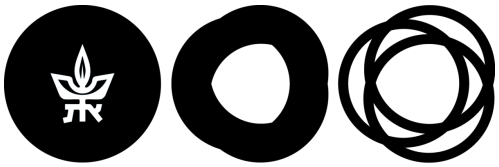


Linguistics Department
The Lester and Sally Entin Faculty of Humanities

Iterated Rationality Models and Conjunctive Readings of Disjunctions

MA thesis submitted by

Alma Frischoff

Under the supervision of

Prof. Roni Katzir Dr. Moshe Bar-Lev

September 2025

Abstract

The literature varies in its views on the nature of Scalar Implicatures (SIs) and their derivation. According to the pragmatic approach, SIs are a pragmatic phenomenon derived by a general reasoning mechanism, affected only by the dynamics of conversation. According to the competing grammatical approach, SIs are logical entailments derived compositionally. In recent years, there has been an increasing interest in a group of pragmatic models, referred to as iterated rationality models (IRMs), which utilize an iterative, often probabilistic, approach to general reasoning. The present work proposes a modification of IRMs that replaces the common naive-speaker assumption with a new perspective of the speaker, based on weighted probabilities. This modification resolves a major issue for such models in deriving conjunctive readings for disjunctions (e.g., free choice inferences), as observed by Franke (2009, 2011), van Rooij (2010) and Fox and Katzir (2021). Then, we show that the proposed model, due to its success in deriving conjunctive readings of disjunctions, enhances the plausibility of a modular IRM, which gives the correct prediction in a variety of cases where previous non-modular IRMs failed.

Contents

1. Introduction

Scalar Implicatures (SIs), such as the inference from 'some' to 'some but not all' in (1), have been extensively studied in recent decades, both theoretically and experimentally.

(1) John did some of the homework
$\rightsquigarrow$ John did some but not all of the homework

In the pursuit of a theory that accounts for SIs, two principal approaches have been developed: (a) the *pragmatic approach*, which posits that SIs are a pragmatic phenomenon arising at the speech act level, governed by conversational principles (Horn 1972, Grice 1989, among others); and (b) the *grammatical approach*, which contends that SIs are logical entailments derived compositionally within the grammar, usually by a covert exhaustivity operator (Fox 2007, Bar-Lev and Fox 2017, 2020, among others). The present work focuses on a specific subset of pragmatic theories, termed as *Iterated Rationality Models* (IRMs; Benz 2006, Benz and Van Rooij 2007, Franke 2009, 2011, Frank and Goodman 2012, Rothschild 2013, Bergen et al. 2016, among others). In recent work, Fox and Katzir (2021) compare the two approaches and provide two arguments supporting a grammatical approach to SIs, based on conjunctive readings of disjunctive sentences. Firstly, they argue that IRMs are unable to derive *conjunctive readings of disjunctions* (CRDs) with more than two disjuncts. Secondly, they argue that IRMs are extremely sensitive to prior probabilities in the case of CRDs, and conclude that SI computation is modular.

The present work proposes to rethink the initial (naive) speaker assumption to deal with CRDs, offering a simple and general solution to this failure of existing IRMs by replacing this assumption with a more nuanced notion of speaker. These results undercut Fox and Katzir's (2021) first argument against IRMs, based on the inability to provide a general derivation of CRDs. The proposed IRM also provides a foundation for resolving several challenges faced by existing IRMs. First, it successfully derives CRDs when they are embedded under a universal quantifier (Universal CRD). This has been a significant challenge for both grammatical and pragmatic theories (see Chemla 2009, van Rooij 2010, Franke 2011), which so far has only been accounted for by the grammatical theory of Bar-Lev and Fox (2017, 2020). Second, the current proposal is crucial to address the issue of sensitivity to priors in IRMs. As noted by Fox and Katzir (2021), while a modular IRM can address this problem, it remains ineffective unless IRMs are capable of deriving CRDs. Therefore, the proposed model changes the plausibility of a modular IRM, which gives the correct prediction in a variety of cases where previous non-modular IRM fail.

The work is structured as follows. I start by describing in Sect. 2 the grammatical and IRM approaches and their prediction for CRDs in more detail, and specifying the challenge faced by IRMs. In Sect. 3, I present my proposal, explain the intuition behind it, and show its predictions. I first show how it overcomes this challenge and then move to the challenge posed by Universal CRDs. In Sect. 4, I present the problem of prior sensitivity, focusing on CRDs and reasoning in reference games, and discuss the role of modularity in addressing the challenge in these cases.

2. Background

2.1 Conjunctive Readings of Disjunctions (CRDs)

Disjunction gives rise to a conjunctive interpretation in multiple configurations where the alternatives are not closed under conjunction (i.e. 'A' and 'B' are alternatives, but 'A and B' is not). Such inferences have been argued to account for the behavior of Warlpiri connectives, (Bowler 2014), disjunction in child language (Singh et al. 2016) and Free Choice disjunctions (FC; Kamp 1974) like (2).

(2) John is allowed to eat an apple or a banana
   a. $\rightsquigarrow$ John is allowed to eat an apple
   b. $\rightsquigarrow$ John is allowed to eat a banana

This observation generalizes to any $n \geq 2$ disjunctions within such configurations, and to k-out-of-n disjunctions, i.e., disjunctions of 'A_1',...,'A_k' given that $A_1, \ldots, A_n$ are salient in the context. In these cases, the disjunction gives rise to a conjunctive inference of the disjuncts and an exhaustive inference that excludes the other salient alternatives. For instance, if there are 3 options for dessert – apples, bananas, and cherries – the inferences are:

(3) John is allowed to eat an apple or a banana
   a. $\rightsquigarrow$ John is allowed to eat an apple
   b. $\rightsquigarrow$ John is allowed to eat a banana
   c. $\rightsquigarrow$ John is not allowed to eat cherries

Such inferences have been argued to be SIs (Kratzer and Shimoyama 2002, Alonso-Ovalle 2005), and therefore we can formulate the following desideratum for theories of SIs:

(4) DESIDERATUM FOR THEORIES OF SCALAR IMPLICATURES: for any $1 \leq k \leq n$, a disjunction 'A_1 or ... or A_k' with A_{k+1},...,A_n as other salient alternatives, derive:
   a. A conjunctive inference of k-out-of-n $\bigwedge\{A_1, \ldots, A_k\}$
   b. An exhaustive inference of n-k-out-of-n $\bigwedge\{\neg A_{k+1}, \ldots, \neg A_n\}$

Focusing on the basic (i.e., 2-way disjunctions such as (2)), CRDs have served as a central criterion for comparing the grammatical and pragmatic approaches: while accounted for under the grammatical theories, they were a major problem for early pragmatic theories, and as such were a key argument for the former in the literature (Fox 2007). However, a prominent body of work on Iterated Rationality Models (IRMs; following the terminology of Fox and Katzir 2021) later offered a pragmatic derivation of these inferences (Franke 2009, van Rooij 2010).

2.2 The grammatical approach and CRDs

Before describing IRMs and their predictions for CRDs, I will briefly review the grammatical approach and how it derives CRDs. According to the grammatical approach, SIs are derived by a covert exhaustivity operator *Exh*, akin to overt *only*, that can be applied at various positions in the parse tree of the assertion. Focusing on Bar-Lev and Fox (2017, 2020) theory, *Exh* is attached to the root position, and assigns a truth value to every alternative in a two-step procedure. In the first step, it negates as many alternatives as possible consistently with the assertion. This is achieved by an *Innocent Exclusion* procedure (Fox 2007), which takes all maximal sets of alternatives that can be negated without contradicting the assertion and negates the alternatives that are members of all such sets. The negated alternatives are referred to as the *Innocently Excludable (IE)* alternatives. In the second step, *Exh* asserts as many of the remaining alternatives as possible consistently with the assertion. This is achieved by an Innocent Inclusion procedure (Bar-Lev and Fox 2017, 2020), which takes all maximal sets of alternatives that can be asserted without contradicting the original assertion when taken together with the negation of all innocently excludable alternatives and asserts the alternatives that are members of all such sets. The asserted alternatives are referred to as the *Innocently Includable (II)* alternatives. By these two steps, we avoid contradictions and arbitrary choices of negated and asserted alternatives.

Consider a 2-disjunct CRD in which 'A or B' is interpreted as 'A and B' (like (2)).[1] In this case, the alternative messages are {'A', 'B', 'A or B'} (see Sauerland 2004, Fox 2007, Fox and Katzir 2011, Trinh and Haida 2015).[2] According to the grammatical theory described here, *Exh* starts with Innocent Exclusion. The maximal sets of alternatives that can be negated without contradicting the assertion 'A or B' are {'A'} and {'B'}. This reflects the fact that if 'A' is negated then 'A or B' entails 'B', and symmetrically if 'B' is negated then 'A or B' entails 'A'. It is impossible for both to be negated because that leads to a contradiction with the assertion. Since no alternative is a member of both maximal sets, *Exh* does not negate any alternative. Then, *Exh* continues with Innocent Inclusion. There is one maximal set, which includes all alternatives {'A', 'B', 'A or B'}. This is because all can be included without contradicting the assertion 'A or B' when taken together with the negation of all the innocently excludable alternatives (which are non-existent in this case). Since this is the only maximal set, all its members are asserted. The result is the inference $A \wedge B$, as desired. Similarly, it generalizes to (4) as desired.

[1] Though in FC disjunctions like (2) the disjunction is embedded under a modal, I omit the modal in this analysis and henceforth – both for ease of presentation and because there are CRDs cases not involving modals (such as the case of Warlpiri connectives, Bowler 2014).

[2] For simplicity I ignore the conjunctive alternatives, focusing on the generalization that CRDs should be derived when the set of alternatives is not closed under conjunction. Their lack does not affect the predictions (see the detailed derivations in Bar-Lev and Fox 2017, 2020).

2.3 Iterated Rationality Models (IRMs)

2.3.1 Non-probabilistic IRM

As a preliminary step, in this subsection I introduce a simple model which involves an iterated rationality reasoning. I will show how this simple IRM derives SIs in (finite) scalar cases, but not CRDs. This challenge will serve as an impetus to incorporate probabilities into the reasoning process, as will be discussed in the next subsection.

Consider the simple scalar case of asserting 'some' with 'all' as its only alternative. For example, recall the sentence in (1) and assume that it only has 'John did all of the homework' as its alterative. Given these alternatives, the relevant epistemic states that the speaker may convey are that John did all of the homework ($\forall$), that he did some but not all of the homework ($\exists \wedge \neg\forall$), and that he did not do any of the homework ($\neg\exists$). We assume that the goal of the speaker in a discourse is to convey her epistemic state to the hearer, and that they are truthful (i.e, adhere to Grice's Maxim of Quality). In the described scenario, the speaker can use the following strategy:

- Step I: If the speaker's epistemic state is $\forall$, they can say 'all'. Hearing the message 'all', the hearer can easily infer that the speaker conveys the meaning of $\forall$, since the message is false in the other states.

- Step II: If the speaker's epistemic state is $\exists \wedge \neg\forall$, they can rely on the hearer's knowledge of Step I to exclude the state of $\forall$ when hearing the message 'some'. Since the only other state of affairs that is compatible with this message is the state of $\exists \wedge \neg\forall$, the hearer can conclude that the meaning conveyed by 'some' is $\exists \wedge \neg\forall$, as desired.

More formally, the model involves a set of *alternatives*, M, which are syntactic elements referred to as *messages* and defined independently by a theory of alternatives. Here, I assume the definition of structural alternatives of Katzir (2007). The alternatives induce a *partition* of the context set into states, Π. It would be useful to describe the (in)compatibility of states and messages with a table, where the columns refer to the alternatives in M, the rows refer to the states in Π, and each cell indicates whether the relevant message is consistent with the relevant state or not, by 1 or 0 respectively. For example, in the 'some'/'all' case, $M = \{$'some', 'all'$\}$ and $\Pi = \{\neg\exists, \exists \wedge \neg\forall, \forall\}$, and it can be described by following table (since the state $\neg\exists$ is inconsistent with both alternatives, I set it aside for the discussion and omit it from the table):

(5) Message-state compatibility ('some'/'all'):

	'some'	'all'
$\exists \wedge \neg\forall$	1	0
$\forall$	1	1

The reasoning process involves the identification of states by messages, based on the following criterion:

(6) STATE IDENTIFICATION (non-probabilistic version): a message $m \in M$ *identifies* a state $t \in \Pi$ if m is true in t and there is no other state $t' \in \Pi$ in which m is true.

The identification applies iteratively: in each iteration, the hearer pairs together messages and states according to this criterion. After being identified, those messages and states are eliminated and the next iteration is applied to the remaining messages and states. Once convergence has been reached, or if no identification and elimination are possible, the process ends. In the current example, the message 'all' is true only in the state $\forall$ (Step I above). Therefore, it identifies this state, as indicated in pink in the table, and can therefore be eliminated. Note that in the first iteration, the message 'some' is true in both $\forall$ and $\exists \wedge \neg\forall$, so it cannot identify any message in the first place. However, after eliminating $\forall$, as indicated in gray, in the second iteration the message 'some' is consistent only with $\exists \wedge \neg\forall$ (Step II above), and therefore identifies this state, as indicated in yellow:

(7) Non-probabilistic identification ('some'/'all'):

	'some'	'all'
$\exists \wedge \neg\forall$	1	0
$\forall$	1	1

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II

The present IRM captures the idea that at each iteration the hearer can infer about the speaker's state in terms of full certainty; that is, in each step there is at least one message that necessarily conveys a specific state. In particular, it means that if there is no message conveying a state with full certainty in the first place, then no identification occurs. This is exactly the prediction for CRDs: consider the basic case of two disjuncts, where the set of alternatives is $M = \{$'A or B', 'A', 'B'$\}$ and the induced partition is $\Pi = \{\neg A \wedge \neg B, A \wedge \neg B, \neg A \wedge B, A \wedge B\}$. Each of the alternatives is compatible with at least two messages, as shown in (8): 'A' is true in both $A \wedge \neg B$ and $A \wedge B$; symmetrically, 'B' is true in both $\neg A \wedge B$ and $A \wedge B$; and 'A or B' is true in $A \wedge \neg B$, $\neg A \wedge B$, and $A \wedge B$. That is, no state can be identified by a message in terms of full certainty. Hence, the IRM developed here does not satisfy the desideratum in (4).

(8) Message-state compatibility ('A'/'B'/'A or B'):

$P(m\|t)$	'A'	'B'	'A or B'
$A \wedge \neg B$	1	0	1
$\neg A \wedge B$	0	1	1
$A \wedge B$	1	1	1

This failure could be taken as motivation to move toward a different identification criterion, which does not require full certainty but rather with respect to the best guess the hearer can make about the speaker's state. A natural way to think about such a criterion is in terms of probabilities. Informally speaking, the hearer infers that the speaker's state is t if this state is the most probable given that the speaker uttered the message m. In the next subsection, I will present a simple probabilistic IRM based on that intuition.

2.3.2 Probabilistic IRM

The IRM literature has typically offered probabilistic models, including the Rational Speech Act model (Frank and Goodman 2012, Goodman and Stuhlmüller 2013, Bergen et al. 2016), Iterated Best Response (Franke 2009, 2011), Bidirectional OT (Blutner 1998, 2000, Jäger 2002), among others. These models have been developed for purposes other than the derivation of SIs and are motivated by independent empirical roles and conceptual arguments for probabilistic reasoning. However, unlike previous pragmatic theories, some IRMs (Franke 2009, van Rooij 2010) derive the conjunctive reading in cases like (2) (see Fox and Katzir 2021), and more generally, have been proposed as theories of SIs.

Following the intuition in the previous subsection, the identification criterion should involve a comparison between the probabilities of states given the utterance of the speaker. Simplifying somewhat, the speaker starts by assigning probabilities to every message $m \in M$, given the epistemic state they are in $t \in \Pi$, $P(m|t)$ – a conditional probability referred to as the *likelihood*. To think about the evaluation of the likelihood, consider again the 'some'/'all' case. In the state $\exists \wedge \neg\forall$, the only consistent message is 'some'. Therefore, if the speaker's state is $\exists \wedge \neg\forall$, they would necessarily utter 'some', meaning, $P(\text{'some'}|\exists \wedge \neg\forall) = 1$. On the other hand, if their state is $\forall$, there are two true messages – 'some' and 'all'. That is, both $P(\text{'some'}|\forall)$ and $P(\text{'all'}|\forall)$ should be positive, but less than 1. Since $\forall$ is consistent with only these messages in the current scenario, these likelihoods should also sum up to 1. The issue at hand is determining the precise probability distribution. A common response is to assume that the speaker has no preference between the true messages that are true in a given state, meaning that the likelihood distributes uniformly among these messages. In the current example, it means that $P(\text{'some'}|\forall) = P(\text{'all'}|\forall) = \frac{1}{2}$. In alignment with the IRM literature, I refer to this assumption as that of a *naive speaker*:[3]

(9) NAIVE SPEAKER: If n messages make t true, then: $P(m|t) = \begin{cases} 1/n & \text{if } m(t) = 1 \\ 0 & \text{otherwise} \end{cases}$

Like the non-probabilistic IRM, it is useful to describe the likelihoods in a table format. Here, the columns correspond to the alternatives, the rows correspond to the states, and each

[3]The notation $m(t) = 1$ indicates that m makes t true

cell displays the probability of the relevant message given the relevant state. The table below, for instance, illustrates the likelihoods a naive speaker assigns in the 'some'/'all' scenario:[4]

(10) Naive speaker's probability assignment ('some'/'all'):

$P(m\|t)$	'some'	'all'
$\exists \wedge \neg\forall$	1	0
$\forall$	$\frac{1}{2}$	$\frac{1}{2}$

Let us now assume that the priors distribute uniformly, meaning that each state has equal probability:

(11) FLAT PRIORS: For any $t \in \Pi$ (Π is finite), $P(t) = \frac{1}{|\Pi|}$

Upon receiving a message, the hearer then has to compare the probabilities of the different states given that message, denoted $P(t|m)$. In other words, we would say that that message m identifies[5] a state t if its conditional probability is the strict maximum compared to all other messages, that is, that $P(t|m) > P(t'|m)$ for any other $t' \in \Pi$. By Bayesian reasoning, this is equivalent to $P(m|t) \cdot P(t) > P(m|t') \cdot P(t')$.[6] That is, we change the identification criterion in (6) to the following probabilistic criterion:

(12) STATE IDENTIFICATION (probabilistic version): A message $m \in M$ *identifies* a state $t \in \Pi$ if for every other $t' \in \Pi$, $P(m|t) \cdot P(t) > P(m|t') \cdot P(t')$.

Under the assumption in (11) of flat priors, the comparison in (12) can be reduced to a comparison of likelihoods. That is, to reason about the intended meaning of an utterance, the hearer could directly compare the speaker's probabilities, $P(m|t)$ for each of the messages instead of $P(t|m)$, in each of the steps. For example, in the 'some'/'all' case, the hearer could compare $P(\text{'some'}|\exists \wedge \neg\forall)$ with $P(\text{'some'}|\forall)$ and $P(\text{'all'}|\exists \wedge \neg\forall)$ with $P(\text{'all'}|\forall)$. Assuming a naive speaker, this identification process ends within one step (see illustration in (13)): 'some' identifies $\exists \wedge \neg\forall$, since $P(\text{'some'}|\exists \wedge \neg\forall) = 1 > \frac{1}{2} = P(\text{'some'}|\forall)$; and 'all' identifies $\forall$ since $P(\text{'all'}|\forall) = \frac{1}{2}$ is the only positive probability given the state 'all', as desired.

[4]For ease of presentation, I omit from the probabilities tables rows in which all probabilities are 0. E.g. in (10), the row of the state $\neg\exists$ is omitted, since its probabilities are all 0 (no message is compatible with that state).

[5]Note that not all IRMs use that notion of identification. Below is a brief discussion of how this way of presentation relates to IRMs in the literature.

[6]According to Bayes rule, $P(t|m) = \frac{P(m|t) \cdot P(t)}{P(m)}$. Therefore, $P(t|m) > P(t'|m)$ can be rewritten as $\frac{P(m|t) \cdot P(t)}{P(m)} > \frac{P(m|t') \cdot P(t')}{P(m)}$. Since we are interested in a relative comparison of probabilities, the denominator $P(m)$ can be omitted from both sides.

(13) Probabilistic identification ('some'/'all'):
Naive speaker's probabilities

	'some'	'all'
$\exists \wedge \neg \forall$	1	0
$\forall$	$\frac{1}{2}$	$\frac{1}{2}$

identified message-state pairs in iteration I

Note that the IRM presented here is a simplified model drawn from Fox and Katzir (2021) – it incorporates the core assumptions of the Iterated Best Response model of Franke (2009, 2011), and appears to be much farther from other IRMs proposed in the literature (such as the Rational Speech Act framework). However, it is sufficient to use this model to generalize about the IRMs approach, since the present work focuses on CRDs. Most of the literature on IRMs cannot derive CRDs and focuses on simple kinds of SIs such as the 'some'/'all' case. As mentioned in Fox and Katzir (2021), the Franke (2009, 2011) and van Rooij (2010) frameworks are the only ones that derive CRDs. Therefore, this simple model offers a relatively clear presentation of the results achieved by IRMs with respect to CRDs.

2.4 IRMs and CRDs: 2 disjuncts

The IRM presented above derives CRDs in the simple case of 2 disjuncts (like (2)). Recall that for the 2-disjunct sentence 'A or B' in a CRD configuration, the alternative messages are 'A', 'B', 'A or B' and the states are $\{\neg A \wedge \neg B, A \wedge \neg B, \neg A \wedge B, A \wedge B\}$. The naive speaker's probabilities are as in (14). Based on them, the hearer identifies in the first step the messages 'A' and 'B' with $A \wedge \neg B$ and $\neg A \wedge B$, respectively (as indicated in pink in table (14)). 'A or B' cannot identify any state at this stage because there is no single state that obtains the highest probability, but rather there is a tie between $A \wedge \neg B$ and $\neg A \wedge B$. However, these states (and their corresponding messages) are eliminated after the first iteration (as indicated in gray). In the second step, 'A or B' is the only remaining message, as is the state $A \wedge B$, so they are necessarily paired together (as indicated in yellow).

(14) Probabilistic identification ('A'/'B'/'A or B'):
Naive speaker's probabilities

$P(m\|t)$	'A'	'B'	'A or B'
$A \wedge \neg B$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\neg A \wedge B$	0	$\frac{1}{2}$	$\frac{1}{2}$
$A \wedge B$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II

2.5 Beyond 2 disjuncts

As argued in Sect. 2.1, CRDs are not limited to 2-way disjunctions, and generalize to every $n \geq 2$ disjunctions and to k-out-of-n disjunctions, and a theory of SIs has to derive them all as formulated in (4). Indeed, this desideratum is obtained within the grammatical approach, in the same way it handles 2 disjuncts (see Fox and Katzir 2021). However, van Rooij (2010), Franke (2011) and later Fox and Katzir (2021) observe that IRMs fail with constructions involving more than 2 disjuncts, even though they appear to be a simple extension. Consider the case of 3 disjuncts, 'A or B or C', which should be interpreted as a conjunction ($A \wedge B \wedge C$) given the desideratum in (4). The alternatives in that case are {'A', 'B', 'C', 'A or B', 'A or C', 'B or C', 'A or B or C'}, and the probabilities assigned by the naive speaker are as in (15). Based on these probabilities, the hearer identifies in the first step the alternative messages 'A', 'B', 'C' with $A \wedge \neg B \wedge \neg C$, $\neg A \wedge B \wedge \neg C$ and $\neg A \wedge \neg B \wedge C$, respectively (as indicated in pink in the table). After eliminating these messages and states (indicated in gray), the remaining states still involve a tie for the first place by multiple messages (see the remaining white cells in the table). Hence, in the second step, no identification is possible and the procedure ends, and the remaining messages do not get the strengthened conjunctive reading.

(15) Probabilistic identification (three disjuncts):
Naive speaker's probabilities

$P(m\|t)$	'A'	'B'	'C'	'A or B'	'A or C'	'B or C'	'A or B or C'
$A \wedge \neg B \wedge \neg C$	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{4}$
$\neg A \wedge B \wedge \neg C$	0	$\frac{1}{4}$	0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{4}$
$\neg A \wedge \neg B \wedge C$	0	0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$A \wedge B \wedge \neg C$	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$A \wedge \neg B \wedge C$	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$\neg A \wedge B \wedge C$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$A \wedge B \wedge C$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$

identified message-state pairs in iteration I cells eliminated after iteration I
□ cells not eliminated after iteration I

Fox and Katzir (2021) observe that this IRM fails twice: neither the full message 'A or B or C' nor the partial 2-way disjunctions when there is a third salient alternative (e.g. 'A or B' when 'C' is salient) identify any state. That is, this IRM fails to satisfy the desideratum in (4). These failures are shared by all existing IRMs, and attempts to fix them at most address the former only (van Rooij 2010, Franke 2011). Fox and Katzir (2021) conclude that these results constitute an argument in favor of the grammatical approach over IRMs.

3. Proposal

3.1 WMS-IRM: IRM with weighted probabilities

The question underlying this work is whether a more nuanced IRM can address the challenge facing existing IRMs. Specifically, it proposes to reexamine the assumption of a naive speaker in (9). Such an assumption – although simple – seems nontrivial in some scenarios. For instance, consider again the of 2-out-of-3 disjuncts case, e.g., the message 'A or B' that should be interpreted as the state $A \wedge B \wedge \neg C$. There are multiple messages that are compatible with this state, in particular 'A or B' and 'A or C'. Since both are true in that state, the probability of a naive speaker uttering them is positive and equal – neither of them is considered better than the other. However, focusing on these 2-disjunct messages, one may expect the message 'A or B' to be better to convey this state than the message 'A or C' – while both disjuncts in the first message are true in the state, the second message has a disjunct which is false in the state (C), and also does not have a disjunct which is true in the state (B). In some intuitive sense then, the first message is a better fit for the state than the second, but this is not reflected in the assignment of probabilities according to the naive speaker assumption.
This example suggests a different type of speaker. Instead of a uniform distribution among all messages compatible with a certain state, I propose using weighted probabilities – which intuitively reflect the degree of "overlap" between messages and states – rather than a binary value indicating only whether they are consistent or not. The main question arising is how such "overlap" should be defined and measured. More formally, we need to define some notion of *weighted message score* (WMS), according to which the probabilities are computed. To answer this question, we will follow the intuition behind the 2-out-of-3 case presented above. We have considered 'A or B' as a better, or more "overlapping" message than 'A or C' with respect to $A \wedge B \wedge \neg C$ because of the consistency of their disjuncts with this state. The more disjuncts that are compatible with a state, the more the whole utterance captures that state and is therefore perceived as corresponding to it. Hence, we would expect that each different disjunct of the message will contribute to the weight, according to its (in)consistency with the given state. Moreover, we would assume that the speaker has no preference among the messages consistent with that state, so that each should contribute equally to the weight. In order to have some measure which increases with the number of the true disjuncts in a state, we need some mechanism by which we can access the disjuncts composing a given disjunction, and then test which of them are true in that state. For that purpose, I adopt the notion of *deletion alternatives*, Alt_{del}, based on Katzir's (2007) definition of alternatives.[7]

[7] It is important to note that we use only deletion alternatives, and not alternatives in general, which include, besides deletion alternatives, alternatives obtained by substitution. Otherwise, by both structural substitution and deletion, all messages with the same number of disjuncts have the same alternatives. For example, the alternatives to 'A or B', 'A or C' and 'B or C' are {'A or B', 'A or C', 'B or C', 'A', 'B', 'C'}. WMS based on such alternatives eliminate the effect of weights on distinguishing between these messages since all are computed based on the compatibility of the same set of alternatives with the same state. In the 2-out-of-3 case, it means that 'A or B', 'A or C' and 'B or C' all have the same weights for $A \wedge B \wedge \neg C$, for example. Consequently, their probabilities are equal. That is, it brings us back to the initial obstacle that probabilities do

According to this definition, the deletion alternatives to a disjunction are the disjunction itself and all partial disjunctions composed of at least one individual disjunct. For instance, Alt_{del}('A or B') ={'A', 'B', 'A or B'} and Alt_{del}('A') ={'A'}. Indeed, such a set contains messages beyond just the individual disjuncts. Hence, the weight of a message m given a state t, $w_t(m)$, will be defined as the number of m's deletion alternatives that are consistent with t (16a) if m is consistent with t (and 0 otherwise). The probabilities of m given t are then distributed based on these weights (16b).

(16) a. WEIGHTED MESSAGE SCORE (WMS):
$w_t(m) = |\{m' : m' \in Alt_{del}(m) \wedge m'(t) = 1\}| \cdot m(t)$

b. WEIGHTED PROBABILITIES: $P(m|t) = w_t(m) / \sum_{m' \in M} w_t(m')$

In other words, the current speaker, unlike the naive speaker, prefers some messages over others to convey a certain state, since they have a greater "overlap" with this state in terms of the number of true deletion alternatives, thus considered better.

3.2 Predictions for CRDs

Before proceeding to the predictions of this WMS-IRM for CRDs, let us start with the simple SI of 'some' strengthened to 'some but not all' (like (1)). The alternatives in this case, {'some', 'all'}, are obtained by substitution. That is, the deletion alternatives for each message include only the message itself. It means that for a given message and state, the weight is 1 or 0 and indicates whether the message is compatible with the state or not. With that binary weighting, the probabilities are similar to those computed by the original IRM (i.e., (10)). As shown in Sect. 2.3, the original IRM derives the strengthened inference as desired. Hence, the WMS-IRM derives it as well.

Now, consider the basic case of CRD with two disjuncts, such as (2). The deletion alternatives to 'A' and 'B' are {'A'} and {'B'} respectively. Therefore, the weights assigned for each of them according to (16a) are 0 or 1, indicating whether they are compatible with a certain state or not. For example, $w_{A \wedge \neg B}$('A') = 1 and $w_{A \wedge \neg B}$('B') = 0. However, Alt_{del}('A or B') ={'A', 'B', 'A or B'}, so its weights are natural numbers on a scale of 0 to 3. For example, $w_{A \wedge \neg B}$('A or B') = 2 (because the deletion alternatives 'A' and 'A or B' are the only ones compatible with that state) and $w_{A \wedge B}$('A or B') = 3. Table (17) summarizes these weights (the last column shows the total weight for each state, according to which the probabilities are computed).

---

not reflect how close messages are to a state if they are messages of the same complexity (i.e., have the same number of disjuncts). However, though it is necessary to use deletion alternatives in this part, it is not trivial since the model already computes all alternatives to the pronounced utterance, as the set of possible messages that induces the set of states. That is, the IRM computes two different sets of alternatives – one for the measure function and one for identifying states. At this point, I simply make this stipulation, leaving a justification of this choice for future work.

(17) Weighted Message Scores ('A'/'B'/'A or B'):

$w_t(m)$	'A'	'B'	'A or B'	Σ
$A \wedge \neg B$	1	0	2	3
$\neg A \wedge B$	0	1	2	3
$A \wedge B$	1	1	3	5

Based on these weights and according to (16b), the probabilities are as in (18). The WMS-IRM still derives a CRD: as in the previous model (see (14)), in the first iteration 'A' identifies $A \wedge \neg B$ and 'B' identifies $\neg A \wedge B$. After elimination, 'A or B' identifies $A \wedge B$ (since they are the only message and state remaining), as desired.

(18) Probabilistic identification ('A'/'B'/'A or B'):
Weighted probabilities

$P(m\|t)$	'A'	'B'	'A or B'
$A \wedge \neg B$	$\frac{1}{3}$	0	$\frac{2}{3}$
$\neg A \wedge B$	0	$\frac{1}{3}$	$\frac{2}{3}$
$A \wedge B$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{5}$

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II

We move on to the 3-disjunct case, in which the previous model failed to achieve the desideratum in both the 2-out-of-3 and 3-out-of-3 cases as we have seen in (15). As before, the only deletion alternative to an individual disjunct is itself, and its weight is therefore binary; and the deletion alternatives to a 2-way disjunction include itself and its individual disjuncts, so its weights are between 0 and 3. As for the full message, Alt_{del}('A or B or C') = {'A', 'B', 'C', 'A or B', 'A or C', 'A or B', 'A or B or C'} so its weights are natural numbers between 0 and 7. Hence, based on (16), the probabilities are as in (19). In the first step 'A', 'B' and 'C' identify $A \wedge \neg B \wedge \neg C$, $\neg A \wedge B \wedge \neg C$ and $\neg A \wedge \neg B \wedge C$ (resp.; indicated in pink). After eliminating these messages and states (as indicated in light gray), in the second step, 'A or B', 'B or C', 'A or C' identify $A \wedge B \wedge \neg C$, $\neg A \wedge B \wedge C$ and $A \wedge \neg B \wedge C$ (resp.; indicated in yellow). After further elimination (indicated in dark gray), 'A or B or C' is the only remaining message, as is the state $A \wedge B \wedge C$, so they are necessarily paired together (indicated in green).

(19) Probabilistic identification (three disjuncts):

Weighted probabilities

$P(m\|t)$	'A'	'B'	'C'	'A or B'	'A or C'	'B or C'	'A or B or C'
$A \wedge \neg B \wedge \neg C$	$\frac{1}{9}$	0	0	$\frac{2}{9}$	$\frac{2}{9}$	0	$\frac{4}{9}$
$\neg A \wedge B \wedge \neg C$	0	$\frac{1}{9}$	0	$\frac{2}{9}$	0	$\frac{2}{9}$	$\frac{4}{9}$
$\neg A \wedge \neg B \wedge C$	0	0	$\frac{1}{9}$	0	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{4}{9}$
$A \wedge B \wedge \neg C$	$\frac{1}{15}$	$\frac{1}{15}$	0	$\frac{1}{5}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{2}{5}$
$A \wedge \neg B \wedge C$	$\frac{1}{15}$	0	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{5}$	$\frac{2}{15}$	$\frac{2}{5}$
$\neg A \wedge B \wedge C$	0	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{1}{5}$	$\frac{2}{5}$
$A \wedge B \wedge C$	$\frac{1}{19}$	$\frac{1}{19}$	$\frac{1}{19}$	$\frac{3}{19}$	$\frac{3}{19}$	$\frac{3}{19}$	$\frac{7}{19}$

identified message-state pairs in iteration I; cells eliminated after iteration I
identified message-state pairs in iteration II; cells eliminated after iteration II
identified message-state pairs in iteration III

In addition to 3-disjunctions, these results generalize to any $n \geq 2$ disjunctions, both for n-out-of-n and k-out-of-n cases,[8] thus achieving the desideratum in (4) and significantly broadening the scope of the model's success. Hence, the grammatical approach has no empirical advantage over this IRM with respect to CRDs.

3.3 Predictions for negation

The WMS-IRM is also well-defined for sentences including negation and derives the required meaning also for negative SIs. Consider the case of conjunction embedded under negation (when the set of alternatives is closed under disjunction, unlike in the FC case):

(20) John did not eat an apple and a banana
$\rightsquigarrow$ John didn't eat an apple or he didn't eat a banana (but he ate one of them)

The alternative messages in that case are {'not A', 'not B', 'not (A or B)', 'not (A and B)'}. Following previous work (Romoli 2012, Trinh and Haida 2015, Breheny et al. 2018, among others), I assume that the deletion alternatives of a message do not include deletions of negations (e.g., Alt_{del}('not (A and B)')={'not A', 'not B', 'not (A and B)'}). The weighted probabilities are therefore as in (21), and the iterative process is as follows: in the first step, 'not A' and 'not B' identify $\neg A \wedge B$ and $A \wedge \neg B$ (resp.). After eliminating these messages and states, in the second step, 'not (A or B)' identifies $\neg A \wedge \neg B$ and 'not (A and B)' identifies $(A \wedge \neg B) \vee (\neg A \wedge B)$, as desired.

[8] Full proof is in Appendix A.1.

(21) Probabilistic identification (conjunction under negation):
Weighted probabilities

$P(m\|t)$	'not A'	'not B'	'not (A or B)'	'not (A and B)'
$\neg A \wedge \neg B$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$
$A \wedge \neg B$	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$\neg A \wedge B$	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
$(A \wedge \neg B) \vee (\neg A \wedge B)$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{3}{7}$

identified message-state pairs in iteration I cells eliminated after iteration I
identified message-state pairs in iteration II

Similarly, the so-called Negative Free Choice (Negative FC; Fox 2007), as in (22), is also derived by the WMS-IRM.

(22) John is not required to eat an apple and a banana

a. $\rightsquigarrow$ John is not required to eat an apple
b. $\rightsquigarrow$ John is not required to eat a banana

The inference from (22) to (22a)-(22b) is semantically parallel to the positive case, i.e., the inference from (2) to (2a)-(2b). As such, their derivation is similar: the weighted probabilities assigned by the speaker are as in (23). In the first step, 'not A' and 'not B' identify $\neg A \wedge B$ and $A \wedge \neg B$ respectively. After elimination, 'not (A and B)' is the only remaining message, as is the state $\neg A \wedge \neg B$, so they are necessarily paired together.

(23) Probabilistic identification (FC under negation):
Weighted probabilities

$P(m\|t)$	'not A'	'not B'	'not (A and B)'
$\neg A \wedge \neg B$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{5}$
$A \wedge \neg B$	0	$\frac{1}{3}$	$\frac{2}{3}$
$\neg A \wedge B$	$\frac{1}{3}$	0	$\frac{2}{3}$

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II

The generalization to any $n \geq 2$ conjuncts follows directly from the generalization in the parallel case of positive CRDs.

3.4 Predictions for Universal CRDs

3.4.1 Background

CRDs embedded under a universal quantifier as in the universal free choice example in (24), are interpreted in a similar manner as their non-embedded counterpart (Chemla 2009).

(24) Every student is allowed to eat an apple or a banana

a. ⇝ Every student is allowed to eat an apple
b. ⇝ Every student is allowed to eat a banana

In the grammatical view, one can derive the inference in (24) locally, by applying *Exh* in the scope of the universal quantifier. This option is not available to pragmatic views such as the one we entertain here, where embedded implicatures are a problem in general. However, Chemla (2009) points out that a local derivation cannot achieve the desired inference in the negative case, as in (25).

(25) No kid is required to eat an apple and a banana

a. ⇝ No kid is required to eat an apple
b. ⇝ No kid is required to eat a banana

Like in CRDs, positive Universal CRDs are semantically equivalent to their negative counterparts. However – and in contrast to negative CRDs – negative Universal CRDs cannot be derived locally: such inferences challenge also grammatical theories that derive implicatures locally, such as the exhaustivity operator proposed by Fox (2007), as a local derivation for (24) is not available for the negative case in (25), and concludes that Universal CRDs must be derived globally. This challenge was recently resolved by Bar-Lev and Fox (2017, 2020), who propose an exhaustivity operator that applies globally (for a brief overview of their proposal, see Sect. 2.2). To the best of my knowledge, no other existing theory of SIs provides a global derivation for Universal CRD inferences.

3.4.2 A (global) grammatical account

Let us first see how the global grammatical mechanism proposed by Bar-Lev and Fox (2017, 2020) derives Universal CRDs. Considering a Universal CRD sentence of a form 'every x (Ax or Bx)' (like (24)), the alternative messages are assumed to be 'every x (Ax)', 'every x (Bx)', 'every x (Ax or Bx)', 'some x (Ax)', 'some x (Bx)' and 'some x (Ax or Bx)' (i.e., the alternative set consists of universal and existential messages, both are parallel to the alternatives generated in the unembedded CRD case). The exhaustivity operator *Exh* is applied globally: it is attached to the root and assigns truth values to the alternatives by the single, two-step procedure of Innocent Exclusion + Innocent Inclusion. In the first step, *Exh* negates all IE alternatives, which, as in Sect. 2.2, refer to all alternatives that are at the intersection of the maximal sets of alternatives that can be *negated* consistently with the prejacent. In our case, the maximal sets of alternatives that can be negated without contradicting the 'every x (Ax or Bx)' are: (i) {'every x (Ax)', 'every x (Bx)'}; (ii) {'every x (Ax)', 'some x (Ax)'}; (iii) {'every x (Bx)', 'some x (Bx)'}. The intersection of these three maximal sets is empty, which means that no alternative is negated by *Exh*. In the second step, *Exh* asserts all the II alternatives, which as in Sect. 2.2, refer to all alternatives that are at the intersection of the maximal sets of alternatives that can be *asserted* consistently with the prejacent. In our case,

there is one maximal set of alternatives that can be asserted without contradicting 'every x (Ax or Bx)', which includes all alternatives (since none of the was previously innocently excluded, and are all consistent with the prejacent). As in the case of unembedded CRD, since there is one maximal set, all of its members are asserted. In particular, the stronger alternatives 'every x (Ax)' and 'every x (Bx)' are asserted. The inference is therefore $\forall x(Ax \wedge Bx)$, as desired.

The negative case 'no x (Ax and Bx)' (like (25)) is parallel to the positive case in terms of the entailment relations between the prejacent and its alternatives. Therefore, it is derived in a similar way. As in the case of unembedded CRDs, this mechanism is number-independent, so it generalizes to (4) with respect to the CRD embedded under the universal quantifier.

3.4.3 Naive speaker IRM

As most pragmatic theories do not derive CRDs, they in particular do not obtain Universal CRDs inferences. The IRMs proposed in the literature that do derive CRDs in the simple case (van Rooij 2010, Franke 2011 do not derive Universal CRDs inferences either.[9] The naive-speaker IRM outlined in this work is no different. As mentioned above, we consider the following alternatives for a Universal CRD sentence of the form 'every x (Ax or Bx)': {'every x (Ax)', 'every x (Bx)', 'every x (Ax or Bx)', 'some x (Ax)', 'some x (Bx)' and 'some x (Ax or Bx)'}. The states induced by these alternatives are:
$\{\neg\exists xAx \wedge \neg\exists xBx$,
$\forall xAx \wedge \forall xBx$,
$\forall xAx \wedge \exists xBx \wedge \neg\forall xBx$,
$\forall xBx \wedge \exists xAx \wedge \neg\forall xAx$,
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \forall x(Ax \vee Bx)$,
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \neg\forall x(Ax \vee Bx)$,
$\forall xAx \wedge \neg\exists xBx$,
$\forall xBx \wedge \neg\exists xAx$,
$\exists xAx \wedge \neg\forall xAx \wedge \neg\exists xBx$,
$\exists xBx \wedge \neg\forall xBx \wedge \neg\exists xAx\}$.

The probabilities assigned by a naive speaker are as in (26). The iterative process begins by identifying the messages 'every A (Ax)', 'every B (Bx)', 'some A (Ax)' and 'some B (Bx)' with the states $\forall xAx \wedge \neg\exists xBx$, $\forall xBx \wedge \neg\exists xAx$, $\exists xAx \wedge \neg\forall xAx \wedge \neg\exists xBx$, and $\exists xBx \wedge \neg\forall xBx \wedge \neg\exists xAx$, respectively. At this stage the messages 'every x (Ax or Bx)' as well as 'some x (Ax or Bx)' are unable to identify any state, since in both there is no single state which has the highest value but rather there is a tie between two states which have the highest value. After eliminating these states and message, 'every x (Ax or Bx)' and 'some x (Ax or Bx)' *do* identify states, but, the message 'every x (Ax or Bx)' wrongly identifies the state

[9] Franke (2011) proposes a fix in one of the assumptions of the Iterated Best Response model, which results in the derivation of Universal CRDs, in both positive and negative cases (24)-(25). This involves ignoring states that are unlikely in the context. Although this stipulation does not seem completely unnatural, it requires further scrutiny, as noted by Franke (2011). Moreover, it does not resolve the challenge of generalization in (4), both for CRDs and Universal CRDs.

$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \forall x(Ax \vee Bx)$, failing to identify the state $\forall xAx \wedge \forall xBx$ needed for deriving a Universal CRD.

(26) Probabilistic identification (Universal CRD):
Naive speaker's probabilities

$P(m\|t)$	'every x (Ax)'	'every x (Bx)'	'every x (Ax or Bx)'	'some x (Ax)'	'some x (Bx)'	'some x (Ax or Bx)'
$\forall xAx \wedge \forall xBx$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$\forall xAx \wedge \exists xBx \wedge \neg\forall xBx$	$\frac{1}{5}$	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
$\forall xBx \wedge \exists xAx \wedge \neg\forall xAx$	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \forall x(Ax \vee Bx)$	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \neg\forall x(Ax \vee Bx)$	0	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$\forall xAx \wedge \neg\exists xBx$	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{4}$
$\forall xBx \wedge \neg\exists xAx$	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{4}$
$\exists xAx \wedge \neg\forall xAx \wedge \neg\exists xBx$	0	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
$\exists xBx \wedge \neg\forall xBx \wedge \neg\exists xAx$	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$

identified message-state pairs in iteration I cells eliminated after iteration I
identified message-state pairs in iteration II cells eliminated after iteration II

Notably, this failure is different from the one in the case of naive speaker with regard to a CRD with 3 disjuncts (like (15)): while in the latest the disjunctive messages do not identify any message (and therefore can be treated as indicating uncertainty or anomaly), the message in this case *does* identify a state, but an incorrect one. That is, there is no tie between states given a certain message, but a strict maximum obtained by an undesired state given that message.

3.4.4 WMS-IRM

We see that for an IRM to derive (24), its probability function needs to deviate from the naive speaker's one, so that the message receives a higher probability given the state $\forall xAx \wedge Bx$ than given $\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \forall x(Ax \vee Bx)$, reflecting the better compatibility of the first state over the other with respect to the assertion. We will see now that the WMS-IRM does exactly that, and the desired derivation is obtained.

The deletion alternatives of the possible messages in the embedded case are similar to those in the non-embedded case. That is, the only deletion alternative of an individual disjunct under a universal/existential quantifier is itself, e.g., Alt_{del}('every x (Ax)') = {'every x (Ax)'}, and its weight is therefore binary; and the deletion alternative of a 2-way disjunction under a universal/existential quantifier includes the disjunction and its individual

disjuncts, each of them embedded under the quantifier, e.g., Alt_{del}('every x (Ax or Bx)') = {'every x (Ax)', 'every x (Bx)', 'every x (Ax or Bx)'}, and its weight is between 0 and 3. Based on that analysis, the weights are as in (27) and the probabilities are as in (28). As with the naive speaker IRM, in the first step the messages 'every A (Ax)', 'every B (Bx)', 'some A (Ax)' and 'some B (Bx)' identify states $\forall xAx \wedge \neg\exists xBx$, $\forall xBx \wedge \neg\exists xAx$, $\exists xAx \wedge \neg\forall xAx \wedge \neg\exists xBx$, and $\exists xBx \wedge \neg\forall xBx \wedge \neg\exists xAx$. The messages 'every x (Ax or Bx)' and 'some x (Ax or Bx)' are still unable to identify any state (there are multiple states achieving the maximum probability). After eliminating the identified states and messages, however, 'every x (Ax or Bx)' *does* identify the desired state $\forall xAx \wedge Bx$, because now its probability is the strict maximum compared to the other (remaining) states.

(27) Weighted Message Scores (Universal CRD):

$w_t(m)$	'every x (Ax)'	'every x (Bx)'	'every x (Ax or Bx)'	'some x (Ax)'	'some x (Bx)'	'some x (Ax or Bx)'	Σ
$\forall xAx \wedge \forall xBx$	1	1	3	1	1	3	10
$\forall xAx \wedge \exists xBx \wedge \neg\forall xBx$	1	0	2	1	1	3	8
$\forall xBx \wedge \exists xAx \wedge \neg\forall xAx$	0	1	2	1	1	3	8
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \forall x(Ax \vee Bx)$	0	0	1	1	1	3	6
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \neg\forall x(Ax \vee Bx)$	0	0	0	1	1	3	5
$\forall xAx \wedge \neg\exists xBx$	1	0	2	1	0	2	6
$\forall xBx \wedge \neg\exists xAx$	0	1	2	0	1	2	6
$\exists xAx \wedge \neg\forall xAx \wedge \neg\exists xBx$	0	0	0	1	0	2	3
$\exists xBx \wedge \neg\forall xBx \wedge \neg\exists xAx$	0	0	0	0	1	2	3

(28) Probabilistic identification (Universal CRD):
Weighted probabilities

$P(m\|t)$	'every x (Ax)'	'every x (Bx)'	'every x (Ax or Bx)'	'some x (Ax)'	'some x (Bx)'	'some x (Ax or Bx)'
$\forall xAx \wedge \forall xBx$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$
$\forall xAx \wedge \exists xBx \wedge \neg\forall xBx$	$\frac{1}{8}$	0	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$\forall xBx \wedge \exists xAx \wedge \neg\forall xAx$	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \forall x(Ax \vee Bx)$	0	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge \neg\forall xBx \wedge \neg\forall x(Ax \vee Bx)$	0	0	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{5}$
$\forall xAx \wedge \neg\exists xBx$	$\frac{1}{6}$	0	$\frac{1}{3}$	$\frac{1}{6}$	0	$\frac{1}{3}$
$\forall xBx \wedge \neg\exists xAx$	0	$\frac{1}{6}$	$\frac{1}{3}$	0	$\frac{1}{6}$	$\frac{1}{3}$
$\exists xAx \wedge \neg\forall xAx \wedge \neg\exists xBx$	0	0	0	$\frac{1}{3}$	0	$\frac{2}{3}$
$\exists xBx \wedge \neg\forall xBx \wedge \neg\exists xAx$	0	0	0	0	$\frac{1}{3}$	$\frac{2}{3}$

identified message-state pairs in iteration I cells eliminated after iteration I
identified message-state pairs in iteration II cells eliminated after iteration II

Based on this result, we can conclude that the WMS-IRM derives the negative case in (25) as well. Since it is a global mechanism and given the semantic parallelism between the positive and negative cases, their derivations are similar. That is, the WMS-IRM overcomes the challenge faced by pragmatic and some grammatical theories of deriving Universal CRDs, in the positive and negative cases, at least in the case of 2 disjuncts. In Appendix A.2, I show that it does so in the case of 3 disjuncts as well.[10]

4. Modularity

4.1 WMS-IRM and prior sensitivity

WMS-IRM undercuts Fox and Katzir (2021)'s argument against IRMs as a theory of SIs by providing a general account of CRDs which handles cases with more than 2 disjuncts and Universal CRDs. However, Fox and Katzir (2021) point out that it is not sufficient for a theory of IRM to overcome the challenge of deriving CRDs in the general case, and present a second argument against IRMs, based on their sensitivity to priors.

An underlying assumption of such a theoretical framework – being pragmatic- and probabilistic-based – is that it incorporates probabilities influenced by world-knowledge and general reasoning. In particular, it means that we need to abandon the initial assumption that the prior probabilities are flat (in (11)). This assumption was crucial for the hearer to compare $P(m|t)$

[10] I do not provide general proof for any $n \geq 2$ disjuncts, however. Unlike CRDs, it is not clear that such a generalization is actually valid and requires a more general characterization of Universal CRDs, which is beyond the scope of the present work.

instead of $P(t|m)$. If it is no longer assumed, the probability $P(t|m)$ can only be reduced to the comparison of $P(m|t) \cdot P(t)$ (given Bayes rule), which means that the priors play a role in the hearer's inference. We will see now that such a dependence leads to wrong predictions for CRDs, even if the priors only slightly deviate from a flat distribution.

Consider again the probabilities assigned by a naive speaker in the case of CRD with 2 disjuncts, repeated here from (14):

(29) Naive speaker's probability assignment ('A'/'B'/'A or B'):

$P(m\|t)$	'A'	'B'	'A or B'
$A \wedge \neg B$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\neg A \wedge B$	0	$\frac{1}{2}$	$\frac{1}{2}$
$A \wedge B$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

If $P(A \wedge \neg B) < \frac{2}{3} \cdot P(A \wedge B)$ then the message 'A' will be incorrectly strengthened to mean $A \wedge B$ (symmetrically, if $P(\neg A \wedge B) < \frac{2}{3} \cdot P(A \wedge B)$, 'B' will incorrectly identify $A \wedge B$). Moreover, not only noticeable skewed priors yield wrong results, but also minute imbalances do: if $P(\neg A \wedge B) < P(A \wedge \neg B)$ (and $(P(A \wedge B) < \frac{3}{2} \cdot P(A \wedge \neg B))$, the symmetry between $\neg A \wedge B$ and $A \wedge \neg B$ with regard to 'A or B' will be violated, and 'A or B' will be incorrectly strengthened to mean $A \wedge \neg B$. This means that even the slightest deviation of the priors from a uniform distribution will lead to wrong results. However, this is a wrong prediction. Assume, for example, that John's father tells him 'You are allowed to eat an apple or a brownie'. The fact that the father cares about his son's health, so the permission to eat an apple seems more reasonable, does not affect John's inference that he can choose between the two desserts. Fox and Katzir (2021) observe that this is a distinct challenge for IRMs (see also Degen et al. 2015, Cremers et al. 2023). Even if there is an IRM that overcomes the first challenge and derives CRDs in the general case, as long as it relies on actual priors – as stems from the pragmatic nature of the theory – the challenge of priors sensitivity remains unresolved.

One approach to this challenge is to weaken the probability sensitivity of IRMs (Franke 2009, Degen et al. 2015, Cremers et al. 2023). However, Fox and Katzir (2021) note that this approach is insufficient for the extreme sensitivity to priors in the case of CRDs. While they do not rule out the possibility that some future weakenings of probability-sensitivity will succeed, the actual weakenings proposed in the literature appear to be inadequate.

Another approach is to entirely discard the use of probabilities and use a non-probabilistic IRM instead. So far, this has not seemed to be a promising approach: as shown in Sect. 2.3.1, a standard version of non-probabilistic identification fails even in the basic case of CRD. One could ask whether the integration of the WMS concept into non-probabilistic settings brings any enhancement. I will show that a non-probabilistic WMS-IRM involving performs better than the one described in Sect. 2.3.1 in terms of CRDs, but is unable to handle Universal CRDs.

According to the non-probabilistic identification criterion in (6), a message identifies a state if there is no other state consistent with that message, where consistency is a binary

property: a message is either consistent with a state or not. But as discussed previously, we could also refer to a scalar property reflecting the consistency of messages *and their parts* (i.e. deletion alternatives) with states – this notion of WMS captured exactly this idea. Therefore, we can propose the non-probabilistic identification criterion in (30), which does not rely on probabilistic factors and is therefore insensitive to priors.

(30) STATE IDENTIFICATION (non-probabilistic version with WMS): a message $m \in M$ *identifies* a state $t \in \Pi$ if for every other $t' \in \Pi$, $w_t(m) > w_{t'}(m)$.

Let us first see that this criterion makes the correct predictions regarding CRDs. Recall the WMSs in the basic case of CRD from (17), repeated in (31). In the first iteration, 'A or B' identifies the state $A \wedge B$ according to the current criterion, because $w_{A\wedge B}(\text{'A or B'}) = 3 > 2 = w_{A\wedge B}(\text{'A'}), w_{A\wedge B}(\text{'B'})$ (as indicated in pink); 'A' and 'B' cannot identify any state at this point because there are multiple states that obtain the highest WMS for each of them. However, after eliminating the state $A \wedge B$ in the first iteration (as indicated in light gray), this tie is resolved. Therefore, in the second step, 'A' identifies A and 'B' identifies B (as indicated in yellow), as desired.

(31) Non-probabilistic identification ('A'/'B'/'A or B'):

Weighted Message Scores

$w_t(m)$	'A'	'B'	'A or B'
$A \wedge \neg B$	1	0	2
$\neg A \wedge B$	0	1	2
$A \wedge B$	1	1	3

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II
cells eliminated after iteration II

Similarly, the desideratum is achieved in the case of 3 disjuncts, as illustrated by Table (32). More generally, the desideratum in (4) is obtained by the non-probabilistic WMS-based identification criterion.[11]

[11] Note that this process is analogous to the probabilistic WMS-based criterion, with the distinction being the order of identification: in the probabilistic version, simpler messages (i.e., with fewer disjuncts) identify their state on earlier steps; in the non-probabilistic variant, more complex messages (i.e., with more disjuncts) identify their state on earlier steps.

(32) Non-probabilistic identification (three disjuncts):

Weighted Message Scores

$P(m\|t)$	'A'	'B'	'C'	'A or B'	'A or C'	'B or C'	'A or B or C'
$A \wedge \neg B \wedge \neg C$	1	0	0	2	2	0	4
$\neg A \wedge B \wedge \neg C$	0	1	0	2	0	2	4
$\neg A \wedge \neg B \wedge C$	0	0	1	0	2	2	4
$A \wedge B \wedge \neg C$	1	1	0	3	2	2	6
$A \wedge \neg B \wedge C$	1	0	1	2	3	2	6
$\neg A \wedge B \wedge C$	0	1	1	2	2	3	6
$A \wedge B \wedge C$	1	1	1	3	3	3	7

identified message-state pairs in iteration I cells eliminated after iteration I
identified message-state pairs in iteration II cells eliminated after iteration II
identified message-state pairs in iteration III cells eliminated after iteration III

While the non-probabilistic WMS-IRM successfully accounts for CRDs regardless of priors, it falls short of deriving Universal CRD. Recall the WMSs for the case of Universal CRD with two disjuncts as shown in (27), repeated below in (33). In the first iteration, 'every x (Ax or Bx)' identifies the state $\forall xAx \wedge \forall xBx$, since this state gets the highest WMS by this message, compared to all other states. No further identification is possible in this step, as there is a tie between multiple states for the other messages. After eliminating 'every x (Ax or Bx)' and $\forall xAx \wedge \forall xBx$, the remaining states still involve a tie. Hence, no identification is possible in the second step and the procedure ends, failing to get the desired meaning to the remaining messages.

(33) Non-probabilistic identification (Universal CRD):

Weighted Message Scores:

$w_t(m)$	'every x (Ax)'	'every x (Bx)'	'every x (Ax or Bx)'	'some x (Ax)'	'some x (Bx)'	'some x (Ax or Bx)'
$\forall xAx \wedge \forall xBx$	1	1	3	1	1	3
$\forall xAx \wedge \exists xBx \wedge \neg\forall xBx$	1	0	2	1	1	3
$\forall xBx \wedge \exists xAx \wedge \neg\forall xAx$	0	1	2	1	1	3
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge\neg\forall xBx \wedge \forall x(Ax \vee Bx)$	0	0	1	1	1	3
$\exists xAx \wedge \neg\forall xAx \wedge \exists xBx \wedge\neg\forall xBx \wedge \neg\forall x(Ax \vee Bx)$	0	0	0	1	1	3
$\forall xAx \wedge \neg\exists xBx$	1	0	2	1	0	2
$\forall xBx \wedge \neg\exists xAx$	0	1	2	0	1	2
$\exists xAx \wedge \neg\forall xAx \wedge \neg\exists xBx$	0	0	0	1	0	2
$\exists xBx \wedge \neg\forall xBx \wedge \neg\exists xAx$	0	0	0	0	1	2

identified message-state pairs in iteration I cells eliminated after iteration I
☐ cells not eliminated after iteration I

We therefore see that the non-probabilistic WMS-IRM successfully addresses the problem of prior sensitivity problem for CRDs but faces a new challenge in accounting for Universal CRD (at least given the current weight function). Furthermore, in Sect. XXX we will see that, like other IRMs, the non-probabilistic WMS-IRM yields wrong predictions in a range of scenarios described by Asherov et al. (2024).

This leads to a third approach, proposed by Fox and Katzir (2021), of IRMs based on a modular architecture. According to that perspective, the source of priors is not external to the mechanism that computes SIs, thus not reflecting actual beliefs about the world. Instead, the priors involved in the probabilistic computation of a modular system are formal constructs defined internally for the system that computes SIs. As such, the priors distribute uniformly and have nothing to do with actual probability assessment. Hence, a modular variation of IRM resolves the challenge of prior sensitivity. What made the notion of a modular IRM problematic in Fox and Katzir (2021) is the perceived inability of IRMs to derive CRD, which meant that an additional exhaustification operation was needed. This left a modular IRM with no real role to play. However, in contrast to other IRMs, the WMS-IRM generates CRDs, making it a plausible basis for a modular IRM that serves as a potential theory for deriving SIs.

Incorporating modularity into the IRM theory raises a question about its pragmatic nature. IRMs have typically been proposed in the literature as a purely pragmatic approach. In particular, the derivation of SIs has been considered to be a result of a general reasoning process. The notion of modularity, which is grammatical in nature, is therefore at odds with the majority of the literature and rather pushes toward a hybrid theory.[12] Note that for a modular WMS-IRM, this is not the only assumption that takes the theory away from its pragmatic nature: the definition of WMS in (16a) neglects the fundamental pragmatic assumption of treating an utterance as a whole. Instead, it adopts the notion of deletion alternatives as defined by Katzir (2007), which is based on the syntactic structures of the assertion.[13] Therefore, a modular WMS-IRM could be considered a viable option for an exhaustification mechanism that relies on both grammatical components and pragmatic concepts.

Of the three potential solutions to the issue of prior sensitivity in IRMs, which emerges when priors are regarded as representing real probabilities, the most convincing is the idea of modular system, as discussed by Fox and Katzir (2021). This explanation is corroborated by various independent studies that present evidence for the modular nature of SI computation, regardless of probabilistic considerations (e.g. Fox and Hackl 2006, Magri 2009). In the next section, I present another evidence for a modular architecture of IRMs.

[12] This is in line, to a certain degree, with several recent studies, such as Franke and Bergen (2020), Champollion et al. (2019), Cremers et al. (2023), which advocate for IRMs (particularly Rational Speech Act models) to involve an encapsulated exhaustification mechanism. However, unlike these theories, which argue for a grammatical derivation of SIs and a pragmatic reasoning to solve disambiguation, a theory of modular IRM argues for an iterative process of probabilistic assessment to derive SIs.

[13] It is possible that other notions of alternatives, which are not based on syntactic structures of messages (e.g., conceptual alternatives, see Buccola et al. 2022). I leave this issue open for future research.

4.2 IRMs and reasoning in reference games

In this section, I briefly outline another empirical issue encountered by IRMs which comes from the domain of reference games, as identified by Asherov et al. (2024), and explain why this is also a problem for the non-modular WMS-IRM. Then, I show that integrating modularity into IRMs addresses this issue, thereby offering further evidence in favor of a modular framework for IRMs.

In their work, Asherov et al. (2024) examine the strengthening of expressions of the form 'Pick the crate with an *x*', where *x* is a single fruit name (like (35)) in a variety of scenarios and show that unlike grammar-based theories, IRMs make wrong predictions, suffering from an overgeneration problem.

As a baseline, consider Scenario A as illustrated in (34) (figures are reproduced from Asherov et al. 2024): crate I is empty, crate II has a banana and nothing else in it, and crate III has a banana and an apple in it. In that scenario, the sentence in (35) is acceptable and can be used to convey that the relevant crate is II.

(34)

Figure 1: Scenario A

(35) Pick the crate with a banana

The acceptability judgment implies that the uniqueness presupposition of the definite article in the utterance is satisfied, that is, the expression 'crate with a banana' refers to exactly one crate. Given that both crates II and III contain a banana, it can be explained if an SI arises, that is, this expression gets the strengthened meaning that has just crate II in its denotation. This SI can be explained by both grammatical and IRM theories. Generally speaking, in the grammatical approach, the expression 'crate with a banana' is close in meaning to 'crate with *only* a banana'. For example, Bar-Lev and Fox's (2017, 2020) theory strengthening proceeds as follows: assuming that the alternatives in that setting are of the form 'crate with an *x*', (i.e., 'crate with an apple', 'crate with a pear' etc.), all of the alternatives to 'crate with a banana' are innocently excludable (i.e., they can be negated consistently with prejacent), and hence the only asserted alternative is the prejacent 'crate with a banana'. The result of affirming banana and negating all other fruits is consistent with crate II only, so the uniqueness presupposition is satisfied as desired.

The IRM approach also arrives at the desired inference, and does it in a similar way it does for the 'some'/'all' scenario: while 'crate with a banana' is compatible with both crates II and III, 'crate with an apple' is unambiguous and conveys only crate III; therefore, 'crate with an apple' could be used to convey directly crate III, which leaves 'crate with a banana' to convey crate II. This strategy is obtained by both non-probabilistic and probabilistic versions of the simple IRM proposed in Fox and Katzir (2021). The non-probabilistic IRM derives it within two steps, as illustrated in table (36). In the first step, the message 'crate with an apple' (abbreviated as 'apple' in the table) identifies the state corresponding to crate III, since it is true in this state and false in the other states. After eliminating this state, in the second step the message 'crate with a banana' (abbreviated as 'banana' in the table) is consistent only with the state corresponding to crate II, and therefore identifies it.

(36) Non-probabilistic identification (Scenario A):
Message-state compatibility

	'apple'	'banana'
crate I	0	0
crate II	0	1
crate III	1	1

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II
cells eliminated after iteration II

The probabilistic IRM can derive the desired meaning within one step. Under the assumption of a naive speaker, the probabilities are as in (37). The message 'crate with an apple' identifies the state corresponding to crate III, since $P(\text{'apple'}|\text{crate III}) = \frac{1}{2}$ is the only positive probability given crate III. The message 'crate with a banana' identifies the state corresponding to crate II at the same iteration, since $P(\text{'banana'}|\text{crate II}) = 1 > P(\text{'banana'}|\text{crate III}) = \frac{1}{2}$.

(37) Probabilistic identification (Scenario A):
Naive speaker's probabilities

$P(m\|t)$	'apple'	'banana'
crate I	0	0
crate II	0	1
crate III	$\frac{1}{2}$	$\frac{1}{2}$

identified message-state pairs in iteration I
cells eliminated after iteration I

In the next step, Asherov et al. (2024) presents Scenario B, which is similar to Scenario A, with the addition of a pear to each of crates I and II:

(38)

Figure 2: Scenario B

Unlike Scenario A, in the current scenario sentence (35) is judged unacceptable in Scenario B because it does not identify a specific crate. The grammatical approach accounts for this judgment: as explained above, the derived meaning of this sentence is similar to that of 'crate with *only* a banana'; however, such a crate does not appear in Scenario B. In contrast, IRMs make the wrong prediction in this scenario, using a strategy similar to the one applied in Scenario A: the message 'crate with an apple' is only consistent with the state corresponding to crate III, and therefore identifies this state, leaving them message 'crate with a banana' to identify the state corresponding to crate II. As shown in the following tables, this is obtained by both the non-probabilistic and probabilistic models:

(39) Non-probabilistic identification (Scenario B):
Message-state compatibility

	'apple'	'banana'	'pear'
crate I	0	0	1
crate II	0	1	1
crate III	1	1	0

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II
cells eliminated after iteration II
identified message-state pairs in iteration III

(40) Probabilistic identification (Scenario B):
Naive speaker's probabilities

$P(m\|t)$	'apple'	'banana'	'pear'
crate I	0	0	1
crate II	0	$\frac{1}{2}$	$\frac{1}{2}$
crate III	$\frac{1}{2}$	$\frac{1}{2}$	0

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II

A possible response to this overgeneration problem in Scenario B is to restrict the probabilistic identification to one iteration.Indeed, a non-iterative model of probabilistic rational

reasoning would yield the correct results for both Scenario A and Scenario B. This is so since unlike Scenario A, where 'crate with a banana' identifies crate II within one step, in Scenario B the identification is achieved only in the second iteration (note that this is not the case with the non-probabilistic IRM, which requires two steps in Scenario A). However, Asherov et al. (2024) point out that this move does not effectively address the problem. Consider Scenario C, which is minimally different from scenario B by moving the pear from crate I to crate III:

(41)

Figure 3: Scenario C

Similarly to Scenario B, the sentence 'crate with a banana' is unacceptable in Scenario C. While it is predicted by the grammatical approach (as there is no crate with only a banana), IRMs incorrectly predict that it is acceptable and identify crate II, even restricted to a single iteration, as shown in (42). More generally, in a scenario where there is an inequality in the number of fruits in a set of crates containing x, it is incorrectly predicted that 'crate with an x' necessarily identifies the crate that has the minimum number of fruits, and does it within one step.

(42) Probabilistic identification (Scenario C):
Naive speaker's probabilities

$P(m\|t)$	'apple'	'banana'	'pear'
crate I	0	0	0
crate II	0	$\frac{1}{2}$	$\frac{1}{2}$
crate III	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

identified message-state pairs in iteration I
cells eliminated after iteration I
identified message-state pairs in iteration II

Asherov et al. (2024) concludes that the kind of strengthening that occurs in reference games is computed grammatically by the covert counterpart of the overt exhaustivity operator 'only' (like *Exh* of Bar-Lev and Fox 2017, 2020) but not by the pragmatic reasoning modeled by different versions of IRMs. Although their argument is based on Fox and Katzir's (2021) simple IRMs, which do not involve any notion of weights, the WMS-IRMs do not improve the picture: assuming that the messages include only a single fruit name, the only deletion alternative to a message is itself. That is, the weight assigned to each message is

binary, and the probabilities are those assigned by a naive speaker. Therefore, both the non-probabilistic and probabilistic WMS-IRMs behave similarly to the naive IRMs, thus leading to the same conclusion.

However, once we assume a modular IRM, the correct results are obtained: in such a system, the entire logical space plays a role in the computation, rather than the states that appear in a specific scenario in such a game. Being pragmatic in nature, non-modular IRMs evaluate messages with respect to the context they are uttered in. The setting of reference games narrows the set of relevant states down to the referents in a game. This is, in effect, another case of a deviation from the flat prior assumption: the prior probabilities are 0 for each state that does not correspond to any of the crates in the given scenario, meaning that $P(t|m) = 0$ if the state t is not part of that scenario. The prior distribution over the states that do appear in that scenario is uniform (since there is no reason to prefer one crate over the other), which leads to a direct comparison of their likelihoods in the identification process, as illustrated in (37). Within a modular IRM framework, on the other hand, the priors are uniform over *all* logically possible states, and therefore, all states play a role in the identification process. In particular, it means that a message can identify a state with no corresponding referent in the game.

Consider, for example, that like in Scenarios B-C, there are three relevant fruits: apple (A), banana (B), and pear (P). Assume that a crate can contain no more than one piece of each fruit. The logically possible states in that context consist of all subsets of $\{A,B,P\}$. If all of these states have the same prior probability, the hearer can directly compare their likelihoods in (43) in the reasoning process. In the first iteration, 'crate with a *banana*' identifies the state corresponding to a crate containing only a banana, because $P(\text{`banana'}|\neg A \wedge B \wedge \neg P) = 1 > P(\text{`banana'}|t')$ for any state $t' \neq \neg A \wedge B \wedge \neg P$. Similarly, 'crate with an *apple*' identifies the crate with only an apple, 'crate with a *pear*' identifies the crate with only a pear. A crate with only a banana does not appear in Scenarios B-C, and therefore the message 'Pick the crate with a *banana*' is predicted to be unacceptable in these scenarios, as desired.

(43) Probabilistic identification (three fruits):
Naive speaker's probabilities

$P(m\|t)$	'apple'	'banana'	'pear'
$\neg A \wedge \neg B \wedge \neg P$	0	0	0
$A \wedge \neg B \wedge \neg P$	1	0	0
$\neg A \wedge B \wedge \neg P$	0	1	0
$\neg A \wedge \neg B \wedge P$	0	0	1
$A \wedge B \wedge \neg P$	$\frac{1}{2}$	$\frac{1}{2}$	0
$A \wedge \neg B \wedge P$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\neg A \wedge B \wedge P$	0	$\frac{1}{2}$	$\frac{1}{2}$
$A \wedge B \wedge P$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

identified message-state pairs in iteration I cells eliminated after iteration I

It is easy to see that the same results generalize for any number n of fruits that are salient in the context. Moreover, it holds for messages of the form 'Pick the crate with an x_1 and ... and an x_k' for any $1 \leq k \leq n$.[14] That is to say, if the IRM is blind to what actual states are apparent, the inferences presented in Asherov et al. (2024) can be derived employing the logic of iterative and probabilistic reasoning applied by IRMs. Hence, changing the IRM framework to be a modular system yields empirical results that were previously unaccountable within various IRMs, thereby aligning it more closely as a potential alternative to other exhaustification mechanisms proposed in the existing literature.

5. Conclusion

In this work, I presented a new conception of a non-naive speaker in the IRM approach, which I referred to as WMS-IRM, that is based on an intuition of 'degree of overlap' between messages and epistemic states, rather than reflecting only whether they are compatible or not. I showed that the WMS-IRM succeeds in deriving CRDs beyond the basic case of 2 disjuncts – in contrast to the IRMs proposed in the literature, as observed by Franke (2009, 2011), van Rooij (2010) and Fox and Katzir (2021). In doing so, this work eliminates Fox and Katzir's (2021) argument against IRMs based on CRDs. This opened the door to adopting their proposal of incorporating a modular architecture into the IRM framework. As Fox and Katzir pointed out, this addresses the problem of IRMs being sensitive to priors. I examined the cases of CRDs and reasoning in reference games (as observed by Asherov et al. 2024), where IRMs produce incorrect predictions due to prior sensitivity, and showed that such failures are effectively eliminated by employing a modular WMS-IRM approach. This change requires a shift in perspective of the nature of IRMs: rather than being a pragmatic framework, it becomes a candidate for an exhaustification mechanism which is based on probabilistic reasoning.

To become a real alternative to existing exhaustification mechanisms and a theory of SIs, more work needs to be done. At the empirical level, all the cases discussed in the present work are accounted for under Bar-Lev and Fox's 2017, 2020 theory. Analysis of additional empirical facts – those explained by other theories of exhaustification and those not – is essential for a further evaluation of the WMS-IRM. For example, extending Spector's (2016) comparison of exhaustification operators to also include WMS-IRM can shed more light on its patterns under various conditions.In addition to this, it is necessary to further reflect on the conceptual motivation of WMS-IRM. While there are conceptual arguments supporting Bar-Lev and Fox's theory, as of now, the conceptual ground for the WMS-IRM remains unclear. Such an investigation will require a deeper understanding of the conceptual motivation behind the basic idea of 'degree of overlap', as well as the various choices the model involves, such as the particular weight function, the use of deletion alternatives, among others.

[14] Intuitively speaking, the more fruits are in a crate, the more messages it is compatible with, and therefore the smaller the probability of a message given that crate is. Hence, the crate with exactly the same fruits as in the message gets the highest probability given that message (in both a naive speaker IRM and a WMS IRM)

References

Alonso-Ovalle, Luis. 2005. Distributing the disjuncts over the modal space. In *North East Linguistic Society (NELS)*, ed. by L. Bateman and C. Ussery, volume 35, 1–12.

Asherov, Daniel, Danny Fox, and Roni Katzir. 2024. Strengthening, exhaustification, and rational inference. To appear in Linguistics and Philosophy.

Bar-Lev, Moshe E., and Danny Fox. 2017. Universal free choice and innocent inclusion. In *Semantics and Linguistic Theory (SALT)*, ed. by Dan Burgdorf, Jacob Collard, Sireemas Maspong, and Brynhildur Stefánsdóttir, volume 27, 95–115.

Bar-Lev, Moshe E., and Danny Fox. 2020. Free choice, simplification, and innocent inclusion. *Natural Language Semantics* 28:175–223.

Benz, Anton. 2006. Utility and relevance of answers. In *Game theory and pragmatics*, 195–219. Springer.

Benz, Anton, and Robert Van Rooij. 2007. Optimal assertions, and what they implicate. a uniform game theoretic approach. *Topoi* 26:63–78.

Bergen, Leon, Roger Levy, and Noah Goodman. 2016. Pragmatic reasoning through semantic inference. *Semantics and Pragmatics* 9.

Blutner, Reinhard. 1998. Lexical pragmatics. *Journal of semantics* 15:115–162.

Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of semantics* 17:189–216.

Bowler, Margit. 2014. Conjunction and disjunction in a language without 'and'. In *Proceedings of the 24th Semantics and Linguistic Theory Conference*, ed. by Todd Snider, Sarah D'Antonio, and Mia Weigand, 137–155.

Breheny, Richard, Nathan Klinedinst, Jacopo Romoli, and Yasutada Sudo. 2018. The symmetry problem: current theories and prospects. *Natural Language Semantics* 26:85–110.

Buccola, Brian, Manuel Križ, and Emmanuel Chemla. 2022. Conceptual alternatives: Competition in language and beyond. *Linguistics and philosophy* 45:265–291.

Champollion, Lucas, Anna Alsop, and Ioana Grosu. 2019. Free choice disjunction as a rational speech act. In *Semantics and linguistic theory*, 238–257.

Chemla, Emmanuel. 2009. Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics* 2:1–33.

Cremers, Alexandre, Ethan G Wilcox, and Benjamin Spector. 2023. Exhaustivity and anti-exhaustivity in the RSA framework: Testing the effect of prior beliefs. *Cognitive Science* 47:e13286.

Degen, Judith, Michael Henry Tessler, and Noah D Goodman. 2015. Wonky worlds: Listeners revise world knowledge when utterances are odd. In *CogSci*.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, ed. by Uli Sauerland and Penka Stateva, 71–120. London: Palgrave Macmillan UK.

Fox, Danny, and Martin Hackl. 2006. The universal density of measurement. *Linguistics and philosophy* 29:537–586.

Fox, Danny, and Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19:87–107.

Fox, Danny, and Roni Katzir. 2021. Notes on iterated rationality models of scalar implicatures. *Journal of Semantics* 38:571–600.

Frank, Michael C., and Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science* 336:998–998.

Franke, Michael. 2009. Signal to act: Game theory in pragmatics. Doctoral dissertation, Universiteit van Amsterdam.

Franke, Michael. 2011. Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics* 4:1–82.

Franke, Michael, and Leon Bergen. 2020. Theory-driven statistical modeling for semantics and pragmatics: A case study on grammatically generated implicature readings. *Language* 96:e77–e96.

Goodman, Noah D, and Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in cognitive science* 5:173–184.

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

Horn, Laurence R. 1972. On the semantic properties of logical operators in English. Doctoral dissertation, University of California, Los Angeles.

Jäger, Gerhard. 2002. Some notes on the formal properties of bidirectional optimality theory. *Journal of Logic, Language and Information* 11:427–451.

Kamp, Hans. 1974. Free choice permission. *Proceedings of the Aristotelian Society* 74:57–74.

Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30:669–690.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from japanese. In *The Third Tokyo Conference on Psycholinguistics*, ed. by Yukio Otsu, 1–25. Hituzi Syobo.

Magri, Giorgio. 2009. A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural language semantics* 17:245–297.

Romoli, Jacopo. 2012. Soft but strong. neg-raising, soft triggers, and exhaustification. Doctoral dissertation, Harvard University.

van Rooij, Robert. 2010. Conjunctive interpretations of disjunctions. *Semantics and Pragmatics* 3:11:1–28.

Rothschild, Daniel. 2013. Game theory and scalar implicatures. *Philosophical Perspectives* 27:438–478.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and philosophy* 27:367–391.

Singh, Raj, Ken Wexler, Andrea Astle-Rahim, Deepthi Kamawar, and Danny Fox. 2016. Children interpret disjunction as conjunction: Consequences for theories of implicature and child development. *Natural Language Semantics* 24:305–352.

Spector, Benjamin. 2016. Comparing exhaustivity operators. *Semantics and Pragmatics* 9:11–1.

Trinh, Tue, and Andreas Haida. 2015. Constraining the derivation of alternatives. *Natural Language Semantics* 23:249–270.

A. Appendix

A.1 WMS-IRM derives CRD: Generalization to n-out-of-k disjunctions ($1 \leq k \leq n$)

This appendix shows that the WMS-IRM (assuming flat priors) achieves the desideratum in (4), by proving Theorem 1. We will use the following notations:
Let $C = \{A_1, ..., A_n\}$ be the individuals that are salient in the context and let us denote for any $m' \in M$ and $s' \in \Pi$:

- $Pos(s') = \{\varphi \in C : \varphi \text{ is a positive conjunct in } s\}$
- $Dis(m') = \{\varphi \in C : \varphi \text{ is a disjunct in } m\}$

Let $m =$'$\mathrm{A_{i_1}}$ or ... or $\mathrm{A_{i_k}}$' be a message and let $s = A_{i_1} \wedge ... \wedge A_{i_k} \wedge \neg A_{i_{k+1}} \wedge ... \wedge \neg A_{i_n}$ be a state, for some $1 \leq k \leq n$.

Theorem 1. m identifies s under WMS-IRM.

To prove this theorem, I will first prove the following lemmas:

Lemma 2. For any $s \neq s' \in \Pi$, $w_s(m) \geq w_{s'}(m)$ and $w_s(m) > w_{s'}(m)$ if $|Pos(s')| = |Pos(s)|$

Proof. By the definition of WMS in (16b), $w_t(m) = |\{m' : m' \in Alt_{del}(m) \wedge m'(t) = 1\} \cdot m(t)$ for any state t. In particular, since all of disjuncts in m are consistent with s, all of deletion alternatives of m are consistent with s. Therefore, $w_s(m) = |Alt_{del}(m)| \geq w_{s'}(m)$. If $|Pos(s')| = |Pos(s)|$, there is an individual $A_{i_j} \in Pos(s)$ such that $A_{i_j} \notin Pos(s')$. Since '$\mathrm{A_{i_j}}$'$\in$ $Dis(m)$, it is also a deletion alternative of m. This alternative makes s' false and therefore:

$$w_{s'}(m) = |\{m' : m' \in Alt_{del}(m) \wedge m'(s') = 1\} \cdot m(s') \leq |Alt_{del}(m) \setminus \{A_{i_j}\}| < |Alt_{del}(m)| = w_s(m)$$

□

Lemma 3. For any $s \neq s' \in \Pi$, if $|Pos(s')| = |Pos(s)|$ then $\sum_{m' \in M} w_s(m') = \sum_{m' \in M} w_{s'}(m')$.

Proof. Assume that $s' = A_{j_1} \wedge ... \wedge A_{j_k} \wedge \neg A_{j_{k+1}} \wedge ... \wedge \neg A_{j_n}$. s and s' are symmetrical: we can define a permutation σ that maps every conjunct A_{i_t} to A_{j_t} and a permutation σ^* over messages by extending σ to each of the disjuncts that compose a message. For every message $m' \in M$, we get $w_s(m') = w_{s'}(\sigma^*(m'))$. Being a permutation of the messages, we get:

$$\sum_{m' \in M} w_s(m') = \sum_{\sigma^*(m') \in M} w_{s'}(\sigma^*(m')) = \sum_{m' \in M} w_{s'}(m')$$

□

Lemma 4. For any $s \neq s' \in \Pi$, if $|Pos(s')| > |Pos(s)|$ then $\sum_{m' \in M} w_{s'}(m') > \sum_{m' \in M} w_s(m')$.

Proof. From Lemma 3 we can assume, without loss of generality, that $Pos(s) \subsetneq Pos(s')$. Hence, there is some positive conjunct in s', A_r, which is a negative conjunct in s. Let m^* be a message containing '$\mathrm{A_r}$' as one of its disjuncts. By the definition of WMS in (16a), we get $w_{s'}(m') \geq w_s(m')$ for any message m^* because any deletion alternative that makes s true also makes s' true and $w_{s'}(m^*) \geq w_s(m^*)$ because the deletion alternative '$\mathrm{A_r}$' only makes s' true. Hence:

$$\sum_{m' \in M} w_{s'}(m') = \sum_{m' \in M \setminus \{m^*\}} w_{s'}(m') + w_{s'}(m^*) > \sum_{m' \in M \setminus \{m^*\}} w_s(m') + w_s(m^*) = \sum_{m' \in M} w_s(m')$$

□

Corollary 5. For any $s \neq s' \in \Pi$, if $|Pos(s')| \geq |Pos(s)|$ then $P(m|s) > P(m|s')$

Proof. There are two cases to consider:

- $|Pos(s')| = |Pos(s)|$: By Lemma 2 $w_s(m) > w_{s'}(m)$ and by Lemma 3 $\sum_{m' \in M} w_s(m') = \sum_{m' \in M} w_{s'}(m')$. Hence:

$$P(m|s) = \frac{w_s(m)}{\sum_{m' \in M} w_s(m')} > \frac{w_{s'}(m)}{\sum_{m' \in M} w_{s'}(m')} = P(m|s')$$

- $|Pos(s')| > |Pos(s)|$: By Lemma 2 $w_s(m) \geq w_{s'}(m)$ and by Lemma 4 $\sum_{m' \in M} w_{s'}(m') > \sum_{m' \in M} w_s(m')$. Hence:

$$P(m|s) = \frac{w_s(m)}{\sum_{m' \in M} w_s(m')} > \frac{w_{s'}(m)}{\sum_{m' \in M} w_{s'}(m')} = P(m|s')$$

□

Lemma 6. Assume that $k \geq 2$ and let $0 < d < k$ be a natural number. Then there are two distinct states $s^* \neq s^{**}$ such that $|Pos(s^*)| = |Pos(s^{**})| = d$ and $P(m|s^*) = P(m|s^{**})$ such that for every state s' with $|Pos(s')| = d$, $P(m|s^*) \geq P(m|s')$. This means that no state with d positive conjuncts gets a higher probability of m than all other states with d positive conjuncts.

Proof. Let $0 \leq d \leq k$ be a natural number. Since $|Pos(s)| = k \geq 2$, there are at least two different subsets of $Pos(s)$ of size $k-d$, denoted by X and Y. Let s^*, s^{**} be states such that $Pos(s^*) = Pos(s) \setminus X$ and $Pos(s^{**}) = Pos(s) \setminus Y$ (that is, similar to s with negation of the individuals in X and Y, respectively). Therefore, $s^* \neq s^{**}$ and $|Pos(s^*)| = |Pos(s^{**})| = d$. m has $2^k - 1$ deletion alternatives, $2^{k-d} - 1$ out of them consist of X element only. Therefore, $w_{s^*}(m) = 2^k - 1 - (2^{k-d} - 1) = 2^k - 2^{k-d}$. Similarly, $2^{k-d} - 1$ out of the deletion alternatives of m consist of Y elements only, and therefore $w_{s^*}(m) = w_{s^{**}}(m) = 2^k - 2^{k-d}$. From Lemma 3 we can conclude that:

$$P(m|s^*) = \frac{2^k - 2^{k-d}}{\sum_{m' \in M} w_{s^*}(m')} = \frac{2^k - 2^{k-d}}{\sum_{m' \in M} w_{s^{**}}(m')} = P(m|s^{**})$$

Let s' be a state with $|Pos(s')| = d$. If $Pos(s') \subset Dis(m)$, then $P(m|s') = P(m|s^*)$ for the same reasoning as shown above. Otherwise, $|Pos(s') \cap Dis(m)| = h < d$. Hence, $2^{k-h} - 1$ out of the deletion alternatives of m consist of only elements in $Dis(m) \setminus Pos(s')$, so $w_{s'}(m) = 2^k - 2^{k-h}$. Given that and Lemma 3, it follows that:

$$P(m|s') = \frac{2^k - 2^{k-h}}{\sum_{m' \in M} w_{s'}(m')} = \frac{2^k - 2^{k-h}}{\sum_{m' \in M} w_{s^*}(m')} < \frac{2^k - 2^{k-d}}{\sum_{m' \in M} w_{s^*}(m')} = P(m|s^*)$$

□

Lemma 7. If m identifies s, then every other message with k disjuncts also identifies its corresponding state in the same iteration.

Proof. Assume that $s' = A_{j_1} \wedge ... \wedge A_{j_k} \wedge \neg A_{j_{k+1}} \wedge ... \wedge \neg A_{j_n}$, such that $s' \neq s$. Let $m' =$'$\mathrm{A_{j_1}}$ or... or $\mathrm{A_{j_k}}$'. By induction on k, we will show that if m identifies s then m' identifies s' in the same iteration.

- $k = 1$: in the first iteration, m identifies s: from Corollary 5, for every $s \neq s'' \in \Pi$, $P(m|s) > P(m|s'')$. Therefore, m identifies s as required. This holds for any message with an individual disjunct because there is a symmetry between the cases (as shown in the proof of Lemma 3). Therefore, m' identifies s' in the first iteration. In particular, it means that if m identifies s, then m' identifies s' in the same iteration.

- Assume that the statement is true for any $d < k$. We will show that it is also true for k. Assume that m identifies s in the ith iteration, and assume by contradiction that m' does not identify s' in that iteration. That is, there is a state s^* that has not been eliminated yet, such that $P(m'|s^*) \geq P(m'|s')$. By Corollary 5, s^* has fewer positive conjuncts than s', meaning $|Pos(s')| > |Pos(s^*)|$. Let us denote $|Pos(s^*)| = d$. There are two cases to consider:

  - If there is a state with d positive conjuncts that was identified before the ith iteration, then by the induction assumption, s^* was also identified in that iteration. Therefore, s^* was eliminated before the ith iteration – a contradiction to the assumption that s^* has not been eliminated.
  - Otherwise, no state with d positive conjuncts has been eliminated up to this point. Following the proof in Lemma 3, $P(m|s) = P(m'|s')$, and there is a state s^{**} with d positive conjuncts such that $P(m|s^{**}) = P(m'|s^*)$ (formally, $s^{**} = \sigma^*(s^*)$, where σ^* is the permutation defined for s and s' in Lemma 3). Therefore, $P(m|s^{**}) \geq P(m|s)$. Since $|Pos(s^{**})| = d$, it has not been eliminated yet. Hence, m cannot s in the ith iteration – contradiction.

  Hence, m' identifies s' in the ith iteration.

□

Now, we can prove Theorem 1:

Proof (Theorem 1). Assume by contradiction that m does not identify s. There are two cases to consider:

- m identifies another state $s' \neq s$: from Corollary 5, $d = |Pos(s')| < |Pos(s)| = k$. From Lemma 7, all other states with positive conjuncts d have not been eliminated up to this point. From Lemma 6, there are two distinct states $s^* \neg s^{**}$ with d positive disjuncts such that $P(m|s^*) = P(m|s^{**}) \geq P(m|s')$. Therefore, m does not identify s' – contradiction.

- m does not identify any state: that is, in each iteration, there is no strict maximum among the probabilities of m given states. From Lemma 7, it means that all other messages with k disjuncts do not identify their corresponding states. In this case, no state containing k positive conjuncts has been identified. If such a state existed, it would imply that another message incorrectly identified that state. However, as shown above, this results in a contradiction. Note that $k \geq 2$: as shown in the proof of Lemma 7, all messages consisting of an individual identify their corresponding states as desired in the first iteration. Therefore, we can assume without loss of generality that for any $d < k$, a message containing d disjuncts identified its corresponding state in some iteration. From Lemma 7, this implies that all messages with d disjuncts have identified their corresponding state in some iteration. Therefore, after these iterations, all remaining states have at least k positive conjuncts. However, in that scenario $P(m|s) > P(m|s')$ for every remaining state s'. Therefore, m identifies s at this step – in contradiction to the assumption that m does not identify any state.

□

A.2 WMS-IRM and Universal CRD with 3 disjuncts

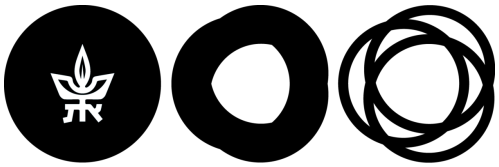
$w_t(m)$	'every x (Ax)'	'every x (Bx)'	'every x (Cx)'	'every x (Ax or Bx)'	'every x (Ax or Cx)'	'every x (Bx or Cx)'	'every x (Ax or Bx or Cx)'	'some x (Ax)'	'some x (Bx)'	'some x (Cx)'	'some x (Ax or Bx)'	'some x (Ax or Cx)'	'some x (Bx or Cx)'	'some x (Ax or Bx or Cx)'	Σ
$\neg\exists xAx \wedge \neg\exists xBx \wedge \neg\exists xCx$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\exists xAx \wedge \neg\exists xBx \wedge \neg\exists xCx$	0	0	0	0	0	0	0	1	0	0	2	2	0	4	9
$\neg\exists xAx \wedge \exists xBx \wedge \neg\exists xCx$	0	0	0	0	0	0	0	0	1	0	2	0	2	4	9
$\neg\exists xAx \wedge \neg\exists xBx \wedge \exists xCx$	0	0	0	0	0	0	0	0	0	1	0	2	2	4	9
$\exists xAx \wedge \exists xBx \wedge \neg\exists xCx \wedge \neg\forall x(Ax \vee Bx)$	0	0	0	0	0	0	0	1	1	0	3	2	2	6	15
$\exists xAx \wedge \exists xBx \wedge \neg\exists xCx \wedge \forall x(Ax \vee Bx)$	0	0	0	1	0	0	2	1	1	0	3	2	2	6	18
$\exists xAx \wedge \neg\exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Cx)$	0	0	0	0	0	0	0	1	0	1	2	3	2	6	15
$\exists xAx \wedge \neg\exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Cx)$	0	0	0	0	1	0	2	1	0	1	2	3	2	6	18
$\neg\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	0	0	0	0	0	1	1	2	2	3	6	15
$\neg\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Bx \vee Cx)$	0	0	0	0	0	1	2	0	1	1	2	2	3	6	18
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx \vee Cx)$	0	0	0	0	0	0	0	1	1	1	3	3	3	7	19
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx \vee Cx) \wedge \neg\forall x(Ax \vee Bx) \wedge \neg\forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	0	0	0	1	1	1	1	3	3	3	7	20
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge \neg\forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	1	0	0	2	1	1	1	3	3	3	7	22
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx) \wedge \forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	0	1	0	2	1	1	1	3	3	3	7	22
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx) \wedge \neg\forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	0	0	1	2	1	1	1	3	3	3	7	22
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge \forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	1	1	0	3	1	1	1	3	3	3	7	24
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge \neg\forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	1	0	1	3	1	1	1	3	3	3	7	24
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx) \wedge \forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	0	1	1	3	1	1	1	3	3	3	7	24
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge \forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	1	1	1	4	1	1	1	3	3	3	7	26
$\neg\exists xAx \wedge \neg\exists xBx \wedge \forall xCx$	0	0	1	0	2	2	4	0	0	1	0	2	2	4	18
$\neg\exists xAx \wedge \exists xBx \wedge \forall xCx$	0	0	1	0	2	2	4	0	1	1	2	2	3	6	24
$\exists xAx \wedge \neg\exists xBx \wedge \forall xCx$	0	0	1	0	2	2	4	1	0	1	2	3	2	6	24
$\exists xAx \wedge \exists xBx \wedge \forall xCx \wedge \neg\forall x(Ax \wedge Bx)$	0	0	1	0	2	2	4	1	1	1	3	3	3	7	28
$\exists xAx \wedge \exists xBx \wedge \forall xCx \wedge \forall x(Ax \wedge Bx)$	0	0	1	1	2	2	5	1	1	1	3	3	3	7	30
$\neg\exists xAx \wedge \forall xBx \wedge \neg\exists xCx$	0	1	0	2	0	2	4	0	1	0	2	0	2	4	18
$\neg\exists xAx \wedge \forall xBx \wedge \exists xCx$	0	1	0	2	0	2	4	0	1	1	2	2	3	6	24
$\exists xAx \wedge \forall xBx \wedge \neg\exists xCx$	0	1	0	2	0	2	4	1	1	0	3	2	2	6	24
$\exists xAx \wedge \forall xBx \wedge \exists xCx \wedge \neg\forall x(Ax \wedge Bx)$	0	1	0	2	0	2	4	1	1	1	3	3	3	7	28
$\exists xAx \wedge \forall xBx \wedge \exists xCx \wedge \forall x(Ax \wedge Bx)$	0	1	0	2	1	2	5	1	1	1	3	3	3	7	30
$\forall xAx \wedge \neg\exists xBx \wedge \neg\exists xCx$	1	0	0	2	2	0	4	1	0	0	2	2	0	4	18
$\forall xAx \wedge \neg\exists xBx \wedge \exists xCx$	1	0	0	2	2	0	4	1	0	1	2	3	2	6	24
$\forall xAx \wedge \exists xBx \wedge \neg\exists xCx$	1	0	0	2	2	0	4	1	1	0	3	2	2	6	24
$\forall xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Bx \wedge Cx)$	1	0	0	2	2	0	4	1	1	1	3	3	3	7	28
$\forall xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Bx \wedge Cx)$	1	0	0	2	2	1	5	1	1	1	3	3	3	7	30
$\forall xAx \wedge \forall xBx \wedge \neg\exists xCx$	1	1	0	3	2	2	6	1	1	0	3	2	2	6	30
$\forall xAx \wedge \forall xBx \wedge \exists xCx$	1	1	0	3	2	2	6	1	1	1	3	3	3	7	34
$\forall xAx \wedge \neg\exists xBx \wedge \forall xCx$	1	0	1	2	3	2	6	1	0	1	2	3	2	6	30
$\forall xAx \wedge \exists xBx \wedge \forall xCx$	1	0	1	2	3	2	6	1	1	1	3	3	3	7	34
$\neg\exists xAx \wedge \forall xBx \wedge \forall xCx$	0	1	1	2	2	3	6	0	1	1	2	2	3	6	30
$\exists xAx \wedge \forall xBx \wedge \forall xCx$	0	1	1	2	2	3	6	1	1	1	3	3	3	7	34
$\forall xAx \wedge \forall xBx \wedge \forall xCx$	1	1	1	3	3	3	7	1	1	1	3	3	3	7	38

$P(m\|t)$	'every x (Ax)'	'every x (Bx)'	'every x (Cx)'	'every x (Ax or Bx)'	'every x (Ax or Cx)'	'every x (Bx or Cx)'	'every x (Ax or Bx or Cx)'	'some x (Ax)'	'some x (Bx)'	'some x (Cx)'	'some x (Ax or Bx)'	'some x (Ax or Cx)'	'some x (Bx or Cx)'	'some x (Ax or Bx or Cx)'
$\neg\exists xAx \wedge \neg\exists xBx \wedge \neg\exists xCx$	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\exists xAx \wedge \neg\exists xBx \wedge \neg\exists xCx$	0	0	0	0	0	0	0	$\frac{1}{9}$	0	0	$\frac{2}{9}$	$\frac{2}{9}$	0	$\frac{4}{9}$
$\neg\exists xAx \wedge \exists xBx \wedge \neg\exists xCx$	0	0	0	0	0	0	0	0	$\frac{1}{9}$	0	$\frac{2}{9}$	0	$\frac{2}{9}$	$\frac{4}{9}$
$\neg\exists xAx \wedge \neg\exists xBx \wedge \exists xCx$	0	0	0	0	0	0	0	0	0	$\frac{1}{9}$	0	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{4}{9}$
$\exists xAx \wedge \exists xBx \wedge \neg\exists xCx \wedge \neg\forall x(Ax \vee Bx)$	0	0	0	0	0	0	0	$\frac{1}{15}$	$\frac{1}{15}$	0	$\frac{1}{5}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{2}{5}$
$\exists xAx \wedge \exists xBx \wedge \neg\exists xCx \wedge \forall x(Ax \vee Bx)$	0	0	0	$\frac{1}{18}$	0	0	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{18}$	0	$\frac{3}{18}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{3}$
$\exists xAx \wedge \neg\exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Cx)$	0	0	0	0	0	0	0	$\frac{1}{15}$	0	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{5}$	$\frac{2}{15}$	$\frac{2}{5}$
$\exists xAx \wedge \neg\exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Cx)$	0	0	0	0	$\frac{1}{18}$	0	$\frac{1}{9}$	$\frac{1}{18}$	0	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{3}{18}$	$\frac{1}{9}$	$\frac{1}{3}$
$\neg\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	0	0	0	0	0	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{5}$	$\frac{3}{18}$	$\frac{2}{5}$
$\neg\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Bx \vee Cx)$	0	0	0	0	0	$\frac{1}{18}$	$\frac{1}{9}$	0	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{3}{18}$	$\frac{1}{3}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx \vee Cx)$	0	0	0	0	0	0	0	$\frac{1}{19}$	$\frac{1}{19}$	$\frac{1}{19}$	$\frac{3}{19}$	$\frac{3}{19}$	$\frac{3}{19}$	$\frac{7}{19}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx \vee Cx) \wedge$ $\neg\forall x(Ax \vee Bx) \wedge \neg\forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	0	0	0	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{7}{20}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge$ $\neg\forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	$\frac{1}{22}$	0	0	$\frac{1}{11}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{3}{22}$	$\frac{3}{22}$	$\frac{3}{22}$	$\frac{7}{22}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx) \wedge$ $\forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	0	$\frac{1}{22}$	0	$\frac{1}{11}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{3}{22}$	$\frac{3}{22}$	$\frac{3}{22}$	$\frac{7}{22}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx) \wedge$ $\neg\forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	0	0	$\frac{1}{22}$	$\frac{1}{11}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{3}{22}$	$\frac{3}{22}$	$\frac{3}{22}$	$\frac{7}{22}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge$ $\forall x(Ax \vee Cx) \wedge \neg\forall x(Bx \vee Cx)$	0	0	0	$\frac{1}{24}$	$\frac{1}{24}$	0	$\frac{1}{8}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{7}{24}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge$ $\neg\forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	$\frac{1}{24}$	0	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{7}{24}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Ax \vee Bx) \wedge$ $\forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	0	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{7}{24}$
$\exists xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Ax \vee Bx) \wedge$ $\forall x(Ax \vee Cx) \wedge \forall x(Bx \vee Cx)$	0	0	0	$\frac{1}{26}$	$\frac{1}{26}$	$\frac{1}{26}$	$\frac{2}{13}$	$\frac{1}{26}$	$\frac{1}{26}$	$\frac{1}{26}$	$\frac{3}{26}$	$\frac{3}{26}$	$\frac{3}{26}$	$\frac{3}{26}$
$\neg\exists xAx \wedge \neg\exists xBx \wedge \forall xCx$	0	0	$\frac{1}{18}$	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	0	0	$\frac{1}{18}$	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{2}{9}$
$\neg\exists xAx \wedge \exists xBx \wedge \forall xCx$	0	0	$\frac{1}{24}$	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	0	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{4}$
$\exists xAx \wedge \neg\exists xBx \wedge \forall xCx$	0	0	$\frac{1}{24}$	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{24}$	0	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{4}$
$\exists xAx \wedge \exists xBx \wedge \forall xCx \wedge \neg\forall x(Ax \wedge Bx)$	0	0	$\frac{1}{28}$	0	$\frac{1}{14}$	$\frac{1}{14}$	$\frac{1}{7}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{3}{28}$	$\frac{3}{28}$	$\frac{3}{28}$	$\frac{1}{4}$
$\exists xAx \wedge \exists xBx \wedge \forall xCx \wedge \forall x(Ax \wedge Bx)$	0	0	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{6}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{7}{30}$
$\neg\exists xAx \wedge \forall xBx \wedge \neg\exists xCx$	0	$\frac{1}{18}$	0	$\frac{1}{9}$	0	$\frac{1}{9}$	$\frac{2}{9}$	0	$\frac{1}{18}$	0	$\frac{1}{9}$	0	$\frac{1}{9}$	$\frac{2}{9}$
$\neg\exists xAx \wedge \forall xBx \wedge \exists xCx$	0	$\frac{1}{24}$	0	$\frac{1}{12}$	0	$\frac{1}{12}$	$\frac{1}{6}$	0	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{4}$
$\exists xAx \wedge \forall xBx \wedge \neg\exists xCx$	0	$\frac{1}{24}$	0	$\frac{1}{12}$	0	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{24}$	$\frac{1}{24}$	0	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
$\exists xAx \wedge \forall xBx \wedge \exists xCx \wedge \neg\forall x(Ax \wedge Bx)$	0	$\frac{1}{28}$	0	$\frac{1}{14}$	0	$\frac{1}{14}$	$\frac{1}{7}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{3}{28}$	$\frac{3}{28}$	$\frac{3}{28}$	$\frac{1}{4}$
$\exists xAx \wedge \forall xBx \wedge \exists xCx \wedge \forall x(Ax \wedge Bx)$	0	$\frac{1}{30}$	0	$\frac{1}{15}$	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{6}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{7}{30}$
$\forall xAx \wedge \neg\exists xBx \wedge \neg\exists xCx$	$\frac{1}{18}$	0	0	$\frac{1}{9}$	$\frac{1}{9}$	0	$\frac{2}{9}$	$\frac{1}{18}$	0	0	$\frac{1}{9}$	$\frac{1}{9}$	0	$\frac{2}{9}$
$\forall xAx \wedge \neg\exists xBx \wedge \exists xCx$	$\frac{1}{24}$	0	0	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{6}$	$\frac{1}{24}$	0	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{4}$
$\forall xAx \wedge \exists xBx \wedge \neg\exists xCx$	$\frac{1}{24}$	0	0	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{6}$	$\frac{1}{24}$	$\frac{1}{24}$	0	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
$\forall xAx \wedge \exists xBx \wedge \exists xCx \wedge \neg\forall x(Bx \wedge Cx)$	$\frac{1}{28}$	0	0	$\frac{1}{14}$	$\frac{1}{14}$	0	$\frac{1}{7}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{3}{28}$	$\frac{3}{28}$	$\frac{3}{28}$	$\frac{1}{4}$
$\forall xAx \wedge \exists xBx \wedge \exists xCx \wedge \forall x(Bx \wedge Cx)$	$\frac{1}{30}$	0	0	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{30}$	$\frac{1}{6}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{7}{30}$
$\forall xAx \wedge \forall xBx \wedge \neg\exists xCx$	$\frac{1}{30}$	$\frac{1}{30}$	0	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{5}$	$\frac{1}{30}$	$\frac{1}{30}$	0	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{5}$
$\forall xAx \wedge \forall xBx \wedge \exists xCx$	$\frac{1}{34}$	$\frac{1}{34}$	0	$\frac{3}{34}$	$\frac{1}{17}$	$\frac{1}{17}$	$\frac{3}{17}$	$\frac{1}{34}$	$\frac{1}{34}$	$\frac{1}{34}$	$\frac{3}{34}$	$\frac{3}{34}$	$\frac{3}{34}$	$\frac{7}{34}$
$\forall xAx \wedge \neg\exists xBx \wedge \forall xCx$	$\frac{1}{30}$	0	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{5}$	$\frac{1}{30}$	0	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{5}$
$\forall xAx \wedge \exists xBx \wedge \forall xCx$	$\frac{1}{34}$	0	$\frac{1}{34}$	$\frac{1}{17}$	$\frac{3}{34}$	$\frac{1}{17}$	$\frac{3}{17}$	$\frac{1}{34}$	$\frac{1}{34}$	$\frac{1}{34}$	$\frac{3}{34}$	$\frac{3}{34}$	$\frac{3}{34}$	$\frac{7}{34}$
$\neg\exists xAx \wedge \forall xBx \wedge \forall xCx$	0	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{10}$	$\frac{1}{5}$	0	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{10}$	$\frac{1}{5}$
$\exists xAx \wedge \forall xBx \wedge \forall xCx$	0	$\frac{1}{34}$	$\frac{1}{34}$	$\frac{1}{17}$	$\frac{1}{17}$	$\frac{3}{34}$	$\frac{3}{17}$	$\frac{1}{34}$	$\frac{1}{34}$	$\frac{1}{34}$	$\frac{3}{34}$	$\frac{3}{34}$	$\frac{3}{34}$	$\frac{7}{34}$
$\forall xAx \wedge \forall xBx \wedge \forall xCx$	$\frac{1}{38}$	$\frac{1}{38}$	$\frac{1}{38}$	$\frac{3}{38}$	$\frac{3}{38}$	$\frac{3}{38}$	$\frac{7}{38}$	$\frac{1}{38}$	$\frac{1}{38}$	$\frac{1}{38}$	$\frac{3}{38}$	$\frac{3}{38}$	$\frac{3}{38}$	$\frac{7}{38}$

identified message-state pairs in iteration I identified message-state pairs in iteration II identified message-state pairs in iteration III
cells eliminated after iteration I cells eliminated after iteration II cells eliminated after iteration III

תקציר

בספרות הבלשנית הוצגו עמדות שונות לגבי טבען של אימפליקטורות סקלריות והאופן שבו הן נגזרות. על פי הגישה הפרגמטית, אימפליקטורות סקלריות הן תופעה פרגמטית הנגזרת ממנגנון היסק כללי, המושפע רק מהדינמיקה של שיחה. על פי הגישה הדקדוקית המתחרה, אימפליקטורות סקלריות הן היסקים לוגיים הנגזרים באופן קומפוזיציונלי. בשנים האחרונות, יש עניין גובר בקבוצת מודלים פרגמטיים, המכונים "מודלים איטרטיביים רציונליים", המציעים מנגנון איטרטיבי, לרוב הסתברות, של היסקים כלליים. העבודה הנוכחית מציעה שינוי של מודלים אלו המחליף את ההנחה הנפוצה של "דובר נאיבי" בפרספקטיבה אחרת לגבי הדובר, המבוססת על הסתברויות משוקללות. שינוי זה פותר בעיה מרכזית עבור המודלים הללו — עליו הצביעו Franke (2009, 2011) , Van Rooji (2010) ו־Fox & Katzir (2021) — בגזירת קריאים קוניונקטיביים של דיסיונקציות (למשל, משפטי Free Choice). לנוכח הצלחתה בגזירת קריאים אלו, ההצעה הנוכחית מאפשרת לדון בארכיטקטורה מודולרית של מודלים איטרטיביים רציונליים. ארכיטקטורה זו, יכולה להסביר מגוון מקרים נוספים בהם נכשלו המודלים הלא מודולריים שהוצעו בעבר.



החוג לבלשנות
בית הספר לפילוסופיה, בלשנות ולימודי מדע

מודלים רציונליים איטרטיביים וקריאים קוניונקטיביים של דיסיונקציות

חיבור זה הוגש כעבודת גמר לקראת התואר
"מוסמך אוניברסיטה" – M.A. באוניברסיטת תל אביב

על ידי
עלמה פרישוף

בהנחיית
פרופ' רוני קציר ד"ר משה בר־לב

ספטמבר 2025