A Computational Pragmatics for Weaseling

An implementation in the RSA-framework

Leander Vignero
KU Leuven, Institute of Philosophy (CLPS)
April 12, 2019
What is weaseling? And what are probabilistic expressions?

Oxford English Dictionary:

1. Achieve something by use of cunning or deceit.
2. *North American* Behave or talk evasively.

Simple, technical definition: misrepresenting one’s degrees of belief/beliefs through vague verbiage.

What are PEs?
Examples: *probably, unlikely, about even, possibly, …*
Life Hack: Avoiding Meetings

meeting enthusiast: Are you coming to the meeting tomorrow?
You: ...
Tell the truth

meeting enthusiast: Are you coming to the meeting tomorrow?
You: No, I won’t come.

Downside: they are quite likely to ask a follow-up question: “Oh, why not?”
Lie outright

meeting enthusiast: Are you coming to the meeting tomorrow?
You: Yes, I will come.

Downside: You break your word. (Great social cost)
Weasel out from under the situation

meeting enthusiast: Are you coming to the meeting tomorrow?
You: Yes, I might come. / I will try to come. / Maybe.

Probabilistic expressions lower your accountability!
Another example

Tetlock’s example (Tetlock et al., 2017):
A forecaster says there is a distinct possibility of $p$ happening.

1. If $p$ happens the forecaster can say: I told you it could happen.
2. If $p$ does not happen, they can say: It was a mere possibility.

The study of probabilistic expressions originated in the intelligence studies (Kent, 1964).
Why should you care?

1. Academic Interest *per se*
2. Modeling/understanding human languages
3. Relation to other issues in social epistemology (specifically testimony)
Some housekeeping

1. I will tell you about the PE-literature and the RSA-literature
2. I will tell you how we can use these ideas to create an idealized language to think about communicating uncertainty and about weaseling
What are PEs?

Probabilistic expressions are used to communicate uncertainty (Kent 1964):
Thinking about these problems

It’s all nice and easy to speculate and philosophize about these things. But it would be even better to have an idealized language to analyze these communicative situations (communication of uncertainty and weaseling).

This is a reinterpretation of the RSA-framework: using it as an idealized prescriptive framework, rather than a descriptive framework.

Why not probabilistic dynamic epistemic logic (probabilistic DEL)?

1. probabilistic DEL is too strong: there is no real room for vagueness.
2. probabilistic DEL does not have a “critical” theory of mind.
The Rational Speech Act Framework (RSA)

Framework from computational cognitive psychology that operationalizes the Gricean idea of cooperation probabilistically.

Advantages:

1. Limited set of utterances.
2. Theory of mind.
3. Picks up on implicatures.
Philosophy of language crash course: Gricean pragmatics

Grice thought of language as a cooperative venture: speakers choose their words in such a way that they are best understood. Listeners, in turn, know this and use this information to further narrow the scope of interpretation.

RSA turns “in such a way that they are best understood” into “in a way that maximizes the probability of being correctly understood”.
Scalar implicature

“The idea is that if the speaker were in a position to make the stronger statement, he would have. Since he did not, he must believe that the stronger statement is not true.” – Wayne Davis

Modeled by Goodman and Stuhlmüller (2013) using an RSA-style model for utterances like none, some and all.

For the philosophers and modal logicians: this also holds for classic modal notions like possibility, necessity.
The literal listener

RSA is a layered model that produces degrees of belief regarding states of the world upon hearing an utterance (Frank & Goodman, 2012). A literal listener that interprets utterances literally:

\[ L_0(s \mid u) \propto P(s)\chi_{[u]}(s). \]  

(1)

In words: the literal listener interprets utterances \( u \) in proportion to their priors with respect to the states \( s \) that belong to the meaning of \( u \). This amounts to Bayesian updating.
The literal listener

Consider the utterances: *some* and *all*.

(a) Prior

(b) $L_0(\cdot \mid \text{some})$
The pragmatic speaker

Utility function that determines how well an utterance succeeds at communicative aims:

\[
\mathcal{U}(u \mid s) = -\log(L_0(s \mid u)).
\] (2)

A (subrational) pragmatic speaker:

\[
S_1(u \mid s) \propto \exp(\lambda \mathcal{U}(u \mid s)).
\] (3)

The pragmatic speaker picks words proportionally to the power of the utilities.
The pragmatic speaker

(a) $\mathbb{U}(\cdot | 6)$

(b) $S_1(\cdot | 6)$
The pragmatic listener

A pragmatic listener:

\[ L_1(s \mid u) \propto P(s)S_1(s \mid u). \]  

(4)

Figure: \( L_1(\cdot \mid \text{some}) \)
Quick Recap

RSA is a layered model that uses Bayesian inference to infer knowledge states of agents that are situated lower in the hierarchy.

For instance: the pragmatic speaker $S_1$ is defined in terms of the literal listener $L_0$ (mediated by the utility function). This is why this framework has such an interesting theory of mind.
Some more housekeeping

The work by Yoon et al. (2016, 2017, 2018) on politeness is a starting point.

1. a variation on Yoon et al. (2016, 2017, 2018)
2. a variation on Bergen and Goodman (2015)
3. a variation on Yoon et al. (2016, 2017, 2018)
4. a personal variation

I will start with two honest speakers and then move to weasels.

I will conclude with one slide on the implementation in Julia.
Hearing the nuance of PEs

Think about the states as probabilities ($0; 0.2; 0.4; 0.6; 0.8, 1$). We need to adjust equation 1 from $L_0(s \mid u) \propto P(s)\chi_{[u]}(s)$ to:

$$L_0(s \mid u) \propto P(s)N(u)(s).$$

(Variation on Yoon et al. (2016, 2017, 2018).)

(a) Normal. (b) With nuance.
Communicating a distribution

We need to update both the listeners and the utility function. We need to assume a set of probability distributions that can be communicated. For now let us assume they are the posteriors of observations.

\[ L_0(s, o \mid u) \propto P(s)P(o \mid s)N(u)(s) \]  

(6)

What really matters is the utility function:

\[ \mathbb{U}(u \mid o) = -KL_D(P(\cdot \mid o) \parallel L_0(\cdot \mid u)) \]  

(7)

(Variation on Bergen and Goodman (2015))
Two small problems

1. Are we always communicating posteriors of well-defined observations?
2. Is there always common belief about the distributions that could be communicated?

Two solutions

1. Finding ways of modeling the communication of other types of distributions.
2. An alternative I propose later on.
Tweeking the utility function

Introduce a new utility function $U_{ego}$, that only concerns non-epistemic utilities. Then the total utility, $U_{total}$, is a convex combination of $U_{epistemic}$ and $U_{ego}$:

$$U_{total} = \beta U_{epistemic} + (1 - \beta)U_{ego}.$$  \hspace{1cm} (8)

$\beta$ represents the honesty of the speaker.

(idea from Yoon et al. (2016, 2017, 2018))

With such utility functions we can try to understand weaseling! Speakers can try to weasel out from under situations, whilst listeners can try to catch them in the act!
An alternative

A four-layered model:

\[ L_0(s \mid u) \propto P(s)N(u)(s) \quad (9) \]

\[ S_1(u \mid s) \propto \exp(\lambda U) \quad (10) \]

\[ L_1(s \mid u) \propto P(s)S_1(u \mid s) \quad (11) \]

\[ S_2(u) \propto \exp(\lambda K L_D(P(\cdot) \mid| L(\cdot \mid u))) \quad (12) \]
Implementation in the Julia programming language

1 procedural implementation.
2 The crux is that we can represent probability distributions as unit vectors.
Conclusion/Upshot

1. I have provided a provisional philosophical analysis of weaseling.
2. I have created a couple of computational models for thinking about communicating uncertainty with PEs.
3. I have created computational models for thinking about weaseling.
4. I have implemented these frameworks in the Julia programming language.
References

References


References


References


Thank you for listening!

e-mail: leander.vignero@kuleuven.be

github: leandervignero