

## Towards Measuring Uncertainty of Discourse Interpretation and Update

Aleksandre Maskharashvili, Department of Linguistics, University of Illinois Urbana-Champaign

Efforts to extend the Montagovian approach to the form-meaning interface from the *sentence-level* to the *discourse-level* proved to be challenging. It was unclear how to interpret a discourse in which sentences and clauses feature an ‘unbound’ pronoun—a pronoun whose antecedent typically lies outside of the sentence or clause in which it appears, because since such pronouns are not bound, they are represented as free variables and therefore, one needs to assign values to them (to free variables) to define a truth-conditional interpretation. Kamp’s seminal work, Discourse Representation Theory (DRT), offers a logical framework for discourse, which successfully addresses these challenges. In DRT, “each new sentence of a discourse is interpreted in the context provided by the sentences preceding it...” (Eijck and Kamp, 1997). DRT uses a pair of (partial) functions  $\langle f_1, f_2 \rangle$ , where  $f_2$  is an extension of  $f_1$  to allow one to interpret the discourse. For the sake of current discussion, we assume that  $f_1$  and  $f_2$  map discourse referents to individuals (entities); and  $f_2$  being an extension of  $f_1$  assigns values to the discourse referents (e.g. unbound pronouns) that the incoming piece of discourse brings with, called a discourse update. Since  $f_2$  extends  $f_1$ , if  $f_2$  were to be contextually restricted to certain values (to derive certain interpretations), so would be  $f_1$  if this restriction concerns the discourse referents from the domain of  $f_1$ . An interpretation can be constructed and then revised if needed: a new piece of discourse may help to disambiguate an interpretation of a previous piece of discourse; in some cases, however, a new piece of discourse might as well inject uncertainty, prompting to revisit our commitments to certain earlier decisions. Consider the following example: (1) “*Al entered the bar and saw Bob sitting at a table. He ran away.*” In (1), does *He* refer to *Al* or *Bob*? We cannot answer that. However, in the following discourse (2) “*Al saw Bob. He ran away. But Al chased him down and caught him.*” the ‘same’ *He* clearly refers to *Bob*, not *Al*. In this case, the discourse update presents a new piece of information allowing us to choose between two possible interpretations of the earlier piece, about which we would have been unsure otherwise. Although a discourse may involve various kinds of uncertainties, we focus on unbound pronouns and their resolution to address the following questions:

- (a) How to quantify the uncertainty involved in the interpretation of the current piece of discourse when it comes to unbound pronouns.
- (b) How do we measure the information gain when it comes to the uncertainty of a discourse update?

We introduce a notion of a probabilistic update in discourse: (unbound) pronouns are not seen just as free variables, but rather as probability distributions over domain entities in the current discourse.<sup>1</sup> For notational purposes, we enumerate pronouns according to the number of the clause (elementary discourse unit) they appear in and by their respective position within that clause. (For example, for the unbound pronouns in the discourse “[*Al saw Bob.*]<sub>1</sub> [*He ran away.*]<sub>2</sub> [*But Al chased him down and he caught him.*]<sub>3</sub>”, we use  $he_{21}$ ,  $him_{31}$ ,  $he_{31}$ , and  $he_{32}$ .) Let us imagine that the current discourse is [*Al saw Bob.*]<sub>1</sub> [*he<sub>21</sub> ran away.*]<sub>2</sub>, then we have two entities in the domain, Al and Bob. The pronoun  $he_{21}$  may resolve to either Al and Bob and since we don’t have a preference between them (at this point) for resolving  $he_{21}$ , we represent this by  $he_{21}:\{Al, Bob\} \rightarrow [0,1]$  such that  $he_{21}(Al) = he_{21}(Bob) = 0.5$ , which is to say that Al and Bob are equally likely to be the antecedent of  $he_{21}$ . Now let us assume that the discourse is updated by “[*But Al chased him down and he caught him.*]<sub>3</sub>” The new piece of discourse induces an update for the probabilistic distribution  $he_{21}:\{Al, Bob\} \rightarrow [0,1]$ . The change in  $he_{21}$  is noticeable since we are now able to resolve the pronoun  $he_{21}$ , as after the update we have:  $he_{21}(Al) = 1$

<sup>1</sup>This is very similar to the approach by Kehler et al. (*Coherence and coreference revisited*, 2008. Journal of Semantics 25 (1):1-44.), who introduced probabilistic, Bayesian models to study pronoun resolution.

and  $he_{21}(\text{Bob}) = 0$ . This is a Bayesian update: Based on a new piece of evidence, we update the prior distribution. In this example (but not in general), in the updated discourse, there's less uncertainty than before. How do we know that? How do we quantify uncertainty? *Entropy* can help to measure that:<sup>2</sup>  $\text{Entropy}(he_{21}(\text{Al})=he_{21}(\text{Bob})=0.5) > \text{Entropy}(he_{21}(\text{Al})=1; he_{21}(\text{Bob})=0)$ .

Yet another measure based on entropy we may employ to quantify ‘the change’ brought by a discourse update is the Kullback–Leibler (KL) divergence between two probability distributions:  $KL(he_{21}, he_{21}^{up})$  where  $he_{21}$  denotes  $he_{21}(\text{Al})=0.5$  and  $he_{21}(\text{Bob})=0.5$ , whereas  $he_{21}^{up}$  denotes  $he_{21}^{up}(\text{Al})=1$  and  $he_{21}^{up}(\text{Bob})=0$ . While entropy measures uncertainty in the current discourse, the KL divergence measures the scale or extent of the update. To put it another way, we can measure the current state of uncertainty involved within the interpretation of the discourse using entropy, whereas we can use the KL divergence to measure how much information change is associated with the discourse update. This motivates for having both measures:  $\text{entrodifff21} = \text{Entropy}(he_{21}) - \text{Entropy}(he_{21}^{up})$  and  $KL(he_{21}, he_{21}^{up})$ , where  $\text{entrodifff21}$  indicates reduction or increase in uncertainty of the interpretation of  $he_{21}$ , whereas  $KL(he_{21}, he_{21}^{up})$  measures the extent of the discourse update. (The above example only concerns one pronoun, though we can generalize it to any finite number of unbound pronouns,  $p_1, \dots, p_n$ , as follows:  $\sum_{k=1}^n \text{entrodifff}_k$ , where  $\text{entrodifff}_k$  is the entropy difference of  $p_k$ , for  $k = 1, \dots, n$ . Similarly, we may combine KL divergences. A careful analysis is needed to not overestimate the uncertainty: generally, the resolution of two pronouns,  $p_1$  and  $p_2$ , is interdependent, and thus, probabilistically they are not independent. Instead of merely summing their entropies, considering other methods of combining them could be more preferable.)

Psycholinguistic experiments to establish whether there is a link between higher/lower entropy of pronoun interpretation and discourse processing times would be interesting to design, as well as, to examine whether the higher/lower information gain through a discourse update correlates with discourse processing times. Here's a sketch of a possible experiment: (Case i) starting with a discourse with a ‘high’ entropy, i.e., high-level of uncertainty, updating it with a discourse update which comes with a ‘large’ KL distance, but the resultant discourse has a ‘low’ entropy, would that be considered easier for speakers to process vs (Case ii) starting with a discourse with a ‘high’ entropy, updating it with a discourse update which has a ‘moderate’ KL distance, and the resultant discourse has a ‘moderate’ entropy?

Discourse structure theories assume that each new piece of discourse connects to the current discourse through a *rhetorical* relation. While rhetorical relations vary in many aspects, they can be divided into two categories: subordinating and coordinating. This dichotomy allows to define *Right Frontier Constraint* (RFC) to restrict the locations in the discourse structure where the new piece can be joined. This would also restrict the magnitude of the discourse update, preventing it from affecting a pronoun interpretation that isn't accessible to it.<sup>3</sup> It would be interesting to characterize rhetorical relations in terms of their capacity to bring with a discourse update that would help to resolve uncertainties related to unbound pronouns (lowering entropy) or, conversely, introduce uncertainties leading to an increased entropy of pronoun resolution.

---

<sup>2</sup>Entropy is highest for a uniform distribution (all values are equally probable), whereas entropy is lowest for deterministic distribution (a single value has a probability 1 and thus, other values have a probability 0).

<sup>3</sup>For specific discourse relations, their roles in generation and interpretation of pronouns are studied in (Kehler et al. *Coherence and coreference revisited*, 2008. *Journal of Semantics* 25 (1):1-44), (Kehler & Rohde, *Prominence and coherence in a Bayesian theory of pronoun interpretation*. *Journal of Pragmatics* 154. 63–7, 2019), and (Liao et al. *Comparing models of pronoun production and interpretation via observational and experimental evidence*. *Glossa: A Journal of General Linguistics*. 2024; 9(1)).