1. Introduction

Evidential expressions in natural language always imply a certain relation between prejacent and evidence. What sort of relation holds between them? What qualifies as the evidence that licenses the use of the evidential? How and by what formal apparatus can we characterize the relation in a necessary and sufficient way? To answer these questions has been one of the central issues in the study of evidentiality.

Inferential evidentials, the focus of this study, are analyzed as encoding a certain inference between prejacent and perceived evidence. The Japanese evidential youda is possibly one of the most intensively studied inferential evidentials, and several attempts have so far been made to characterize the inferential link it encodes. However, a full satisfactory account has not yet been accomplished. Different authors have highlighted different aspects of the inference patterns of youda, but, as will be shown, all of them end up with a partial empirical coverage, each facing counterexamples that go beyond their characterizations.

This study attempts a unified semantics that capture all the possible inference patterns of youda, using the recent development of probabilistic causal models called causal Bayesian networks \cite{Pearl2000,SpirtesGlymourScheines2000}. In Section 2, we will organize the empirical landscape of youda with several novel observations. In Section 3, we will introduce the definitions of the model and update system upon which our semantics will be built. In section 4, we will show our truth condition of youda and give demonstrations of how it will work out. In Section 5, we conclude.

2. Data

\cite{McCreadyOgata2007} (henceforth M&O) first highlighted that the essential meaning of inferential evidentials is a probability update: evidence and prejacent are in a relation in which the former makes the speaker more confident about the latter, with the help from the speaker’s world knowledge. Their insight is reflected in the contrast between (1) and (2). In (1) seeing the wet street is able to increase the speaker’s credence toward it having rained.
In contrast, the probability update is unsuccessful in [2] on the assumption that there is no inferential link between rain and sprinkler: the fact that the sprinkler was on does not probabilistically affect whether it indeed rained or not.

We take M&O’s insight to be empirically correct: successful use of inferential evidential always comes with a probability increase. However, increase of the prejacent’s probability is necessary but not sufficient for the felicity of inferential evidentials, and M&O thus do not account for the restriction to certain inferences. Indeed, many have so far argued that the inference involved in inferential evidentials is one from effect to cause (Willet 1988, Davis and Hara 2014 a.o.). In the case of Japanese youda, Davis and Hara (2014) argue against M&O by showing the oddness of [3] which instantiates the ‘cause-to-effect’ inference. They argue that M&O would wrongly predict that [3] is fine because the conditional probability of the street getting wet given rain is estimated to be very high, and passing the probability threshold after update is M&O’s sole requirement.

While we agree that ‘effect-to-cause’ underlies most uses of inferential evidential, we argue that this characterization is empirically too restrictive. We present two cases in which the relation between evidence and prejacent cannot be reduced to a simple ‘effect-cause’ one. The first is the use in which multiple observations are necessary to infer the prejacent. Assume that in [4] and [5] rain and sprinkler are supposed to be the only reasonable causes of the wet street and their occurrence is equally likely, and that you are initially uncertain both whether it rained or not and whether the sprinkler was on or not. Assume also that rain and sprinkler are causally independent of each other.

Your first observation ‘the wet street’ creates three possibilities: (i) rain caused it alone, (ii) sprinkler caused it alone, or (iii) rain and sprinkler caused it together. Even though there is a ‘effect-cause’ relation between wetness and rain, your observation of the former is not sufficient to infer the latter: the possibility (ii) still cannot be neglected based on the mere observation of the wet street, because of the equal likelihood of rain and sprinkler. So you cannot conclude only with this evidence that it actually rained. Once you obtain the second
observation that the sprinkler was not used, you can eliminate the possibility (ii) and (iii) and safely infer that it did rain, and thus saying (5) is judged to be felicitous. Here what leads you to this conclusion is the combination of the wet street and the turned-off sprinkler. It is unclear whether such set of evidence can be said to establish an effect-cause relation to the prejacent.

The second case is the use in which no causal domination can be established between evidence and prejacent. Imagine a region where rain occurs much more frequently than thunder, and rain without thunder often occurs while thunder without rain is very rare. As shown in (6) the observation of thunder licenses the use of youda. The inference is reasonable because observing thunder makes the cooccurrence of rain highly probable, given the extreme rarity of rain-less thunder. Note crucially that thunder cannot be the cause of rain. Due to the non-past on the verb and the non-stative aspect of the sentence, the raining event is expected to happen in the future: it is physically or metaphysically impossible for a future event to cause a past or present event. The evidential claim in (6) should rather be a prediction of a future event based on the perceived evidence and knowledge about probabilistic correlations. Note also that if the evidence and the prejacent of (6) are switched as in (7), the use of youda becomes infelicitous: given its high frequency, the observation of rain does not help one confidently infer the high expectation of thunder.

(6) (Hearing the sound of thunder) (7) (Looking at falling raindrops)

<table>
<thead>
<tr>
<th>Japanese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>ame-ga fur-u youda.</td>
<td>‘It seems that it will rain soon.’</td>
</tr>
<tr>
<td>rain-NOM fall-PRES YOUDA</td>
<td></td>
</tr>
</tbody>
</table>

The asymmetries (4)-(5) and (6)-(7) thus constitute counter-cases against the ‘effect-to-cause’ characterization of inferential evidentials. The contrast between (4)-(5) underlines the significance of high credence toward prejacent proposition; an evidential inference fails even if the causal relation between prejacent and evidence is salient. Indeed, this has been already pointed out in the preceding literature, leading some to advocate the ‘abduction’ or ‘best-explanation’ approach (e.g. Takubo 2009, Krawczyk 2012), in which an inferential evidential is licensed only if the updated probability of the prejacent given the evidence is the highest among possible causal explanations. However, such approach would fail to account for the asymmetry between (6)-(7) in which the prejacent and the evidence are anything but an explanation to each other. The current dataset calls for a treatment which minimally rules out the cases like (2)-(3)-(4) and (7) while allowing all the rest. In the following sections, we will present our analysis of Japanese youda that fully captures this empirical coverage.

3. **Evidential update in a probabilistic causal model**

This section provides a preliminary formal set-up upon which our semantics of Japanese inferential evidential youda is to be built. While we share M&O’s fundamental intuition that probability updates lie at the core of inferential evidentials, what is desired is an explicit, schematic representation of the world knowledge against which the constraints on
probability update and causal relation are formally definable. For this purpose, we base our semantics on the recent development of probabilistic causal models, generally known as causal Bayesian networks (Pearl 2000, Spirtes, Glymour, and Scheines 2000, a.o.). Following other linguistic applications in the literature (e.g. Lassiter 2017b), we define a model by using concepts from intensional and degree semantics.

(8) A probabilistic causal model \( c \) is a triple \( \langle W, \mathcal{C}, \mu \rangle \), where \( W \) is a set of possible worlds, \( \mathcal{C} = \langle U, \Leftrightarrow \rangle \) is a causal structure where \( U \) is a set of finite partitions on \( W \) and \( \Leftrightarrow \) a directed acyclic graph on \( U \). \( \mu \) is a probability measure which maps each proposition to a value in [0,1].

We remark the correspondences between the current model and the statistical terminologies. The set of worlds \( W \) corresponds to the sample space in statistics. Each possible world is considered a sample or outcome of a certain event. Propositions (sets of worlds) are the analog of events (sets of outcomes) in statistics. The partitions in \( U \) correspond to variables (sets of events) in the statistical sense. In intensional semantics, they are identified to question denotations (sets of propositions). Notationally, we use \( 'x, y, ...' \) as notations of propositions, \( 'X, Y, ...' \) as variables, \( 'X, Y, ...' \) as subsets of \( U \).

For instance, when \( \mathcal{C} \) represents \( U \Rightarrow X \Leftarrow Y \Leftarrow Z \), \( pa(X)_{\mathcal{C}} = \{U, Y\} \) and \( de(X)_{\mathcal{C}} = \{X\} \), while \( pa(Z)_{\mathcal{C}} = \emptyset \) and \( de(Z)_{\mathcal{C}} = \{X, Y, Z\} \).

The probability measure \( \mu \) is countably additive, namely \( \mu(W) = 1 \) and for countable sequences \( x_1, x_2, ... \) of pairwise disjoint events \( x_i, \mu(\bigcup_i x_i) = \sum_i \mu(x_i) \). The conditional probability \( \mu(x|y) \) is defined as \( \mu(x \cap y) / \mu(y) \) when \( \mu(y) \neq 0 \), undefined otherwise. Importantly, when \( \mu \) is associated with a causal network \( \mathcal{C} \), it is assumed to obey the (causal) Markov condition: each variable \( X \) in \( \mathcal{C} \) is probabilistically independent of \( U \setminus (de(X)_{\mathcal{C}} \cup pa(X)_{\mathcal{C}}) \), given \( pa(X)_{\mathcal{C}} \). Satisfying this condition ensures that \( \mathcal{C} \) and \( \mu \) combine to reflect a real causal situation.

We also provide a definition of evidential update in the current model. As is standardly assumed in the literature, the use of evidentials always presupposes or implicates the existence of some evidence. Such evidence can be sometimes multiple, as was illustrated by the example (5). We therefore assume that an evidential relation is established between a prejacent and a set of evidence, rather than a single piece of evidence. An evidence set \( \mathcal{E} \) is defined as a set of propositions that have been observed by the speaker. Every time a new
Toward a unified account of Japanese evidential youda

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observation is made, it is added to the speaker’s evidence set. While we pos

observer
(11) Probability-Increase constraint:
Given a prejacent proposition \( p \) and evidence set \( \varepsilon \), \( \mu_\varepsilon(p) > \mu_\emptyset(p) \).

In the current model, the Markov condition is assumed to provide the set of probabilistic (in)dependencies between variables. However, as no one would argue against, evidential inference is an epistemic activity: agents are assumed to be non-omniscient beings who know only a limited number of truths about the world. Crucially, knowing or not knowing the values of certain variables are known to trigger an observational correlation between causally independent variables. As long as the constraint (11) inspects change of probability in one’s epistemic state, a criterion is needed that captures which observation could turn on a spurious probabilistic correlation in the given causal network. To this end, we make use of the \( d \)-separation criterion proposed by Pearl (1988), which exhausts the conditional (in)dependencies in the causal network given by the directed acyclic graph. The \( d \)-separated variables are assumed to be probabilistically independent conditional on the set of variables \( Z \) (see also Verma and Pearl 1988).

(12) A path \( p \) in \( C \) is said to be \( d \)-separated by a set of variables \( Z \) iff

a. \( p \) contains a chain \( 'I \Rightarrow M \Rightarrow J' \) or a fork \( 'I \Leftarrow M \Rightarrow J' \) such that \( M \in Z \), or

b. \( p \) contains a collider \( 'I \Rightarrow M \Leftarrow J' \) such that \( de(M) \cap Z = \emptyset \)

Here a path \( p \) is a sequence of consecutive edges in the causal network. In causal chains \( 'I \Rightarrow M \Rightarrow J' \) or causal forks \( 'I \Leftarrow M \Rightarrow J' \), \( I \) and \( J \) are dependent but become independent once we condition on (i.e. know the value of) the middle variable \( M \): once we observe the value of \( M \), learning about \( I \) has no effect on the probability of \( J \) and vice versa. In our case, the fork structure instantiates the relation between rain and thunder in (6) and (7) under the assumption of common cause principle (Reichenbach 1956), according to which two correlated variables always share a common cause affecting both. So as long as we are ignorant of the factor causing both thunder and rain, they remain dependent on each other. Once we detect and condition on it, they become \( d \)-separated and therefore independent of each other. In causal colliders \( 'I \Rightarrow M \Leftarrow J' \), conditioning works in the opposite way: \( I \) and \( J \) are independent but become dependent once we condition on the middle variable \( M \) or any of its descendants. For instance in (5), once the evidence ‘the wet ground’ is obtained, whether it rained or not and whether the sprinkler was on or off are no longer \( d \)-separated, and the value of sprinkler (i.e. ‘off’) starts affecting the probability of rain. Importantly, chain, fork and collider (i.e. three-node networks) are building blocks of every causal network. The \( d \)-separation criterion works not only for the simplest cases shown here but also for more complicated networks which involve more than three nodes.

The second component of our semantics is the constraint on causal relation which minimally filters out cause-to-effect inferences like [3]. This constraint can be easily defined in the current model by referring to the descendant variables in the causal network. The Non-Descendant constraint [13] simply requires that the prejacent variable is not among the causal descendants of any of the evidence variables.
(13) Non-Descendant constraint:
Given a prejacent $p$ and evidence set $\varepsilon$, for every $e \in \varepsilon$, $P \notin de(E \varepsilon)$ (where $p \in P$ and $e \in E$, and $P$ and $E$ are causal variables in $\varepsilon$).

The following toy diagram illustrates the situation in which one only has rain (bold-faced) in her evidence set. By the Non-Descendant constraint, the dashed variables are excluded from the possible prejacents of the evidential proposition.

(14) A toy causal diagram and Non-Descendant constraint given ‘$\varepsilon = \{\text{rain}\}$’

A digression: M&O and many other works on inferential evidentials and epistemic modals (Matthewson, Davis, and Rullmann 2007, von Fintel and Gillies 2010, a.o.) posit the indirect evidence requirement as one of the crucial presuppositions. It is intended to exclude the case in which the speaker uses ‘it rained’ as prejacent while she already knows its truth. In our semantics, such ban on the use of direct evidence appears to be covered by the Non-Descendant constraint: insofar as the descendant relation is reflexive as defined in (9), the variable of rain in the current case is filtered out from the possible domain of prejacent by the Non-Descendant constraint together with its ‘proper’ descendants, as shown by the dashed line on rain in (14). This implies that directly observed propositions (namely propositions added to evidence set) cannot be used as a potential prejacent in the current model, the same result obtained by the indirect evidence requirement. We leave open any possibility that such ban on direct evidence proposed by other works is required from an independent conceptual ground. Here we assume that our Non-Descendant constraint implies the equivalent effect.

The final piece of our semantics is the threshold value. While M&O argues that the threshold for youda (and other evidentials) is 0.5, the actual force of youda is even stronger. Consider the example in (15). For youda to be felicitously uttered, the speaker needs to engage in reiterated trials until the resultant distribution convinces her that $p$ is highly probable. We thus assume that the threshold for youda is as high as that for a weak necessity modal such as ‘must’ (see Lassiter 2016).

(15) Coin A is a fair coin, and Coin B is biased with 70% chance of head and 30% chance of tail. One of the two coins was randomly picked and flipped $n$ times. ($\mu(A) = \mu(B) = .5$, $\mu(\text{Head}|A) = .5$, $\mu(\text{Head}|B) = .7$)

<table>
<thead>
<tr>
<th>CoinB-wo tukatta youda.</th>
<th>tosses</th>
<th>heads</th>
<th>Prob. for CoinB</th>
<th>Felicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘It seems they used Coin B.’</td>
<td>(i) 1</td>
<td>1</td>
<td>.5833...</td>
<td>#</td>
</tr>
<tr>
<td></td>
<td>(ii) 10</td>
<td>7</td>
<td>.6948...</td>
<td>??</td>
</tr>
<tr>
<td></td>
<td>(iii) 100</td>
<td>69</td>
<td>.9993...</td>
<td>OK</td>
</tr>
</tbody>
</table>

Toward a unified account of Japanese evidential youda
Mizuno & Yang

The truth condition of youda is the combination of (11), (13) and the threshold value for the updated probability of \( p \) given \( \varepsilon \). We treat the two constraints as presuppositions which need to be satisfied for the evidential sentence to be defined. Following Lassiter (2011, 2016, 2017a), we set a context-dependent threshold \( \theta \) for the updated probability to pass.

(16) \( [p\text{-youda}]^{c, \varepsilon} \) is defined in \( c \) given the non-empty evidence set \( \varepsilon \) only if \( p \) and \( \varepsilon \) satisfy both (11) and (13) in \( c \). When defined, \( [p\text{-youda}]^{c, \varepsilon} \) is true in \( c \) given \( \varepsilon \) iff \( \mu_{\varepsilon}(p) \geq \theta \) (\( \theta \) is a high probability threshold).

We illustrate how our truth condition explains the asymmetries between (4) and (5) and (6) and (7). For the former case, assume the following model: The causal structure represents \( \text{Sprinkler} \Rightarrow \text{Wet?} \Leftrightarrow \text{Rain?} \), the threshold \( \theta = .97 \), and the prior distribution is \( \mu(s) = \mu(r) = .2 \), \( \mu(w|s, r) = .999 \), \( \mu(w|s, \bar{r}) = .99 \), \( \mu(w|\bar{s}, r) = .99 \), \( \mu(w|\bar{s}, \bar{r}) = .005 \).

(17) replicates the probability update in (4), in which the speaker makes the unsuccessful inference to ‘it rained’ with the evidence ‘the street is wet’. Although this inference satisfies both the Probability-Increase constraint and the Non-Descendant constraint, the upshot probability ‘.5510...’ is far below the threshold.

(17) \[
\mu_{\{w\}}(r) = \mu(r|w) = \frac{\mu(w \cap r)}{\mu(w)} = \frac{\mu(s)\mu(r)\mu(w|s, r) + \mu(\bar{s})\mu(r)\mu(w|\bar{s}, r)}{\mu(w)} + \mu(s)\mu(\bar{r})\mu(w|s, \bar{r}) + \mu(\bar{s})\mu(r)\mu(w|\bar{s}, r) + \mu(\bar{s})\mu(\bar{r})\mu(w|\bar{s}, \bar{r}) = .5510... < \theta
\]

(18) replicates the probability update in (5). The evidence set reinforced by the additional evidence ‘the sprinkler was off’ raises the probability of ‘it rained’ above the threshold. The truth condition in (16) thus vindicates the judgment asymmetry between (4) and (5).

(18) \[
\mu_{\{w, \bar{s}\}}(r) = \mu(r|w, \bar{s}) = \frac{\mu(r \cap w \cap \bar{s})}{\mu(w \cap \bar{s})} = \frac{\mu(s)\mu(\bar{r})\mu(w|s, r) + \mu(\bar{s})\mu(\bar{r})\mu(w|\bar{s}, r)}{\mu(w)} = .9801... > \theta
\]

Our semantics also predicts the felicity of (19), in which the prejacent is a disjunction of two propositions. Since rain and sprinkler are the only reasonable explanations in the current context, it is safe to infer that at least one of them caused the wet street. As shown by the calculation in (20), the upshot probability surpasses the threshold (note that rain and sprinkler here are not assumed to be mutually disjoint).

(19) (Looking at the wet street)
\[
\text{rain-NOM fell or otherwise sprinkler-NOM was used YOUDA}
\]
‘It seems that it rained or the sprinkler was used.’
Turning to (6) and (7), assume that the causal structure represents \([\text{Rain}] \leftarrow X \Rightarrow \text{Thunder}\), the threshold \(\theta = .97\), and the prior distribution is \(\mu(x) = .099\), \(\mu(r|x) = \mu(t|x) = .99\), \(\mu(r|\bar{x}) = .224\), \(\mu(t|\bar{x}) = .002\). Here a certain climate variable \(X\) is the common cause of rain and thunder, and its value is assumed to be still unknown. Because of this, learning either one of them can affect the probability of the other by the definition of \(d\)-separation. Since rain and thunder do not cause each other, the inferences in both directions escape the Non-Descendant constraint constraint. The only difference is their updated probabilities. As shown by the calculations, there is a huge gap between (21) and (22), and only the former surpasses the threshold probability. The truth condition in (16) successfully rules out the inference from rain to thunder in the current situation.

\[
\begin{align*}
\mu_{\{w\}}(r \cup s) &= \mu(r \cup s|w) \\
&= \mu(r|w) + \mu(s|w) - \mu(r \cap s|w) \\
&= \mu(r \cap w)/\mu(w) + \mu(s \cap w)/\mu(w) - \mu(r \cap s \cap w)/\mu(w) \\
&= .5510... + .5510... - .1110... = .9911... > \theta
\end{align*}
\]

(20)

5. Conclusion

While we focused on defining the possible inference patterns of \(\text{youda}\), we did not deal with the issue of speaker’s commitment highlighted by Davis and Hara (2014), who pointed out that \(p\text{-youda}\) can be continued by the negation of \(p\). Although we believe that the characterization of inference and the presence of commitment are independent issues, the proposed semantics does not account for when and how the speaker lacks a commitment to the prejacent. Working out a full-fledged semantics that also predicts the presence/absence of commitment is our future task.

References


Mizuno & Yang


Teruyuki Mizuno, Muyi Yang
teruyuki.mizuno@uconn.edu, muyi.yang@uconn.edu