A Satisficing Account of Conjunction Fallacies*

Ash Asudeh
Department of Linguistics
University of Rochester

December 4, 2020
Workshop on Natural Language Meaning

1 Overview

- This talk is about a topic of long-standing interest in the psychology of reasoning: *conjunction fallacies*.
- Tversky and Kahneman (1983) propose that conjunction fallacies arise because experimental participants employ a *representativeness* heuristic.
  - Representativeness must be based on a relation of *similarity*.
- We propose instead that the conjunction fallacies arise because experimental participants employ a *satisficing* heuristic (Simon 1956).
  - Satisficing is based on a relation of *optimization*, relative to some defined optimum.
- We believe satisficing to be more satisfying, because there are well-defined ways to compute optima.
- A feature of previous literature on this problem has been the lack of a compositional analysis of *and* that predicts the correct pattern of experimental results.
- We offer such an analysis, based on a *monad* built around *tropical semirings* (Simon 1988).

Computational tools

- We have developed a couple of implementations that you can play with:
  1. A Jupyter Notebook, which is useful as an exercise tool to play with the framework and to try things out.
     `https://github.com/gianlucagiorgolo/nls-with-ct-exercises`
  2. A theorem prover, does what it says on the tin: It proves theorems based on sequents that you provide and displays any proofs it finds.
     `http://www.sas.rochester.edu/lin/sites/asudeh/tp.html`

*This talk is based on joint work with Gianluca Giorgolo, in particular chapter 6 of Asudeh and Giorgolo (2020). Any errors in this talk are my own.

1.1 Outline

- Section 2 lays out the more general project developed in Asudeh and Giorgolo (2020).
- Section 3 lays out the background on conjunction fallacies and issues of compositionality raised by previous approaches.
- Section 4 lays out the intuitions behind our model.
- Section 5 discusses how monads can model uncertainty as probability.
- Section 6 discusses how a previous monadic solution that we have proposed (Giorgolo and Asudeh 2014) needs to appeal to general Gricean reasoning to fully explain conjunction fallacies.
- Section 7 introduces tropical semirings, explains how this predicts the pattern of experimental results without appeal to Gricean pragmatics, and shows that the result captures the duality of conjunction and disjunction.
- Section 8 concludes.

2 Our project

- The *Enriched Meanings* project (Asudeh and Giorgolo 2020) is about using a concept from category theory and functional programming — *monads* — to model some murkier aspects of natural language meaning.[1]
- Monads have been successfully used in the semantics of programming languages to characterize certain classes of computation (Moggi 1989, Wadler 1992, 1994, 1995).
- Our original inspiration came from Shan (2001), who sketched how monads could potentially be used to offer new solutions to certain problems in natural language semantics and pragmatics.
- Before I answer the question of what we mean by *enriched meanings*, it is useful to build up some common ground about semantic theory.
  - We adhere to the fundamental assumption of formal semantics, the *Principle of Compositionality* (PoC), which requires that there be some basic stock of meanings that can be combined to yield the meanings of larger expressions that are built out of the basic meanings. The non-basic expressions can then be composed into yet larger expressions, and so on.
    - Language is productive/generative and it is also interpreted. So something like PoC simply must be true. It seems axiomatic.
  - The basic meanings must be stored somewhere, so let's assume there's a lexicon of some kind.
    - PoC does not place particularly substantive restrictions on the nature of lexical meanings, other than requiring them to support composition.

[1] For other research in this vein, see: Bumford (2015, 2017a,b), Bumford and Barker (2013), Charlow (2014, 2019a, 2018, 2019b), Cohn-Gordon (nd), Giorgolo and Unger (2009), van Eijck and Unger (2010), Maršik (2016), Shan (2010), Unger (2011), White, Needle, Charlow, and Bumford (2017). Monads are also related to *continuations* (Wadler 1994), which are the formal tool used in a rich body of work by Chris Barker and Ken Shan (Barker 2002, 2012, 2018, 2019, Barker and Shan 2006, 2008, 2014), and also in their collaboration with Raffaella Bernardi (Barker, Bernardi, and Shan 2010).

- There has often been a tendency in formal semantics to generalize lexical meanings to the worst case.
  - For example, based on the necessity of treating *some* nominals as generalized quantifiers, Montague (1973) famously treated *all* nominals as generalized quantifiers.
  - Similarly, in order to capture the semantics of intensional verbs, Montague treated simpler, extensional verbs equivalently, but had their lexical entries disregard the intensional parameters. In other words, because some verbs need intensional parameters, all verbs have them.
  - Yet another example is the use of assignment functions to interpret variables: Every interpretation is relativized to an assignment function, whether what is being interpreted contains variables or not. Again, because assignment functions are needed in some cases, they are there in all cases.

- If we want linguistic semantics to be a linguistically and philosophically sound science, we should avoid generalizing to the worst case.

  - Serious steps were taken in this direction through the development of *type shifting rules* (Partee and Rooth 1983, Partee 1986), which allow certain meanings to be listed in a simpler form lexically but to be shifted to a more complex form if compositionally necessary.
    - For example, a proper name can be listed lexically as denoting an individual, but can be shifted to a generalized quantifier if necessary in composition (e.g., when conjoined with a generalized quantifier built around a quantificational determiner).
    - Moreover, some type shift rules are logically justified (van Benthem 1991), including the shift from an individual to a generalized quantifier.
    - However, this doesn't hold for all type shifting, such as shifting from the type of a predicative adjective to that of an attributive adjective.
    - Moreover, type shifting is not formally restricted, unless we restrict it to, for example, only logically justified type shifting.
    - Restrictions instead come from linguistic observation; i.e, shifts are motivated by whatever happens in natural language.
    - But this means that type shifting is descriptive, not explanatory, since the only restrictions come from the very data it seeks to explain, rather than from some formally defined theoretical framework for mapping simple meanings to complex meanings.

- This is where enriched meanings come in. An enriched meaning is precisely the result of mapping some input space of objects and relations to a richer space.

  - To do this in a principled way, we use monads.

- Much like type shifting, enriched meanings allow us to avoid generalizing basic meanings to the worst case, but in a more principled way.

- This is particularly useful in formally capturing some aspects of meaning that seem to straddle the boundary between semantics and pragmatics, without destroying the boundary, which seems conceptually and empirically necessary.

- Furthermore, we believe that *formal* analysis of semantic and pragmatic phenomena is necessary to properly understand the phenomena.

- Thus, enriched meanings provide a principled way to capture a mixture of simpler and more complex meanings and also provide empirical benefits in shedding new light on tricky phenomena at the semantics–pragmatics boundary.
- Conjunction fallacies, today's topic, are an instance of such phenomena.

3 Background

- Tversky and Kahneman (1983) noticed that, in a task asking for ratings of the relative likelihoods of different propositions being true, the majority of experimental participants consistently rated the likelihood of the conjunction of two propositions being true as higher than the likelihood that one of the conjoined propositions is true.
- One of their examples is the well-known 'Linda paradox' or 'Linda effect'.
  - Experimental participants were given statements like (1).

    (1) Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.
  - As part of the experimental task, they where asked to rank the probability that various statements were true of Linda; the resulting ranking for the relevant cases is given in (2), where $>$ indicates "ranked likelier than" (by experimental participants).

    (2) Linda is active in the feminist movement. [F(eminist)M(ovement)]

    $>$Linda is a bank teller and is active in the feminist movement. [B(ank)T(eller)&FM]

    $>$Linda is a bank teller. [BT]
- The context is designed to bias towards the label **Feminist** and it is unproblematic and unsurprising that the relevant proposition (FM: *Linda is active in the feminist movement*) is ranked most likely, but the result that the joint probability BT&FM is ranked higher than BT is interesting, and constitutes an instance of conjunction fallacy: A conjunction of two propositions is reported by subjects to be more probably than the probability of one of the two propositions on its own.
- This result is problematic given probability theory, which holds that the likelihood of the conjunction of two events cannot be greater than the likelihood of either of the two events. Formally, for any two events A and B, the standard law for joint probability is:

  (3) $P(A \text{ and } B) \leq P(A), P(B)$
- The Linda effect has been replicated by various studies that have investigated different ways in which this apparently fallacious response can be elicited (among others, Yates and Carlson 1986, Tentori et al. 2004).
- Tversky and Kahneman (1983) explained these results in terms of a notion of representativeness.
  - They claimed that the observed responses are due to the fact that people do not operate in terms of probabilistic reasoning, but instead use a *representativeness heuristic*.
  - According to Tversky and Kahneman (1983), experimental participants check the degree of correspondence between the events whose likelihood they are judging and a certain model of reality and select as likely those events that are more representative of what the model predicts as being the more likely event.

- Representativeness tends to covary with frequency but not necessarily.

- A crucial point of Tversky and Kahneman's analysis is that this heuristic operates on the conjunction as a whole, or as they put it:

  > [T]he judged probability (or representativeness) of a conjunction cannot be computed as a function (e.g., product, sum, minimum, weighted average) of the scale values of its constituents. (Tversky and Kahneman 1983: p. 305)

  - This is essentially a claim that the conjunction is non-compositional, which is rather problematic for any attempt to integrate their observations with linguistic theories of meaning composition.

- Aspects of Tversky and Kahneman's explanation has been challenged by a number of researchers.
  - Hertwig and colleagues have proposed that conjunction fallacies are not real errors, but rather emerge because of the intrinsic ambiguity of linguistic operators such as the conjunction *and* (Hertwig and Gigerenzer 1999, Mellers et al. 2001, Hertwig et al. 2008).
  - Another important contribution of this line of research has been the demonstration that in certain contexts conjunction fallacies do not arise so easily.
    - This is particularly true of contexts in which experimental participants are somehow primed to reason in terms of frequencies.
      1. When presented with a scenario that explicitly introduces frequencies
      2. When required to explicitly express their judgements regarding the likelihood of different events in terms of numerical estimates (rather than only doing so implicitly by ordering the events in terms of likelihood)
      3. When the "stakes are raised", for example in betting scenarios
  - However, Hertwig et al.'s explanation is itself linguistically and philosophically problematic, since it posits an otherwise unmotivated ambiguity in the logical operator *and*.
    - This runs against the *Modified Occam's Razor* (a.k.a. *Grice's razor*): 'Senses are not to be multiplied beyond necessity' (Grice 1978: 118).'
    - Kripke (1977: 268) similarly counsels against positing ambiguity (unless all else fails), characterizing it as 'the lazy man's approach'.

- It would be preferable, then, to have a compositional account of conjunction fallacies (*contra* Tversky and Kahneman) that posits no ambiguity for *and* (*contra* Hertwig et al.).

- We try to reconcile these different points of view on the basis of the data reported in the literature and our goal of maintaining compositionality and unambiguity.

4 The intuitions behind our model

- Our model starts from the assumption that people employ multiple strategies when evaluating the likelihood of conjoined uncertain events, an assumption shared with Yates and Carlson (1986).
  - However, Yates and Carlson (1986) propose a less principled "signed summation model" that is not based on probability theory, unlike ours.
  - They also assume that unrelated computational processes underpin the different strategies.
- We show that it is possible to assume a single uniform process that computes the likelihood of the conjunction of two events from their two relative likelihoods.
  - This uniform process may use different but related representations of uncertainty, expressed as alternative algebraic structures, yielding different results.
  - We can explain the results reported in the literature in terms of an algebraic structure known as a *semiring*.
  - This mathematical object is at the heart of *monads*.
- This yields more than one possibility. How is the choice between them made?
- We assume that the observation made in the literature (Hertwig and Gigerenzer 1999) that conjunction fallacies arise only under specific conditions and can be cancelled under other conditions is explained in terms of cognitive/computational economy
  - In fact, the same computational structure, the monad, can be used together with a number of different underlying semirings, one of them being the probability semiring.
  - We predict that, in general, if people are presented with a task where there are "no stakes" they will reason based on a representation corresponding to a semiring defined over a relatively simple set with generally simple operations.
  - Using this strategy will normally lead people to make overconfident estimates, which may result in conjunction fallacies.
  - If, on the other hand, people are forced to evaluate the consequences of their judgements, such as in the context of a gaming scenario, or if explicitly primed to think in frequentist terms, then they can switch to a more complex representation, with properties that better approximate those of probability theory.
  - In such circumstances, people would not necessarily commit conjunction fallacies.
  - Crucially, in our model, logical operators such as *and* or *or* maintain their core logical meaning, while the probabilistic behaviour is determined by the context in which they operate.
  - In this sense, with respect to Hertwig et al. (2008)'s analysis, we move the ambiguity to the context rather than assuming that a word like *and* has multiple meanings.
  - As I noted above, this is in fact a problem with Hertwig et al. (2008)'s analysis, which effectively treats *and* as ambiguous, i.e. truly polysemous.
- Our approach is similar in spirit to the standard Gricean approach (Grice 1975), although it also differs in some important ways.
  - Like the Gricean view, we assume that logical words such as *and* have a non-ambiguous core logical meaning.

  - In the standard Gricean approach the core meaning is further elaborated by *conversational implicatures* — additional meanings that are not explicitly stated by the speaker but inferable by the speaker based on conversational goals (Grice's Cooperative Principle and conversational maxims) and real world knowledge.
  - Crucially, conversational implicatures are determined by context and can be cancelled by explicitly stating that they do not hold.
    - For instance, it is a consequence of the standard Gricean view that a numeral like *two* is usually understood to mean "exactly two", as in *Wilgefortis has two daughters*, but the implicature can be cancelled, as witnessed by examples like *Wilgefortis has two daughters, in fact she has three*.

- Similarly, in our approach the core meaning of a word like *and* is its logical, non-probabilistic meaning.
  - Instead of relying on implicatures defined by the context, we have a grammatical process (as part of the logic of composition) that can lift this core meaning to the probabilistic level if other linguistic expressions introduce uncertainty in the interpretation process.
  - For instance, if *and* is used to conjoin two uncertain propositions, then the conjunction itself will become uncertain.
  - The relevant role of context in our model is not to permit the derivation of related additional meanings, but instead to select which specific mode of uncertainty is going to be used.
  - Notice that this does not introduce a source of non-compositionality in our system: Both modes of uncertainty — pure probability and the simpler approximation — function completely compositionally.
  - The limited form of ambiguity that may be attributed to our system stems entirely from the choice of the mode of uncertainty, which is motivated on cognitive grounds.
  - In contrast, in Hertwig et al. (2008)'s analysis, *and* is treated as truly lexically ambiguous, and its final interpretation requires context as an additional parameter.

- Lastly, I want to stress that our approach, in its current form, is limited to the explanation of the conjunction of two events, as expressed by the conjunction of two propositions.
  - The related phenomenon of *concept conjunction*, as discussed for instance by Hampton (1988), shares a number of similarities with conjunction fallacies, but at the same time displays some crucial differences, at least at the linguistic level.
  - For example, in the case of noun-noun compounds, such as *school furniture*, the denotation of the expression is not constructed by simply intersecting the denotations of its component parts, but the interpretation function operates in non-trivial ways on these meanings (Jackendoff 2010).
  - This is not the case for the conjunction of events (propositions, in our case), where the result is not an entirely novel type of event, but is indeed the boolean meet of the two sub-events.

5 Monads and uncertainty

- As a preface to what follows, we need to familiarize ourselves with the related notion of *semirings*.
- A semiring is:
  - Some set A,
  - With two distinguished elements 0 and 1, and
  - Two binary operations $+$ and $\cdot$,
  - Such that the following axioms are satisfied, for all x, y and $z \in A$:

(4) $(x + y) + z = x + (y + z)$

(5) $x + y = y + x$

(6) $x + 0 = 0 + x = x$

(7) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

(8) $x \cdot 1 = 1 \cdot x = x$

(9) $x \cdot 0 = 0 \cdot x = 0$

(10) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$

(11) $(x + y) \cdot z = (x \cdot z) + (y \cdot z)$

- In the case of the probability semiring, A is the real interval $[0, 1]$, with 0 and 1 representing the two identity elements and $+$ and $\cdot$, defined respectively as addition ($+$) and multiplication ($\cdot$).
- A monad can intuitively think of them as a way to map between different types of objects, in particular as a way to map simple objects into more complex ones and relations between the simple objects to relations between the more complex ones.
- The fact that probability distributions form a monad in the category of measurable spaces was an early discovery in category theory (Lawvere 1962, Giry 1982).
- In our case, we can view the the probability monad maps from semantic types to probability distributions over the inhabitants of the types.
  - For example, if we lift a proper name like *Kim* using the probability monad, we would get a probability distribution over individuals; i.e., the semantic type e would be mapped to $\Diamond e$, where the latter is a probability distribution over inhabitants of e.
  - If we lift a predicate like *slept* (i.e., an arrow), we would get a function from probability distributions over individuals to probability distributions over truth values; i.e., the semantic type $e \to t$ would be mapped to $\Diamond e \to \Diamond t$.
- A monad is a particular type of mathematical structure, a type of endofunctor that is equipped with two natural transformations that encode the notions of *embedding* and *joining/composing/combining* (this intuition is from Kuhnle 2013).
  - Monads are thus good at modelling embedding relations and also at compositionally combining results.
  - Related to this is their ability to map one space of objects and relations to a more complex space of objects and relations while preserving properties of the input space.

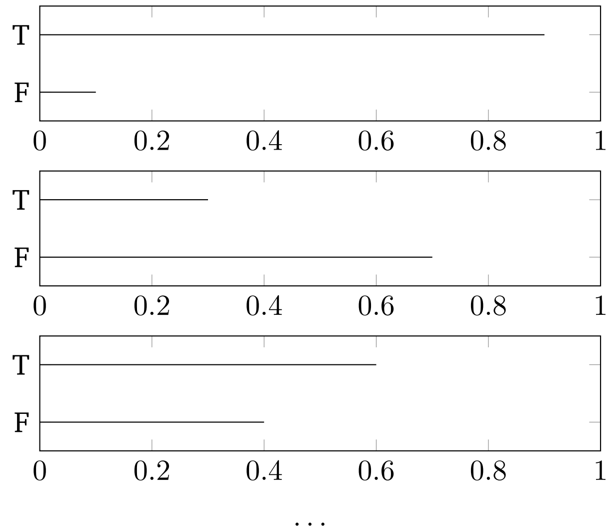


Figure 1: Some inhabitants of $P(t)$

  - See Asudeh and Giorgolo (2020: chapter 3) for background on category theory, including endofunctors, natural transformations, and monads.

- Formally our monad is defined by the triple $\langle P, \eta, \star \rangle$.
  - P is a functor that operates in the way described above: It maps our semantic types to probability distributions over the inhabitants of the type and lifts the mappings between the types so that they operate between probability distributions.
    - For instance, in the case of the type t of truth values the inhabitants of P are going to look like the distributions shown in Figure 1.
  - *Unit*, η, is an operation that creates a special type of probability distribution that corresponds to a categorical judgement, i.e. assigning the entire probability mass to a single element of a semantic type.
    - Basically this is a way to integrate certainty into our model of uncertainty. Formally we define it as follows (where we represent probability distributions as functions from types to the interval $[0,1]$, or more generally to any base set of a semiring):

      (12) $\eta(x) = \lambda y. \begin{cases} 1 \text{ if } y = x \\ 0 \text{ otherwise} \end{cases}$

      where x represents an element of any semantic type (e.g. an individual, a truth value, a collection of individuals)
    - For instance in the case of η_t, its application to the truth value T will result in the distribution shown in Figure 2.
  - *Bind*, $\star$, is how we combine uncertainties.
    - Its definition is based on the definition of joint probability.
    - Recall that the joint probability of two events is computed as follows:

      (13) $P(A \text{ and } B) = P(A) \cdot P(B|A)$
    - It is important to notice that joint probability, as in (13), is not what we are actually trying to model in the case of two conjoined propositions (e.g., *Linda is a bank teller and Linda is active in the feminist movement*).

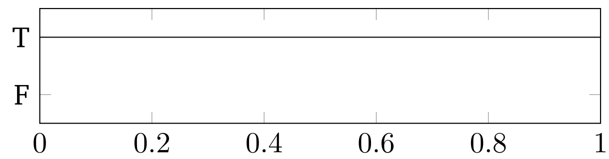


Figure 2: $\eta_t(T)$

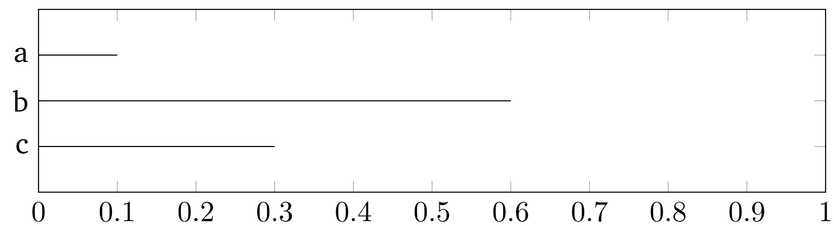


Figure 3: Example distribution over $\{a, b, c\}$.

  - What we want is derived from joint probability, but is in fact somewhat coarser, and is at the core of the process in our model that constructs the probability of an event from the linguistic elements that describe it.
  - This coarsened notion of joint probability, which we will discuss in what follows, arises as the byproduct of the interactions of the uncertainties associated with the composed expressions.
- There is also another difference between the way the *bind* operation is used and what is standardly meant by joint probability in (13).
  - While the joint probability of two events gives us the likelihood of both atomic events occurring together, *bind* returns the probability distribution of what we can consider another truly *atomic* event.
  - In other words, *bind* effectively combines the results of several relevant joint probabilities.
  - This will become clearer in looking at some examples below, but first we define *bind* as in (14).

    (14) $m \star k = \lambda y. \sum_{x \in S} m(x) \cdot k(x)(y)$
    where S is the set of elements measured by the probability distribution m
  - *Bind* takes two arguments, a probability distribution (m) and a function (k) from elements of the domain of the probability distribution m to probability distributions over elements of a (possibly different) set.
  - The resulting probability distribution is obtained by collecting all possible ways in which we can obtain the various results from this second set, and by scaling them (hence the use of $\cdot$) using the likelihood that they emerge from the first distribution.
  - The results are collected together using summation ($\sum$).
  - Let's look at an example to better understand the workings of *bind*. Assume that the first argument of *bind* is represented by the probability distribution over some type A with three inhabitants, $\{a, b, c\}$, as represented in Figure 3.
- The second argument of *bind* is a function that maps each element of $\{a, b, c\}$ to a probability distribution over some type, let's say the type of truth values.
  - For instance we could have the function in Figure 4.
- The standard joint probability would give us the distribution shown in Figure 5.

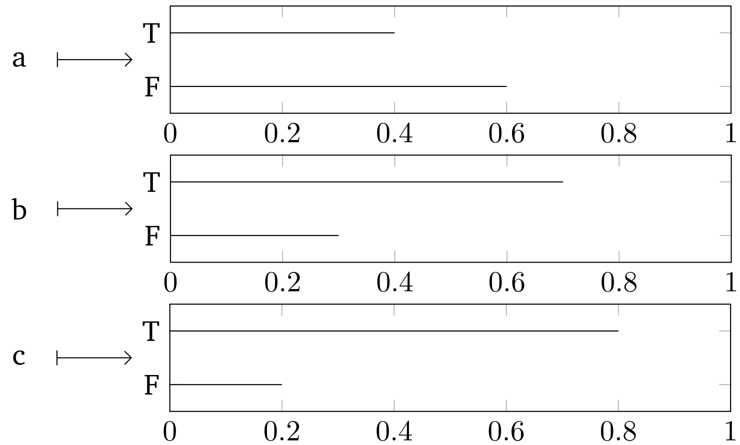


Figure 4: Example of a function from $\{a, b, c\}$ to probability distributions over truth values.

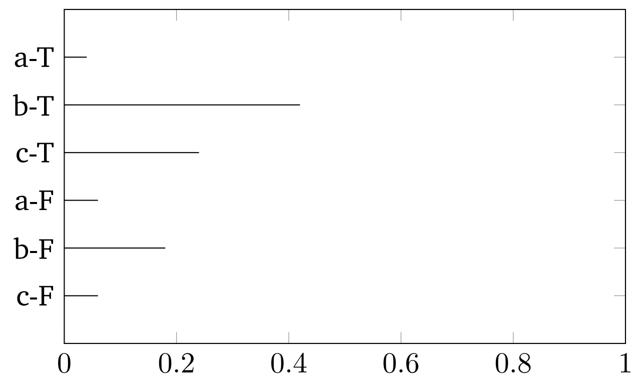


Figure 5: Pure joint probability for $\{a, b, c\}$ and $\{T, F\}$.

- But *bind* actually generates a distribution probability over the final type of its second argument, so the result we get is instead the one shown in Figure 6.
- What *bind* returns for T in Figure 6 is the sum of a-T, b-T and c-T in Figure 5; similarly for F and a-F, b-F, and c-F. You can check this by eye.

- Thanks to these proofs we can substitute any semiring for the probability semiring and still use our monadic calculus to derive the meaning of complex expressions.
- Giorgolo and Asudeh (2014) offered a semiring that we called the *the one semiring* as a candidate solution for the problem of conjunction fallacies.
- We used the results of Yates and Carlson results, which are summarized in Table 1.

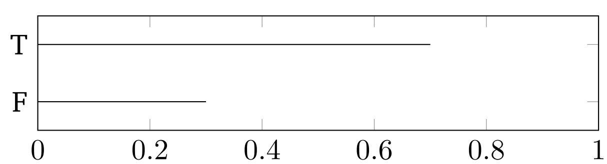


Figure 6: The result of applying *bind* to its two arguments

Likelihood of event A	Likelihood of event B	Observed rating	Intuition
U(nlikely)	U	$P(A \text{ and } B) \leq P(A), P(B)$	Probabilistically correct (No Linda Effect)
U	L(ikely)	$P(A) \leq P(A \text{ and } B) \leq P(B)$	Probabilistically incorrect (Standard Linda Effect)
L	L	$P(A), P(B) \leq P(A \text{ and } B)$	Probabilistically incorrect (Strong Linda Effect)

Table 1: Results reported by Yates and Carlson (1986).

- The first row shows the result of the conjunction of two events that are unlikely.
  - For example, given the Linda context in (1) above, the two component events could be *Linda is pro-life* and *Linda is a bank teller* and their conjunction *Linda is a pro-life bank teller*, which we take as shorthand for *Linda is pro-life and Linda is a bank teller*.
  - Yates and Carlson observe that the conjoined event is deemed by experimental participants to be (even) less likely than the component events. This is probabilistically correct, as it upholds the standard law for the joint probability of two events in (3). The probabilistically correct result therefore displays no Linda effect and we label it as such.
- The second row shows the combination of an unlikely event, such as *Linda is a bank teller*, with a likely event, such as *Linda is a feminist*.
  - Yates and Carlson observed that participants rate the unlikely event as less likely than the conjunction of the two events and the conjunction in turn as less likely than the likely event.
  - This is probabilistically incorrect, as it does not uphold (3).
  - In other words, this is Tversky and Kahneman's standard Linda effect and we label it as such.
- The third row shows a scenario that Yates and Carlson tested but that Tversky and Kahneman did not.
  - It is the combination of a likely event, such as *Linda is leftwing*, with another likely event, such as *Linda is a feminist*, such that the conjunction would be *Linda is a leftwing feminist*, which we take as shorthand for *Linda is leftwing and Linda is a feminist*.
  - Yates and Carlson observed that participants rate the conjunction of the likely events as (even) more likely than each likely event on its own.
  - This is again probabilistically incorrect, given (3). We label this a *Strong Linda Effect*, as it features two errors of the kind observed by Tversky and Kahneman (1983).

- Giorgolo and Asudeh (2014) used a simple discrete set as the base for the semiring: {I(mpossible), U(unlikely), P(ossible), L(ikely), C(ertain)}. I and C correspond to 0 and 1 respectively.
  - The only additional condition that needs to be imposed so that I and C behave as boolean values is that, for all x in our set, $x + C = C$.
  - There are sixteen possible well-behaved semirings that we can define for this set.
- Giorgolo and Asudeh (2014) showed that of these semirings, only one semiring both correctly models conjunction and predicts the Yates and Carlson (1986) pattern of results.

$+$	I	U	P	L	C
I	I	U	P	L	C
U	U	U	P	L	C
P	P	P	P	L	C
L	L	L	L	L	C
C	C	C	C	C	C

$\cdot$	I	U	P	L	C
I	I	I	I	I	I
U	I	U	U	U	U
P	I	U	P	P	P
L	I	U	P	L	L
C	I	U	P	L	C

Table 2: The one semiring.

- The one semiring is shown in Table 2.
  - This semiring is not homomorphic to the probability semiring, meaning that we cannot reproduce its behaviour using probability theory.
  - This semiring is just a min-max semiring, if we impose the strict order I $<$ U $<$ P $<$ L $<$ C, where $\cdot$ is min and $+$ is max.

6 A Gricean explanation?

- Recall that the Linda scenario is as follows, where $>$ indicates "ranked likelier than" (by experimental participants, according to the original work by Tversky and Kahneman 1983):

(15) Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.
Linda is active in the feminist movement. [F(eminist)M(ovement)]
$>$ Linda is a bank teller and is active in the feminist movement. [B(ank)T(eller)&FM]
$>$ Linda is a bank teller. [BT]

- The ranking that Tversky and Kahneman observed is FM $>$ BT&FM $>$ BT.
- Probability theory would instead dictate the ranking BT,FM $\geq$ BT&FM.
- As noted above, Giorgolo and Asudeh (2014) developed a semiring that we called the *the one semiring* to tackle this problem.
  - The one semiring, which models a lazy probabilistic evaluation in the absence of stakes, instead yields the ranking FM $>$ BT&FM, BT.
- We just need to break the tie between BT&FM and BT in order to capture the ranking that Tversky and Kahneman actually observed.
- How could Gricean reasoning do this?
  - The context that Tversky and Kahneman provide for the Linda example (*mutatis mutandis* for their other stimuli) is such that *Linda is a bank teller* flouts the Maxim of Relation (Relevance): It is not relevant to the prior context that Linda is a bank teller, whereas it is potentially relevant that she is a feminist.

- Thus, the statement that Linda is a bank teller creates an unnecessary implicature that would cause participants to reconsider the context in an aberrant way, perhaps calculating that the experimenters for some bizarre reason believe that bank tellers are generally interested in "issues of discrimination and social justice", but this likely clashes with the real world knowledge of most competent adults, who have no reason to believe that bank tellers should particularly care about such things.

- Thus, the one semiring gives us the ranking FM $>$ BT&FM and the Maxim of Relation gets us BT&FM $>$ BT.
  - In other words, FM $>$ BT&FM is a kind of entailment, while BT&FM $>$ BT is a kind of implicature.
  - This is analogous to the standard Gricean explanation of exclusive *or*: Using only inclusive *or* and *and*, we derive $a \vee b \wedge \neg(a \wedge b)$, where the first conjunct is an entailment and the second is a scalar implicature.

- Asudeh and Giorgolo (2020) pursued a different strategy.
  - We concluded that a Gricean explanation based around the Maxim of Relation is explanatorily weak, since what counts as relevant is hard to define formally or to capture systematically (they boil down to the same thing, in our view), Relevance Theory notwithstanding (Sperber and Wilson 1986, 1995).
  - We therefore thought it profitable to investigate whether there was some other structure that might yield the conjunction fallacy pattern, under reasonable assumptions.
  - This does invalidate Gricean explanations in other domains, for us.

7 Let's get tropical

- It is clear that experimental participants are not in general using the probability semiring, otherwise the Linda effect would not be observed, but it consistently has been (see Lu 2016 for a useful recent review).

- The one semiring does not fully capture the Linda effect — it requires an appeal to Gricean reasoning.

- Here we take a different approach and instead use the family of semirings commonly known as *tropical semirings* (Simon 1988, Pin 1998).
  - These have previously found linguistic applications in weighted finite-state automata and transducers (Kuich and Salomaa 1986), which have been used for morphological analysis, machine translation, and speech recognition and synthesis, among other applications (Mohri and Sproat 1996, Mohri et al. 2008, Mohri 2009).

- Tropical semirings can be defined using various base sets.
  - Here we choose to use the set of real numbers $\mathbb{R}$ as the base.
  - Having chosen a base set, we can either construct a *min-plus* semiring or a *max-plus* semiring.
    - But the *min-plus* and *max-plus* semirings are isomorphic, which means that we can freely choose either one.

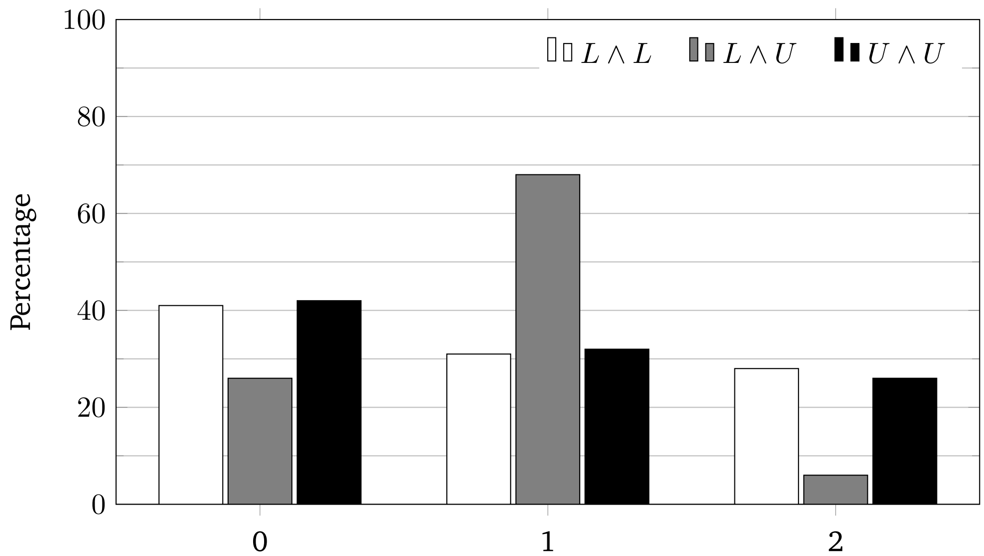


Figure 7: Percentages of zero, single, and double conjunction fallacies by likelihood of component propositions (see Lu 2016: 522, Fig. 3).

 – In fact, given the equivalence of the min-plus and max-plus semirings, there is no motivation for the assumption that speakers necessarily use one instead of the other.
 – We therefore assume that speakers freely use both.

- This will be important in explaining the pattern of data reported in the literature, as shown in Figure 7, which is based on Lu (2016: 522, Fig. 3).
  - This bar graph shows the proportions of different kinds of conjunction fallacies that experimental participants made.
- A zero on the x-axis of conjunction errors in Figure 7 indicates that the conjunction of the two component propositions was ranked as less likely than either of the component propositions on its own, i.e. that things are as they should be according to probability theory.[2]
  - This is what we called *No Linda Effect* in Table 1 above.
- A one on the x-axis is the classic Linda effect, in which the conjunction is ranked as more likely than one of the component propositions, but less likely than the other.
  - What we called a *Standard Linda Effect* in Table 1 would be an instance of this.
- A two on the x-axis indicates that the conjunction is ranked as more likely than both of the two component propositions.
  - This is what we called a *Strong Linda Effect* in Table 1.
- The results can be summarized as follows.

[2]Technically it's "less likely than or equal in likelihood to". I'll use "less likely than" or "more likely than" as shorthand for "less likely than or equal in likelihood to" and "more likely than or equal in likelihood to", respectively. It's less cumbersome.

- ○ Given a conjunction of a likely proposition (e.g., *Linda is a feminist*, in light of the context in the Linda scenario) and an unlikely proposition (e.g., *Linda is a bank teller*), the vast proportion of responses were conjunction fallacies (nearly 70%), and most of the remaining responses were probabilistically correct (about 25%).
- ○ There is thus a clear peak at one conjunction error.
- ○ The remaining data is particularly interesting.
  - If instead two likely propositions are conjoined (e.g., *Linda is a feminist* and *Linda is leftwing*, in the Linda scenario), a slightly larger proportion of responses were without conjunction errors (just over 40%) and the remaining responses were split almost evenly between one error and two errors (just over and just under 30%, respectively).
  - The pattern for the conjunction of two unlikely propositions (e.g., *Linda is a bank teller* and *Linda is pro-life*, in the Linda scenario) is virtually identical.
- ○ There are three key things to notice in these results with respect to a conjunction of two likely or two unlikely propositions.
  1. There is a small peak in both cases at zero conjunction errors.
  2. Over a quarter of responses for both cases make two conjunction errors, which does *not* happen for the actual Linda effect cases (where such errors occur in only about 5% of responses).
  3. The proportion of one conjunction error and two conjunction errors in both cases is almost the same (about 30%).

- The two take home points are these.
  1. It seems that experimental participants are somewhat confused about how to rank the conjunction of two likely or two unlikely propositions relative to the component propositions.
     - ○ Thus we observe nearly a random distribution of responses among zero, one, and two conjunction errors.
  2. However, only the responses at one conjunction error and two conjunction errors seem to be fully randomly distributed with respect to each other: Each of these two error types has about the same proportion of responses.
     - ○ There is a small peak at zero errors, which indicates that somehow the conjunction of two likely or two unlikely propositions a) is a bit more likely to cause zero errors than it is to cause one error or two errors; and b) is less likely to lead to an error than the Linda case of a conjunction of a likely and an unlikely proposition.

- In order to explain how tropical semirings give us a way to model the data trends shown in Figure 7, we have to define them and see how they can be used to represent uncertainty for our purposes.

- We start with the *min-plus* semiring. The semiring is defined on the set $\mathbb{R} \cup \{+\infty\}$, the set of real numbers together with positive infinity.
  - ○ The two operations are defined as follows:

    (16) $x + y = \mathbf{min}(x, y)$

    (17) $x \cdot y = x + y$

    where **min** is a function that returns the smallest of two numbers. The unit of $+$, 0, is $+\infty$, and the unit of $\cdot$, 1, is 0.

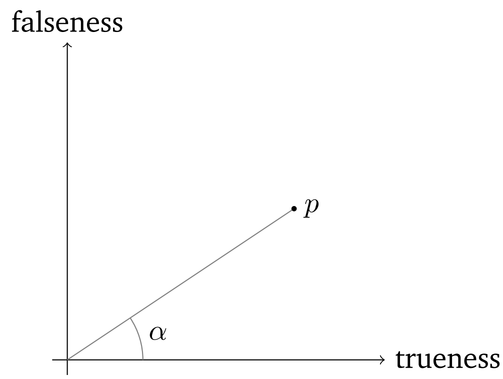


Figure 8: The likelihood of the proposition p represented as a point in the trueness/falseness space.

- The *max-plus* semiring is specified in a similar way.
  - The semiring is defined on the set $\mathbb{R} \cup \{-\infty\}$, the set of real numbers together with negative infinity. The two operations are defined as follows:

    (18) $x + y = \mathbf{max}(x, y)$

    (19) $x \cdot y = x + y$

    where **max** is a function that returns the greatest of two numbers. The unit of $+$, 0, is $-\infty$, and the unit of $\cdot$, 1, is again 0.
- When a proposition is *lifted* in the monadic space of tropical semirings it becomes a function from truth values to elements of its base set ($\mathbb{R} \cup \{+\infty\}$ for the *min-plus* semiring and $\mathbb{R} \cup \{-\infty\}$ for the *max-plus* semiring).
  - This means that we can represent the function in its tabular form as a two-coordinate vector — one coordinate representing the value assigned to T(true) and the other the value assigned to F(alse). This in turns means that we can represent a proposition graphically as a vector in the "trueness/falseness" space, as shown in Figure 8.
  - The "trueness" component (the projection of the vector on the "trueness" axis) of the vector represents some measure of confidence in the fact that p is true, while the "falseness" component represents the same for the fact that p is false.
  - We can get a single measure of the likelihood of p by checking the angle α formed by the vector and the "trueness" axis: an angle of $0°$ represents total confidence in the truth of the proposition, an angle of $45°$ represents a completely uncertain judgement, while an angle of $90°$ represents absolute falsehood.

- To bootstrap the analysis we assign some fixed values to the atomic propositions that we are going to conjoin.
  - The values for the components of the vectors are arbitrary and reflect the general assumptions about the likelihood of the atomic propositions used in the analysis of the Linda effect.
  - In our case a likely proposition will be associated with a vector that forms an angle smaller than $45°$, while an unlikely proposition will form an angle larger than $45°$.
  - An example is shown in Figure 9.

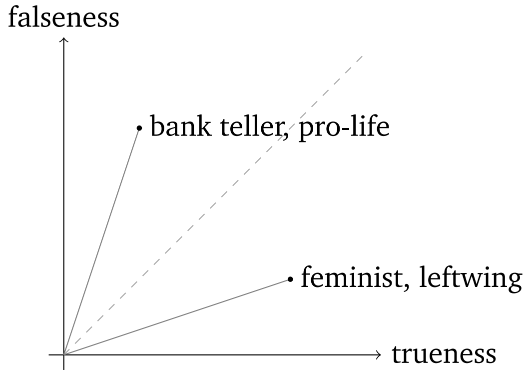
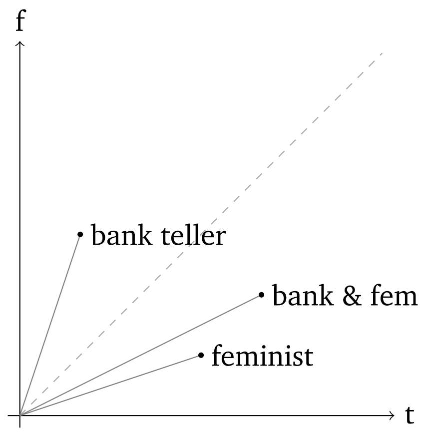
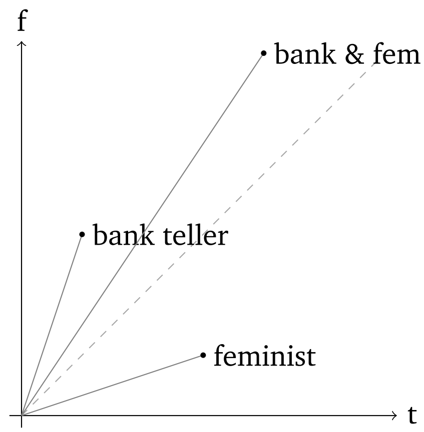


Figure 9: Examples of likely and unlikely propositions in the trueness/falseness space.

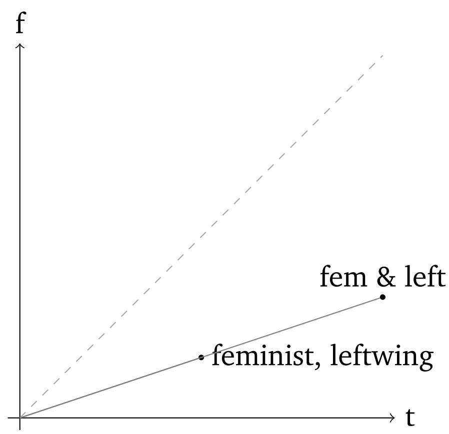
- We noted above that, given the equivalence of the *min-plus* and *max-plus* tropical semirings, there is no reason to assume that people use only one and not the other.
  - We therefore assume that people have both tropical semirings at their disposal and may use either one or even both.
    - In other words, we assume that people have cognitive equivalents of both semirings.
  - In the Linda scenerio, we assume as part of the experimental conditions, that the proposition *Linda is a bank teller* is unlikely and the proposition *Linda is a feminist* is likely.
    - This is the same assumption that Tversky and Kahneman (1983) make.
  - We therefore assign each some value in the monadic space — on the false side of 45° in the case of the unlikely proposition and on the true side in the case of the likely proposition — and keep the assumption constant across both tropical semirings.
  - In the case of the conjunction of the unlikely and likely propositions — which leads to the observed Linda effects/conjunction fallacies in the psychological literature — the two semirings yield identical ordering results, as shown in Figure 10a and Figure 10b.
    - For both semirings, the unlikely proposition that *Linda is a bank teller* comes out as more false than the conjunction of this unlikely proposition with the likely proposition that *Linda is a feminist* and this conjunction in turn comes out as more false than the likely proposition on its own.
    - Thus, both semirings capture the Linda effect.
- So how does this explain the results reported in Figure 7?
  - Notice that Figure 7 (and a similar graph in Yates and Carlson 1986) is reporting total number of responses and thus summing across all experimental participants.
  - The mass of responses clearly peaks at one conjunction error (nearly 70%) in the case of the conjunction of an unlikely and a likely proposition, with a smaller peak at zero conjunction errors (about 25%).
    - We assume that the small mass of responses at two conjunction errors is noise.
  - There are two possible explanations for this, both of which converge on the same result.
    1. If an individual in any particular response exercises the choice to use the *min-plus* or *max-plus* semiring, they will commit a single conjunction error (i.e., a Linda effect).



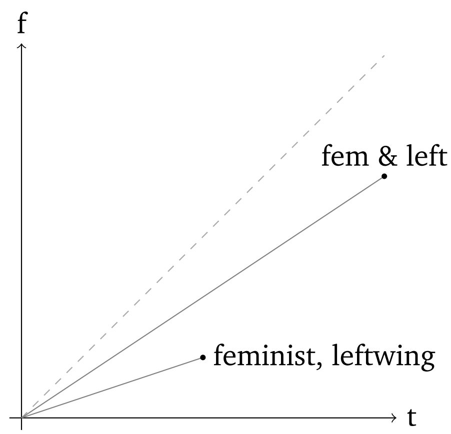
(a) Conjunction of an unlikely and a likely proposition, given the *min-plus* tropical semiring.



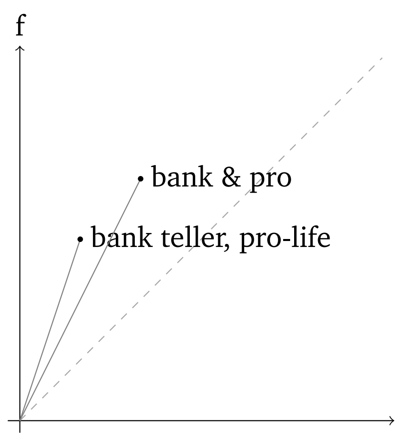
(b) Conjunction of an unlikely and a likely proposition, given the *max-plus* tropical semiring.



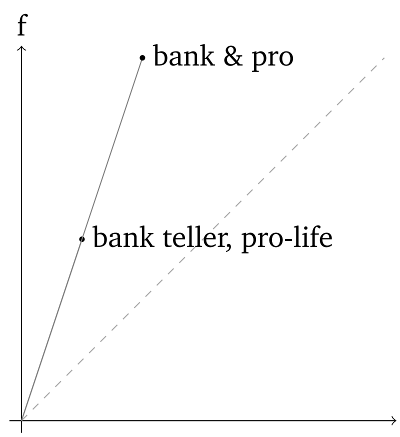
(c) Conjunction of two likely propositions, given the *min-plus* tropical semiring.



(d) Conjunction of two likely propositions, given the *max-plus* tropical semiring.



(e) Conjunction of two unlikely propositions, given the *min-plus* tropical semiring.



(f) Conjunction of two unlikely propositions, given the *max-plus* tropical semiring.

Figure 10: Conjunction of propositions, given the *min-plus* (left column) or *max-plus* (right column) tropical semirings.

    - Even if an individual were to use both semirings, the results in fact converge, as explained above; this again results in a single conjunction error.
    - It is only if an individual instead uses the probability semiring that they will not commit a conjunction error.
        - Thus, of the three semirings in play, two will lead to a Linda effect (one conjunction error) — whether each semiring on its own or both together — and one will lead to probabilistically accurate reasoning (zero conjunction errors).
    - It is thus entirely unsurprising according to our tropical model that by far the greatest mass of responses is gathered at one conjunction error when an unlikely proposition is conjoined with a likely one.
    - It is similarly unsurprising that there is another small peak at zero conjunction errors, because one of the three semirings in play delivers that result.

- Let us now turn to the other two cases:
    1. The conjunction of two likely propositions
    2. The conjunction of two unlikely propositions.

- Here the two tropical semirings yield divergent results.

- Once again, we start by assigning the two component propositions some appropriate point in the monadic space.
    - For example, given the Linda scenario, we place the two likely propositions *Linda is a feminist* and *Linda is leftwing* in the true part of the monadic space, as shown in Figure 10c and Figure 10d.
    - Similarly, we place the two unlikely propositions *Linda is a bank teller* and *Linda is pro-life* in the false part of the monadic space, as shown in Figure 10e and Figure 10f.

- But now notice that the *min-plus* and *max-plus* semirings make different predictions.
    - The *min-plus* semiring collapses the order of two likely propositions with their conjunction: All three have the same angle in Figure 10c.
        - However, it treats the conjunction of two unlikely propositions as more true than either proposition on its own, as shown in Figure 10e.
    - In contrast, the *max-plus* semiring collapses the order of two unlikely propositions with their conjunction: All three have the same angle in graph Figure 10f.
        - However, it treats the conjunction of two likely propositions as more false than either proposition on its own, as shown in Figure 10d.

- This results in two possibilities for explaining the distribution of conjunction errors for the Likely/Likely and Unlikely/Unlikely conjunctions in Figure 7.
    - In both cases, about 40% of the responses in Figure 7 make zero conjunction errors, just over 30% make one conjunction error, and just under 30% make two conjunction errors.
    - In effect, this is a near random distribution, but with a bit more of the probability mass associated with zero conjunction errors.

1. The first possible explanation is that experimental partipants have a preference for the tropical semiring that makes the greatest distinction: This would mean preferring the *max-plus* semiring in the case of a conjunction of two likely propositions and preferring the *min-plus* semiring in the case of a conjunction of two unlikely proposition.
   - This would still correctly predict the (near) tie between Likely/Likely and Unlikely/Unlikely at zero conjunction errors in Figure 7.
   - And it also allows for the almost uniform distribution between Likely/Likely and Unlikely/Unlikely in the other cases (one and two conjunction errors in Figure 7).
2. The second possible explanation is that experimental participants try both semirings, but in the face of the inconsistent results that this delivers (i.e., the *min-plus* and *max-plus* semirings give inconsistent answers for the Likely/Likely case and also for the Unlikely/Unlikely case), at least some participants back off to proper probabilistic reasoning and use the probability semiring, such that more responses occur in the zero conjunction error part of the response mass for both Likely/Likely and Unlikely/Unlikely.

   - Both of these explanations would thus predict the distribution of data, but using different mechanisms. Therefore there are distinct empirical predictions at play here which could somehow be tested experimentally.

- The tropical model also behaves well logically, as illustrated in Figure 11.
  - The key observation is that — despite the fact that conjunction does not behave absolutely logically (likewise disjunction), which is after all necessary to model the Linda effect — conjunction and disjunction still form duals.
  - We use the likely proposition that *Linda is a feminist* and the unlikely proposition that *Linda is a bank teller* to demonstrate this.
  - The dashed line in Figure 11 once again represents total uncertainty as to the truth/falsity of a proposition.
  - The *feminist* proposition is assigned a vector in the true space, so below the dashed line, whereas the *bank teller* proposition is assigned the symmetrical vector in the false space, so above the dashed line.
  - The *min-plus* semiring treats the conjunction of *bank teller* and *feminist* as more true than false; this vector is labelled "bank $\wedge_{min}$ fem" in Figure 11.
    - We also observe that the *min-plus* semiring treats the disjunction of *bank teller* and *feminist* as more false than true and that, crucially, it assigns the disjunction the symmetrical vector in the false space; this vector is labelled "bank $\vee_{min}$ fem" in Figure 11.
  - The *max-plus* semiring behaves as the dual of the *min-plus* semiring in the positive quadrant and therefore, in contrast, treats the conjunction of *bank teller* and *feminist* as more false than true (see the vector labelled "bank $\wedge_{max}$ fem" in Figure 11) and the disjunction of *bank teller* and *feminist* as more true than false (see the vector labelled "bank $\vee_{max}$ fem" in Figure 11).
    - Moreover, the two vectors are likewise symmetrical.

- In sum, the tropical model seems to give a structured and interesting explanation of the Linda effect, without falling back on general Gricean pragmatics, as was the case for the one semiring.

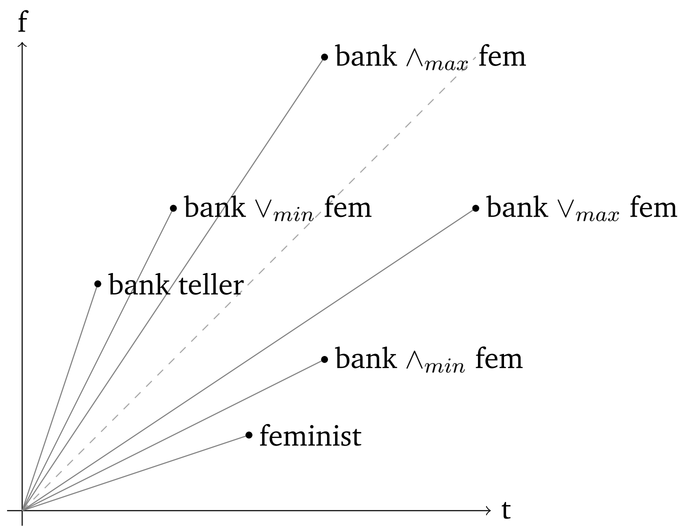


Figure 11: Conjunction and disjunction symmetries, given the *min-plus* or *max-plus* tropical semirings.

- We should stress that we do not believe that the tropical model should be in the business of predicting actual observed *frequencies* of responses, but rather should simply predict and explain the *pattern* of responses.
- In other words, the tropical model is a grammatical model that constitutes a *competence theory* of the Linda effect, which we believe is the correct aim in the cognitive sciences (Pylyshyn 1973).

8 Conclusion

- Tversky and Kahneman's work on conjunction fallacies has been very influential in cognitive science.
  - They explain the Linda effect as stemming from a *representativeness* heuristic.
- But subsequent work has shown the landscape to be more complex.
  - People do not tend to make conjunction fallacies if stakes are introduced or if they are prompted to reason probabilistically, e.g. to think about frequencies (Yates and Carlson 1986).
  - This suggests that there may be factors involved in the Linda scenario that lead people to relax their probabilistic competence in the absence of a specific need to use it.
- We think this kind of trade-off receives a better explanation in terms of the *satisficing* heuristic (Simon 1956) than in terms of representativeness: People use a simpler and cognitively less taxing method when there is no particular cost to making errors or when accuracy is not emphasized, but use the cognitively more taxing method when there are such costs or accuracy is emphasized.
- The psychological models that have previously been proposed for conjunction fallacies are nevertheless unified in their non-compositionality: Conjunction (*and*) is not treated as a standard compositional logical operator.

- This is, in our view, a serious defect which we have sought to remedy by offering a compositional treatment instead.
- In particular, we have developed a model in which the behaviour of natural language *and* is closely connected to the basic logical meaning of conjunction as a propositional connective, which upholds a key aspect of Grice's program (Grice 1989) and takes seriously his Modified Occam's Razor: 'Senses are not to be multiplied beyond necessity' (Grice 1978: 118).

References

Ash Asudeh and Gianluca Giorgolo. *Enriched Meanings: Natural Language Semantics with Category Theory*. Oxford University Press, Oxford, 2020.

Chris Barker. Continuations and the nature of quantification. *Natural Language Semantics*, 10(3): 211–242, 2002.

Chris Barker. Quantificational binding does not require c-command. *Linguistic Inquiry*, 43(4):614–633, 2012.

Chris Barker. Negative polarity as scope marking. *Linguistics and Philosophy*, 41(5):483–510, 2018.

Chris Barker. NL_λ as the logic of scope and movement. *Journal of Logic, Language and Information*, 28(2):217–237, 2019.

Chris Barker and Chung-chieh Shan. Types as graphs: Continuations in Type Logical Grammar. *Journal of Logic, Language and Information*, 15(4):331–370, 2006.

Chris Barker and Chung-chieh Shan. Donkey anaphora is in-scope binding. *Semantics and Pragmatics*, 1(1):1–46, June 2008.

Chris Barker and Chung-chieh Shan. *Continuations and Natural Language*. Oxford University Press, Oxford, 2014.

Chris Barker, Raffaella Bernardi, and Chung-chieh Shan. Principles of interdimensional meaning interaction. In Man Li and David Lutz, editors, *Proceedings of SALT 20*, pages 109–127, 2010.

Johan van Benthem. *Language in Action: Categories, Lambdas, and Dynamic Logic*. North-Holland, Amsterdam, 1991. Reprinted in 1995 by the MIT Press, Cambridge, MA.

Dylan Bumford. Incremental quantification and the dynamics of pair-list phenomena. *Semantics and Pragmatics*, 8(9):1–70, 2015.

Dylan Bumford. Split-scope definites: Relative superlatives and Haddock descriptions. *Linguistics and Philosophy*, 40(6):549–593, 2017a.

Dylan Bumford. *Split-Scope Effects in Definite Descriptions*. PhD thesis, New York University, 2017b.

Dylan Bumford and Chris Barker. Association with distributivity and the problem of multiple antecedents for singular *Different*. *Linguistics and Philosophy*, 36(5):355–369, 2013.

Simon Charlow. *On the Semantics of Exceptional Scope*. PhD thesis, New York University, 2014.

Simon Charlow. Post-suppositions and semantic theory. Ms., Rutgers University. To appear in *Journal of Semantics*, 2018.

Simon Charlow. The scope of alternatives: Indefiniteness and islands. *Linguistics and Philosophy*, pages 1–46, 2019a. First online: August 26, 2019. Early access version and pagination.

Simon Charlow. Where is the destructive update problem? *Semantics and Pragmatics*, 12(10):1–17, 2019b. Early access version and pagination.

Reuben Cohn-Gordon. Monad transformers for natural language: Combining monads to model effect interaction. Ms., Stanford University, nd.

Jan van Eijck and Christina Unger. *Computational Semantics with Functional Programming*. Cambridge University Press, Cambridge, 2010.

Gianluca Giorgolo and Ash Asudeh. One semiring to rule them all. In *Proceedings of the 36th Annual Cognitive Science Society Conference*, pages 116–121, 2014.

Gianluca Giorgolo and Christina Unger. Coreference without discourse referents: A non-representational DRT-like discourse semantics. In B. Plank, T. Kim Sang, and T. Van de Cruys, editors, *Computational Linguistics in the Netherlands 2009*, number 14 in LOT Occasional Series, pages 69–81. LOT, 2009.

Michele Giry. A Categorical Approach to Probability Theory. In *Categorical Aspects of Topology and Analysis, Proc. int. Conf., Ottawa 1981, Lect. Notes Math. 915, 68-85 (1982).*, 1982.

H. Paul Grice. Logic and conversation. In Peter Cole and Jerry L. Morgan, editors, *Syntax and Semantics: Speech Acts*, volume 3, pages 41–58. Academic Press, San Diego, CA, 1975.

H. Paul Grice. Further notes on logic and conversation. In Peter Cole, editor, *Syntax and Semantics: Pragmatics*, volume 9, pages 113–127. Academic Press, San Diego, CA, 1978.

H. Paul Grice. *Studies in the Way of Words*. Harvard University Press, Cambridge, MA, 1989.

James A. Hampton. Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(1):12–32, 1988.

Ralph Hertwig and Gerd Gigerenzer. The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, 12:275–305, 1999.

Ralph Hertwig, Björn Benz, and Stefan Krauss. The conjunction fallacy and the many meanings of *and*. *Cognition*, 108:740–753, 2008.

Ray Jackendoff. The ecology of english noun-noun compounds. In *Meaning and the Lexicon: The Parallel Architecture 1975–2010*, pages 413–451. Oxford University Press, Oxford, 2010.

Saul Kripke. Speaker's reference and semantic reference. *Midwest Studies in Philosophy*, 2:255–276, 1977.

Alexander Kuhnle. Modeling uncertain data using monads and an application to the sequence alignment problem. Master's thesis, Karlsruhe Institute of Technology, 2013.

Wener Kuich and Arto Salomaa. *Semirings, Automata Languages*. Springer-Verlag, Berlin, 1986.

Francis William Lawvere. The category of probabilistic mappings. Seminar handout, 1962.

Yong Lu. The conjunction and disjunction fallacies: Explanations of the Linda problem by the equate-to-differentiate model. *Integrative Psychological and Behavioral Science*, 50(3):507–531, 2016.

Jirka Maršik. *Effects and Handlers in Natural Language*. PhD thesis, Université de Lorraine, 2016.

Barbara Mellers, Ralph Hertwig, and Daniel Kahneman. Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, 12:269–275, 2001.

Eugenio Moggi. Computational lambda-calculus and monads. In *LICS*, pages 14–23. IEEE Computer Society Press, 1989.

Mehryar Mohri. Weighted automata algorithms. In Manfred Droste, Werner Kuich, and Heiko Vogler, editors, *Handbook of Weighted Automata*, pages 213–254. Springer-Verlag, Berlin, 2009.

Mehryar Mohri and Richard Sproat. 231–238. In *34th Meeting of the Association for Computational Linguistics: Proceedings of the Conference*. Association for Computational Linguistics, 1996.

Mehryar Mohri, Fernando Pereira, and Michael Riley. Speech recognition with weighted finite-state transducers. In Jacob Benesty, M. Mohan Sondhi, and Yiteng Huang, editors, *Springer Handbook of Speech Processing*, pages 559–584. Springer-Verlag, Berlin, 2008.

Richard Montague. The proper treatment of quantification in ordinary English. In Jaakko Hintikka, Julian Moravcsik, and Patrick Suppes, editors, *Approaches to Language*, pages 221–242. Reidel, Dordrecht, 1973. Reprinted in Montague (1974: 247–270).

Richard Montague. *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press, New Haven, CT, 1974. Edited and with an introduction by Richmond H. Thomason.

Barbara H. Partee. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof, editors, *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, volume 8 of GRASS, pages 115–143. Foris, Dordrecht, 1986. Reprinted in Partee (2004: 203–230).

Barbara H. Partee. *Compositionality in Formal Semantics: Selected Papers by Barbara H. Partee*. Blackwell, Oxford, 2004.

Barbara H. Partee and Mats Rooth. Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, editors, *Meaning, Use, and Interpretation of Language*. Walter de Gruyter, Berlin, 1983.

Jean-Eric Pin. Tropical semirings. In Jeremy Gunawardena, editor, *Idempotency*, pages 50–69. Cambridge University Press, Cambridge, 1998.

Zenon W. Pylyshyn. The role of competence theories in cognitive psychology. *Journal of Psycholinguistic Research*, 2(1):21–50, 1973.

Chung-chieh Shan. Monads for natural language semantics. In Kristina Striegnitz, editor, *Proceedings of the Sixth ESSLLI Student Session*, pages 275–288, 2001.

Chung-chieh Shan. The character of quotation. *Linguistics and Philosophy*, 33(5):417–433, 2010.

Herbert A. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63 (2):129–138, 1956.

Imre Simon. Recognizable sets with multiplicities in the tropical semiring. In Michal P. Chytil, Václav Koubek, and Ladislav Janiga, editors, *Mathematical Foundations of Computer Science 1988*, pages 107–120, Berlin, Heidelberg, 1988. Springer Berlin Heidelberg.

Dan Sperber and Deirdre Wilson. *Relevance: Communication and Cognition*. Blackwell, Oxford, 1 edition, 1986.

Dan Sperber and Deirdre Wilson. *Relevance: Communication and Cognition*. Blackwell, Oxford, 2 edition, 1995.

Katya Tentori, Nicolao Bonini, and Daniel Osherson. The Conjunction Fallacy: A Misunderstanding about Conjunction? *Cognitive Science: A Multidisciplinary Journal*, 28(3):467–477, 2004.

Amos Tversky and Daniel Kahneman. Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review*, 90(4):293–315, 1983.

Christina Unger. Dynamic semantics as monadic computation. In Manabu Okumura, Daisuke Bekki, and Ken Satoh, editors, *New Frontiers in Artificial Intelligence — JSAI-isAI 2011*, pages 68–81, 2011.

Philip Wadler. Comprehending monads. *Mathematical Structures in Computer Science*, 2:61–78, 1992.

Philip Wadler. Monads and composable continuations. *Lisp and Symbolic Computation*, 7(1):39–56, 1994.

Philip Wadler. Monads for functional programming. In Johan Jeuring and Erik Meijer, editors, *Advanced Functional Programming*, pages 24–52. Springer, Berlin, 1995.

Michael White, Jordan Needle, Simon Charlow, and Dylan Bumford. Parsing with Dynamic Continuized CCG. In *Proceedings of the 13th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+ 13)*, pages 71–83, 2017.

J. Frank Yates and Bruce W. Carlson. Conjunction errors: Evidence for multiple judgment procedures, including "signed summation". *Organizational Behavior and Human Decision Processes*, 37: 230–253, 1986.