

The Varieties of Relevance: Over-informative answers and the case for keeping relevance strong*

Ido Benbaji-Elhadad

Omri Doron

Massachusetts Institute of Technology University of Massachusetts Amherst

Abstract We offer a conceptual evaluation of the different relevance constraints argued in the formal pragmatic literature to regulate the admissibility of utterances in discourse. We distinguish strong *negative* formulations, which prohibit speakers from conveying certain kinds of information, from weaker *positive* ones, which require that speakers convey a certain minimal amount of information. We then illustrate that, contra what is often implicitly assumed, both formulations can be cast in probabilistic and in non-probabilistic terms. While each approach suffers from an over- or under-generation problem, we show that the over-generation challenge is more detrimental to the positive approaches than the under-generation challenge is to the negative ones. We thus propose a modification of negative relevance that overcomes its under-generation problem by having the admissibility of an utterance evaluated “dynamically” relative to a particular chunk of hypothesized discourse, rather than to a single question under discussion. We conclude by debating the virtues and vices of introducing probabilities into the mix, offering some new considerations against it.

1 Introduction

What are the formal constraints on meanings of utterances in discourse? A common assumption in the formal pragmatic literature is that these constraints pertain to the logical relation between the meaning of an utterance and two hypothesized parameters: a set of worlds reflecting the shared background knowledge of discourse participants, termed *the context set* (Stalnaker 1977 and onward), and a partition of this set into equivalence classes, or *cells*, reflecting the properties of the world that discourse participants are interested in determining, termed *the subject matter, issue, or question under discussion* (Groenendijk & Stokhof 1984; Lewis 1988a,b; Roberts 1996/2012). In particular, utterances should “add” information to the context set (namely, eliminate worlds from it) and should somehow “address” the conversational subject matter.

Pretty much everyone subscribing to this framework seems to agree on the general picture. However, when it comes to what it means to “address” an issue, things get hairy. Many linguists and philosophers working on discourse dynamics, have initially taken a restrictive view, due to Lewis (1988a,b), that a proposition p must be entailed or contradicted by each cell in a given issue for it to count as “addressing” it. On this view, the constraint – often termed *Relevance* – boils down to a “purely negative requirement. . . it is just the absence of any irrelevant (i.e., not specifically asked for) information” (von Fintel & Heim 2021: 152): p must not convey any information that distinguishes between worlds in the same cell.

But only a sliver of utterances which speakers judge to be felicitous seem to be relevant in this restrictive sense. Based on this apparent under-generation problem, Roberts (1996/2012) (building on Groenendijk & Stokhof (1984)) proposed a *positive* relevance requirement, demanding

* This paper has been resubmitted to *Journal of Semantics*. Both authors contributed equally.

that a proposition eliminate at least one cell from the issue relative to which it was proffered, allowing it to distinguish between worlds in the cells with which it is compatible. This conception of relevance is logically independent from the Lewisian, negative one. However, in combination with the independently-motivated requirement that utterances eliminate worlds from the context set (sometimes called *Informativity*), the negative approach is strictly stronger than the positive one. Although they both end up demanding that utterances eliminate at least one *cell* from their respective issues, the positive view takes it to be a sufficient condition for an utterance to be relevant, while the negative view demands in addition that the proposition should not cut through cells.

But cell-elimination itself turned out to suffer from under-generation. This motivated a shift towards an even *weaker* positive formulation of relevance, which makes use of probabilities. In the context of scientific methodology, Carnap (1950) famously defined relevance of information i to an entertained hypothesis h , as the extent to which i raises our “degree of confirmation” of h . Thus, for Carnap, “addressing” the issue raised by h (namely, the question of “*whether h?*”) was, essentially, a matter of degree. Adapting this approach to theorizing about language, Büring (2003) (a.o.) suggests that a proposition p be considered relevant relative to an issue if it is assigned a (non-negligible) positive value by Carnap’s metric relative to one of the issue’s cells (see also: Russell 2012; Feinmann 2020). On this conception, of course, relevance consists of the positive requirement that an utterance shift the probability distribution over cells in a given partition; nothing prevents utterances from making any finer-grained distinctions between worlds in the same cell.

It is often (at least implicitly) assumed that the negative Lewisian approach is at odds with probabilistic relevance, because the latter is *necessarily* positive. This, we intend to show in the first part of this paper, is a misconception. We argue that, once we take a different approach to the introduction of probability, particularly *vis-a-vis* the makeup of propositions, we can reproduce the distinction between positive and negative conceptions for *probabilistic* relevance as well. That is, relevance may vary along two dimensions that are, in principle, orthogonal to each other; positive *vs* negative, and probabilistic *vs* non-probabilistic.

In the second part of the paper, we then defend the negative approach. First, we challenge the methodology often used to motivate the weakening of relevance to begin with. In particular, we show that even the most permissive positive formulation on the market suffers from some degree of under-generation and necessitates an auxiliary mechanism to capture the complexity of the empirical landscape; namely, what is often called *QUD-accommodation*. We argue that the availability of this mechanism heavily confounds the standard empirical arguments in the debate over relevance. Without controlling for QUD-accommodation, one cannot interpret examples of question-answer pairs as supporting one formulation over the other. We therefore propose a novel way to control for QUD accommodation, building on Doron & Wehbe 2022.

We then move on to the core of our argument in favor of negative relevance, which is based on the distribution of *over-informative* sentences; i.e., declarative utterances that eliminate a cell from their corresponding issue, but cut through another cell. These sentences are useful in teasing apart positive from negative relevance, since they are predicted to be relevant by the former approach but not by the latter. The empirical picture that arises turns out to be problematic for both approaches – while some over-informative sentences are judged felicitous, others are judged infelicitous. This means that the negative approach to relevance suffers from an under-generation problem, while

the positive one suffers from an over-generation problem. We argue, however, that the negative approach can overcome this challenge upon a particular modification, which is not available for the positive one. The proposed modification, which we term *dynamic relevance*, has interesting implications for the nature of the process by which we evaluate relevance.

Our arguments can be taken to tease apart positive and negative approaches, while remaining neutral about the introduction of probabilities. However, in the third and final part of this paper we discuss some conceptual considerations and tentative empirical arguments in favor of a non-probabilistic Lewisian approach.

The paper has another, more general goal; namely, to explicitly spell out the foundational assumptions underlying different formulation of relevance, particularly the probabilistic ones. Our hope is that by doing so we could encourage a more direct engagement between proponents and opponents of the probabilistic approach, since the introduction of probabilities into our definition of relevance has been relatively under-debated in recent years, especially given that a consensus on this issue has eluded the field thus far.

As mentioned, the paper is organized into three parts. The first is devoted to illustrating that there is a coherent notion of relevance that is negative, yet probabilistic. To that end, Section 2 introduces the negative vs positive distinction, and the data behind the shift in the literature, from the most restrictive negative formulation, towards the most permissive positive one. Section 3 then makes explicit some implicit assumptions behind the introduction of probabilistic notions of relevance, and illustrates how, on a particular view of the role of probabilities, *negative* probabilistic relevance is possible. In the second part of the paper, we set aside the question of introducing probabilities, concluding that it is orthogonal to the positive vs. negative debate, and turn to our general case in favor of negative relevance: Section 4 provides a conceptual argument against the methodological tools underlying the rejection of negative approaches. In Section 5, we present data that is over-generated by positive approaches and under-generated by negative ones, ruling out potential solutions for both in Section 6. Finally, we propose our own modification of negative relevance that evades its under-generation problem in Section 7. Having defended a negative conception of relevance, we return in the third part of the paper, in Section 8, to the debate regarding the introduction of probabilities, citing some potential arguments against it. Section 9 concludes.

2 The gradual weakening of *Relevance*

Utterances, as most famously articulated by Stalnaker (1978), are made with respect to a *common ground*, defined as a set of propositions that discourse participants take to be shared knowledge among themselves. This set in turn derives the conversational *context set* c , defined in (1), that constrains which declarative utterances are admissible in a given discourse. For instance, the infelicity of the discourse in (2) is attributed by Stalnaker (1978) to a *non-redundancy* or *informativity* condition, preventing speakers from uttering what is already entailed by the context set (3).

- (1) $c = \bigcap \{p \mid p \text{ is assumed to be shared knowledge by discourse participants}\} = \{w \mid \text{what discourse participants take for granted in } w\}$
- (2) As Jane and Sue watch TV, Jane suddenly says: #*The TV is on*.

- (3) **Informativity (INF):** A declarative utterance ϕ is felicitous given c only if $\llbracket \phi \rrbracket \cap c \neq \emptyset$.

There is, however, another form of infelicity – exemplified in (4) – that cannot be attributed to INF but which nevertheless seems to involve a violation of the rules governing how agents should behave in discourse. The general intuition *vis-a-vis* (4) is that the infelicity of Jane’s utterance is due to its unrelatedness to the *topic of conversation*; Sue and Mira are discussing football, as Jane brings up her car problems out of the blue.¹

- (4) **Context:** Jane, Mira, and Sue argue about a football match that just ended. Mira and Sue did not see Jane in a while, and do not know anything that happened to her lately.
- a. Sue: The referee should not have called that foul.
 - b. Mira: What are you talking about! It was an obvious foul!
 - c. Jane: #My car broke down the other day.

On-topic-ness is traditionally cashed out by enriching our model of discourse, which so far contains only a context set c , with a *Question Under Discussion* (QUD) – a simple version of which is defined in (5). The QUD partitions the context set into disjoint, non-empty subsets, each corresponding to a possible answer, and serves to define the task that discourse participants are faced with at a given point of the conversation; namely, the task of determining which, among the cells of the QUD, is home to the actual world.

- (5) A model of discourse is a pair $\langle c, \mathcal{Q} \rangle$, such that:
- a. c is the discourse *context set* (as defined in (1)).
 - b. $\mathcal{Q}$, the QUD, is a *partition* over c ; i.e., a set of propositions $\{q_1, \dots, q_n\}$ such that:
 - (i) $c = q_1 \cup q_2 \cup \dots \cup q_n$
 - (ii) $\forall i \leq n [q_i \neq \emptyset]$
 - (iii) $\forall i, j \leq n [i \neq j \rightarrow q_i \cap q_j = \emptyset]$

Assuming a QUD is present even in the absence of an overt question is useful, as then infelicities like Jane’s utterance in (4) could be modeled as a violation of a constraint on *question-answer pairs* requiring that every declarative utterance in discourse *address* the implicit or explicit question relative to which it was uttered (Collingwood 1940/2001; Carlson 1983/2012; Groenendijk & Stokhof 1984; Lewis 1988a; Roberts 1996/2012; Farkas & Bruce 2010 a.m.o.). If in the case of (4) the implicit QUD is assumed to be *was there a foul?* then Jane’s utterance is infelicitous as it simply does not bear on this question.

This kind of constraint on declarative utterances is often termed *relevance* in the formal pragmatic literature. However, no consensus has emerged as to what notion of *addressing the QUD* underlies relevance. For David Lewis, an utterance addresses a QUD if it is contextually equivalent to the union of a subset (possibly trivial or empty) of the QUD (Lewis 1988a,b), giving rise to the relevance constraint in (6).

¹ As discussed by Grice (1975), by uttering a completely unrelated sentence, Jane might be signaling in (4) that she wishes to change the topic of conversation. Thus, flouting a maxim of conversation is not always infelicitous, as it can have an intended effect. Crucially, however, this requires that there *be* a maxim of conversation flouted in (4), and what exactly this maxim is is the subject of our inquiry here.

- (6) **Negative Relevance (NR):** A declarative utterance ϕ is only felicitous given a model $\langle c, \mathcal{Q} \rangle$ if $\exists \mathcal{Q}' \subseteq \mathcal{Q}: \llbracket \phi \rrbracket \cap c = \bigcup \mathcal{Q}'$.

Notice that NR does not rule out examples like (2) above, which are entailed by the context set; namely, contextually equivalent to the union of the entire QUD. NR is thus not a generalization of INF; to account for the full range of infelicitous examples, utterances should be subject to both constraints. While we can lump the two together to form the constraint in (7), sometimes termed *strong relevance* in the literature (e.g., Križ & Spector 2021; Haslinger 2023), which requires utterances to express propositions that are equivalent to a proper subset of the QUD, this is a mere convenience for presentational purposes. Given that Stalnaker’s INF is independently required, the Lewisian formulation isolates the further contribution of *relevance* on top of INF in a way that (7) obfuscates: While INF informs a speaker of what she must *do*, namely, ensure that worlds are eliminated from the context set, *relevance* instructs her what *not* to do, i.e., ensures that she does not harm the integrity of cells in the QUD in the process of world-elimination. As von Stechow & Heim (2021) stress, “it is important to see that ‘relevance’ in this technical sense is a purely negative requirement, so to speak: it is just the absence of any irrelevant (i.e., not specifically asked for) information.” We thus term the Lewisian formulation *negative* in (6).

- (7) **NR+INF:** A declarative utterance ϕ is only felicitous given a model $\langle c, \mathcal{Q} \rangle$ if $\exists \mathcal{Q}' \subsetneq \mathcal{Q}: \llbracket \phi \rrbracket \cap c = \bigcup \mathcal{Q}'$.

As von Stechow & Heim also point out, this rendition of relevance seems to under-generate a wide range of utterances. For instance, assuming that the overtly asked question is the QUD (an assumption that we will scrutinize in Section 4), none of the answers in (8) is relevant to the question relative to which they are proffered in the Lewisian sense.

- (8) Is it raining?
 a. Heavily.
 b. Well, John came in dripping wet.
 c. We will be staying mostly indoors.

In (8-a) the speaker is conveying *too much information* to satisfy NR+INF; her answer not only rules out the possibility that it is raining, but also distinguishes worlds in which it is raining *heavily* from the remaining worlds in which it is raining. In (8-b)-(8-c) the speaker is conveying *too little information* to satisfy NR+INF; both answers can be true or false regardless of whether it is, in fact, raining, and so by definition fail to eliminate any cells of the partition while also distinguishing among worlds in a way that cuts across cells.²

To account for the felicity of over-informative answers such as (8-a), Groenendijk & Stokhof (1984) and Roberts (1996/2012) allow for greater flexibility in which utterances are admissible. According to their more permissive notion of relevance, termed *positive* in (9), an utterance must express a proposition that *entails* a proper subset of the QUD; i.e., eliminates at least one of its cells.

² For (8-b), we assume that it is not common ground that if John came in dripping wet it must be raining (otherwise the sentence is over-informative in the same sense of (8-a)).

- (9) **Positive Relevance (PR):** A declarative utterance ϕ is only felicitous given a model $\langle c, \mathcal{Q} \rangle$ if $\exists \mathcal{Q}' \subsetneq \mathcal{Q}: \llbracket \phi \rrbracket \cap c \subseteq \bigcup \mathcal{Q}'$.

We are now in a position to be somewhat more explicit regarding our classification of relevance formulations as *negative* or *positive*. We hope that, on an intuitive level, the distinction is clear: positive formulations inform the speaker of the minimal amount of information her utterance must convey for it to be relevant, whereas negative formulations *prevent* an utterance from conveying certain kinds of information. More formally, the distinguishing property between them is defined in (10), which builds on a related notion: To check whether an utterance ϕ satisfies some *positive* relevance constraint, it is sufficient to check whether ϕ is relevant to a subset of the QUD. *Negative* constraints cannot be evaluated in this way: we can determine whether ϕ is relevant relative to some partitioned context, only by observing the partition in its entirety.

- (10) a. A relevance constraint is *negative* if for every utterance ϕ , ϕ is relevant to a model $\langle c, \mathcal{Q} \rangle$ iff it is relevant to every model $\langle c', \mathcal{Q}' \rangle$ where $\mathcal{Q}' \subseteq \mathcal{Q}$ (and $c' = \bigcup \mathcal{Q}'$).
b. A relevance constraint is *positive* if it is not negative; i.e., if there exist an utterance ϕ and two models $\langle c', \mathcal{Q}' \rangle \subset \langle c, \mathcal{Q} \rangle$ s.t. ϕ is relevant to $\langle c, \mathcal{Q} \rangle$ but not to $\langle c', \mathcal{Q}' \rangle$.

PR is strictly weaker than NR+INF: given a discourse model, the set of utterances that satisfy PR is a superset of those that satisfy NR+INF. And while we group PR within the broader family of relevance constraints, note that it could plausibly be thought of as a generalization of INF itself. If we think of INF as the requirement that a speaker rule out some possibilities with their utterance, then PR reduces to INF with a more course-grained distinction among possibilities; now construed as equivalence-classes rather than as particular worlds. In other words, while an utterance can satisfy INF without satisfying NR, and *vice versa*, PR renders INF superfluous: All discourse moves ruled out by the latter, i.e., those that do not eliminate *worlds*, are trivially ruled out by the former, as they also do not eliminate *cells*.

Of course, while PR generates *more* than NR+INF – e.g., (8-a) can now be considered relevant – it still suffers from some under-generation problems apparently plaguing NR+INF. In particular, (8-b)-(8-c) which both convey propositions that are compatible with both cells of the partition imposed by the QUD, are still ruled out as irrelevant. For now, we set (8-c) aside (though we return to it in section 4). The reason (8-b) seems acceptable as a response to the question “*is it raining?*” is that, pre-theoretically, it provides us with information that renders one answer more plausible than the other. Granted, that John came in dripping wet does not settle the question; he may have gone swimming and forgot a towel. However, it is more likely that he arrived wet because it is raining outside and our knowledge of this should bias us towards an affirmative answer to the question.

Notice that we had to appeal to probabilistic language, to terms like “likely” and “plausible”, to explain our intuitions regarding (8-b). Thus, to capture this intuition, more recent conceptualizations of relevance introduce probabilities into the definition. These conceptualizations all build on the Carnapian approach mentioned in the previous section (Carnap 1950), which could be anachronistically rephrased as follows: A proposition p is relevant to the question of “*whether q?*” if knowing that p would shift the probability we assign in our epistemic model to q (and, hence, would shift the probability assigned to $\neg q$, assuming that we are rational beings).

The various technical implementations of this idea all introduce a novel discourse parameter: a probability distribution P mapping subsets of logical space, i.e., propositions, to values between 0 and 1. ϕ is then not required to eliminate cells in $\mathcal{Q}$ to be considered relevant, but rather, *knowing that* ϕ must “alter” P in a particular way – with the approaches differing as to which way that is.³ For instance, one version renders ϕ relevant relative to $\langle c, \mathcal{Q}, P \rangle$ only if $\llbracket \phi \rrbracket$ “shifts the probabilistic weights among the propositions denoted by $\mathcal{Q}$ ” (Büring 2003: 517).⁴ An alternative, proposed by Van Rooy (2004), models relevance as *entropy reduction*, a measure of the degree to which an answer decreases uncertainty *vis-a-vis* the resolution of $\mathcal{Q}$. Russell (2012) dispenses with a categoric notion of relevance altogether, in favor of a gradient notion according to which the relevance measure of ϕ is the difference between the probability P assigns to an answer $q \in \mathcal{Q}$ before the context is update with $\llbracket \phi \rrbracket$ and the measure assigned to q after update has occurred (i.e., $P(q|c) - P(q|c \cap \llbracket \phi \rrbracket)$), resulting in some utterances being “more relevant” to $\mathcal{Q}$ than others.

We discuss the introduction of probabilities in depth in the next section. For now we make two comparative observations. First, the probabilistic notions gestured at above, while more permissive than PR in that they render non-resolving answers like (8-b) relevant, all share a crucial property with PR: they constitute positive constraints in that they demand propositions to contain some sufficient amount of information to be relevant (and therefore they satisfy the definition in (10)). Second, the shift in the literature from the strongest formulation of relevance, which we call here *Lewisian* and describe as *negative*, through Roberts’s weaker *positive* notion, towards probabilistic approaches more recently, embodies the gradual weakening of the constraint: a move motivated at every step by perceived under-generation, and intended to cover more empirical ground.

Next, we try to dispel some assumptions regarding both of these observations. First, we argue that, being positive is not a necessary feature of probabilistic formulations, and that there is a coherent negative formulation that is nevertheless probabilistic. Second, we argue that the methodology behind the gradual weakening of relevance is flawed: overt questions do not necessarily determine the QUD, and thus examples that consist of overt question-answer pairs may be interpreted in multiple ways, and do not by themselves constitute sufficient evidence in favor of said weakening.

3 Probabilistic relevance is not at odds with the Lewisian approach

Probabilistic models of relevance generally assume a probability function P over subsets of W as one of the discourse parameters. This function assigns a probability (namely a value between 0 and 1) to each cell in the given QUD, and thus tracks the probabilistic consequences of accepting different pieces of information by discourse participants. In those models, relevance is a constraint on the relation between this probability function before and after the acceptance of a given proposition as common knowledge. However, the introduction of probabilities raises some non-trivial issues; particularly, questions regarding the nature of that probability distribution, what it is supposed to represent, and its relation to the rest of our discourse parameters. How one answers these foundational questions plays a role in how, exactly, probabilities are to be incorporated into our

³ Cf. Warstadt & Agha 2022 for a useful overview of the different approaches.

⁴ For more formal renditions of the approach, see Hawkins & Goodman 2017; Feinmann 2020. This formulation is often the one used in the linguistic literature; e.g., Hyska 2015; Benbaji & Doron 2023.

model, and this, in turn, shapes the theoretical space of possibilities regarding the conceptualization of relevance. We thus begin with some brief remarks on these more general issues.

3.1 Some foundational questions

On the Stalnakerian picture, recall, common knowledge is represented by a set of worlds, c , which is updated throughout the conversation as new propositions are accepted via set intersection: c is intersected with any accepted proposition to yield a new context set. As far as we know, the relation between c and the probability distribution parameter P , as well as the role of propositions in the process of context update within a probabilistic framework, is rarely discussed explicitly in the context of the relevance debate.⁵ As we see it, on one possible approach, the introduction of a probability parameter P does not require any alterations to the classical Stalnakerian notion of context update. Accepted propositions are still sets of worlds that intersect with c . On this view, at any point in a conversation, the shape of the probability distribution P is simply determined by what discourse participants have accepted as true up to that point, represented by c . The relation between c and P is deterministic, and cannot be manipulated directly by the update of a proposition.⁶

How can relevance be defined within such a system? For concreteness, assume our discourse model contains, in addition to context set c and the QUD $\mathcal{Q}$, a probability function P , which indicates the credences discourse participants assign to propositions that are neither entailed nor contradicted by the context set. Since c itself determines the shape of P , once a new proposition $\llbracket\phi\rrbracket$ is introduced, c will determine different probability distributions, before and after it is intersected with $\llbracket\phi\rrbracket$. Relevance can then be defined as a condition on the relation between two distributions, P_c and $P_{c \cap \llbracket\phi\rrbracket}$; i.e., the prior probabilities P assigns to subsets of the context (and in particular to the cells in $\mathcal{Q}$), and the posterior probabilities it assigns to them after c has been updated with $\llbracket\phi\rrbracket$.

While existing proposals vary with respect to the relation they attribute to P_c and $P_{c \cap \llbracket\phi\rrbracket}$, they can all generally be thought of as demanding that the difference between P_c and $P_{c \cap \llbracket\phi\rrbracket}$ exceed a certain threshold. It is easy to see how this requirement is compatible with a positive formulation of relevance. This is the case, for example, in [Büring 2003](#) (and in the variation on it in [Feinmann 2020](#)), where a relevant answer ϕ is required to shift the probabilistic weights among the cells of

⁵ One exception is a short note by [Feinmann \(2020: 85\)](#), who suggests that given P , the Stalnakerian context set is simply superfluous. [Feinmann's](#) logic, as we understand it, is as follows: A probability distribution over logical space is a richer notion than a context set. Any information conveyed by the latter can be conveyed by the former, *and then some*. In particular, if P assigns any proposition *true* in the Stalnakerian context set c to 1, and any proposition *false* in c to 0, it exhaustively captures everything that c has to offer. Presumably (though [Feinmann](#) is not explicit about this), propositions whose truth is not determined by c are assigned to probabilities between 1 and 0, intended to reflect the credences of discourse participants. We think that maintaining a context parameter c is useful, at least for presentational purposes, even if [Feinmann](#) is correct that certain conceptualizations of relevance can do without it. Therefore, in what follows we construct a model of discourse with c and P , that retains the basic insights from [Feinmann's](#) discussion.

⁶ Note that, if one assumes that the new discourse parameter P corresponds to a single probability distribution, then this approach assumes what philosophers call “the uniqueness hypothesis”; i.e., the view that “rational probabilities are constrained to the point of uniqueness by one’s evidence” ([Hájek 2023](#)). However, we can avoid this strong hypothesis if we so desire, as long as we take P to correspond to a *set* of probability distributions. Then, while the relation between c and P is still deterministic, there need not be a unique probability distribution compatible with any given body of knowledge; rather, c only has to determine which distributions are ruled out.

2. This requirement is positive both in an intuitive sense – it instructs the speaker of the minimal thing they must do for their utterance to be felicitous – and also in the formal sense defined in (10); one can examine a *subset* of the partitioned context to verify that an utterance meets this constraint.

Negative formulations of relevance, however, require something else altogether: They generally demand that utterances not differentiate between worlds within the same cell in the partitioned context. On the Lewisian, non-probabilistic version, utterances must not eliminate worlds from a cell while leaving their other cellmate worlds untouched. The introduction of probability provides us with another way to define what it means to differentiate between cellmates, as we can now demand that utterances not make a certain subset of worlds within a cell more probable than any another subset in the same cell. To our understanding, this is virtually the only possible way to define a notion of relevance that is both probabilistic and negative within this kind of a system, in which utterances cannot directly manipulate the probability parameter.

While consistent, this notion of relevance is very strong. In fact, it is at least as strong than non-probabilistic NR, which, combined with INF, is the strongest constraint we have entertained thus far. Consider an utterance which violates non-probabilistic NR; i.e. it eliminates a subset of worlds from a certain cell in 2, while keeping the other subsets intact. The probability of the subset of eliminated worlds is 0, while the probability of the rest of the cell’s subsets is necessarily greater than 0 (otherwise they would be eliminated as well). Therefore, this utterance violates the probabilistic relevance constraint discussed above, which is thus at least as strong as NR. Recall, however, that the main motivation for introducing probabilities in the first place was to fix the apparent *under-generation* of non-probabilistic relevance. It is therefore clear that a notion of negative probabilistic relevance that rules out, at the very least, the same utterances ruled out by its non-probabilistic counterpart, does not deliver what it was intended to deliver to begin with.

Next, however, we would like to offer an alternative way of incorporating probability distributions into our model of discourse that allows us to formulate a relevance constraint which tracks more closely to the Lewisian negative approach, but is nevertheless weaker than NR. The key would be to redefine the notion of context update, so that propositions can directly alter the shape of the probability parameter. To do so, we adopt a pragmatic model that is homomorphic to the Stalnakerian one, with probability functions replacing worlds as primitive elements. The model is, essentially, a naive implementation of a proposal by Swanson (2006) and Moss (2015a,b, 2018) – greatly simplified for presentational purposes – who suggest that sentences have sets of probability functions as their semantic value, supplanting the classical notion of content with a probabilistic one.⁷ The context set *c* can then itself be thought of as a set of probability functions. Thus, any constraint which, in the Stalnakerian model, blocks the update of *c* with certain sets of *worlds* exhibiting particular properties, can be recast in the model we adopt below to block update with certain sets of *probability functions* that exhibit those same properties, *including NR*.

Before we continue, a caveat: As will be discussed in Section 8, given the argument we present in the second part of this paper in favor of negative relevance, we are generally inclined to prefer the Lewisian, non-probabilistic version, rather than the probabilistic one we provide next. As the reader will likely take notice, and as Moss (2015b) herself notes, the probabilistic alternative requires

⁷ Other relevant inspirations include the insights on probability operators in Yalcin 2010 and Lassiter 2011, which nevertheless differ quite radically in their use of probability functions from the proposal we adopt here.

a rather dramatic shift from standard truth-conditional semantics. Our argument in favor of this alternative will thus be a conditional one: To the extent that one adopts a probabilistic model to begin with, the model cited below, paired with a negative relevance constraint, is the one that should be adopted. We will not be concerned with defending such a model any further in the present work.

3.2 Probabilistic negative relevance

To define negative relevance in a probabilistic framework, we follow Swanson (2006); Moss (2015b) in baking probability into the semantics of declarative utterances themselves. One can do so by defining a new primitive domain, $\mathcal{D}_p$, of probability functions over worlds; i.e., measure functions from the power set of logical space, $\mathbf{P}(\mathbb{W})$ to real numbers in $[0,1]$. Such functions will play in our probabilistic semantics the same role that worlds do in classical semantics. But while any world fully determines the truth and falsity of every proposition (for any proposition, it is either in its extension or not), a probability function does not; it can assign some propositions to 1, others to 0, and yet others to values in between. These functions can thus be thought of as representing a possible state of the world, given our epistemic uncertainty about whether certain propositions obtain; namely, the state in which the likelihood that any proposition p obtains is $P(p)$.

This view of the role of probabilities assumes that a given set of evidence, and in particular our shared knowledge, does not deterministically decide the probability of all propositions whose truth is left open by this set; i.e., two distinct possible states can agree on which propositions are true or false, but differ in the probabilities assigned to propositions whose truth has not been determined.

With these assumptions, we can define the *p-context* as the set of all probability functions compatible with the shared knowledge of discourse participants (11).

- (11) **Probabilistic context set:** $\tilde{c} = \{P \in \mathcal{D}_p : | P \text{ is compatible with the shared knowledge of discourse participants}\}.$

Since we assume here that the notion of meaning is inherently relativized to probability, the meanings that sentences denote are themselves probabilistic. We posit that declarative sentences denote functions from probability values – *degrees* between 0 and 1 – to sets of probability functions. These functions, which we term *p-propositions*, intuitively state that the probability of a certain set of worlds is equal to its degree argument (12). In sentences that lack any overt probabilistic particles, we assume that the *d*-argument is saturated by an implicit argument whose value is 1. Thus, the utterance *it is raining*, denotes the set of all probability functions that map the proposition that it is raining to 1, as exemplified in (13).⁸

⁸ There is a choice point here regarding how to treat the closure of the degree argument in sentences without an overt probability operator. For simplicity, we assume here that it is done by plugging in the value 1, the result of which is that simple sentences like *it's raining* convey that it is completely certain that it's raining (similarly to the classical model). Another approach, in the spirit of Lassiter (2011), has that argument saturated by a default “positive form” operator, which maps the *p*-proposition to the set of probability functions which assign rain to a probability value higher than a certain threshold. On this view, the sentence *it's raining* conveys that the chance that it is raining is above some contextually-determined value. This might explain, for example, the contribution of operators like *certainly*. We refer the reader to Lassiter's work for a more detailed discussion.

(12) **Probabilistic propositions:**

- a. We call $\tilde{p}$ a *p(probabilistic)-proposition* if it is in $\mathcal{D}_{\langle d, \langle p, t \rangle \rangle}$, where d is a probability value (a number between 0 and 1) and p is a probability function.
- b. Given a classical proposition p , its probabilistic extension is the function $\tilde{p}$ that takes a probability value d and returns the set of probability functions which assign p this value: $\tilde{p} = \lambda d. \lambda P. P(p) = d$

(13) $\llbracket \text{It's raining} \rrbracket = \lambda P. P(\{w \mid \text{It's raining in } w\}) = 1$

This notion of p-proposition *subsumes* the classical notion, as it is strictly more expressive. Given a classical proposition p , its truth conditions can be expressed by feeding the value 1 to its probabilistic counterpart $\tilde{p}$, yielding the set of probability functions that assign 1 to p . But of course, plugging in other values to $\tilde{p}$, or quantifying over them, gives us a range of meanings that cannot be expressed in the classical system. This additional expressive power allows us to analyze probability operators like *certainly*, *probably*, *maybe*, etc. as directly quantifying over the d -argument of p-propositions (Moss 2015b). A toy implementation is given in (14-a)-(14-b).⁹

- (14) a. $\llbracket \text{Maybe it's raining} \rrbracket = \lambda P. P(\{w \mid \text{It's raining in } w\}) > 0$
b. $\llbracket \text{It's probably raining} \rrbracket = \lambda P. P(\{w \mid \text{It's raining in } w\}) > 0.5$

Given these semantics for propositions, we can keep to the same process of context update via intersection of the proposition with the context, as in the classical Stalnakerian model. We can also now define the QUD simply as a *p-question* that partitions the p-context. More specifically, we assume that questions in this probabilistic framework are sets of p-propositions. The partition a p-question imposes on $\tilde{c}$ divides it to cells such that the probability functions in each cell assign the same probability to each p-proposition in the question (15).¹⁰

(15) **Probabilistic question:**

- a. A *p-question* is a set of p-propositions.
- b. Given a p-question $\tilde{q}$ containing n p-propositions $\tilde{p}_1, \dots, \tilde{p}_n$, and a p-context $\tilde{c}$, we define the partition imposed on $\tilde{c}$ by $\tilde{q}$ as follows:

For any $d_1, \dots, d_n$ n -tuple of probability values we define a cell:

$$q_{d_1, \dots, d_n} = \{P \in \tilde{c} \mid P \in \tilde{p}_1(d_1) \wedge \dots \wedge P \in \tilde{p}_n(d_n)\}$$

We define the partition $\tilde{Q}$ as the set of non empty cells: $\tilde{Q} = \{q_{d_1, \dots, d_n} \mid q_{d_1, \dots, d_n} \neq \emptyset\}$

⁹ The definitions for the probability operators below are syncategorematic, but of course, we can isolate the semantic contribution of the operators as follows:

- (i) a. $\llbracket \text{maybe} \rrbracket = \lambda \tilde{p}_{\langle d, \langle p, t \rangle \rangle} \lambda P_p. \exists n > \mathbf{0} : \tilde{p}(n)(P)$
b. $\llbracket \text{probably} \rrbracket = \lambda \tilde{p}_{\langle d, \langle p, t \rangle \rangle} \lambda P_p. \exists n > \mathbf{0.5} : \tilde{p}(n)(P)$

¹⁰ This definition ignores issues of granularity in the probability values $d_1, \dots, d_n$. If we allow these values to range over any real numbers between 0 and 1, the partitions we end up deriving may be very fine grained. This would mean that virtually any non-trivial proposition eliminates a cell from the QUD. A discussion on whether this is a welcome result or not is beyond the scope of this paper, but we note that a natural way to raise the bar of what constitutes eliminating a cell is to define some granularity on the values of $d_1, \dots, d_n$ that would restrict their density.

We have thus defined an embedding of the classical Stalnakerian pragmatic model within an enriched model, whose primitive objects are probability functions, rather than worlds. It should come as no surprise, then, that we can reproduce the pragmatic constraints posited in the old model, within our modified one. Specifically, we can define both negative and positive notions of relevance along the lines of the non-probabilistic definitions; i.e., NR (6) and PR (9). These are given in (17)-(18), alongside a modified version of Informativity (16), for completeness.

- (16) **Probabilistic Informativity (p-INF):** A p-proposition $\tilde{p}$ is *informative* relative to a p-context set $\tilde{c}$ if $\tilde{c} \cap \tilde{p} \neq \tilde{c}$
- (17) **Probabilistic Negative Relevance (p-NR):** A p-proposition $\tilde{p}$ is relevant given a p-model $\langle \tilde{c}, \tilde{\mathcal{Q}} \rangle$ if $\exists \tilde{\mathcal{Q}}' \subseteq \tilde{\mathcal{Q}}: \tilde{p} \cap \tilde{c} = \bigcup \tilde{\mathcal{Q}}'$.
- (18) **Probabilistic Positive Relevance (p-PR):** A p-proposition $\tilde{p}$ is relevant given a p-model $\langle \tilde{c}, \tilde{\mathcal{Q}} \rangle$ if $\exists \tilde{\mathcal{Q}}' \subsetneq \tilde{\mathcal{Q}}: \tilde{p} \cap \tilde{c} \subseteq \bigcup \tilde{\mathcal{Q}}'$.

This model does not suffer from the same issue that prevented us from defining a useful negative notion of relevance in a probabilistic model in which utterances cannot directly alter the shape of the probability parameter; namely, that any proposition that shifts the probability of certain sets of worlds without eliminating a cell, must also rule out some worlds, thus violating NR.

To see this, imagine a p-context $\tilde{c}$, partitioned by the polar question *is it raining?* Let us assume that $\tilde{c}$ contains no information about the chances of rain; namely, for every probability value d , there exists $P \in \tilde{c}$ such that $P(\{w \mid \text{It's raining in } w\}) = d$. The partition we end up with is such that for each possible probability value d , it contains a cell containing all probability functions that assign d to $\{w \mid \text{It's raining in } w\}$. It is easy to see that all sentences in (19) are relevant to given this partition, according to both p-NR and p-PR (17)-(18).

- (19) a. It's raining.
- b. Maybe it's raining.
- c. It's probably raining.

Each of these sentences states that the probability of rain is equal to, or greater than, a certain value (1 in (19-a), 0 in (19-b), and 0.5 in (19-c)), and thus eliminates all and only those cells in which $P(\{w \mid \text{It's raining in } w\})$ returns a lower value, leaving all other cells intact. Notice that while the sentence in (19-a) is negatively-relevant in the classical model as well, (19-b) and (19-c) are not. This demonstrates that p-NR is a slightly weaker notion than NR, and thus ameliorates at least some of the apparent under-generation challenges leveled at the Lewisian negative constraint.

However, it might seem at this point that p-NR is still too strong, since – unlike p-PR – it does not admit a discourse like (8-b), repeated in (20). Assuming for simplicity that (20-b) contextually entails that the probability of rain is, say, greater than 0.5, that answer still eliminates from each cell all probability functions that assign a low enough probability to John coming in wet; and there must be such functions, otherwise the sentence would violate p-INF. It is thus not equal to a union of cells, violating p-NR. This is a reincarnation of the kind of data that motivated the shift to positive conceptions of relevance – namely, over-informative utterances – within our modified probabilistic discourse model: (20-b) is over-informative given the partition imposed by the p-question *is it*

raining? in the same way that *it is raining heavily* in (8-a) is over-informative given the partition imposed by the non-probabilistic, classical denotation of this question. In the next section, we provide a general argument against the shift towards positive formulations of relevance on the basis of such answers, and propose a way to explain them even on the assumption that relevance is a purely negative constraint, whether probabilistic or not.

- (20) a. Is it raining?
 b. Well, John came in dripping wet.

4 Against the gradual weakening of relevance

Let us take stock: We show that relevance constraints may vary along two orthogonal dimensions. They can be positive or negative, and probabilistic or not. We can thus in principle entertain four distinct relevance constraints, two negative and two positive, formulated in a probabilistic or non-probabilistic model. These constraints are partially ordered by strength. Given a particular context and QUD, each of the constraints carves out a distinct set of propositions that it renders relevant such that any utterance that is negatively relevant in a non-probabilistic sense (i.e., that satisfies NR) also satisfies all other formulations, and any utterance satisfies p-PR if it satisfies PR or p-NR, which are themselves unordered with respect to each other.

As an illustration, consider the answers in (21-a)-(21-f), to the question in (21). These answers are all relevant given p-PR, as they all serve to shift the probabilistic weights among possible answers to the question. However, only some of them do so by eliminating cells of *worlds* in the Stalnakerian model, (21-a)-(21-d) and are thus relevant according to PR, and among those only some eliminate cells of worlds without cutting through the cells they do not eliminate (21-a)-(21-b), satisfying NR as well. The rest shift probabilistic weights by providing information that entails certain probability thresholds for certain answers. This is done either by directly quantifying over probability functions (21-e), satisfying p-NR, or by providing information that indirectly shifts the probability of certain cells (21-f), satisfying only p-PR. The inclusion relations among sets of relevant propositions, given the QUD in (21), are demonstrated in Figure 1.

- (21) Q: Who (among Mary and John) is coming to the party?
 a. John.
 b. Mary.
 c. John and probably Mary.
 d. Mary is driving there.
 e. Probably John.
 f. Mary is tired.

The convergence of much of the recent literature on positive probabilistic notions of relevance is largely motivated by the fact that (21-a)-(21-f) all seem felicitous as responses to (21), suggesting that the weakest formulation of relevance is required to avoid under-generation. The methodology that has led to that conclusion, which is exemplified by the dialectic in previous sections, can be described as follows: A certain formulation of relevance is challenged by an apparent counterexam-

QUD – (21): “Who is coming to the party?”

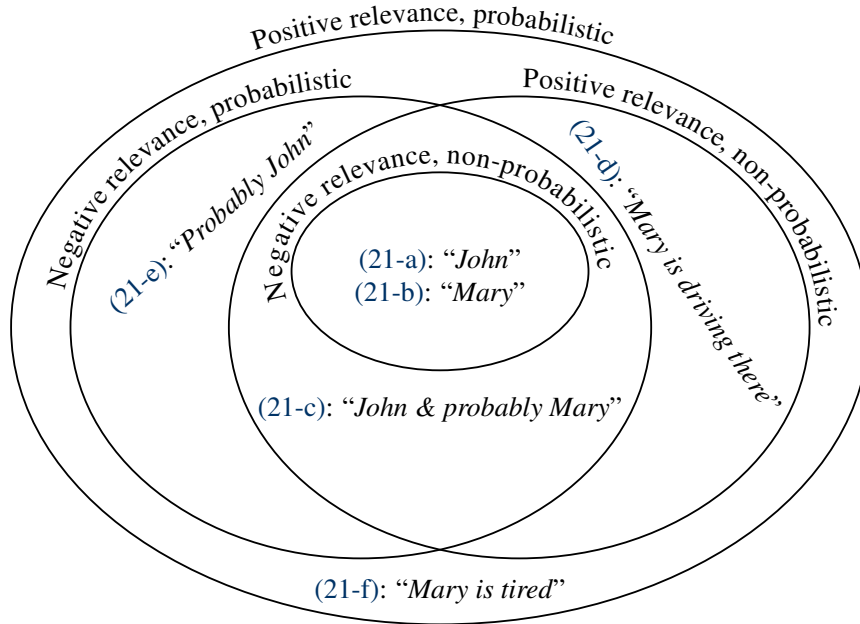


Figure 1 Inclusion relations among the sets of *relevant propositions*, given a particular QUD, as determined by the different formulations of relevance

ple, in the shape of a question-answer pair which does not satisfy that formulation but is still judged to be felicitous; in response, the proposed formulation is weakened. This is repeated until we end up with the weakest formulation in our hierarchy.

However, is probabilistic positive relevance weak *enough* to generate all felicitous question-answer pairs? Unfortunately, the same kind of counterexamples that have been used to target stronger formulations, can be hurled at the weakest one as well. If the methodology behind the weakening of relevance is sound, then, it seems that probabilistic positive relevance, too, must be weakened. But if we do not want to find ourselves weakening *ad absurdum* we should, perhaps, conclude that there is something faulty in that methodology.

Consider, for instance, the answers in (22-b).

- (22) a. **Context:** We are planning a trip to Washington DC, I wonder what kind of clothes I should pack.
 b. Is it going to rain?
 (i) We’ll be spending most of the time indoors.
 (ii) We’ll only go if it doesn’t.

These obviously cut across cells in the partition imposed by the overtly-asked question, and therefore do not satisfy NR with respect to it. They also do not eliminate any of its cells; both answers are compatible with both rain and non-rain scenarios, and therefore do not satisfy PR either. Finally, they also do not seem to change the prior probability we attribute of each of the cells; the knowledge

of where we are planning to hang out during our trip, or what are the conditions for us embarking on it to begin with, is completely orthogonal to the question of whether it will rain, and is thus not even probabilistically-relevant to the question here. Nevertheless, these seem to be felicitous answers to the overtly asked question. Thus, by the logic described above, we should now weaken our notion of relevance even more.

Of course, proponents of the weaker formulations of relevance are well aware of such examples, which can be accounted for *without* weakening relevance any further, simply by rejecting our implicit assumption that the overtly-asked question *just is* the QUD relative to which the answer is uttered. In fact, there is a good conceptual reason to doubt this assumption: Since we assume that a QUD is present even in the absence of an overtly-articulated question, we must already posit a process by which discourse participants accommodate an implicit QUD relative to which discourse is held. But then there is no obvious reason to prevent this process, whatever it is, from playing a role even when a question *was* explicitly articulated (Beaver & Clark 2009; Roberts 1996/2012).

Indeed, the availability of presupposition accommodation teaches us that a similar process is possible for the common ground (Lewis 1979; Stalnaker 1977), and given the inherent uncertainty in discourse about the topic of conversation, there is no reason to expect it not to be available for QUDs as well. In the case of, say, (22), this means that the addressee is required to accommodate a QUD along the lines of *where will we spend our time on this trip?* or *will the trip take place rain or shine?* as relative to these accommodated QUDs, the answers in (22-b-i)-(22-b-ii) satisfy even the strongest formulation of relevance; i.e., NR.

But if a mechanism of QUD accommodation is available in principle, then the debate over the correct formulation of relevance reduces to a debate over which felicity judgments are based purely on relevance proper and which are based on relevance, when supplemented with QUD accommodation. More generally, as noted by Beaver & Clark (2009: 39), any formulation of discourse constraints can be predictive only if the possibility of accommodation is constrained. Any argument for one position or other is unsound if it relies solely on observations regarding the felicity of question-answer pairs, suggesting that the QUD cannot be shifted — contrary to fact.¹¹ Formulating the general constraints on QUD accommodation, however, is a difficult task that goes far beyond the scope of this paper. Rather, in the rest of this paper we do something slightly less ambitious: We provide a diagnostic that allows us to determine that QUD-accommodation is blocked in certain cases. These cases must thus reveal something about the correct formulation of relevance, and indeed, we use this data to motivate a novel version of negative relevance.

11 Agha & Warstadt (2020) suggest that non-sentential “fragment” answers to overtly asked questions (in the sense of Merchant 2005) block accommodation of an implicit QUD and must be relevant relative to the overt question that they follow. The argument is based on an observation by Jacobson (2016), according to which fragment answers to *which*-NP questions must denote elements in the extension of NP, while corresponding full sentential answers are not required to dominate expressions denoting such elements. Based on this diagnostic, Agha & Warstadt attempt to delineate a set of question-answer pairs that can be taken to bear directly on the formulation of relevance, given the impossibility of accommodation. This is, indeed, the general methodology we think one should follow when attempting to reason about relevance, but we think it is misapplied in this case; in work under preparation, we examine the status of fragments and show that Jacobson’s data can only be used to motivate a ban on accommodation, once one is already committed to the probabilistic notion of negative relevance defined in (17). It should also be noted that Jacobson herself provides a semantic account of the facts that has nothing to do with accommodation, or lack thereof.

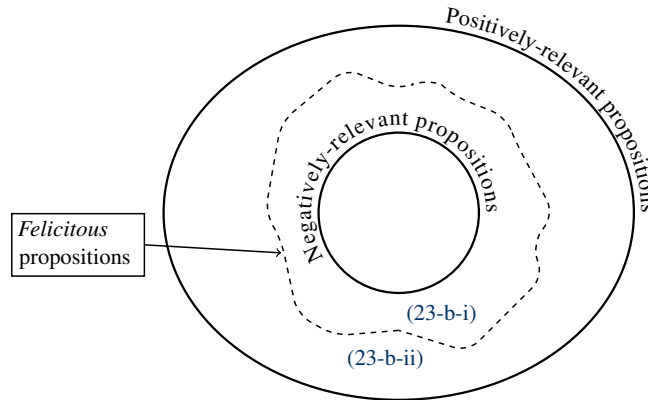


Figure 2 Desideratum of our conceptualization of *relevance*

5 The challenge from over-informative answers

The empirical observation at the heart of the rest of the paper is that some but not all over-informative answers are judged to be felicitous. Consider, for instance, the scenario in (23).

- (23) **Context:** Paula visits an art gallery and sees a painting by an obscure 14th Century Spanish artist. She thinks that no one but herself could possibly be interested in this painting, and figures she has some time to go home and reflect on whether she wants to buy it. A few days later, she returns to the art gallery.
- a. Paula: Can I buy the painting?
 - b. Art dealer:
 - (i) Someone who collects obscure 14th Century Spanish art just bought it.
 - (ii) #Someone whose date of birth is in September just bought it.

All formulations of relevance discussed above, deliver a wrong result for (23). Both responses in (23-b) satisfy PR and p-PR, as they both entail the *no*-answer to Paula's question. On the other hand, both violate NR and p-NR, as they are not *equivalent* to the *no*-answer or to any answer that uniformly changes its probability (in the sense defined for p-NR). As a matter of fact, however, (23-b-i) seems felicitous in the given context, while (23-b-ii) seems infelicitous. *Prima facie*, then, none of our formulations are good candidates for the relevance constraint; the positive formulations over-generate, whereas the negative ones under-generate.

For ease of exposition, given that negative and positive formulations pattern together with respect to this challenge, our discussion below of how it lends support to negative relevance focuses exclusively on non-probabilistic formulations, comparing PR to NR and ignoring p-PR and p-NR. We return to the probabilistic formulations in Section 7.2 to illustrate that our argument successfully distinguishes between them as well, thus serving as a general argument for the negative approach.

As illustrated in Figure 2, (23) arguably suggests that a new conceptualization of relevance is needed; one that is weaker than NR so as to admit (23-b-i), but stronger than PR so as to exclude (23-b-ii). But is such a notion of relevance really required? Perhaps the challenge can be resolved

while maintaining PR or NR and appealing to some auxiliary assumptions to resolve over- or under-generation? We can think of these constraints as functions, whose input consists of a pair of utterance and discourse model, $\langle \phi, \langle c, \mathcal{Q} \rangle \rangle$, that output a felicity judgment; “felicitous” if ϕ addresses $\mathcal{Q}$, and “infelicitous” otherwise. There are then in principle two strategies we could employ to save PR or NR. We can adopt a more permissive function, PR, that generates felicity judgments even for infelicitous utterances, but supplement it with additional constraints that rule out the infelicitous data. Alternatively, we can adopt the less permissive one, NR, but then, given that under-generation problems cannot be avoided by positing additional constraints, we would need to reevaluate our assumptions about the input of the function itself; namely, about the pair $\langle \phi, \langle c, \mathcal{Q} \rangle \rangle$.

The problem with the former strategy, in which relevance *proper* is taken to admit both (23-b-i) and (23-b-ii), is that as far as we know no constraint that can filter out the infelicitous case has been proposed in the literature. We thus set this possibility aside for the purpose of this paper.¹²

Alternatively, if we take (23) as evidence that NR is the relevance function governing discourse, the felicity of (23-b-i) is the curious fact that requires explanation. As discussed in the previous section, given that QUD-accommodation is in principle available, perhaps (23-b-i) is felicitous as it satisfies NR relative to an accommodated question distinct from the overtly articulated one, whereas a similar process is unable to salvage (23-b-ii). Accommodation, after all, is exactly the kind of process that “tinkers” with the input to the relevance function, and can thus be used to solve under-generation problems: It allows us to feed that function a discourse model whose question parameter is not the overtly asked question, but some other contextually salient one. We show next, however, that this route cannot be appealed to in order to defend NR, by ruling out QUD-accommodation for principled reasons in (23-b-i). Thus, an alternative formulation of relevance – one that carves out the set of felicitous propositions as represented in Figure 2 – is indeed required.

6 No QUD accommodation to the rescue

A QUD-accommodation account of our example in (23) rests on the following two assumptions: First, (23-b-i) is felicitous because it satisfies NR relative to some accommodated QUD. Second, it is harder – or perhaps even impossible – to accommodate a QUD relative to which (23-b-ii) satisfies NR. On this view, the difference between these answers boils down to constraints limiting which QUDs may be accommodated. To justify the latter of the two, of course, this approach must be couched within a broader account of QUD accommodation that explains why we can accommodate some questions but not others. However, we will not pursue this endeavor here because we reject

¹² There is a potentially principled point against this strategy: If the factor differentiating (23-b-i) from (23-b-ii) is that only the former intuitively relates to the *topic of conversation*, then – to the extent that our intuitions in this matter are any guide – it is the constraint which enforces *on-topic*-ness, namely, relevance, that should account for the disparity between them, not some other discourse constraint.

Relatedly, note that Feinmann (2020: 107) posits, for independent reasons, a general ban on over-informative answers within a probabilistic account of relevance. He argues that while all probabilistically positively relevant propositions are, indeed, generated by the correct formulation of relevance (something akin to p-PR, formulated in a slightly different manner), those like (23-b-i)-(23-b-ii) are both ruled out independently by a constraint against answers that are proper subsets of a particular cell in the partition. Felicitous over-informative answers are generated in this system, again, by appealing to the possibility of QUD-accommodation. Thus, the argument we provide below against QUD-accommodation in (23-b-i) also applies to Feinmann’s proposal.

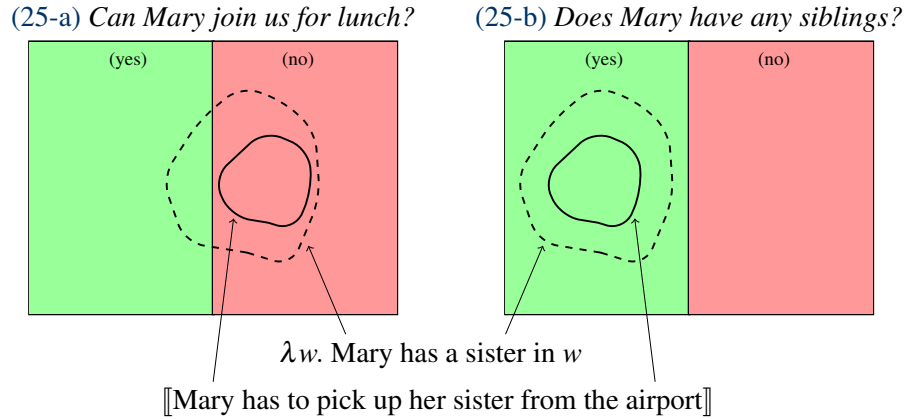


Figure 3 Non/resolution of a question by the presupposition of its answer

the former assumption; namely, that there is an implicit question that can be accommodated that would make the art dealer's felicitous response (23-b-i) satisfy NR.

Let us assume by contradiction that (23-b-i) indeed involves QUD-accommodation. What is the question that Paula is expected to accommodate, that would render (23-b-i) NR-compliant? Some potential candidates that come to mind, provided in (24), all presuppose a negative answer to the overtly-asked question, which, as we will show next, makes them problematic candidates for the accommodated-QUD. In order to be as charitable as possible towards the QUD-accommodation account, however, we will not confine our pool of candidates to any particular set of questions. Instead, we consider all possible partitions of c , capitalizing on an independent generalization that can serve to rule out potential implicit QUDs.

- (24) a. Why did the buyer choose this painting in particular?
 b. A person with what artistic interests bought the painting?

6.1 Heim's generalization

In class notes, Irene Heim observes that an utterance ϕ that presupposes p is infelicitous as a response to a question $\mathcal{Q}$ if p resolves the question raised by $\mathcal{Q}$; i.e., if p eliminates all cells in $\mathcal{Q}$ but one. This is exemplified in (25), where the same answer is a felicitous response to the question in (25-a), but not to that in (25-b) – presumably because the presupposition of the answer that the speaker has a sister settles the question in (25-b), but not in (25-a) – see Figure 3 for an illustration.

- (25) **Context:** A does not know anything about Mary's family.
- a. A: Can Mary join us for lunch?
 B: She has to pick up her sister from the airport.
- b. A: Does Mary have any siblings?
 B: #She has to pick up her sister from the airport.

Doron & Wehbe (2022) point out that Heim's generalization can be derived from the interaction

of different pragmatic constraints; NR+INF (7), and what is commonly known as *Stalnaker's bridge*¹³ (Stalnaker 1977, 1978), defined in (26). To see that, let us briefly introduce the process of known as *presupposition accommodation*.

- (26) **Stalnaker's bridge (SB):** A declarative utterance ϕ is only felicitous given a context c if c entails the presupposition of ϕ .

As Stalnaker observes, SB may be satisfied in one of two ways. First, an utterance's presupposition may already be taken for granted by discourse participants. Second, an utterance's presupposition may become common knowledge "post-hoc" if upon hearing a potentially non-SB-compliant utterance, cooperative discourse participants are willing to incorporate its presupposition into the common ground before update is performed (cf. discussion in von Fintel 2008). This latter way is dubbed *presupposition accommodation*.

Now, notice that as long as presupposition accommodation occurs *before* the evaluation of NR+INF (7), Heim's observation falls out.¹⁴ As illustrated in Figure 3, a question inquiring whether Mary has siblings is resolved by the presupposition of the answer, that Mary has a sister, which leaves only the *yes*-cell in the partition. Given that NR+INF demands that the proposition denoted by the sentence be equivalent to a proper subset of the QUD, the only way for a sentence whose presupposition resolves the QUD to be relevant is to denote a contextual contradiction; we assume that contradictions are ruled out by an independent constraint, and thus that (25-b) is infelicitous.

Doron & Wehbe's (2022) account of Heim's observation makes an interesting prediction pertaining to the distinction between *full* and *partial* resolution of a question, defined in (27). A proposition fully resolves a question if it is only compatible with one of its cells, and partially resolves it if it is compatible with some but not all of them. As illustrated in Figure 3, Heim's original data involves a state of affairs in which the presupposition of an utterance is compatible only with one cell of the question relative to which the utterance was made.¹⁵ As stated above, in such cases relevance cannot be satisfied after presupposition accommodation.

- (27) Given a proposition p and a partition $\mathcal{Q}$,
- a. p fully resolves $\mathcal{Q}$ only if $|\{q \in \mathcal{Q} : q \cap p \neq \emptyset\}| = 1$
 - b. p partially resolves $\mathcal{Q}$ only if $|\mathcal{Q}| > |\{q \in \mathcal{Q} : q \cap p \neq \emptyset\}| > 1$.

However, when a proposition p is compatible with several cells of a partition, utterance ϕ that presupposes p (ϕ_p in short) could still in principle be relevant after p is accommodated. In this case, schematized in Figure 4, accommodation of p narrows down the context set in a way that rules out some cells of the partition, but critically leaves more than one of them alive. ϕ_p can still satisfy NR if it is equivalent to a proper subset of the cells compatible with p . It is thus predicted that a question can be *partially* resolved by the presupposition of its answer, as long as the answer *asserts* a proposition that is relevant relative to the partially-resolved question.

¹³ A term due to von Fintel (2008).

¹⁴ We assume that question-answer pairs which display the pattern in (25) are ones in which the overtly-asked question is unambiguously the QUD. As we will see, not all question answer pairs obey Heim's generalization, and those that do not seem to be the ones in which QUD-accommodation is available.

¹⁵ Heim's original data involves polar questions, which (if resolved at all) are always resolved in full.

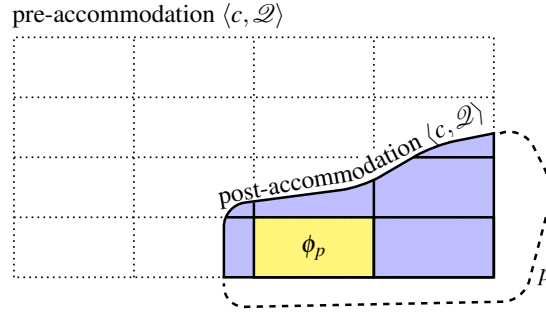


Figure 4 A question that is partially resolved by the presupposition of its answer

The prediction is indeed borne out, as illustrated by the examples in (28)-(29). In (28), for instance, B’s answer asserts that B has two sisters, and presupposes that she has one brother. Thus, the presupposition of B’s answer rules out only those cells of A’s questions in which B has zero or more than one brothers. Full resolution of the question is achieved only by combining the presuppositional and assertive content of B’s answer, which seems completely felicitous. Similarly in (29), B’s utterance presupposes that John came to the party; a presupposition that does not fully resolve the question as it says nothing of who else came. The information that no one else came is contributed by B’s assertion which therefore satisfies NR+INF. Heim’s observation is thus reduced to the order of evaluation of pragmatic constraints, giving rise to the generalization in (30).

- (28) a. A: How many siblings do you have?
b. B: Apart from my brother, I have two sisters.
- (29) a. A: Who came to the party?
b. B: Only John came. (Benbaji-Elhadad & Doron 2024)
- (30) **Heim’s Generalization (HG):** An utterance ϕ with presupposition p is only felicitous given a model $\langle c, \mathcal{Q} \rangle$ if p does not *fully* resolve $\mathcal{Q}$.

It is important to stress that the generalization in (30) regulates the relations between an utterance and the QUD relative to which it is presumed to be uttered, which can crucially differ from the overt question immediately preceding the utterance. This can be illustrated by constructing cases in which the generalization is apparently violated if evaluated relative to an overt question, but satisfied when context makes salient an alternative question relative to which the constraint is adhered to. For instance, in the context of (32-a), (31) is felicitous intuitively because, while Jim’s answer fully settles Paula’s question with its presupposition that mugs and beer glasses exist in the store, Jim seems to be answering what he takes to be the question Paula is actually interested in; namely, *where are the mugs and beer glasses?* The context in (32-b) does not similarly make salient a question not resolved by the presupposition of Jim’s utterance, since, in that context, Paula’s interests are narrowly focused on the existence of mugs and beer glasses.

- (31) Paula: *Do you sell mugs and beer glasses?*
Jim: *The mugs are in isle 3, the beer glasses — in the basement.*

- (32) Two contexts:
- a. **Context #1:** Paula is a consumer looking to purchase mugs and beer glasses in a shop owned by Jim. (discourse (31) ok)
 - b. **Context #2:** Paula is a tax auditor. Income from the sales of mugs and beer glasses is taxed differently from the sales of other products in Jim’s store, and Paula needs to find out whether Jim sells them. (discourse (31) odd)

A similar case is presented in (33)-(34). Given the context in (34-a), where the jurors are expected to deliver a verdict, there is no salient question that can be accommodated that would not be settled by the factual presupposition of “regret”. On the other hand, there is a very natural alternative question that Paula’s lawyer is expecting the judge to accommodate in scenario (34-b); namely, *what are Paula’s attitudes towards her bad deed, which might justify lenient sentencing?*

- (33) Judge: Did Paula murder John?
Addressee: She regrets doing so very much, your honor.

- (34) Two contexts:
- a. **Context #1:** Paula is under trial for the murder of John. The jurors reached a guilty verdict and have also independently formed the opinion that Paula regrets her actions very much, having inspected her behavior throughout the trial. The judge calls on the jury to deliver their verdict. (discourse (33) odd)
 - b. **Context #2:** Paula is found guilty of murdering John. The judge is discussing sentencing with her lawyer, who goes on and on about Paula’s outstanding personality in a way that seems to cast doubt of the jurors’ verdict. The judge, frustrated, wants the lawyer to acknowledge his client is the guilty. (discourse (33) ok)

These observations tie Heim’s generalization to the (possibly implicit) QUD and, we suggest next, provides us with a way to delineate the set of candidates for an accommodated QUD, once it is observed that the constraint has been violated.

The argument will go roughly as follows: First, we will show that when we modify the art dealer’s felicitous answer in (23-b-i), so that Paula’s explicitly formulated question (23-a) is fully resolved via a presupposition, the result is no longer felicitous; suggesting that HG has been violated. Given the analysis of HG above, this means that the QUD that is accommodated to render (23-b-i) NR-compliant must be fully resolved by the presupposition of the utterance. We will show that any QUD that meets this condition suffers from some fatal problem that rules it out as a candidate for the accommodated QUD, hence ruling out an account of (23-b-i) via QUD accommodation.

6.2 Diagnosing implicit QUDs with Heim’s generalization

Consider the variations in (35)-(36), inspired by Aravind, Fox & Hackl 2023, on the art dealer’s felicitous answer (23-b-i). They encode the same information as that conveyed by (23-b-i), but differ in the division of labor between assertion and presupposition. The cases in (35) involve a cleft structure that triggers an existential presupposition (i.e., *it is x that P*, presupposes *there is an x*

that *P*; Delin 1992; Rooth 1999).¹⁶ Those in (36) contain a definite description, which presupposes existence and uniqueness (i.e., *the P*, presupposes that a contextually unique *x* satisfies *P*). Thus, the presuppositions of (35)-(36) both settle Paula's explicitly-formulated question in (23-a): (35-a) and (36-a) presuppose that someone bought the painting, and (35-b) and (36-b) presuppose that Paula cannot buy it. The infelicity of these answers in our art gallery context, then, seems to be an instance of HG violation, in which a question is fully settled via presuppositional content.

- (35) a. #It was someone who collects 14th-C. Spanish art who bought the painting.
b. #It is because a 14th-C. Spanish art collector bought the painting that you cannot buy it.
- (36) a. #The person who bought the painting collects 14th-C. Spanish art.
b. #The reason you cannot buy the painting is that a 14th-C. Spanish art collector did.

Now, recall that our goal in this section is to rule out an alternative account of the data in (23) that adopts NR and appeals to QUD-accommodation to explain the felicity of NR-violating (23-b-i). If we are correct in attributing the oddness of (35)-(36) to a violation of HG, the implicit QUD involved in a QUD-accommodation account of (23-b-i) must satisfy the conditions in (37).

- (37) Conditions on the accommodated QUD in (23):
a. (23-b-i) must be equivalent to a subset of its cells (to satisfy NR).
b. The presuppositions of (35)-(36) must fully resolve it (to account for the HG violation).

These conditions rule out all but a particular class of partitions as candidates for the implicit QUD. First, any partition that in addition to a single cell equivalent to (23-b-i) contains at least one other cell compatible with a “no”-answer to Paula's original inquiry violates (37-b), as the presuppositions of (35-b)-(36-b) are essentially equivalent to the “no”-answer. This rules out the naturally-sounding English questions in (24), which all already presuppose the “no”-answer, but also the not-so-naturally-sounding question in (38) which, in addition to a cell corresponding to a positive answer to Paula's question, further partitions the “no”-cell.¹⁷

- (38) What is the proposition *p* such that, if I cannot buy the painting, *p* holds and explains why?

In fact, the only partitions that meet both conditions in (37) are those schematized in Figure 5. In addition to at least one cell compatible with a “yes”-answer to Paula's question, such partitions all contain exactly one “no”-cell, and in all of them that single “no”-cell is equivalent to (23-b-i). The disjunctive question in (39) exemplifies such a partition. Partitions of this kind satisfy (37-a) by *assumption*. They also satisfy (37-b), because (35)-(36) both presuppose the “no”-answer to Paula's explicit-question. After presupposition accommodation, then, such partitions contain a single cell, resulting in the violation of HG by the presuppositional examples.

- (39) Can I buy the painting or did someone who collects obscure 14th-C. Spanish art buy it?

¹⁶ For simplicity, we ignore the exclusive inference triggered by clefts, itself possibly presuppositional.

¹⁷ These cases exemplify what Isaacs & Rawlins (2008) call tripartitions in their work on conditional questions, which informs the account of relevance we offer in the following section.

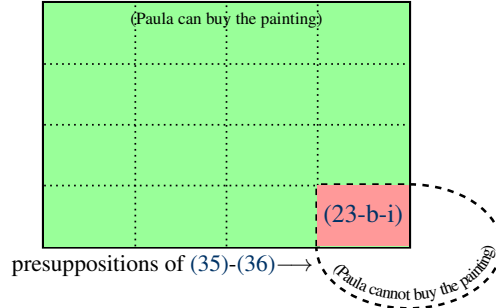


Figure 5 Partitions satisfying the conditions in (37)

Such partitions, however, are ruled out for an independent reason. While they do not presuppose a “no”-answer to Paula’s question, they do trigger the presupposition in (40) according to which either the “yes”-answer or (23-b-i) holds.¹⁸ But in any context in which (40) is taken for granted, (23-b-i) violates *global redundancy* principles governing discourse.

(40) Either Paula can buy the painting, or a collector of obscure 14th-C. Spanish art bought it.

We have already encountered one anti-redundancy principle in this paper, namely INF. This constraint rules out utterances that express what is already taken for granted. However, it has been observed that even utterances that express propositions *not* already taken for granted may exhibit symptoms of redundancy. If discourse participants in (41) do not take for granted that Mary is pregnant, (41) complies with INF but is still hopelessly deviant. Similarly, it is infelicitous for Jill to assert (42-a) in a context which establishes that only Spanish art collectors were in a position to buy the painting, even though her utterance does not express a proposition already taken for granted.

(41) #Mary is expecting a daughter and she is pregnant.

(42) **Context:** An art gallery invites collectors of 14th-C. Spanish art to an exclusive viewing of a painting. Paula and Jill know only Spanish art collectors were invited, but Paula does not know whether the painting was sold or stored away after the viewing. Jill says to Paula:

a. #Someone who collects 14th-C. Spanish art bought the painting.

Examples like these motivate theories of redundancy that rule out utterances based on competition with structurally-simpler, contextually-equivalent alternatives.¹⁹ On such theories, an utterance like (41) is ruled out because it has an equivalent structurally simpler alternative, derived by substituting a constituent in (41) with its sub-constituent; namely, (43). Similarly, (42-a) is infelicitous because (given that its context already entails that all potential buyers of the painting are collectors of Spanish art) it is equivalent to the structurally simpler (44). But then, if in (23) the response in (23-b-i) is uttered with respect to an implicit question that presupposes (40) (i.e.,

¹⁸ In fact, we can achieve a QUD of this kind even without appealing to QUD-accommodation. Instead, if Paula is allowed to accommodate the *proposition* expressed in (40), her explicitly formulated question already partitions the context into worlds in which she can buy the painting, and those in which (23-b-i) is true.

¹⁹ Meyer 2013; Katzir & Singh 2014 who build on Schlenker 2007, 2009; see overview in Mayr & Romoli 2016.

that Paula cannot buy the painting only if a collector of 14th-C. Spanish art bought it), then it too has a contextually-equivalent, structurally-simpler alternative; namely, (44). Thus, a partition presupposing (40) cannot be the implicit QUD relative to which (23-b-i) is evaluated, as this would result in a violation of global redundancy constraints.

(43) Mary is expecting a daughter.

(44) Someone bought it.

To summarize our argument in this section, we show that QUD-accommodation cannot solve the art gallery puzzle by demonstrating that each partition relative to which the felicitous utterance in (23-b-i) satisfies NR suffers from some fatal problem which rules it out as a candidate for the accommodated QUD. We divided those candidate partitions to two classes, eliminating each independently: (i) those that have more than one cell compatible with the “no”-answer to Paula’s question; and (ii) those that have only one cell compatible with the “no”-answer. Partitions in class (i) are ruled out as they do not account for our observation that responses which presuppose a “no”-answer to Paula’s question are infelicitous in the art gallery context. Partitions in class (ii) are ruled out because they inevitably render the felicitous response in (23-b-i) globally redundant. We thus reject an account of (23) that appeals to accommodation of an implicit QUD.

Before turning to the positive part of our proposal, we think a general note on this way of probing for the accommodated QUD is in order: As we have seen, QUD-accommodation can sometime serve to circumvent potential violations of HG; e.g., (31) and (33) are HG-compliant by virtue of the fact that the answers’ presuppositions do not fully resolve an accommodated QUD, and despite the fact that they *do* fully resolve the preceding overtly articulated questions. That (35)-(36) do seem to violate HG in the art gallery context (23), suggests that the QUD that would render them compliant with NR cannot be accommodated. This is what allows us to use HG in our argument against rescuing NR via QUD-accommodation. We are unable, at present, to provide a principled explanation as to why accommodation of certain QUDs is blocked in some contexts, including (23), but not in others. Our ambition here was more narrow; namely, to show that even in a context that restricts QUD accommodation in this way, we can still felicitously utter certain over-informative answers and are thus in need of an alternative resolution of the tension between the felicity of certain over-informative answers and a negative conception of relevance.

7 Negative relevance nonetheless

The strategy adopted above to resolve the art gallery puzzle takes NR as the correct formulation of relevance. This forces us to explain why the felicitous over-informative answer in (23-b-i) is, in fact, felicitous despite violating NR relative to the overtly articulated question (23-a), given that QUD-accommodation seems to be blocked independently. We suggest that a different kind of accommodation can help us rescue NR. The gist of the proposal is as follows: While the difference between the felicitous and infelicitous answers in the art gallery context is due to the availability or unavailability of accommodation, it is entire *segments of hypothesized discourse* that need to be accommodated, rather than just one QUD. Next, we spell-out what we mean by accommodation of entire discourse segments, and formulate a version of NR, which we term *dynamic*, that is sensitive

to such accommodation. We then return to our *probabilistic* notion of negative relevance, illustrating how it, too, can be saved by dynamicizing relevance in the sense discussed below.

7.1 The proposal

To get an intuitive sense of our proposal, suppose that by uttering (23-b-i) in response to Paula's question, the art dealer, instead of just saying *no*, is addressing a series of follow-up questions they assume Paula might ask given the answers they intend to provide. Possibly, the sequence of question-answer pairs in (45) is what they have in mind, and their response in (23-b-i) does not correspond only to an answer of Paula's explicitly-formulated question, but rather conveys the totality of information expressed by all the answers in (45).

- (45) *Can I buy the painting? No. Why not? Someone bought it. Why did they buy it? They collect Spanish art from the 14th-century.*

Importantly, each question-answer pair in this hypothetical sequence satisfies NR. Thus, our formulation of relevance will assume that some representation of a hypothesized discourse like (45) is part of the input to the relevance function, and demand that (i) an utterance convey the totality of information expressed by the answers in the hypothesized sequence relative to which it is evaluated; and (ii) every question-answer pair in this sequence satisfy NR as originally defined in (7).

Let us turn now to the art dealer's infelicitous answer (23-b-ii). Can Paula be expected to accommodate, in the context of (23), a similar discourse segment relative to which this answer is relevant in the sense gestured at above? This segment would have to contain a question relative to which it is negatively relevant that the person who bought the painting was born in September. Intuitively, this kind of exchange sounds quite odd. But more specifically, it is highly *unpredictable* from the point of view of Paula, who is expected to accommodate it. In other words, the art dealer is making a leap in the conversation by guessing what Paula's follow up questions will be, but to make that leap, he should presumably be convinced to a large enough degree that his guess is correct. We argue that the sequences in (46), as opposed to (45), cannot provide this level of conviction.

- (46) a. *Can I buy the painting? No. Why not? Someone bought it. What month were they born in? September.*
 b. *Can I buy the painting? No. Why not? Someone bought it. When is their birthday? September.*

Of course, the notion of predictability invoked here is based on somewhat vague intuitions.²⁰ However, there is a class of examples that may allow us a more direct access to the property that sets (45) apart from (46). We assume that the considerations that govern the availability of the accommodation of discourse chunks are similar to those that govern the accommodation of a QUD in the lack of an overt question. If this is the case, we should expect utterances which require an accommodated sequence like (45)-(46) to be just as felicitous or infelicitous as the corresponding explicit sequences of declarative utterances, in which the QUDs remain implicit.

²⁰ One way to go about explicating it is to analyze it within a model of cooperation between rational agents. While this seems to us a productive route to take, a detailed analysis of this sort is beyond the scope of this paper.

This is demonstrated in (47): The sequence in (47-b) patterns with the felicitous over-informative answer in (23-b-i), while the sequence in (47-c) patterns with the infelicitous one in (23-b-ii).

- (47) a. Can I buy the painting?
 b. No. Someone bought it. She's a collector of Spanish art from the 14th century.
 c. No. Someone bought it. #She was born in September.

The manifest oddness of the sequence in (47-c) is due to the last sentence, which provides information about the birth month of the buyer, requiring the addressee to accommodate a QUD which is highly unpredictable in this context. This, we argue, reflects the hidden process which takes place when evaluating the relevance of over-informative answers like in (23).²¹

Before providing a more formal account of a notion of relevance that appeals to these accommodated chunks of discourse, two notes are in order. First, we do not intend to provide an account of the difference in acceptability between (47-b) and (47-c). As far as we are concerned, this will be the role of a comprehensive future theory of discourse structure. Any such theory which rules out follow-up questions in a way that teases apart the sequences above could be plugged into the system we propose below.²² Rather, we argue that the difference in acceptability between (23-b-i) and (23-b-ii) is reducible to the difference in predictability between different follow-up questions, and that the way we formulate *relevance* must be able to take this difference into account.

Second, the idea that a pragmatic theory requires models of discourse that are more complex than the simple pair $\langle c, \mathcal{Q} \rangle$ assumed so far is far from novel. In fact, it has been entertained in a myriad of different ways. Roberts (1996/2012) posits models in which the QUD consists of a set of questions, “QUD-stacks”, rather than a single one. Kaufmann (2000), and Isaacs & Rawlins (2008) who build on his work, similarly replace our simple notion of context with a stack of states, consisting of the Stalnakerian context c and contexts derived from it by provisionally accepting further propositions *not* taken for granted. According to Farkas & Bruce (2010), after a question has been posited, our model of discourse contains a representation of the *potential context* derived by accepting each of the different answers to the question, and Onea (2016) suggests that *potential questions* that come up after a proposition has been expressed should also be included.

Our account of relevance is informed by each of these proposals, as the constraint we formulate

21 In some cases, a sequence of overt question-answer pairs sounds *less* natural than the corresponding sequence of declaratives, where the questions are left for the addressee to accommodate, as demonstrated by (i) below. We do not fully understand what makes the question *where were they from?* marked when it is uttered overtly, but seems to be easily accommodated. Notice that nothing we say in this paper rules out the existence of independent constraints on overtly-asked questions, and indeed, such a constraint might be at play here – for example, a social taboo on inquiring about certain personal information of a third party. We thank an anonymous reviewer for pointing this out to us.

(i) a. *Can I buy the painting?* No. Someone bought it. They were from the U.S.
 b. *Can I buy the painting?* No. *Why not?* Someone bought it. ??*Where were they from?* The U.S.

22 The desideratum for such a theory could be cast in terms of the notion of *potential questions*, due to Onea 2016: A sequence is predicted to be felicitous if each of its utterance raises a potential question that its succeeding utterance addresses, in Onea's sense. A theory of potential questions should predict that in the art gallery context, the fact that someone bought the painting licenses the question of why they bought it, but not of when they were born. As far as we can tell, Onea defines the technical tools one would need to spell out such a theory.

makes reference to a structured representation of potential sequences of *contexts*, *questions* and *answers* that can be accommodated by discourse participants. For ease of exposition, however, we try in what follows to remain agnostic as to the correct formal representation of discourse. Instead, we stick with the simple model $\langle c, \mathcal{Q} \rangle$ and formulate the minimal objects required for relevance to account for our data in (23) *within the definition of the constraint*. We note that given the richness of the proposals in the literature, such objects can be straightforwardly posited as part of the model itself by adopting tools already on the market.

Finally, we turn to our formal proposal, provided in (48), according to which a declarative utterance ϕ is *dynamically* negatively relevant given $\langle c, \mathcal{Q} \rangle$ if there is a sequence of potential questions, the first of which is $\mathcal{Q}$, and a series of answers to them, such that each answer expresses a *negatively relevant* proposition (in the sense of (6)) relative to its corresponding question.

- (48) **Dynamic Negative Relevance (d-NR):** A declarative utterance ϕ is only felicitous given a discourse model $\langle c, \mathcal{Q} \rangle$ if there exists a sequence of discourse models $\langle c_1, \mathcal{Q}_1 \rangle, \dots, \langle c_n, \mathcal{Q}_n \rangle$ such that $\langle c_1, \mathcal{Q}_1 \rangle = \langle c, \mathcal{Q} \rangle$, and a sequence of utterances $\phi_1, \dots, \phi_n$ such that:
- a. For every $1 \leq i \leq n-1$: $c_{i+1} = c_i + \phi_i$, namely the update of c_i with proposition ϕ_i which is required to obey constraints on update, particularly *Stalnaker's bridge* (26).
 - b. For every $1 \leq i \leq n$: ϕ_i satisfies NR relative to $\langle c_i, \mathcal{Q}_i \rangle$.
 - c. For every $1 \leq i \leq n$: $\mathcal{Q}_i$ is predictable enough given preceding discourse so that the addressee could be expected to accommodate it as a QUD.
 - d. $c + \phi = c + \phi_1 + \dots + \phi_n$.

We term this conceptualization *dynamic*, as it requires in (48-d) that the result of updating the original context c with ϕ be the same as updating c with each of the answers in the sequence, one at a time and in particular order.²³ Clause (48-d) meaningfully interacts with the clause preceding it (48-c), which requires that the accommodated sequence be predictable, in a way that allows us to account for the difference between the art dealer's potential responses in (23-b).

The sequence presented in (45), which we have claimed corresponds to the felicitous answer in (23-b-i), indeed satisfies the definition in (48) above. First, each answer in it satisfies NR relative to its preceding question, complying with (48-b). Second, each of its QUDs is predictable, as indicated by the felicity of (47-b), satisfying (48-c). Finally, sequentially updating any initial context set with the answers in the sequence inevitably results in the same context set as the one we get by a wholesale update of that initial context with the proposition denoted by (23-b-i), satisfying (48-d).

On the other hand, no sequence of questions and answers satisfies this definition for the infelicitous answer in (23-b-ii). In this case there is crucial tension between clauses (48-c) and (48-d); any sequence which encapsulates the information conveyed by the entire answer necessarily involves unpredictable discourse moves, as indicated by the infelicity of (47-c). One can thus characterize the infelicity of (23-b-ii) as stemming from the listener's inability to make sense of how the art dealer imagines the conversation continuing.

This new approach can still make sense of the observation presented in the previous section;

²³ While this bears some intuitive similarity to the framework of *dynamic semantics*, we are not committed to its ontological assumptions; specifically, that utterances themselves denote functions on contexts, and that context update can be done incrementally during the semantic derivation.

namely, that answers whose presuppositions resolve Paula’s overtly asked question, like (35)-(36), are infelicitous in the context of (23). Let there be an answer ϕ presupposing that Paula cannot buy the painting. Above, we have shown that Paula’s explicit question, inquiring whether she can buy the painting, is in fact the initial $\mathcal{Q}$ relative to which the relevance of succeeding utterances must be evaluated (this is not necessarily the general case; $\mathcal{Q}$ may well be itself an accommodated QUD). By (48-d), the update $c + \phi$ must be well-defined, and specifically must not violate SB (26). Therefore, for ϕ to constitute a felicitous utterance, c must entail that Paula cannot buy the painting. But if that is the case, then no answer can satisfy NR relative to $\langle c_1, \mathcal{Q}_1 \rangle$ as required by (48-b), given that c_1 entails a complete answer to $\mathcal{Q}_1$. The “trick” that does the job for us here is that while we allow NR to be evaluated only on the local updates within the accommodated discourse sequence, we require SB to be satisfied also globally, by the overtly expressed utterance.²⁴

This approach may also help us explain why a consensus regarding relevance has eluded the field thus far. While d-NR is a categorical constraint (any utterance either satisfies it or it does not, given particular parameters), it may nevertheless *seem* as if some utterances are *more relevant* than others. This is because the process by which sequences are accommodated is itself heavily context dependent, and the intuition that relevance itself is a gradient notion, reported in e.g., Russell 2012 might be attributed to the fact that some sequences are more easily accommodated than others.

For instance, suppose that in the art gallery context, instead of uttering the infelicitous (23-b-ii), the art dealer would have uttered (49). If the context we imagine is such that Paula has no interest in asking about the buyer’s appearance, this response seems just as bad as (23-b-ii); a judgment explained on our account by the difficulty to accommodate a sequence that includes a question inquiring about the facial hair of the buyer, e.g., (50). However, (49) would become felicitous – or at the very least much less odd – if we suppose the context is set up so that Paula and the art dealer take for granted that, if the painting has already been purchased by someone else, Paula will try to hunt down the buyer and pressure them to sell it to her. In that case, it seems that the art dealer can convey information about the appearance of the buyer in their articulated answer, rendering (49) felicitous. This intuition is predicted by clause (48-c) in our definition of d-NR; in such a context, the sequence in (50) can be reasonably expected by the art dealer to be accommodated.

(49) Someone with thick facial hair just bought it.

(50) *Can I buy the painting?* No. *Why not?* Someone bought it. *What did they look like?* They had thick facial hair.

While we agree that in the art gallery context, (23-b-i) seems better than (49), which is itself better than (23-b-ii), we derive these relative felicity judgments not by making relevance itself gradient, but rather, by appealing to the relative ease or difficulty of sequence-accommodation: It is fairly easy to accommodate a sequence in which a follow up question inquires about the buyer’s motivations for acquiring the painting, less so one including a question about the buyer’s appearances, and it is hopelessly difficult to accommodate a sequence that contains information

²⁴ Note that Heim’s observation in (25), which constitutes the backdrop to our discussion regarding (35)-(36), involves a violation of NR that can be resolved via d-NR: Answering whether Mary can join for lunch by saying that she has to pick up her sister from the airport, provides more information than the inquirer asked for. If we assume that a sequence that includes a *why*-question can be accommodated, d-NR is satisfied in this case.

about the date of birth of the buyer.

Finally, the dynamic mechanism we introduce may also explain why the debate between NR and PR, in particular, rages on. If one accepts d-NR, then despite the fact that the notion of relevance governing discourse is negative, the process of sequential evaluation of NR in accommodated sequences results in the admittance of utterances that seem to obey PR and *not* NR. NR and PR only come apart when a discourse sequence satisfying our definition in (48) cannot be accommodated, and given that, as discussed above, this accommodation process is highly context-dependent, it is not surprising that many in the literature find PR to be an acceptable approximation of relevance.

7.2 Back to probabilistic relevance

Recall that in Section 3.2, we faced a problem similar to the one in (23), which motivated our notion of dynamic relevance. We showed that while it is possible to formulate a probabilistic negative relevance constraint (p-NR, (17)) that is more permissive than its non-probabilistic counterpart, this constraint still seems to under-generate certain felicitous question-answer pairs. These are the discourses in (19)-(20), repeated below in (51). Assuming that the overtly asked question is the QUD, (51-b) is relevant according to p-NR, while (51-c) is irrelevant. The fact that both are judged felicitous may be viewed as evidence against p-NR.

- (51) a. Is it raining?
 b. It's probably raining.
 c. Well, John came in dripping wet.

Before we continue, it is worth stressing again that the introduction of probability into our relevance constraints is only useful under the assumption that no QUD-accommodation occurs in examples like (51). In section 6 we made an effort to show that the QUD in our art dealer context is indeed the overtly-asked question. As it turns out, the methodology that we used to diagnose the QUD in the art dealer case, namely, observing which presuppositions create infelicity when accommodated, is not fine-grained enough for the probabilistic cases in (51). For instance, a possible QUD that would make (51-b) relevant according to NR in its non-probabilistic formulation is *what are the chances that it's raining?* Since any complete answer to the overtly-asked question is also a complete answer to this potentially-accommodated QUD, a test based on informativity-driven infelicity – i.e., Heim's generalization (30) – cannot rule it out. In the following section, we explore some further empirical diagnostics that might shed light on which QUD (51-b) is supposed to answer. Currently, though, we concede that the choice of whether to analyze (51-b) as stemming from probabilistic relevance or from QUD-accommodation is somewhat arbitrary.

Let us assume for the purpose of our discussion here that relevance is indeed probabilistic, and that (51-b) is evaluated relative to the overtly-asked question. The issue of (51-c) remains; since it is not relevant according to p-NR, should its felicity be taken as a counterexample to p-NR? No, as dynamic relevance gives us an account of (51-c) within the p-NR view, in the same way it accounts for felicitous over-informative answers, e.g., (23-b-i), within the NR view. More specifically, the relevance of (51-c) is evaluated with respect to a hypothesized sequence along the lines of (52). The additional information that this sentence provides is a relevant answer to a natural followup question

inquiring about the source of the information conveyed by the speaker, and it is therefore felicitous.

(52) *Is it raining? Probably. How do you know? Because John came in dripping wet.*

One piece is still missing, however. If the felicity of (51-c) indeed depends on accommodation of a followup sequence, we expect to find cases that are over-informative in the same way (namely asymmetrically entail a negatively-relevant probabilistic answer), but seem infelicitous as they do not correspond to any natural followup sequence. To make our case even stronger, we should present a probabilistic counterpart of our art dealer example; i.e., a minimal pair of over-informative answers that differ only in the availability of such a sequence. This is attempted in (53).

(53) **Context:** Colleagues are trying to schedule a Zoom call to discuss an important work issue.

- a. A: Will Mary be able to have the call today?
- b. B: She flew to Japan yesterday until June 11th.
- c. B: #She flew to Japan yesterday on a Boeing 747.

Both answers in (53) asymmetrically entail that Mary will probably not attend the call, and so without dynamic relevance, we would expect them to pattern in the same way. In fact, however, (53-c) is significantly degraded relative to (53-b). Dynamic relevance provides us with an explanation: It is again about the accessibility of accommodation of each of the corresponding sequences. While the sequence in (54-a) (corresponding to (53-b)), and specifically the question *when will she be back?* is natural and predictable, the last question in the sequence in (54-b) (corresponding to (53-c)) is quite unnatural. To diagnose this notion of predictability in a clearer way, we can again look at the corresponding sequences of declarative utterances where the QUDs are left for the addressee to accommodate (55). The infelicity of the last sentence in indicates that indeed, the potential follow-up question is unpredictable, and therefore difficult to accommodate.

(54) a. *Will Mary be able to have the call today?* Probably not. *Why not?* Because she flew to Japan. *When will she be back?* June 11th.
b. *Will Mary be able to have the call today?* Probably not. *Why not?* Because she flew to Japan. *what type of plane did she fly on?* Boeing 747.

(55) a. Will Mary be able to have the call today?
b. Probably not. She flew to Japan. She'll be back by June 11th.
c. Probably not. She flew to Japan. #It was on a Boeing 747.

8 Probabilistic vs non-probabilistic negative relevance

We hope to have motivated two points in this paper thus far: First, that there is a coherent notion of probabilistic relevance that is negative in the Lewisian sense in that it does not make any distinctions among primitive elements residing in the same cells in a partition. And, second, that negative notions of relevance are better than positive ones, since the apparent under-generation challenge they face is less detrimental than the over-generation challenge faced by their positive counterparts. As we note in the introduction, these points do not by themselves distinguish between the probabilistic and non-probabilistic conception of negative relevance; both can avoid over-generation once we

accept that they can be sequentially evaluated relative to an accommodated segment of discourse. Next, however, we discuss some tentative conceptual and empirical considerations in favor of non-probabilistic NR (7) over its probabilistic counterpart p-NR (17), even on the assumption that both can be dynamicized as suggested above.

Recall that, even though we have not presented an argument against the introduction of probabilities *per se*, we did argue that to the extent that they are introduced into the definition of relevance in the first place, this should be via adopting a model in which the notion of content is itself probabilistic, with probability measures replacing worlds as the primitive elements that make up propositions. If we are correct, then, proponents of probabilistic relevance are committed to a rather radical departure from truth-conditional semantics, where content is a set of worlds.

But, given that an overtly asked question does not always constitute the QUD relative to which an answer's relevance is evaluated – committing everyone to some version of QUD-accommodation (cf. our discussion in Section 4) – it is unclear whether p-NR and NR make different empirical predictions at all. If both generate the same data, differing only in which question-answer pairs are felicitous due to the evaluation of relevance relative to the overt question, and in which an intermediate accommodation step is required, there is an argument from Occam's razor to be made against the probabilistic version of the constraint; it gives us very little – perhaps even nothing at all – at a high theoretical cost.²⁵ Thus, on the conceptual side, the way probabilities must be introduced for the result to correctly predict the distribution of over-informative answers, requires more theoretical scaffolding than proponents of probabilities are usually willing to stomach.

Finally, if we can develop some tests that indicate when accommodation is appealed to and when it is not, we should also be able to tease p-NR and NR apart on empirical grounds as well. In particular, if we are correct that, indeed, in the art gallery context, the dealer is expecting Paula to accommodate a hypothesized chunk discourse, then we expect the interaction between them, and more generally any interaction in which interlocutors' utterances are felicitous by virtue of accommodation, to show certain hallmarks of accommodation.

It is a general property of discourse that not every expectation that something be accommodated must be respected (Szabó 2006; von Fintel 2008). In the domain of presuppositions, this is illustrated by the observation in Shanon (1976), termed the *Hey, wait a minute!* test by von Fintel (2004), that an addressee may “legitimately complain” when the speaker presupposed something not previously taken to be an established fact. This diagnostic sets apart presupposed and asserted content; the addressee's complaint cannot target a speaker's assertion, as illustrated in (56).

- (56) A: The mathematician who proved Goldbach's Conjecture is a woman.
 B: Hey, wait a minute! I had no idea that someone proved it.
 B': #Hey, wait a minute! I had no idea it was a woman. (von Fintel 2004)

We can develop an equivalent test for accommodation of QUDs or discourse segments. In particular, the idea is that, when a speaker responds to an explicitly articulated question $\mathcal{Q}$ using an utterance that does not satisfy NR relative to $\mathcal{Q}$, their addressee may indicate their refusal to engage in accommodation by uttering: *I didn't ask you $\mathcal{Q}'$, I asked you $\mathcal{Q}$!*, where $\mathcal{Q}'$ is the question, or

²⁵ Of course, if like Swanson (2006) and Moss (2015a,b, 2018), one takes the probabilistic view of content underlying p-NR to be independently motivated, then this economy argument against p-NR is obviated.

part of a sequence thereof, that should be accommodated, relative to which ϕ *does* satisfy NR. We call this the *I Didn't Ask You*, or IDAY, test. Take, for instance, the discourse in (57), partially repeated from (22), which we take to be a paradigmatic case of QUD-accommodation, given that B's response is not relevant relative to the overtly asked question even on the weakest version of the constraint.

- (57) A: Is it going to rain?
B: We'll be spending most of our time indoors.
A: I didn't ask you where we will be spending our time, I (only) asked you whether it is going to rain!

Since B expects A to accommodate a different question relative to which their utterance satisfies NR, we expect that A should be able to resist accommodation and, in fact, A can do so.

Another case that we argue must involve accommodation is (31), repeated in (58), where the presupposition of Jim's answer fully resolves Paula's question. Given the apparent violation of Heim's constraint which disallows this method of question resolution (30), (58) is felicitous only due to accommodation of a question *not* resolved by the presupposition; something along the lines of *where are the mugs and beer glasses?* And as predicted, Paula can resist accommodating the question Jim is expecting her to, as below. Importantly, an inquirer cannot similarly resist when her addressee responds to her question by providing, say, a partial answer rather than an exhaustive one. This suggests that what is resisted really are expectations of accommodation.

- (58) Paula: Do you sell mugs and beer glasses?
Jim: The mugs are in isle 3, the beer glasses — in the basement.
Paula: I didn't ask you *where* the mugs and beer glasses are, I (only) asked *whether* you sell them!
- (59) A: Who is coming to the party?
B: Mary is.
A: #I didn't ask you whether Mary is coming, I asked you who is coming!

Now, recall that NR and p-NR generate the same data, differing only on when accommodation – either of a QUD or of a sequence – must be appealed to. While both agree that accommodation is required when “John came in dripping wet” is used to address the explicit question “is it raining?”, NR also requires accommodation when “probably” is used to answer this question, while p-NR does not. Interestingly, applying the diagnostic above suggests that, indeed, accommodation is appealed to in such cases and can be resisted. We propose this as a *tentative* argument in favor of non-probabilistic negative relevance; tentative, as further work beyond the scope of this paper is required to establish that IDAY really serves to diagnose accommodation in the question domain.

- (60) A: Is it raining?
B: Probably.
A: I didn't ask you what are the chances that it is raining, I asked you whether it was! (If you don't know, just say so instead of wasting my time.)

Before concluding, we would like to address how the data presented above interacts with a radically different approach to relevance, which some take to be the “null hypothesis”. The view, whose articulation here is due to Benjamin Spector (p.c.), is that – given a view of discourse as rational cooperation – the felicity of a declarative depends directly on the extent to which the information it conveys helps discourse participants promote their goals; e.g. to answer the questions being discussed or make decisions that depend on the answers.²⁶ This view rejects the notion of relevance as a categorical constraint, and attempts to reduce the patterns we discuss throughout this paper to domain-general rational considerations. It is therefore more parsimonious, in some natural sense, than the other views we present here. Let us term this view Relevance Reductionism (RR).

The way RR could account for the art dealer contrast (23) is through an the interaction of the pressure described above to contribute as much useful information as possible with an opposing pressure to be as economical as possible in one’s choice of words.²⁷ The general bridging principle is that a sentence is judged infelicitous if there exists an alternative sentence that satisfies both of these pressures to a higher degree. In the case of the felicitous response by the art dealer (23-b-i), any simplification of the sentence would diminish the amount of useful information that it conveys. In the case of the infelicitous response, however, the RC *who was born in September* could be deleted, resulting in a more economical sentence that contains as much useful information. The infelicity of (23-b-ii) is thus explained by the availability of a more optimal alternative.

The debate between RR and categorial approaches to relevance ultimately boils down to the question of the role of relevance in the cognitive architecture. Is it a part of a language-specific process, or a derivative result of domain-general cognition? A conclusive answer to this question is naturally beyond the scope of this paper. However, at least two points made in this paper can be used as arguments against RR. First, the RR view does not seem to predict the results of the IDAY test presented above; if all we demand from utterances is that they optimize some trade-off between informativity and economy, it is unclear why, for example, partial answers (59) should pattern differently from probabilistic answers (60). The challenge that the IDAY test poses for RR can be further demonstrated by comparing two utterances – one satisfying NR, the other over-informative – that are (presumably) equally economical. This is given in (61)–(62) below. Since *I visited France* and *I visited Paris* are equally economical by any reasonable metric, and (61) is strictly weaker than (62), we expect the latter to be at least as optimal as the former. The fact that the IDAY rejection is felicitous with (62) but not with (61) is therefore completely unexpected in the RR view.

- (61) A: Which countries did you visit this summer?
 B: I visited France.
 A: #I didn’t ask you whether you visited France, I asked you which countries you visited!
- (62) A: Which countries did you visit this summer?
 B: I visited Paris.
 A: I didn’t ask you which cities you visited, I asked you which countries you visited!

26 This view is embedded in the Rational Speech Act framework, when it incorporates QUDs; e.g., Kao, Wu, Bergen & Goodman 2014; Lassiter & Goodman 2017; Spector 2017; Degen 2023.

27 These pressures correspond vaguely to Grice’s (1975) Maxims of quantity and manner, respectively. To be as charitable as possible to the RR approach, we abstract away from the exact definitions of *useful* and *economical*.

Second, it is not clear to us how RR can predict the observation in section 6, that the art dealer’s felicitous response to Paula (23-b-i) is rendered infelicitous once we repackage the information in it so that Paula’s overtly asked question is resolved via a presupposition. This is exemplified in (63), which conveys *the same information*, and is *more concise than* (23-b-i). Recall that we reduce the infelicity of (23-b-i) to a bleeding relation between presupposition accommodation and relevance; but if relevance itself is reduced to domain-general considerations, it is not obvious to us that it should be sensitive to the difference between an accommodated presupposition and assertion.

(63) #The buyer collects Spanish art from the 14th century.

Finally, recall that the way RR accounts for the felicity of (23-b-i) is by assuming that the information it conveys beyond the “no” answer to Paula’s question, namely that someone bought the painting and that they were a collector of Spanish art from the 14th century, is useful, which makes the art dealer’s response more informative than its more economical alternatives (such as *no*). Thus, even if relevance considerations could in principle be made to distinguish between assertion and presupposition under RR, the assertion in (63) must itself be considered useful to at least some degree the art gallery context, and the sentence is predicted wrongly to be felicitous. Beyond these points, we leave for future research a more detailed discussion of RR.

9 Conclusion

This paper consists of four main points. First, we survey some of the prominent formulations of relevance in the linguistic literature, and analyze their logical properties and relation to each other. Based on this survey, we argue that formulations of relevance may vary with respect to two properties: positive vs. negative, and probabilistic vs. non-probabilistic. This gives rise to a four-cell paradigm of formulations, but interestingly, standard approaches only map onto three of them, with the negative probabilistic one left empty. We point out that this is not a necessary result of the logic of relevance, and propose a negative probabilistic formulation of relevance (p-NR), based on a modification to our semantic system developed in Moss (2015b). The fact that p-NR demands a departure from standard assumptions regarding truth-conditional semantics makes it clear why such a constraint is not common in the literature. We nevertheless show that it is consistent, and indeed natural, given a particular view of the role that probability plays in the way we represent meaning.

Second, we criticize the methodology often used in the debate over relevance; namely, relying on the felicity of question-answer pairs to diagnose for the correct formulation of the constraint. We argue that such data is confounded by the fact that relevance is evaluated relative to the QUD, which may differ from an overtly-asked question. To overcome the confound, we present a novel way to probe for the QUD of a given declarative utterance in a given context. Our method relies on what we term *Heim’s Generalization* (HG) – an observation by Irene Heim that ties the infelicity of certain presuppositional sentences to certain logical properties of the QUD. More concretely, HG states that a declarative sentence is infelicitous if its presupposition fully resolves the QUD, a pattern that stems from the assumption that relevance is evaluated *after* presupposition accommodation has taken place. We show that this allows us to reduce the space of possible QUD’s, and in some cases narrow it down to only one candidate. We hope that this method will be put to use in the future for

other cases which require that we control for the QUD.

Third, we propose an account of relevance that relies in its core on the classic negative formulation due to Lewis (1988a,b), combined with a novel view of the way relevance is evaluated. The account is motivated by minimal pairs consisting of two over-informative sentences, one of which is felicitous while the other is infelicitous. This is a surprising pattern, since all formulations of relevance discussed in this paper systematically predict over-informative sentences to pattern uniformly. However, while positive formulations of relevance seem unable to delineate the felicitous set of over-informative sentences, we show that negative ones can predict the facts once we assume relevance is evaluated “dynamically” on accommodated chunks of discourse. We thus provide an argument in favor of negative formulations and against positive ones, while claiming that the basic architecture of relevance is different from what has previously been assumed.

Finally, we consider the second dimension relative to which relevance formulations may vary; namely, probabilistic vs. non-probabilistic. While our argument in favor of negative relevance does not rule out its probabilistic variant, we suggest that there are tentative conceptual and empirical considerations against it. First, the costly theoretical scaffolding required to even formulate a useful probabilistic notion of negative relevance, paired with the fact that the non-probabilistic version can generate much of the same data given the availability of QUD accommodation, provides us with an economy argument against the introduction of probabilities. Second, we take first steps towards developing an empirical diagnostic for accommodation which suggests that, indeed, we have evidence that accommodation is appealed to when sentences only satisfy probabilistic negative relevance relative to the question that precedes them.

The claim that the strongest formulation of relevance – namely, NR – is, fundamentally, the version that plays a role in evaluating felicity of utterances is, we think, a welcome one for independent reasons. The tension between different conceptions of relevance is particularly pertinent for processes that are sensitive to formal alternatives, and many researchers rely on NR as the deciding factor determining which formal alternatives to a given utterance are considered in the calculation of exclusive inferences; e.g., Magri (2009), Fox & Katzir (2011), Križ & Spector (2021), Bar-Lev (2021, 2023). The assumption that NR is the right notion of relevance is problematized by examples like the ones presented in this paper, which feature felicitous utterances that do not, straightforwardly, satisfy NR. Our proposed *dynamic relevance* can serve to smooth out this tension, since we show that one can maintain a version of NR while still predicting examples that NR under-generates to be felicitous, thus strengthening the theoretical basis that this literature relies on.

Many issues remain. The empirical picture regarding both under-informative and over-informative sentences remains somewhat unclear to us. We rely in this paper on a handful of examples that we take as representative of some broad generalizations involving the distribution of these kinds of utterances, but it seems to us very likely that our discussion ignores other important generalizations. A more systematic investigation of the empirical landscape is required, and we hope that some of the tools developed here could be of help. Moreover, our discussion of the different formulations of relevance is based on a four-class classification, as discussed above. But, of course, one can draw many finer-grained distinctions within each class of formulations, which we ignore here for simplicity. We leave that for future research as well, in the hope that this paper could serve as a basis for an explicit debate, which is largely missing from current literature.

References

- Agha, Omar & Alex Warstadt. 2020. Non-resolving responses to polar questions: A revision to the qud theory of relevance. In *Proceedings of Sinn und Bedeutung*, vol. 24 1, 17–34.
- Aravind, Athulya, Danny Fox & Martin Hackl. 2023. Principles of presupposition in development. *Linguistics and Philosophy* 46(2). 291–332.
- Bar-Lev, Moshe E. 2021. An implicature account of homogeneity and non-maximality. *Linguistics and Philosophy* 44. 1045–1097.
- Bar-Lev, Moshe E. 2023. Obligatory implicatures and the relevance of contradictions. *Manuscript, TAU*.
- Beaver, David I & Brady Z Clark. 2009. *Sense and sensitivity: How focus determines meaning*. John Wiley & Sons.
- Benbaji, Ido & Omri Doron. 2023. Adversative only is only only. In *Proceedings of Sinn und Bedeutung*, vol. 27, 64–80.
- Benbaji-Elhadad, Ido & Omri Doron. 2024. How is presupposition accommodation constrained, again? In *Proceedings of the Amsterdam Colloquium*, 22–28.
- Büring, Daniel. 2003. On d-trees, beans, and b-accent. *Linguistics and philosophy* 26. 511–545.
- Carlson, Lauri. 1983/2012. *Dialogue games: An approach to discourse analysis*. Springer Science & Business Media.
- Carnap, Rudolf. 1950. *Logical foundations of probability*. Chicago, IL: University of Chicago Press.
- Collingwood, Robin George. 1940/2001. *An essay on metaphysics*. Oxford University Press.
- Degen, Judith. 2023. The rational speech act framework. *Annual Review of Linguistics* 9(1). 519–540.
- Delin, Judy. 1992. Properties of it-cleft presupposition. *Journal of semantics* 9(4). 289–306.
- Doron, Omri & Jad Wehbe. 2022. A constraint on presupposition accommodation. In *23 rd Amsterdam Colloquium*, 405.
- Farkas, Donka F & Kim B Bruce. 2010. On reacting to assertions and polar questions. *Journal of semantics* 27(1). 81–118.
- Feinmann, Diego. 2020. *Lexical illusions, (non-) maximality, and invisible gaps*: University of Sheffield PhD dissertation.
- von Fintel, Kai. 2004. Would you believe it? the king of france is back! (presuppositions and truth-value intuitions). In Marga Reimer & Anne Bezuidenhout (eds.), *Descriptions and Beyond*, Oxford: Oxford University Press. doi:10.1093/oso/9780199270514.003.009.
- von Fintel, Kai. 2008. What is presupposition accommodation, again? *Philosophical perspectives* 22. 137–170.
- von Fintel, Kai & Irene Heim. 2021. *Intensional Semantics*. MIT lecture notes.
- von Fintel, Kai & Irene Heim. 2021. Intensional semantics. *Unpublished lecture notes*.
- Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural language semantics* 19. 87–107.
- Grice, Herbert P. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Speech acts*, 41–58. New York: Academic Press.
- Groenendijk, Jeroen Antonius Gerardus & Martin Johan Bastiaan Stokhof. 1984. *Studies on the*

- Semantics of Questions and the Pragmatics of Answers*: Univ. Amsterdam PhD dissertation.
- Haslinger, Nina. 2023. *Pragmatic constraints on imprecision and homogeneity*: University of Göttingen PhD dissertation.
- Hawkins, Robert & Noah Goodman. 2017. Why do you ask? the informational dynamics of questions and answers .
- Hyska, Megan Alexandra. 2015. *Discourse-level information structure and the challenge of metadiscursives*: PhD dissertation.
- Hájek, Alan. 2023. Interpretations of Probability. In Edward N. Zalta & Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University Winter 2023 edn.
- Isaacs, James & Kyle Rawlins. 2008. Conditional questions. *Journal of Semantics* 25(3). 269–319.
- Jacobson, Pauline. 2016. The short answer: Implications for direct compositionality (and vice versa). *Language* 92(2). 331–375.
- Kao, Justine T, Jean Y Wu, Leon Bergen & Noah D Goodman. 2014. Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences* 111(33). 12002–12007.
- Katzir, Roni & Raj Singh. 2014. Hurford disjunctions: embedded exhaustification and structural economy. In *Proceedings of Sinn und Bedeutung*, vol. 18, 201–216.
- Kaufmann, Stefan. 2000. Dynamic context management. *Formalizing the dynamics of information* 171–188.
- Križ, Manuel & Benjamin Spector. 2021. Interpreting plural predication: Homogeneity and non-maximality. *Linguistics and Philosophy* 44. 1131–1178.
- Lassiter, Daniel. 2011. *Measurement and modality: The scalar basis of modal semantics*. Ph. D. thesis, New York University.
- Lassiter, Daniel & Noah D Goodman. 2017. Adjectival vagueness in a bayesian model of interpretation. *Synthese* 194. 3801–3836.
- Lewis, David. 1979. Scorekeeping in a language game. *Journal of philosophical logic* 8. 339–359.
- Lewis, David. 1988a. Relevant implication. *Theoria* 54(3). 161–174. doi:[10.1111/j.1755-2567.1988.tb00716.x](https://doi.org/10.1111/j.1755-2567.1988.tb00716.x).
- Lewis, David. 1988b. Statements partly about observation. *Philosophical papers* 17(1). 1–31.
- Magri, Giorgio. 2009. A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural language semantics* 17. 245–297.
- Mayr, Clemens & Jacopo Romoli. 2016. A puzzle for theories of redundancy: Exhaustification, incrementality, and the notion of local context. *Semantics and Pragmatics* 9. 7–1.
- Merchant, Jason. 2005. Fragments and ellipsis. *Linguistics and philosophy* 27. 661–738.
- Meyer, Marie-Christine. 2013. *Ignorance and grammar*: Massachusetts Institute of Technology PhD dissertation.
- Moss, Sarah. 2015a. Credal dilemmas. *Noûs* 49(4). 665–683.
- Moss, Sarah. 2015b. On the semantics and pragmatics of epistemic vocabulary. *Semantics and pragmatics* 8. 5–1.
- Moss, Sarah. 2018. *Probabilistic knowledge*. Oxford University Press.
- Onea, Edgar. 2016. *Potential questions at the semantics-pragmatics interface*. Brill.
- Roberts, Craige. 1996/2012. Information structure: Towards an integrated formal theory of prag-

- matics. *Semantics and pragmatics* 5. 6–1. doi:[10.3765/sp.5.6](https://doi.org/10.3765/sp.5.6).
- Rooth, Mats. 1999. Association with focus or association with presupposition? *Focus: Linguistic, cognitive, and computational perspectives* 232–244.
- Russell, Benjamin. 2012. Probabilistic reasoning and the computation of scalar implicatures. *Unpublished doctoral dissertation, Brown University* .
- Schlenker, Philippe. 2007. Transparency: An incremental theory of presupposition projection. In *Presupposition and implicature in compositional semantics*, 214–242. Springer.
- Schlenker, Philippe. 2009. Local contexts. *Semantics and pragmatics* 2. 3–1.
- Shanon, Benny. 1976. On the two kinds of presuppositions in natural language. *Foundations of language* 14(2). 247–249.
- Spector, Benjamin. 2017. The pragmatics of plural predication: Homogeneity and non-maximality within the rational speech act model. In *Proceedings of the Amsterdam Colloquium*, 435–444.
- Stalnaker, Robert. 1977. Pragmatic presuppositions. In *Proceedings of the Texas conference on performatives, presuppositions, and implicatures*. Arlington, VA: Center for Applied Linguistics, 135–148. ERIC.
- Stalnaker, Robert C. 1978. Assertion. In *Pragmatics*, 315–332. Brill.
- Swanson, Eric Eric Peter. 2006. *Interactions with context*: Massachusetts Institute of Technology PhD dissertation.
- Szabó, Zoltán Gendler. 2006. Comments on kai von fintel’s paper. Unpublished manuscript, revised version of comments given at ohio workshop on presupposition accommodation. <https://bpb-us-w2.wpmucdn.com/campuspress.yale.edu/dist/d/1148/files/2015/10/On-Presupposition-Accommodation-2bdedla.pdf>.
- Van Rooy, Robert. 2004. Utility, informativity and protocols. *Journal of philosophical logic* 33. 389–419.
- Warstadt, Alex & Omar Agha. 2022. Testing bayesian measures of relevance in discourse. In *Proceedings of Sinn und Bedeutung*, vol. 26, 865–886.
- Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11). 916–937.