

(6) Let s_χ be defined as $\{w' \in s \mid [\![\chi]\!]^{c,s,w'} = 1\}$, the χ-subset of s. Then:

    a. $[\![\neg\phi]\!]^{c,s,w} = 1$ iff $[\![\phi]\!]^{c,s,w} = 0$

    b. $[\![\phi \wedge \psi]\!]^{c,s,w} = 1$ iff $[\![\phi]\!]^{c,s,w} = 1$ and $[\![\psi]\!]^{c,s_\phi,w} = 1$

    c. $[\![\phi \vee \psi]\!]^{c,s,w} = 1$ iff $[\![\phi]\!]^{c,s,w} = 1$ or $[\![\psi]\!]^{c,s_{\neg\phi},w} = 1$

    d. $[\![\phi \to \psi]\!]^{c,s,w} = 1$ iff $[\![\phi]\!]^{c,s,w} = 0$ or $[\![\psi]\!]^{c,s_\phi,w} = 1$

Note that the information state parameter sometimes changes depending on which clause is being evaluated, in a way which takes account of what was encountered earlier.[5,6] Presupposition projection facts are generated by the usage condition for complex sentences: a complex sentence should not be used unless the semantic presuppositions of all of its atomic parts are fulfilled relative to the local information state in which they are evaluated.[7]

(7) **Usage constraint** (complex sentences): If ϕ is a (possibly) complex sentence with atomic parts $q_1, \ldots, q_n$ having semantic presuppositions $\underline{q}_1, \ldots, \underline{q}_n$ occurring in local information states $s_1, \ldots, s_n$, then ϕ should not be used unless $s_1 \subseteq \underline{q}_1 \wedge \cdots \wedge s_n \subseteq \underline{q}_n$.

As the reader may check, this account generates predictions about presupposition projection which are equivalent to Heim's (1983) dynamic semantics supplemented by Beaver's (2001) asymmetric treatment of *or*. For example, *If John is a diver, he will bring his wetsuit* is predicted by (7) to be inappropriate unless the local information state of the consequent entails *John has a wetsuit*. Checking (6d), we see that this state is $s_{(John\ is\ a\ diver)}$, the set of worlds in the global information state in which the antecedent is true; so the sentence is predicted to be infelicitous unless John has a wetsuit in every s-world in which he is a diver. This constraint on the local state $s_{(John\ is\ a\ diver)}$ is equivalent to the constraint that the global state s must satisfy the material conditional *John is a diver $\supset$ John has a wetsuit*. The latter is, of course, the same presupposition that Heim predicts.

---

5 For simplicity, I treat *if* as a material conditional. Klinedinst & Rothschild (2012) give a more complicated restrictor analysis à la Kratzer 1986, Yalcin 2007, but the difference does not affect the main issues here.

6 As Klinedinst & Rothschild (2012) also note, it may not be necessary to stipulate lexically how connectives influence the information state parameter: see Rothschild (2011), Schlenker (2009) for proposals deriving these effects. This is indeed a desirable feature of a pragmatic account of presupposition projection, but I will not pursue this connection here.

7 I am glossing over some non-trivial technical details involving the implementation of (7), since they do not affect the issues of theoretical interest.

The equivalence between Heim's (1983) and Klinedinst & Rothschild's (2012) predictions in this example and others relies on a feature which will be important in what follows (since my own proposal differs crucially in this respect). A conjunction of usage conditions of the form $s_{r_1} \subseteq \underline{q}_1 \wedge \cdots \wedge s_{r_n} \subseteq \underline{q}_n$ can always be rewritten as a single condition which directly constrains the global information state: s must satisfy the conjunction of material conditionals $r_1 \supset q_1 \wedge \cdots \wedge r_n \supset q_n$, and the latter will be equivalent to the semantic presupposition that Heim's theory generates. This means that it is not possible to distinguish empirically between Heim's theory — in which both atomic sentences and complex sentences carry semantic presuppositions, and a single usage constraint applies to both — and Klinedinst & Rothschild's, in which only atomic sentences have semantic presuppositions and projection facts are generated by the way that the usage constraint interacts with the definitions of connectives. In a theory like this, then, it does not do any harm to think of complex sentences as carrying semantic presuppositions: even though technically they do not, the usage conditions for complex sentences give us something equivalent to what we would derive if we were to apply the usage condition for simple sentences applied to a complex semantic presupposition.

3.2 Motivations for transition to probabilities

Recently various authors have noted phenomena involving the gradability of epistemic modals such as *possible, likely,* and *certain* and the inferences that they license that are difficult to account for if information states are modeled as sets of worlds (Lassiter 2010, 2011, Swanson 2006, Yalcin 2005, 2007, 2010). These authors argue that the problematic data can be explained if the information states relevant to the semantics of epistemic modals are not sets of worlds, but have the richer structure of a probability measure.

(8)   An information state s is a **probability measure** on a set of worlds W if and only if $\Phi \subseteq \mathcal{P}(W)$ is a set of propositions (sets of worlds) which contains W and is closed under Boolean operations (e.g., $\mathcal{P}(W)$ if $|W|$ is finite); and

a. $s : \Phi \rightarrow [0, 1]$;

b. $s(W) = 1$;

c.   For all A and $B \in \Phi$: if $A \cap B = \emptyset$, then $s(A \cup B) = s(A) + s(B)$.

In the simplest case (when the cardinality of *W* is not too great), a probability measure can be thought of as a set of epistemically possible worlds supplemented by a measure function which is required to sum to 1. Technically, then, probability is a straightforward enrichment of the set-based conception of information states.

If epistemic modals rely on probabilistic information states as these authors claim, then we may reasonably suppose that the information states relevant to theories of presupposition are also probabilistic in form. Further arguments in support of a probabilistic account of information states come from psychology and artificial intelligence, two fields which — like linguistic semantics and pragmatics — have a strong interest in the structure of information and how it is represented and processed by humans. Recent psychological work on higher-level cognition suggests that probabilistic theories of reasoning, learning and decision-making improve upon traditional logical approaches in numerous respects (Chater, Tenenbaum & Yuille 2006, Griffiths, Kemp & Tenenbaum 2008).[8] Likewise, in modern artificial intelligence probabilistic models are widely thought to provide the best available format for models of learning and reasoning given the noisy and inconsistent input with which realistic agents must cope (Pearl 1988, Russell & Norvig 2010).

Obviously, none of this tells us conclusively how the best theory of presupposition should be structured. It would be possible (though unparsimonious) to hold that epistemic modals and presuppositions rely on the bodies of information with different basic structures. Likewise, it could in principle be that the best overall theory of information processing in human cognition is not the one that we should use in formal pragmatics. I don't want to overstate the case for probabilistic models in the theory of presupposition on the basis of indirect evidence, then; but given all of this motion within closely related areas of cognitive science and more recently within formal semantics itself, we ought to take very seriously the possibility that the information states relevant to a formal theory of presupposition are also probabilistic.

---

8 Though there are classic arguments that humans do not reason probabilistically (see especially Kahneman, Slovic & Tversky 1982), the results on which these arguments rely have been show in many cases to admit of alternative explanation, either as artifacts of experimental design (e.g., Gigerenzer 1991) or the semantics of verbal stimuli and the pragmatics of the experimental situation (cf. among others, Hilton 1995, Tenenbaum & Griffiths 2001, Lassiter 2011: ch.4). Overall, the evidence that probability plays a crucial role in the representation and processing of information is strong and growing.

4 Proposal: A probabilistic satisfaction theory

4.1 Main idea

Suppose, then, that speakers come to a conversation equipped with probabilistic information states, and that these states are relevant to determining both the felicity of presuppositional expressions and the way in which presuppositions project. How can this be implemented with minimal modification to a satisfaction theory like the one we outlined briefly above? To get an intuition about the direction we are headed, consider the fact that presuppositions are backgrounded information, and that speakers must make choices about how to divide their utterances into foregrounded and backgrounded information. A plausible condition on cooperative conversation is that information should not be backgrounded unless it can reasonably be taken for granted. Now, basic arithmetic and tautologies aside, people really know very little with certainty; yet we must background some information in order to communicate efficiently. In order to decide whether p can be taken for granted, then, speakers must judge whether the information favoring p is sufficiently strong. If it is not, it should not be treated as uncontroversial; if it is, it can sometimes be presupposed.

From a probabilistic perspective — more specifically, from a Bayesian perspective, where probabilities are interpreted as degrees of belief — the requirement that the evidence for a conclusion be "sufficiently strong" translates into a requirement that the probability of the conclusion be high enough, where the meaning of "sufficient/enough" is determined by some contextual factors. Naming the relevant parameter θ, this reasoning suggests replacing the entailment-based usage condition (5) above with the probabilistic condition (9):

(9) **Usage constraint** (atomic sentences): Let p be an atomic sentence which carries the semantic presupposition $\underline{p}$. Then a speaker should not utter p unless $pr(\underline{p})$ meets or exceeds a high threshold θ according to her epistemic state, and she believes that her audience also assigns $\underline{p}$ at least probability θ.

On this account, the fact that a presuppositional expression p has been used will not necessarily lead to the inference that its semantic presupposition $\underline{p}$ holds in **all** epistemically possible worlds. Instead, the inference is that (the speaker believes that) $\underline{p}$ has high probability, i.e. that it holds throughout some subset of W which is distinguished by the fact that the

actual world is extremely likely to lie somewhere in this set. I will not have much to say about the difficult pragmatic/psychological question of how likely something must be before it can reasonably be taken for granted. For present purposes it should be sufficient to think of the value of θ as a feature of the conversational scoreboard (Lewis 1979), sensitive to numerous factors including conversational stakes. However θ is determined in a particular conversation, we can learn a lot about the predictions of the probabilistic theory without knowing its value, and so I will leave it as a free parameter for now.

It will be important for us to maintain a clear distinction between semantic presuppositions, pragmatic usage constraints, and pragmatic presuppositions. I assume as usual that semantic presuppositions — the division of the informational contribution of expressions into foregrounded and backgrounded content — are triggered by lexically encoded features of particular expressions, and that atomic sentences entail what they semantically presuppose.[9] Pragmatic presuppositions, as I will use the term, are commitments that a speaker takes on in the course of *using* an expression that carries a semantic presupposition. The pragmatic presupposition that a speaker makes by uttering a simple sentence in a context will be that she is obeying the usage constraint in (15). We can expect listeners to draw probabilistic inferences involving the semantic presupposition as a result of the fact that the speaker has made such commitments, but only as a secondary inference which will depend on listeners' beliefs about the speaker's cooperativity, reliability, etc. For example, the sentence *Sam's dog has fleas* semantically presupposes that Sam has a dog; when a speaker chooses to utter this sentence, she pragmatically presupposes that the usage condition in (9) is fulfilled, i.e. that pr(Sam has a dog) $\geq \theta$, and a listener can be expected to recognize this commitment and draw whatever inferences he considers appropriate.[10]

Next, I adopt from Klinedinst & Rothschild (2012) the definitions of the connectives in (6), reinterpreting information states as probability measures.

---

9 The first assumption is standard and simplifies matters, but it is not crucial for me and may well be partially or wholly incorrect as e.g. Romoli (2011), Simons (2001), Stalnaker (1974) argue.

10 Note that the way that I have stated this makes direct reference to the epistemic states of speakers and listeners, without the mediation of a common ground. This is partly because the usual conception of common ground as level-infinity mutual belief is excessively restrictive in a probabilistic framework, and I do not know how best to relax the concept or indeed whether it makes a difference for my account. Perhaps the notion of common ground discussed in Clark & Marshall 1981 would be of use here.

For notational convenience I use *pr* rather than *s* as a variable over information states. We also have to redefine the subscripted information states used in the definitions of connectives so that they are probability measures. The natural way to do this is to treat them as conditional probability measures:

(10) $pr_\phi =_{df}$
the function pr' such that, for any proposition ψ, $pr'(\psi) = \frac{pr(\phi \wedge \psi)}{pr(\phi)}$.

Conditional probabilities maintain the essential function that subscripted information states had in Klinedinst & Rothschild's (2012) theory: in some cases we do not want to place restrictions on the whole information state, but only on some subpart of it, as determined by what occurred earlier in the sentence. The conditional probability measure pr_ϕ restricts attention to the portion of W in which ϕ holds, and so a conditional probability statement involving pr_ϕ does not place any constraints on what goes on outside the ϕ-region. (I will also frequently write $pr_\phi(\psi)$ as $pr(\psi|\phi)$.) The connectives are now defined as:

(11) a. $[\![\neg\phi]\!]^{c,pr,w} = 1$ iff $[\![\phi]\!]^{c,pr,w} = 0$

b. $[\![\phi \wedge \psi]\!]^{c,pr,w} = 1$ iff $[\![\phi]\!]^{c,pr,w} = 1$ and $[\![\psi]\!]^{c,pr_\phi,w} = 1$

c. $[\![\phi \vee \psi]\!]^{c,pr,w} = 1$ iff $[\![\phi]\!]^{c,pr,w} = 1$ or $[\![\psi]\!]^{c,pr_{\neg\phi},w} = 1$

d. $[\![\phi \rightarrow \psi]\!]^{c,pr,w} = 1$ iff $[\![\phi]\!]^{c,pr,w} = 0$ or $[\![\psi]\!]^{c,pr_\phi,w} = 1$

Constraints on the appropriate use of presuppositional expressions in atomic sentences were already redefined in (9); the pragmatic constraint affecting complex sentences (12) is related to the atomic constraint in (9) basically as the corresponding rules were in the version of Klinedinst & Rothschild's (2012) theory presented above.

(12) **Usage constraint** (complex sentences): Let ϕ be a possibly complex sentence. Appropriate use of ϕ requires that, for any atomic part p of ϕ with local information state pr', (9) is satisfied (taking pr in the definition of (9) to be the local information state pr' of p).

There is a subtle but important difference that arises here between previous satisfaction theories and the probabilistic account. As noted above, only atomic clauses have semantic presuppositions in Klinedinst & Rothschild's (2012) theory, but the usage conditions for complex sentences place restrictions on local information states which are systematically equivalent to certain conditions that we could place directly on global information states. As a result, we could reasonably think of the latter as the semantic presup-

positions of complex sentences, derived from the semantic presuppositions of their parts by a projection mechanism.

The situation is quite different for the probabilistic theory. The usage conditions associated with a complex sentence ϕ by (12) are a conjunction of separate usage conditions of the form $pr_1(\underline{q_1}) \geq \theta \wedge pr_2(\underline{q_2}) \geq \theta \wedge \cdots \wedge pr_n(\underline{q_n}) \geq \theta$, where $\underline{q_1} \ldots \underline{q_n}$ are the semantic presuppositions associated with the atomic clauses of ϕ and $pr_1 \ldots pr_n$ are the local probability measures associated with clauses $q_1 \ldots q_n$ as determined by (11). This conjunction of conditions on probability measures will not in general be equivalent to the condition that any proposition X has probability θ.[11] As a result, the usage conditions of a complex sentence generated by (12) cannot be thought of as equivalent to placing a single usage condition on some semantic presupposition of the complex sentence: it is no longer theoretically innocent to pretend that the theory associates complex sentences with semantic presuppositions. I want to highlight this feature, since it will be important in our account of the proviso problem:

> Only atomic sentences have semantic presuppositions. Complex sentences ϕ are associated with complex pragmatic usage constraints placing conditions on certain conditional probabilities, but these are not systematically equivalent to the condition that any particular proposition has high probability.

If there were some proposition X which is required to have high probability by complex sentence ϕ, we could simply find out what X is and call it the semantic presupposition of ϕ. But unless the information state has certain special features (which will occupy us a good deal in what follows), there usually won't be one: instead, felicitous use of ϕ requires (by 12) that certain conditional probabilities be high, and the overall effect of imposing these conditions on an information state pr will depend on various detailed facts about pr.

As a result of all this, the probabilistic account that I am offering differs from previous satisfaction theories in that there is no formula which will take us from the usage conditions for a complex sentence ϕ to a proposition or an English sentence which expresses the "presupposition" of ϕ. Instead, we need to look at the overall form of the relevant information state in order to determine what effects the usage conditions of a complex sentence will

---

11 This follows from the proof given in footnote 18 in §5.3 below.

have, and there will in many cases be no English sentence or proposition that we can use to accurately summarize the constraint that is placed on an information state by the use of a complex sentence. However, in certain special cases there will be an obvious English rendition: in particular, when the form of an epistemic state guarantees that the usage conditions of a sentence are satisfied if and only if some proposition ψ has probability θ, it is not unreasonable to think of ψ as the "presupposition" of ϕ (though it is technically incorrect). We will encounter a number of such cases in what follows, including in the simple cases discussed in the next subsection.

4.2 Examples: Presuppositions of simple and complex sentences

By (11), the initial clause of a complex sentence is always evaluated with respect to the global information state. As a result, when (only) the initial clause of a compound sentence carries a semantic presupposition, the usage conditions and pragmatic presuppositions associated with this compound sentence by (12) are the same that (9) would give us if this clause were uttered as an independent sentence. For example, *Sam's dog has fleas, and he is concerned* shares the pragmatic presupposition of *Sam's dog has fleas*: $pr(\text{Sam has a dog}) \geq \theta$. Consider now (13-14):

(13)   Sam has a dog, and his dog has fleas.

(14)   Sam has a dog, and his cat has fleas.

Abbreviate the first clause of (13) by q. By (11b), q is evaluated relative to a local information state which is identical to the global information state pr, and the second clause is evaluated relative to the local information state pr_q. By the usage condition for complex sentences in (12), the sentence is used appropriately only if the semantic presupposition of the second clause has probability greater than θ relative to its local information state pr_q. This means that the sentence is not appropriate unless

$$pr_q(\text{Sam has a dog}) \geq \theta \quad \text{i.e.} \quad \frac{pr(\text{Sam has a dog} \ \wedge \ \text{Sam has a dog})}{pr(\text{Sam has a dog})} \geq \theta$$

which is trivial: the left side of this inequality will always equal 1, and so the inequality holds for any value of θ.[12] The correct prediction is that using this sentence does not require a speaker to make any non-trivial pragmatic presuppositions.

---

12 Unless $pr(q) = 0$, in which case $pr_q(\cdot)$ is undefined; this does not affect the point here.

On the other hand, appropriate use of (14) requires the speaker to commit to

$$pr_q(\text{Sam has a cat}) \geq \theta \quad \text{i.e.} \quad \frac{pr(\text{Sam has a dog} \wedge \text{Sam has a cat})}{pr(\text{Sam has a dog})} \geq \theta$$

Now this is non-trivial: what it says is that, restricting attention to the portion of *pr* in which *Sam has a dog* is true and normalizing so that this portion has measure 1, the probability that John has a cat is at least θ. Roughly, this condition can be glossed by saying that, if we were to find out for certain that Sam has a dog, the probability that he has a cat once we have incorporated this evidence into our beliefs would meet or exceed the threshold θ.[13]

4.3 (In)dependence relations among propositions

Pragmatic presuppositions in this theory frequently express conditions on conditional probabilities: a speaker who utters (14) does not directly commit himself to assigning high credence to any particular proposition, but he does commit himself to assigning high credence to the conditional probability statement just described. Crucially, what further ramifications this commitment has for his information state — and the commitments that a listener must take on if she accepts it — will depend on which *dependencies* hold among propositions in their respective epistemic states. To illustrate the notion of a dependency, consider the two tiny graphical models below, representing two kinds of probabilistic belief states that a language user might bring to bear on evaluating an utterance of (28). (For technical reasons not crucial here, I suppress the **cat** node that you might expect to see in these models.)
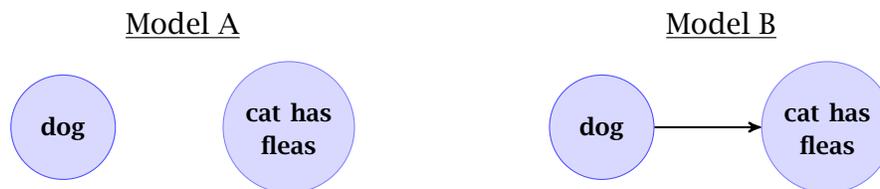
Here propositions are represented as nodes in a graph and edges indicate

---

[13] The gloss is instructive, but not totally accurate; it fails in some cases in which the evaluated proposition makes reference to the speaker's beliefs, cf. van Fraassen 1980, Weirich 1983. Thanks to the *S&P* editors for this point.

probabilistic dependencies between propositions.[14] These models represent two different possible theories among the relationships between the probabilities of the propositions represented in the graph. If two nodes *A* and *B* are not connected, then they are independent, meaning that learning about the probability or truth-value of one will never affect the estimated probability of the other. Independence requires that the following (equivalent) definitions hold:

(15)   *A* and *B* are **independent** iff $pr(A) = pr(A \mid B)$, i.e. if $pr(B) = pr(B \mid A)$.

The models differ in whether **dog** is connected to **cat has fleas**. In Model A, it is not; this model thus represents a class of probability distributions in which the proposition *Sam has a dog* is independent of *Sam's cat has fleas*. This is an accurate representation of an agent's belief state if she believes that whether or not Sam has a dog has no bearing on how likely it is that his cat will have fleas. As a result, the conditional probability of **cat has fleas** given **dog** is the same as the unconditional probability of **cat has fleas**, whatever this may be.

In Model B, on the other hand, **dog** and **cat has fleas** are connected, indicating that these propositions are not thought to be independent. This might represent the epistemic state of an agent who believes that having a dog may affect the probability of one's cat having fleas (e.g., if dogs are thought to carry fleas and spread them to cats). For such an agent the conditional probability of **cat has fleas** given **dog** will not in general be equal to the unconditional probability of **cat has fleas**: learning about one will typically influence the estimated probability of the other.

The use of graphical models of this type is widespread for (at least) two somewhat different types of reasons. One is computational efficiency: representing dependencies explicitly can make inferences much more efficient, since nodes which are independent of x can be ignored in calculating the probability of x (or querying the conditional probability of x given some

---

14 See e.g. Koller & Friedman 2009, Pearl 1988, 2000 for more on graphical models and their motivations and uses. The direction of arrows represents the direction of causal influence, though I will not make use of this feature. Graphical models have been very influential in psychology and AI, but the framework has important expressive limitations, being essentially a propositional language; in particular, the fact that graphical models lack the resources to reason about existence or non-existence of objects (cf. Milch et al. 2007) means that they are not expressive enough to formalize a probabilistic theory of presupposition, which frequently needs to make reference to statements about the existence of objects. However, the core insight that qualitative independence relations are a fundamental organizing principle for probabilistic models seems to be secure.

assumptions). In large models this can lead to substantial improvements. More importantly for us, graphical models can be seen as a hypothesis about an important kind of knowledge that intelligent agents have: in addition to detailed quantitative knowledge about probabilities and conditional probabilities, agents know that certain propositions are independent of certain other propositions and use this knowledge to reduce their information processing load. Borrowing an example from Pearl (2000), people confidently judge that the price of beans in China is unrelated to the amount of traffic in Los Angeles, even though they may have little confidence in their ability to estimate the value of either of these variables. Although the details of the graphical models formalism and its limitations are not overly crucial for our purposes, we will make extensive use of one motivating idea borrowed from this theory: an important part of the cognitive representation of information is our knowledge of qualitative probabilistic relations between propositions, and in particular of independence relations.

5 Probabilistic account of the proviso problem

5.1 The core problem

The theory sketched in §4 and the notion of independence give us everything we need to explain the core data surrounding the proviso problem. The basic problem that we began with was the contrast between (1) and (2): why do conditionals sometimes give rise to presuppositions that are well-described by conditional sentences (e.g. (16)), and sometimes not (e.g. (17))?

(16)   If John is a diver, he'll bring his wetsuit on vacation.   (= (2))

    a.   ⤳ If John is a diver, he has a wetsuit.

    b.   ⤳̸ John has a wetsuit.

(17)   If Theo's wife hates sonnets then his manager does too.   (= (1))

    a.   ⤳̸ If Theo's wife hates sonnets then he has a manager.

    b.   ⤳ Theo has a manager.

The crucial difference between the examples (as many authors have noted before me) is that the truth of the antecedent p does not seem to be relevant to the truth of the semantic presupposition $\underline{q}$ of the consequent q. The present theory allows us to explain why this feature affects the felt presuppositions so sharply. Starting with (17), the usage constraints that

I proposed in §4.1 predict that appropriate use of this sentence requires that the speaker believe that the following condition hold, constraining the conditional probability of *q given the antecedent p*:[15]

(18) pr(Theo has a manager | Theo's wife hates sonnets) $\geq \theta$

In the probabilistic language that we are employing, the fact that we naturally assume that *Theo's wife hates sonnets* is not relevant to *Theo has a manager* corresponds to an assumption that these propositions are probabilistically independent. If this is right, then the equation in (19) holds:

(19) pr(Theo has a manager) $=$
                         pr(Theo has a manager | Theo's wife hates sonnets)

The validity of the following argument guarantees that a probability measure in which (19) holds will also satisfy (20c).

(20)   a. pr(Theo has a manager | Theo's wife hates sonnets) $\geq \theta$
        b. pr(Theo has a manager) $=$
                         pr(Theo has a manager | Theo's wife hates sonnets)
        c. $\therefore$ pr(Theo has a manager) $\geq \theta$

Not coincidentally, (20c) is the same pragmatic presupposition that our theory would assign to the simple sentence *Theo's manager hates sonnets*: there is a high probability that Theo has a manager. This is, I claim, the essential reason why the conditional sentence *If Theo's wife hates sonnets then his manager does too* has the same felt presupposition as the simple sentence *Theo's manager hates sonnets*.

Why, then, do we feel that the "real" presupposition of (17) is the unconditional *Theo has a manager* (17b) and not the conditional *If Theo's wife hates sonnets then he has a manager* (17a)? After all, both of these sentences must presumably have high probability in any situation in which the pragmatic presupposition is appropriate. Of course, (17) is no different from the simple sentence *Theo's manager hates sonnets* in this respect: given that we are in an epistemic state which satisfies the independence condition (19), the latter, too, is appropriate only if pr(Theo has a manager | Theo's wife hates sonnets) $\geq \theta$, as can be seen by inverting (20a) and (20c). Nevertheless, there are two differences between the proposed renditions of these sentences' shared pragmatic presupposition. First, the conditional is logically weaker, and general pragmatic considerations (along the lines of Grice's (1989) Maxim of Quantity)

---

15 I will ignore the irrelevant presupposition of the antecedent of (17).

lead us to prefer the strongest description available.[16] The second difference is that the use of a conditional sentence typically implicates that the antecedent of the conditional is relevant to the consequent. This inference is precisely contrary to the independence assumption that people bring to bear in interpreting (17), and so the conditional paraphrase of these usage conditions is at best seriously misleading. For both of these reasons, (17b) is a better description than (17a) of what must (probably) be the case for either the simple sentence *Theo's manager hates sonnets* or the complex (17) to be uttered appropriately.

None of this reasoning applies, however, to the pragmatic presuppositions associated with a typical utterance of (16). The usage conditions for this sentence require similarly that

(21) pr(John has a wetsuit | John is a diver) $\geq \theta$,

but the independence assumption is clearly inappropriate here: learning whether John is a diver will typically influence the estimated probability that he owns a wetsuit. Without the independence assumption, the reasoning does not go through since the premise corresponding to (20b) is false, and so (21) is not equivalent to pr(John has a wetsuit) $\geq \theta$ in an information state in which independence does not hold. In such cases the usage conditions associated with the utterance are typically naturally rendered using a conditional sentence *If p then q*.[17]

An analysis along these lines seems to be appropriate in other clear cases of conditional presuppositions in the literature as well: the conditional probability of the semantic presupposition of the consequent $\underline{q}$ given the antecedent p is greater than the unconditional probability of $\underline{q}$. Moreover, in disputed cases we can detect the effect directly, e.g. the following example from Beaver 1999:

(22)   If Jane takes a bath, Bill will be annoyed that there is no more hot water.

---

16 Note that we are now talking about the pragmatics involved in our theoretical description of the issues, and not about the pragmatics of the speech situations that we are analyzing. The probabilistic usage constraints that the present theory makes use of are a different sort of object entirely than English sentence meanings, and obeying these constraints presumably does not require our probabilistic agents to render them into English.

17 The precise story of why the latter rendition *is* appropriate will of course depend on our assumptions about the meaning of indicative conditionals. One possibility is that the latter have truth-conditions which are similar to the pragmatic usage condition in (21), cf. the papers by Egré & Cozic (2011), Kratzer (1986) discussed briefly in fn.19 below.

As [Beaver (1999)](#) and [van Rooij (2007)](#) discuss, whether or not you hear this sentence as presupposing that there is indeed no more hot water seems to depend on your background assumptions about whether one person's taking a bath can influence the hot water situation for subsequent bathers. If you do not expect such connections, then it is natural to assume that *Jane takes a bath* and *there is no more hot water* are probabilistically independent, leading to an inference that there is (probably) no more hot water by the same reasoning as given for (17) above. For those of us who are familiar with such situations, the conditional inference feels more appropriate. The existence of such misunderstandings and disagreements is entirely expected from the current standpoint, as different speakers (and theorists) may come to a conversation with different assumptions about dependencies among propositions.

5.2 Semi-conditional presuppositions

The next question that we need to explain is why (23) (repeated from 4) seems to presuppose something that can be paraphrased as (23b) rather than (23a) or (23c).

(23) If John is a diver and wants to impress his girlfriend, he'll bring his wetsuit on vacation.

    a. ↝̸ If John is a diver and wants to impress his girlfriend, he has a wetsuit.

    b. ↝ If John is a diver, he has a wetsuit.

    c. ↝̸ John has a wetsuit.

This result follows straightforwardly from the previous discussion and a few plausible assumptions about the probabilistic dependencies that this example leads us to assume. Let p = *John is a diver*, q = *John wants to impress his girlfriend*, and $\underline{r}$ = *John has a wetsuit*. The usage condition predicted by the present theory is:

$$pr_{(p \wedge q)}(\underline{r}) \geq \theta \quad \text{or, equivalently,} \quad \frac{pr(p \wedge q \wedge \underline{r})}{pr(p \wedge q)} \geq \theta$$

(23) gets its particular effect from two qualitative assumptions about relevant probability distributions that are natural to assume here. First, p and q are presumably independent: whether John is a diver has no bearing on whether he wants to impress his girlfriend. Second, although p and $\underline{r}$ are not

independent, it is reasonable to assume that they are *jointly* independent of q: that is, whether John owns a wetsuit is related to whether or not he is a diver, but neither of these events has anything to do with his relationship with his girlfriend. Joint independence of q and $(p \wedge \underline{r})$ means that

$$pr(p \wedge \underline{r}) = \frac{pr((p \wedge \underline{r}) \wedge q)}{pr(q)} \text{ or equivalently, } pr((p \wedge \underline{r}) \wedge q) = pr(p \wedge \underline{r}) \times pr(q).$$

Using this equation and the fact that p and q are independent as well, we can rewrite the usage condition of (23) as:

$$\frac{pr(p \wedge q \wedge \underline{r})}{pr(p \wedge q)} \geq \theta \quad \text{iff} \quad \frac{pr(p \wedge \underline{r}) \times pr(q)}{pr(p) \times pr(q)} \geq \theta$$

Canceling $pr(q)$ gives us (24):

(24) $pr(\underline{r} \mid p) \geq \theta$, i.e. $pr(\text{John has a wetsuit} \mid \text{John is a diver}) \geq \theta$

(24) is the same usage condition that the theory associates with *If John is a diver, he will bring his wetsuit on vacation* in §5.1; and, for the same reasons, it is well-paraphrased by (23b). In short: as long as the assumption of joint independence is appropriate for a particular example of this type, semi-conditional presuppositions are what we expect.

5.3   Two types of conditional presuppositions

As we saw in §1, a problem for any account of the proviso problem based on strengthening is to explain why the mechanism which strengthens the presupposition of (1)/(17) (*If Theo's wife hates sonnets then his manager does too*) does not also apply to (3), repeated here as (25).

(25) Sam knows that if Theo's wife hates sonnets then he has a manager.

     a. ⤳ If Theo's wife hates sonnets then he has a manager.

     b. ⇸ Theo has a manager.

There are, I suggest, two important differences between (17) and (25), one theoretical and one empirical. Theoretically, it turns out that the probabilistic theory offered here simply *couldn't* assign the same semantic presupposition to (17) and (25), as long as we make the standard assumption that semantic presuppositions are propositions. Empirically, (17) and (25) differ in that they do not trigger the same intuitive independence assumptions: (25) strongly implies that whether Theo's wife hates sonnets is relevant to whether he has

a manager, a fact which prevents the pragmatic reasoning described above for (17) from going through regardless of what the semantic presupposition of (25) is.

Starting with the technical point, the "conditional presupposition" of (17) is a pragmatic condition requiring that a speaker who utters this sentence must believe that the conditional probability of *Theo has a manager* given *Theo's wife hates sonnets* is at least θ (and that her audience does the same). It is natural when encountering this sentence out of the blue to assume that these propositions are independent, in which case this condition is, for reasons now familiar, equivalent to the simpler condition that $pr(\textit{Theo has a manager}) \geq \theta$. (25), on the other hand, is treated in this theory (for better or worse) as an atomic sentence, and it carries a *semantic* presupposition triggered by the verb *knows* in combination with the clause that it embeds. It isn't clear what this semantic presupposition is, in part because there is no consensus about the truth-conditions of indicative conditionals. Fortunately, we don't need to know which proposition is semantically presupposed by (25) to get a sense of how and why (17) and (25) differ.

Call the proposition that (25) denotes p, and its semantic presupposition $\underline{p}$. Our usage conditions indicate that an utterance of p will be infelicitous unless $pr(\underline{p}) \geq \theta$. Whatever $\underline{p}$ is, this usage condition is equivalent to the condition that we assigned to (17) only if the semantic presupposition $\underline{p}$ is a proposition whose probability is systematically equal to the conditional probability of the consequent A of the embedded conditional, given the antecedent B. But it can't be, because it is not possible, for arbitrary A and B, to find a proposition whose unconditional probability is systematically equal to the conditional probability $pr(A|B)$.[18] It is extremely difficult, in partic-

---

18 By "systematically", I mean that the equality is non-accidental and is preserved under conditionalization; essentially, that it will continue to hold under various suppositions and updates. For reductio, fix A and B and let X be the mystery proposition whose unconditional probability is equal to the conditional probability of A given B. Since the equality is maintained under conditionalization, we can ask in particular what happens if we suppose or learn that B; $pr(A|B)$ is unchanged, but our supposition about stability under conditionalization requires that $pr(A|B)$ be equal to $pr(X|B)$, which therefore also equals $pr(X)$. This means that X and B are independent; the probability of B has no influence on the probability of X. In particular, $pr(X|\neg B)$ is also equal to $pr(X)$ and to $pr(A|B)$. For any further C, our supposition requires that $pr(X|C) = pr(A|B \wedge C)$. If we take $C = \neg B$ here, however, we have a contradiction: $pr(X|C)$ is still equal to $pr(X)$, but $pr(A|B \wedge \neg B)$ is undefined. So there can be no proposition X that systematically has probability equal to $pr(A|B)$, and in particular, as long as we assume that the semantic presuppositions of sentences like (25) are propositional in form, the usage conditions derived from them will not be equivalent to the

ular, to associate English indicative conditionals with such propositions.[19] Whatever the semantic presuppositions that are triggered by factives which embed indicative conditionals are, they cannot — on our assumptions — have a probability which is systematically equal to the probability of the conditional consequent given the antecedent. This point defuses Geurts' objection from the non-equivalence of the felt presuppositions of (17) and (25): the probabilistic account does not predict the same usage conditions for these two sentences, and so the fact that they are felt to be different is not an obvious cause for concern.

This is admittedly not a fully satisfying resolution of the problem. We know now that the theory proposed here does not (and could not) systematically assign sentences with the forms of (17) and (25) the same usage conditions, but we do not know what usage conditions it *does* predict for (25). While I do not have a complete answer to offer here, there is an empirical difference between these examples that is important here: to my ear, at least, (17) and (25) do not bring to mind the same independence assumptions. That is, someone who utters (25) would normally be taken to indicate that they believe that there *is* a relevant connection between whether Theo's wife hates sonnets and whether he has a manager. This is, presumably, related to the fact that the conditional sentence also gives rise to this pragmatic inference when it is not embedded in a factive, as in (26).

ones that (I argued above) sentences like (17) receive. Note, by the way, that it may be possible to avoid this issue by denying that semantic presuppositions have to be propositions, cf. Yalcin 2011.

19 See Hájek & Hall 1994, Lewis 1976 among others, who show that on standard assumptions a conditional with the requisite properties cannot be defined in a non-trivial probability space. To my knowledge, the only proposal that avoids the triviality results without denying either that indicative conditionals denote propositions or that *if* is a connective is Rothschild's (2010) trivalent theory. Theories which deny propositional status to indicative conditionals include Edgington 1986, 1995, Kaufmann 2001, 2005, 2009, Stalnaker & Jeffrey 1994. If combined with a theory of presupposition that can make sense of non-propositional semantic presuppositions, these analyses might be able to revive Geurts' objection as applied to my probabilistic account of presupposition; however, the empirical point involving intuitions about independence discussed just below would still hold.

Another possibility is to analyze conditionals not as connectives but as devices of domain restriction affecting the interpretation of overt or covert epistemic modals (Egré & Cozic 2011, Kratzer 1986). The latter approach is the most popular semantics for conditionals among linguists, but note that it is really a change of subject with respect to the question at hand: the restrictor account does not tell us what the probabilities of sentences expressing indicative conditionals are, but how the *truth-conditions* of such sentences depend on certain conditional probabilities.

(26) a. If Theo's wife hates sonnets then he has a manager.
⤳ The speaker believes that the issue of whether Theo's wife hates sonnets is relevant to the issue of whether he has a manager.

b. Sam knows that if Theo's wife hates sonnets then he has a manager.
⤳ Same inference as (26a), plus an inference that Sam believes this too.

In fact implicatures frequently survive embedding in factives, and are taken to indicate the shared beliefs of the speaker and matrix subject. For example:

(27) a. Jane is annoyed that you ate some of her cookies.
⤳ Speaker and Jane both believe that you didn't eat all of her cookies.

b. Bill's car has broken down, but he realizes that a gas station is nearby.
⤳ Speaker and Bill both believe that the gas station is likely to be useful to Bill in resolving his predicament (i.e. is open, has gas, etc.).

There are interesting issues around where this inference comes from and why it remains in embeddings, but we don't need a complete theory for present purposes. What is important is simply that, since (26) strongly suggests that the antecedent and the consequent of the embedded conditional are not independent, the explanation given above for the felt presupposition of (17) would not apply even if the theory *did* generate the same usage conditions for both: that explanation relied crucially on an assumption of probabilistic independence which is not appropriate in this example.[20]

---

20 A further objection to satisfaction theories due to Geurts (1996) is the fact that sentences like (28) can be read as implying that the presupposition of the consequent is true. The theory I have given, like other satisfaction theories, predicts only a trivial presupposition.

(28) If all the boys left, then the janitor won't have noticed that Fred left.

a. OK: ⤳ If all the boys left, Fred left. (= ⊤, if Fred is one of the boys)
b. OK: ⤳ Fred left.

Geurts (1996: 286) argues that such sentences "can be read either as presupposing or as not presupposing that [the semantic presupposition of the consequent is true], and the satisfaction theory accounts only for the latter possibility". This argument is somewhat tendentious, though: nothing about the example forces us to conclude that this (rather weak) inference is presuppositional in nature. My suspicion is that the inference in (28b), when it arises, is a pragmatic inference with a different source (essentially as van Rooij (2007: fn.8) argues).

6   Unconditional inferences without independence

Schlenker (2011), following unpublished work by Raj Singh, suggests an explanation of the proviso problem which was a source of inspiration for the present account. His theory also makes use of probabilistic independence, although the independence condition is presented as a separate layer on top of a theory which treat information states as sets of worlds and which generates and selects among multiple "potential" presuppositions. In addition to avoiding *ad hoc* mechanisms of this sort, the theory proposed here derives support from a new empirical observation: as I will show, there are examples with the form of (16) and (17) in which the felt presupposition is unconditional even though the crucial independence assumption is clearly not appropriate. It is not clear how to deal with these cases in previous independence-based strengthening accounts, but it is possible to account for them within the present theory.

Consider a conditional *if p then q*, where q has semantic presupposition $\underline{q}$. If p and $\underline{q}$ are not probabilistically independent, then there are two possibilities: either $pr(\underline{q}|p) > pr(\underline{q})$ or $pr(\underline{q}|p) < pr(\underline{q})$. Interestingly, all of the examples that we have considered where independence is not appropriate — and most of the ones that appear in the literature on this topic — are of the former type: knowing that p is true will tend to render $\underline{q}$ more likely. A crucial step in the reasoning was that the stronger epistemic condition $pr(\underline{q}) \geq \theta$ is not licensed because the argument in (20) is not sound (the second premise is false). However, the argument is also not sound in cases of non-independence in which $pr(\underline{q}|p)$ is *less* than $pr(\underline{q})$, and so we might expect to find usage conditions that are best rendered in English as conditionals. (28) suggests that this expectation is not borne out, though.

(28)   If Sam is begging in the streets, he ought to sell his mansion.

    a.   ↝̸ If Sam is begging in the streets, he has a mansion.

    b.   ↝ Sam has a mansion.

It seems unlikely that p and $\underline{q}$ are independent here: instead, the probability that Sam has a mansion is presumably much reduced if we assume that he is a beggar, and so $pr(\underline{q}|p) < pr(\underline{q})$. Nevertheless, the most natural paraphrase of the presupposition of (28) is (28b). A similar example is (29).

(29)   If the grass has not been mowed in months, Bill's gardener will do it
       soon.

    a.   ↝̸ If the grass has not been mowed in months, Bill has a gardener.

b. ⤳ Bill has a gardener.

These examples are problematic for a theory which relies on a strengthening mechanism triggered by probabilistic independence, since the mechanism should not be operative in this case. However, on the present theory we have an explanation: (30a) and (30b) together entail (30c).

(30) a. $pr(\underline{q}|p) \geq \theta$

b. $pr(\underline{q}) > pr(\underline{q}|p)$

c. ∴ $pr(\underline{q}) \geq \theta$

(30c) represents the same usage condition that we would associate with the simple *Sam ought to sell his mansion* and *Bill's gardener will mow the grass soon.*

As long as we are in an epistemic state which licenses the second premise of (30), we can account for the appearance of unconditional inferences in these examples in the same way that we did for examples in which p and q are independent.[21] As in §4, the story goes roughly: multiple English sentences must receive high probability if the usage conditions are fulfilled, and we prefer logically stronger renditions as long as they do not lead to misleading secondary inferences. In (28) and (29) the preferred (b) renditions are indeed logically stronger. The (a) examples also give rise to misleading implicatures here, though for a different reason than in the case that we saw earlier (17). There the problem was that the conditional paraphrase gave rise to a misleading relevance implicature. Here, relevance holds but the use of a conditional leads to a different undesirable implicature: conditional perfection. That is, the conditional paraphrases (28a) and (29a) naturally lead to an inference that the consequent fails if the antecedent does, much as (31a) implicates (31b) (see e.g. Geis & Zwicky 1971, Horn 1972, 2000).

(31) a. If you mow the lawn, I'll give you $5.

b. If you don't mow the lawn, I won't give you $5.

Nothing like this inference is associated with the probabilistic presupposition $pr(\underline{q}|p) \geq \theta$. Here again, the possible paraphrase in terms of a conditional

---

21 I am making the non-trivial assumption here that (30b) is a piece of qualitative knowledge about probabilities of events that we possess as part of our understanding of (in the case of (28)) wealth and poverty. It would be interesting to see whether unconditional inferences would arise if we could find parallel cases in which (30b) holds systematically but accidentally; I don't know of any clear examples.

sentence fails not only because it is logically weaker than an available alternative, but also because it introduces extraneous inferences which are not appropriate.

An apparent problem here is that this line of reasoning would seem to predict that any English sentence which denotes a proposition known to have greater probability than $pr(q|p)$ will be a good candidate for the felt presupposition of (28) and (29): both of them would then include among the pragmatic presuppositions that they evoke trivialities such as "I am now breathing" and "Paris is the capital of France". I don't know whether this prediction is so bad, but there is a way to manage it if so. The proposition that Paris is the capital of France, though it is indeed highly probable in the epistemic state of any moderately informed individual who utters (28) in compliance with the usage conditions given in §4.1, has this probability regardless of whether these usage conditions are fulfilled. The proposition that Sam has a mansion, on the other hand, cannot fail to have high probability in any epistemic state in which the usage conditions of this sentence and the qualitative constraints that we are assuming are both fulfilled. We may suppose, then, that trivial inferences of this type are ignored because they would hold whether or not the usage conditions associated with the sentence were true.

7 Further issues

This section deals with a few additional points that seem particularly pressing for the theory proposed here. There are of course many more detailed issues that I am not able to address here.[22]

7.1 Counterfactuals

An apparent problem for the theory proposed here is that the proviso problem also arises in counterfactuals.

---

[22] In particular, I do not have anything very illuminating to say about presuppositions in attitude contexts (see e.g. Geurts 1998, Heim 1992), beyond the general observation that my theory predicts (correctly) that defeasible assumptions about agents' competence on specific topics should be relevant to whether a presupposition projects beyond an attitude verb. Many of the detailed problems discussed in the literature on this topic depend heavily on assumptions about the semantics of attitude verbs about which I have grave doubts on independent grounds (see Lassiter 2011).

(32) If Sam were a diver, he would bring his wetsuit on vacation.

    a. ⤳ If Sam were a diver, he would have a wetsuit.

    b. ⤱ Sam has a wetsuit.

(33) If Sam were kind, he would bring his wife on vacation.

    a. ⤱ If Sam were kind, he would have a wife.

    b. ⤳ Sam has a wife.

Standard accounts of counterfactuals do not have any mechanism for determining probabilities in counterfactual contexts, and so it is not clear how to apply the probabilistic model in these cases.

Rather than being a problem for the probabilistic account of presuppositions, though, I submit that this is a problem for standard theories of counterfactuals: these theories owe us an account of counterfactual probabilities together with a set of rules for determining these from ordinary probabilities and the information in the antecedent of the counterfactual. One independent reason to think this is that epistemic modals can occur in counterfactuals, including gradable epistemic modals of the type that motivated Lassiter (2010, 2011), Yalcin (2010) to posit a probabilistic semantics for epistemic modals.

(34) If it had rained last night, the grass would possibly/probably/very likely/more likely than not/almost certainly have gotten wet.

If *possibly*, *very likely*, etc. are operators which place conditions on probability measures as these authors argue, a treatment of (34) would presumably require some sort of counterfactual probabilities.

Even though theories of counterfactuals popular in linguistic semantics do not give us any way to make sense of this idea, there is a well-developed formal semantics for counterfactuals which does — Pearl 2000. This theory has been extremely influential in philosophy, psychology, computer science, and beyond, but its impact has not been great in linguistic semantics (though Kaufmann (2005), Schulz (2007, 2011) do draw inspiration from Pearl's work). Very briefly, the idea is that counterfactuals are evaluated by modifying a graphical model of probabilistic dependencies to ensure that the antecedent is true, and redistributing probabilities in a way consistent with this. If Pearl's or some other probabilistic account of counterfactuals is viable, we have a straightforward line of attack on sentences with epistemic modals in counterfactuals as in (34). I suspect that such a treatment would also allow

us to treat instances of the proviso problem in counterfactual contexts such as (32)-(33) exactly as we did their indicative counterparts above: (32) and (33) differ in whether the antecedent and the semantic presupposition of the consequent are dependent, with concomitant effects on the probabilistic effects that follow from the pragmatic presupposition.[23]

7.2   Global accommodation

An important issue that I have said little about is global accommodation. How is it that a speaker can sometimes appropriately use a sentence whose usage constraint $pr(\underline{q}) \geq \theta$ is satisfied in her personal probability distribution, despite knowing that her listener's epistemic state does not satisfy this constraint? According to Lewis (1979), principles of charitable interpretation lead listeners to accommodate presuppositions automatically when they are not already common ground: "straightaway that presupposition springs into existence, making what you have said acceptable after all". Beaver & Zeevat (2007) point out that this formulation may be too permissive, suggesting as it does that accommodation is in general easy and free; some presuppositions are clearly more difficult to accommodate than others. Borrowing Beaver & Zeevat's example, a reader of a novel would presumably not balk at (35) even if the author has not said anything which entails its factive presupposition, as long as there is nothing in the context which renders this presupposition implausible given some reasonable resolution of the pronoun "they".

(35)   I knew they would show no mercy.

   The intriguing fact about this example is how easy it is to make accommodation less acceptable, or completely unacceptable, with minimal

---

23 As Schlenker (2011) notes, the proviso problem also arises in quantified sentences such as (35).

(35)   If I grade their homeworks, few of my students will realize that they are incompetent.

   a.   ⇸ If I grade their homeworks, all of my students are incompetent.
   b.   ⤳ All of my students are incompetent.

It is not obvious how to account for such examples, since we would seem to need some way to assign probabilities to open sentences. One possibility — which I will only sketch briefly here — would be to generalize the theory to treat not just probabilities but more generally the *expected values* of functions, of which probability is a special case when the function is of type $\langle s, t \rangle$. Assuming that *I grade x's homework* and *x is incompetent* are independent for each x, this account would predict the presupposition "Almost all of my students are incompetent" for (35), which is slightly weaker than (35b) but fairly plausible, I think.

modifications to the information contained in the context. If (35) is embedded in a context like (36), accommodation is so easy that the presupposition is hardly noticeable.

(36)   About a dozen men in dark cloaks were approaching, carrying swords. I knew they would show no mercy.

In a context like (37), however, the presupposition is harder to swallow, and (38) is just bizarre.

(37)   About a dozen men in suits were approaching, carrying briefcases. I knew they would show no mercy.

(38)   About a dozen small children were approaching, carrying flowers. I knew they would show no mercy.

The move toward probabilistically structured information states makes available to us a well-motivated set of tools from information theory for reasoning precisely about the dynamics of information (Cover & Thomas 1991, MacKay 2003). I suggest that we can, at least as a first approximation, quantify the influence of prior knowledge on the availability of presupposition accommodation using a measure of information content known as *surprisal*:

(39)   The surprisal $I(\phi)$ of a proposition ϕ relative to a probability measure *pr* is defined as

$$I(\phi) = \log_2 \frac{1}{pr(\phi)} = -\log_2 pr(\phi).$$

As you might expect, the surprisal of ϕ under *pr* is a measure of how surprised someone whose information state is given by *pr* would be to learn that ϕ is true. $I(\phi)$ is zero if $pr(\phi)$ is 1 and increases as the probability of ϕ decreases, approaching ∞ as $pr(\phi)$ goes to 0. Surprisal has the right form for a measure of the difficulty of accommodating ϕ: accommodation is free if $pr(\phi) = 1$, impossible if $pr(\phi) = 0$, and harder for ψ than for ϕ if $pr(\psi) < pr(\phi)$. The latter feature explains the contrast in (36)-(38): the context set up by the story makes the crucial presupposition much less likely in (37) than in (36), and even less so in (38). The less likely a presupposition is in context, the higher its surprisal and the less available it is for accommodation.[24]

----

24 On an intuitive level, surprisal is not too different from the measure of the cost of accommodation in terms of unexpectedness suggested by Beaver (2001: 269): "I have sometimes surmised that this cost might be measured in millimetres, a cost of, e.g., 2 mm. correponding to a surprisingness which would cause raising of the eyebrows by this amount".

This is not yet the full story, to be sure. In some contexts accommodation seems to be easy even if the accommodated material does not have high prior probability, as long as it is relatively unimportant or uncontroversial: for instance, the presupposition of *my pet ferret* could be easy to accommodate even though few people keep ferrets as pets, as long as this point is not contentious or directly relevant to the topic of conversation. To account for the fact that presuppositions can be uncontroversial even when they do not have high prior probability we would need to move beyond a simple probabilistic model to a model that incorporates information about the speaker's and listener's preferences and goals as well as the Question Under Discussion (Roberts 1996) and other information about the structure of the conversation. As van Rooij (2003, 2004) shows, information-theoretic reasoning of the type that we have just been engaged in can be seen as a special case of decision-theoretic reasoning when agents are indifferent among the various possibilities. I suspect that this sort of enriched account may be able to deal with the clear counter-examples to the measure of ease of accommodation suggested here, but will leave a detailed exploration of the issue to future work.

7.3 Local accommodation

One important issue that I have not addressed involves cases in which presuppositions disappear unexpectedly, for example:

(40) My pet ferret is not at the vet's — I don't have a pet ferret.

(41) [Sign posted at a store entrance] You must put out your cigarette before entering.

Examples like this are problematic for many theories, including most varieties of dynamic semantics. There are numerous possible accounts, and I do not think that we are necessarily in worse shape than other satisfaction theories in this respect. One possibility is that a probabilistic implementation of a pragmatic derivation of the dynamics of connectives along the lines of Schlenker (2009) might be able to exhibit sufficient sensitivity to global facts about the discourse goals of interlocutors to make this option available in some cases. However, it remains to be seen what such a theory would look like, and local accommodation remains as an important challenge for my account just as for other satisfaction theories.

8   Conclusion

The proviso problem has been taken to be a serious objection to satisfaction theories of presupposition. In response satisfaction theorists have proposed a number of additional mechanisms in order to account for the apparent fact that conditional presuppositions only arise when the antecedent is relevant to the consequent, and unconditional presuppositions arise otherwise. However, these mechanisms have often been stipulative, and the fact that DRT predicts a preference for global accommodation of presuppositions on independent grounds has been seen as an important point in its favor, as Geurts (1996) argues.

A probabilistic account of presupposition of the type proposed here is technically close to previous satisfaction theories but it is able to predict the contrast from (1-2) without adding any special-purpose machinery to the basic theory of presupposition. The proposed derivation of usage conditions for complex sentences from the semantic presuppositions of their atomic parts yields conditional probability statements which, if certain independence relations hold, are equivalent to other, unconditional probability statements. This approach makes the novel and correct empirical prediction that conditional sentences should give rise to unconditional presuppositions not only when the antecedent and the presupposition of the consequent are probabilistically independent, but also when the conditional probability of the consequent presupposition given the antecedent is *less* than its unconditional probability. In addition to improving on previous satisfaction-based accounts of the proviso problem by avoiding stipulative conditional strengthening mechanisms, then, the present theory has improved empirical coverage.

The account generalizes readily to semi-conditional presuppositions like (4), which are quite difficult to account for within either previous satisfaction theories or DRT. The probabilistic theory also offers an explanation of the divergence between conditional presuppositions and presupposed conditionals illustrated by (3) and suggests promising new lines of attack on several other difficult problems such as instances of the proviso problem in counterfactuals.

References

Beaver, David. 1999. Presupposition accommodation: A plea for common sense. In Lawrence S. Moss, Jonathan Ginzburg & Maarten de Rijke (eds.),

Logic, language and computation, vol. 2, 21-44. CSLI. https://webspace.utexas.edu/dib97/itallc.pdf.

Beaver, David. 2001. *Presupposition and assertion in dynamic semantics*. CSLI.

Beaver, David & Henk Zeevat. 2007. Accommodation. In Gillian Ramchand & Charles Reiss (eds.), *Oxford handbook of linguistic interfaces*. Oxford University Press.

Chater, Nick, Joshua B. Tenenbaum & Alan Yuille. 2006. Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences* 10(7). 287-291. http://dx.doi.org/10.1016/j.tics.2006.05.007.

Clark, Herbert H. & Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (eds.), *Elements of discourse understanding*, 10-63. Cambridge University Press.

Cover, Thomas M. & Joy A. Thomas. 1991. *Elements of information theory*. Wiley.

Cresswell, M.J. 2002. Static semantics for dynamic discourse. *Linguistics and Philosophy* 25(5). 545-571. http://dx.doi.org/10.1023/A:1020834910542.

Dekker, Paul J. E. 2010. There is something about *might*. In Maria Aloni, Harald Bastiaanse, Titiku de Jager & Katrin Schulz (eds.), *Logic, language and meaning: 17th Amsterdam Colloquium, Amsterdam, The Netherlands, December 16-18, 2009, revised selected papers*, 244-253. Springer. http://dx.doi.org/10.1007/978-3-642-14287-1_25.

Edgington, Dorothy. 1986. Do conditionals have truth conditions? *Crítica: Revista Hispanoamericana de Filosofía* 18(52). 3-39. http://www.jstor.org/pss/40104380.

Edgington, Dorothy. 1995. On conditionals. *Mind* 104(414). 235. http://dx.doi.org/10.1093/mind/104.414.235.

Egré, Paul & Mikaël Cozic. 2011. If-clauses and probability operators. *Topoi*. 1-13. http://dx.doi.org/10.1007/s11245-010-9087-y.

van Fraassen, Bas C. 1980. Review of Brian Ellis, *Rational Belief Systems*. *Canadian Journal of Philosophy* 10(3). 497-511.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition, and logical form*. Academic Press.

Geis, Michael L. & Arnold M. Zwicky. 1971. On invited inferences. *Linguistic Inquiry* 2(4). 561-566. http://www.jstor.org/stable/4177664.

Geurts, Bart. 1996. Local satisfaction guaranteed: A presupposition theory and its problems. *Linguistics and Philosophy* 19(3). 259-294. http://dx.doi.org/10.1007/BF00628201.

Geurts, Bart. 1998. Presuppositions and anaphors in attitude contexts. *Linguistics and Philosophy* 21(6). 545–601. http://dx.doi.org/10.1023/A:1005481821597.

Geurts, Bart. 1999. *Presuppositions and pronouns.* Elsevier.

Gigerenzer, Gerd. 1991. How to make cognitive illusions disappear: Beyond "heuristics and biases". *European Review of Social Psychology* 2(1). 83–115. http://dx.doi.org/10.1080/14792779143000033.

Grice, H. Paul. 1989. *Studies in the ways of words.* Harvard University Press.

Griffiths, Thomas L., Charles Kemp & Joshua B. Tenenbaum. 2008. Bayesian models of cognition. In R. Sun (ed.), *Cambridge handbook of computational psychology*, 59–100. Cambridge University Press.

Hájek, Alan & Ned Hall. 1994. The hypothesis of the conditional construal of conditional probability. In Ellery Eels & Brian Skyrms (eds.), *Probability and conditionals: Belief revision and rational decision*, 75–112. Cambridge University Press.

Heim, Irene. 1983. On the projection problem for presuppositions. *Proceedings of the Second West Coast Conference on Formal Linguistics.* 114–125.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221. http://dx.doi.org/10.1093/jos/9.3.183.

Hilton, Denis J. 1995. The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin* 118(2). 248–271. http://dx.doi.org/10.1037/0033-2909.118.2.248.

Horn, Laurence R. 1972. *On the semantic properties of logical operators in English.* UCLA PhD thesis.

Horn, Laurence R. 2000. From if to iff: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics* 32(3). 289–326. http://dx.doi.org/10.1016/S0378-2166(99)00053-3.

Kahneman, Daniel, Paul Slovic & Amos Tversky. 1982. *Judgment under uncertainty: Heuristics and biases.* Cambridge University Press.

Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1(1-3). 181–194.

Kaufmann, Stefan. 2001. Probabilities of conditionals. *Semantics and Linguistic Theory (SALT)* 11. 248–267. http://www.elanguage.net/journals/index.php/salt/article/view/11.248/1711.

Kaufmann, Stefan. 2005. Conditional predictions: A probabilistic account. *Linguistics and Philosophy* 28(2). 181–231. http://dx.doi.org/10.1007/s10988-005-3731-9.

Kaufmann, Stefan. 2009. Conditionals right and left: Probabilities for the whole family. *Journal of Philosophical Logic* 38(1). 1–53. http://dx.doi.org/10.1007/s10992-008-9088-0.

Klinedinst, Nathan & Daniel Rothschild. 2012. Connectives without truth-tables. *Natural Language Semantics.* http://dx.doi.org/10.1007/s11050-011-9079-5.

Koller, Daphne & Nir Friedman. 2009. *Probabilistic graphical models: Principles and techniques.* The MIT Press.

Kratzer, Angelika. 1986. Conditionals. *Chicago Linguistics Society* 22(2). 1–15.

Lassiter, Daniel. 2010. Gradable epistemic modals, probability, and scale structure. *Semantics and Linguistic Theory (SALT) 20.* 197–215. http://elanguage.net/journals/index.php/salt/article/view/20.197.

Lassiter, Daniel. 2011. *Measurement and modality: The scalar basis of modal semantics.* New York University PhD thesis. http://www.stanford.edu/~danlass/Lassiter-diss-Measurement-Modality.pdf.

Lewis, David. 1976. Probabilities of conditionals and conditional probabilities. *Philosophical Review* 85(3). 297–315. http://dx.doi.org/10.2307/2184045.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8(1). 339–359. http://dx.doi.org/10.1007/BF00258436.

MacKay, David J.C. 2003. *Information theory, inference, and learning algorithms.* Cambridge University Press.

Milch, Brian, Bhaskara Marthi, Stuart Russell, David Sontag, Daniel L. Ong & Andrey Kolobov. 2007. BLOG: Probabilistic models with unknown objects. In Lise Getoor & Ben Taskar (eds.), *Introduction to statistical relational learning,* 373–398. The MIT Press.

Pearl, Judea. 1988. *Probabilistic reasoning in intelligent systems: Networks of plausible inference.* Morgan Kaufmann.

Pearl, Judea. 2000. *Causality: Models, reasoning and inference.* Cambridge University Press.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics.* 91–136. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.8440&rep=rep1&type=pdf.

Romoli, Jacopo. 2011. The presuppositions of soft triggers aren't presuppositions. *Semantics and Linguistic Theory (SALT)* 21. 236–256. http://elanguage.net/journals/salt/article/view/21.236.

van Rooij, Robert. 2003. Questioning to resolve decision problems. *Linguistics and Philosophy* 26(6). 727–763. http://dx.doi.org/10.1023/B:LING.0000004548.98658.8f.

van Rooij, Robert. 2004. Utility, informativity and protocols. *Journal of Philosophical Logic* 33(4). 389–419. http://dx.doi.org/10.1023/B:LOGI.0000036830.62877.ee.

van Rooij, Robert. 2007. Strengthening conditional presuppositions. *Journal of Semantics* 24(3). 289–304. http://dx.doi.org/10.1093/jos/ffm007.

Rothschild, Daniel. 2010. Do indicative conditionals express propositions? *Noûs*. http://dx.doi.org/10.1111/j.1468-0068.2010.00825.x.

Rothschild, Daniel. 2011. Explaining presupposition projection with dynamic semantics. *Semantics and Pragmatics* 4(3). 1–43. http://dx.doi.org/10.3765/sp.4.3.

Roussarie, Laurent. 2009. What might be known: Epistemic modality and uncertain contexts. *Semantics and Linguistic Theory (SALT) 19*. 379–394. http://www.elanguage.net/journals/index.php/salt/article/view/19.22.

Russell, Stuart & Peter Norvig. 2010. *Artificial intelligence: A modern approach*. Prentice Hall.

van der Sandt, Rob A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9(4). 333–377. http://dx.doi.org/10.1093/jos/9.4.333.

Schlenker, Philippe. 2009. Local contexts. *Semantics and Pragmatics* 2(3). 1–78. http://dx.doi.org/10.3765/sp.2.3.

Schlenker, Philippe. 2011. The proviso problem: A note. *Natural Language Semantics*. 19(4). 395–422. http://dx.doi.org/10.1007/s11050-011-9072-z.

Schulz, Katrin. 2007. *Minimal models in semantics and pragmatics: Free choice, exhaustivity, and conditionals*. Institute for Logic, Language & Computation, University of Amsterdam PhD thesis. http://dare.uva.nl/document/53840.

Schulz, Katrin. 2011. "If you'd wiggled A, then B would've changed": Causality and counterfactual conditionals. *Synthese* 179(2). 239–251. http://dx.doi.org/10.1007/s11229-010-9780-9.

Simons, Mandy. 2001. On the conversational basis of some presuppositions. *Semantics and Linguistic Theory (SALT) 11*. 431–448. http://www.hss.cmu.edu/philosophy/simons/simons.SALT2001.pdf.

Singh, Raj. 2007. Formal alternatives as a solution to the proviso problem. *Semantics and Linguistic Theory (SALT) 17*. http://elanguage.net/journals/index.php/salt/article/view/17.264/1854.

Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophical Logic* 2(4). 447–457. http://dx.doi.org/10.1007/BF00262951.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton K. Munitz & Peter Unger (eds.), *Semantics and philosophy*, 197–213. New York: New York University Press.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics 9: Pragmatics*. Academic Press.

Stalnaker, Robert & Richard Jeffrey. 1994. Conditionals as random variables. In Ellery Eels & Brian Skyrms (eds.), *Probability and conditionals: Belief revision and rational decision*, 31–46. Cambridge University Press.

Swanson, Eric. 2006. *Interactions with context*. MIT PhD thesis. http://dx.doi.org/1721.1/37356.

Tenenbaum, Joshua B. & Thomas L. Griffiths. 2001. The rational basis of representativeness. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*. 1036–1041. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.27.5273&rep=rep1&type=pdf.

Veltman, F. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25(3). 221–261. http://dx.doi.org/10.1007/BF00248150.

Weirich, Paul. 1983. Conditional probabilities and probabilities given knowledge of a condition. *Philosophy of Science*. 82–95. http://dx.doi.org/10.1086/289091.

Yalcin, Seth. 2005. Epistemic modals. *MIT Working Papers in Linguistics* 51.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. http://dx.doi.org/10.1093/mind/fzm983.

Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11). 916–937. http://dx.doi.org/10.1111/j.1747-9991.2010.00360.x.

Yalcin, Seth. 2011. Nonfactualism about epistemic modality. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*. Oxford University Press.

Daniel Lassiter
Computation & Cognition Lab
Department of Psychology
Stanford University
420 Jordan Hall, 450 Serra Mall
Stanford, CA 94305, USA
danlassiter@stanford.edu