9.19: Computational Psycholinguistics, Pset 6
due Thursday 30 November 2023

16 November 2023

Quantifier monotonicity

Recall that one can define a notion of monotonicity for sentential operators and partially applied quantifiers. We saw in class *via* an example that the quantifier *all* was downward-monotone w.r.t its restrictor, and upward-monotone w.r.t. its nuclear scope. Here is another example, showing that *and* (a logical operator) is upward monotone w.r.t. its second argument (=second conjunct):

- We consider a sentence of the form *John is in Berlin and P*, which corresponds to the partial application of *and* to its first argument, *John is in Berlin*.
- We consider P_1: *Mary is in Paris* and P_2: *Mary is in France*. We have $P_1 \implies P_2$.
- Subsitituting P with P_1 and P_2 in *John is in Berlin and P*, we get that *John is in Berlin and Mary is in Paris* entails *John is in Berlin and Mary is in France*.
- So *and* leaves entailment patterns unchanged in its second argument: it is upward monotone w.r.t. its second argument.
- We could show the same thing *mutatis mutandis* for *and*'s first argument.

1. Using the same methodology (i.e. two statements or predicates P_1 and P_2 s.t. $P_1 \subseteq P_2$ or $P_1 \implies P_2$), determine the monotonicity of the following operators w.r.t. the argument labeled P.
   - It's not the case that P
   - If P, then Mary will be happy
   - If it's raining, then P
   - Some P are happy
   - Some students P

- At least 3 P are happy
- At least 3 students P
- Exactly 3 P are happy
- Exactly 3 students P

2. We have seen in class that some specific lexical items like *some* and *or*, trigger pragmatic inferences called *scalar implicatures*. For instance, *some* usually ends up meaning *some but not all*, and *or* ends up meaning *or but not and* ("exclusive" *or*). Let's consider the sentence *John ate some of the cookies*. Uttered out of the blue, it suggests John did no eat *all* of the cookies. What happens if you embed this sentence under:

   - It's not the case that ...
   - If ... then Mary will be happy

   Given your answers to 1., what kind of generalization can be drawn regarding the availability of scalar implicatures?

Context dependence in adjectives

We have talked about several types of modifying adjectives in class. If an adjective is used to modify a set-denoting expression X (semantic type `e->t`, e.g., a noun, or a noun that is already modified by another adjective), then the type of the adjective can be distinguished based on the nature of the relationship between the meaning of the expression before modification, i.e. $[\![X]\!]$, and the meaning of the expression after modification, i.e., $[\![\text{Adj } X]\!]$. These types include:

- **Intersective** adjectives, where the meaning of the adjective is a set (semantic type `e->t`) that can be characterized independently of what it modifies, and the denotation of a modified expression is $[\![\text{Adj } X]\!] = [\![\text{Adj}]\!] \cap [\![\text{X}]\!]$. An example is *red shirt*, where what it means for something to be red can be characterized more or less independently of what the something is, and for something to be a *red shirt* it is necessary and sufficient for it to be both red and a shirt.

- **Gradable** adjectives, where the meaning of the adjective constrains some property of the referent to be in a certain range of values, where the property is picked out by the adjective in a way that doesn't depend on the modified expression, but the requisite range of values depends both on the adjective's meaning and on a **comparison class**, typically provided by the modified expression. The denotation of a modified expression $[\![Adj\ X]\!]$ is the subset of $[\![X]\!]$ whose property value falls within the requisite range. A example is *big mouse*, where the property picked out by the adjective is size, the range of values is that the size must be above some threshold determined by the comparison class, and the comparison class is (probably) the set of mice. As a result, for something to be a *big mouse* it must be a mouse, and (slightly informally), it must be big, for a mouse. The term **subsective** adjective is sometimes used to include gradable adjectives, since the denotation of the modified expression is always a subset of the denotation of the original expression.

- Adjectives like *possible* and *alleged*, which are sometimes called "non-intersective" or "epistemic". The hallmark of such adjectives is that $[\![\text{Adj } X]\!]$ is not necessarily a subset of $[\![X]\!]$. For example, an *alleged criminal* is not necessarily a criminal.

- **Anti-intersective** adjectives, whose hallmark is that $[\![\text{Adj } X]\!] \cap [\![X]\!] = \emptyset$. For example, a *fake Rolex* cannot be a Rolex.

1. Explain what **context dependence** means for adjective meaning. Which of the above types of adjectives is the least context dependent?

2. Now consider the adjective–noun combination *beautiful dancer*. The adjective is ambiguous in a certain way pertaining to the its meaning contribution it composes with the noun: effectively, the adjective can have two different types of meanings. Describe that ambiguity. Do either of these types fall into one of the above four classes?

3. The adjectives *typical*, *average*, and *skillful* behave in a way that is similar to one of the types of meanings that *beautiful* can have. Which of these meanings is it?

Adjective embedding with BERT

This exercise is intended to be done on Colab here:

This exercise is intended to be completed on Colab, using this notebook: https://colab.research.google.com/drive/15Xee4ZLbVJzqMDR5sDJUdtFycNzHALDa?usp=sharing. Please carefully read and follow all instructions in the notebook.

Rational Speech Acts model.

The Rational Speech Acts model (Goodman & Frank, 2016) has a number of variants, but here is the original version used in Frank and Goodman (2012), where r is a referent (or more generally a meaning) and u is an utterance:

$$L_0(r|u) \propto \begin{cases} 1 & \text{if } r \text{ is compatible with the literal meaning of } u \\ 0 & \text{otherwise} \end{cases}$$

$$S_i(u|r) \propto L_{i-1}(r|u)$$

$$L_i(r|u) \propto S_i(u|r)P(r)$$

where $P(r)$ is a prior distribution over referents (or more generally meanings). (In the supplementary information to Frank and Goodman (2012), L_0 is referred to as $\tilde{w}_C(o)$.) Sometimes you will see the speaker and listener "functions" applied to each other, e.g., $S_1 = S(L_0)$, $L_1 = L(S(L_0))$, and so forth, emphasizing the functional and recursive nature of this model of pragmatic inference as a special case of theory of mind.

1. For the context presented in Figure 1A of Frank and Goodman (2012), compute $S(L_0)$ and $L(S(L_0))$[1] for that paper's model (the *Rational Speech-Act theory*—RSA) under a uniform referent prior $P(r_S)$, assuming the set of alternative utterances is `blue`, `green`, `circle`, `square`.

2. Consider a variant of their model in which the referent prior is incorporated into the L_0 level. For this model, the literal listener L_0's distribution is determined by ruling out referents inconsistent with the literal meaning of w and renormalizing the prior over the remaining referents:

$$L_0(r|u) \propto \begin{cases} P(r_S) & \text{if } r \text{ is compatible with the literal meaning of } u \\ 0 & \text{otherwise.} \end{cases}$$

   Compute $S(L_0)$ and $L(S(L_0))$ for this revised listener under the referent prior $P({\color{blue}\square}) = 0.1$, $P({\color{blue}\bigcirc}) = P({\color{green}\square}) = 0.45$.

References

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

[1] i.e., for each of these functions produce an exhaustive list of values for all 12 logically possible referent/utterance pairs, preferably presented as a matrix.