Computational pragmatics:
Introducing the Rational Speech Act framework

Brandon Waldon

Georgetown University

COSC/LING 5402 - April 24, 2024

Can you reach the top shelf?

… please? … or should you be working the register?

Pragmatics

Same string, many possible interpretations depending on *context*.

> When a diplomat says *yes*, he means ‘perhaps‘;
> When he says *perhaps*, he means ‘no‘;
> When he says *no*, he is not a diplomat. — Voltaire

Context: common-ground information, social dynamics, beliefs about speaker intentions/identity, ‘norms’ of linguistic interaction...

Pragmatics: study of speaker meaning and its relationship to context.

"My friend has glasses"

If I had meant...

,

... there were 'better' things I could have said.

So, I probably meant

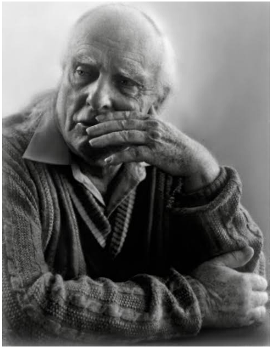
(But this is obviously quite context-dependent!)

Grice's approach

Pragmatic inference results from counterfactual reasoning about *alternative utterances* a speaker could have produced.

Reasoning assumes that speakers are rational and cooperative.

Assumptions expressed as *maxims* of conversation.



Maxim of *quality*: don't say things that you know to be false (or for which you lack evidence).

Maxim of *quantity*: say all you need to say (and no more).

Grice (1975)

“My friend has glasses”

(a) *Contextual premise*: ↑ these are S’s three relevant possible referents.

(b) *Contextual premise*: it is mutual, public information that S has complete knowledge re: who they intend to refer to.

(c) Assume S is being cooperative insofar as they are obeying the maxims of Quality and Quantity.

(d) Then S will assert what is maximally relevant, informative, and true.

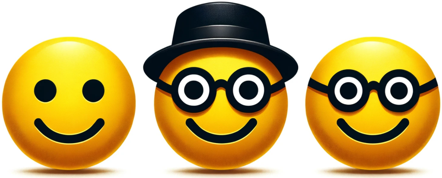
(e) By (a), *My friend has a hat* (`hat`) is more informative (and just as relevant) as *My friend has glasses*.

(f) Therefore, S must lack sufficient evidence to assert `hat`.

(g) By (b), S must lack evidence for `hat` because it’s false.

Inspired by Chris Potts’ handout on conversational implicature:
`https://web.stanford.edu/class/linguist130a/materials/ling130a-handout-02-20-implicature.pdf`

The Classic View of Pragmatics



"My friend has glasses" implicates ¬`hat`.

Meaning treated as categorical: implicatures computed or not.

Inferences arise categorically if premises are met.

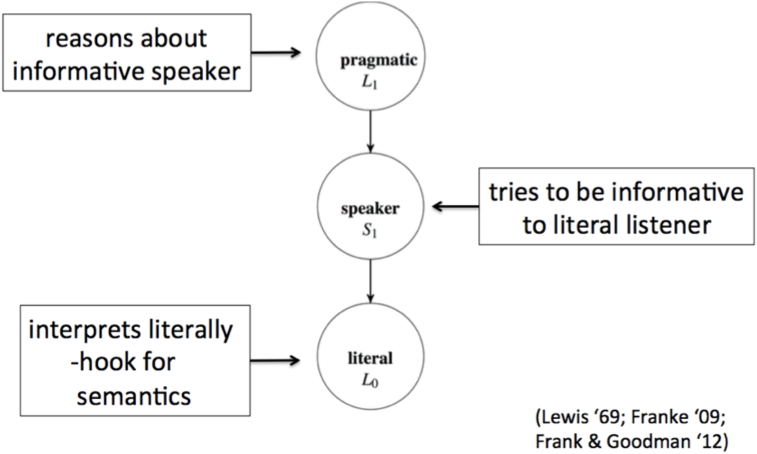
Gradience (and population-level variation!) ignored, or accommodated by exceptions or processing contraints.

Degen (2023)

The 'probabilistic' turn in pragmatics

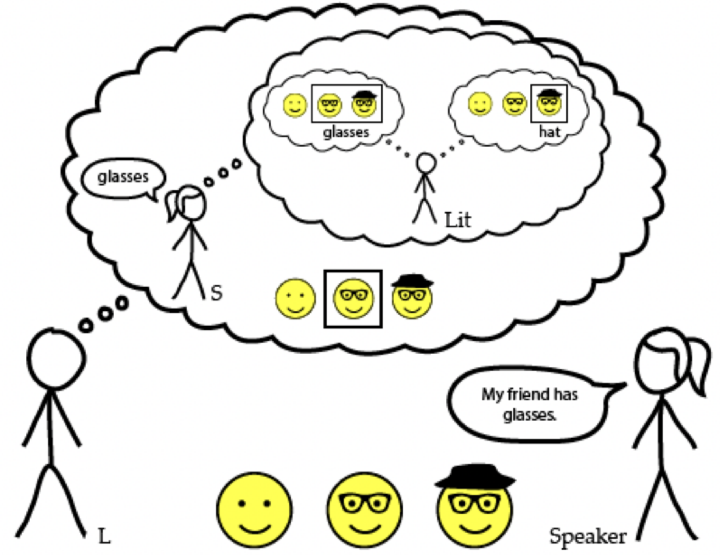
- Probabilistic pragmatics: computational approach integrating insights from formal semantics, psycholinguistics, cognitive science
- Key advances:
  - Formalizes general principles of conversation (including Grice's 'maxims')
  - Treats language production and interpretation as probabilistic, boundedly rational processes
  - Allows linguistic knowledge to interact with communicative pressures and world knowledge
- Rational Speech Act (RSA) framework: influential probabilistic approach

Modeling pragmatic inference: the RSA framework



Goodman and Frank (2016)

Modeling pragmatic inference: the RSA framework



Goodman and Frank (2016)

The basic 'vanilla' RSA model

- Language use as signaling game between speaker and listener
- Defined over utterances U and meanings S
- Semantic foundation: literal meaning $[\![\cdot]\!] : U \rightarrow S \rightarrow \{0, 1\}$
- Recursive probabilistic production and interpretation rules:
  - Literal listener Pr_{lit}: interprets according to literal semantics
  - Pragmatic speaker $Pr_{speaker}$: reasons about Pr_{lit}, balances informativeness and cost
  - Pragmatic listener Pr_L: infers $Pr_{speaker}$'s meaning w/ Bayes' rule

Bayes' rule: for a hypothesis H and data d:

$$Pr(H|d) \propto Pr(d|H) * Pr(H)$$

For Pr_L, Hypotheses are elements of S (intended meanings), $d$ata are elements of U (observed utterances).

Pragmatic interpretation in 'vanilla' RSA

$S = \{$  $\}$ $U = \{$ "hat", "glasses", "smiling" $\}$

'Literal' listener: $Pr_{lit}(s|u) = \dfrac{[\![u]\!]^s}{\sum_{s' \in S} [\![u]\!]^{s'}}$

'Literal' listener			
"hat"	0	0	**1**
"glasses"	0	**1**	**1**
"smiling"	**1**	**1**	**1**

Pragmatic interpretation in ‘vanilla’ RSA

$S = \{$  $\}$ $U = \{$ “hat”, “glasses”, “smiling” $\}$

‘Literal’ listener: $Pr_{lit}(s|u) = \dfrac{[\![u]\!]^s}{\sum_{s' \in S} [\![u]\!]^{s'}}$

‘Literal’ listener			
“hat”	0	0	1
“glasses”	0	0.5	0.5
“smiling”	0.33	0.33	0.33

Pragmatic interpretation in ‘vanilla’ RSA

$S = \{$  $\}$ $U = \{$ “hat” , “glasses” , “smiling” $\}$

‘Literal’ listener: $Pr_{lit}(s|u) = \frac{[\![u]\!]^s}{\sum_{s' \in S} [\![u]\!]^{s'}}$

‘Literal’ listener			
“hat”		0	
“glasses”		0.5	
“smiling”		0.33	

Pragmatic interpretation in 'vanilla' RSA

$S = \{$  $\}$ $U = \{$"hat", "glasses", "smiling"$\}$

'Pragmatic' speaker: $Pr_{speaker}(u|s) = \dfrac{Pr_{lit}(s|u)}{\sum_{u' \in U} Pr_{lit}(s|u')}$

'Pragmatic' speaker	"hat"	"glasses"	"smiling"
	0	0	1
	0	0.5	0.33
	1	0.5	0.33

Pragmatic interpretation in 'vanilla' RSA

$S = \{$  $\}$ $U = \{$"hat", "glasses", "smiling"$\}$

'Pragmatic' speaker: $Pr_{speaker}(u|s) = \frac{Pr_{lit}(s|u)}{\sum_{u' \in U} Pr_{lit}(s|u')}$

'Pragmatic' speaker	"hat"	"glasses"	"smiling"
	0	0	1
	0	0.60	0.40
	0.55	0.27	0.18

Pragmatic interpretation in ‘vanilla’ RSA

$S = \{$  $\}$ $U = \{$ “hat”, “glasses”, “smiling” $\}$

‘Pragmatic’ listener: $Pr_L(s|u) = \frac{Pr_{speaker}(u|s)}{\sum_{s' \in S} Pr_{speaker}(u|s')}$

‘Pragmatic’ listener			
“hat”	0	0	0.55
“glasses”	0	0.60	0.27
“smiling”	1	0.40	0.18

Pragmatic interpretation in ‘vanilla’ RSA

$S = \{$  $\}$ $U = \{$ “hat”, “glasses”, “smiling” $\}$

‘Pragmatic’ listener: $Pr_L(s|u) = \dfrac{Pr_{speaker}(u|s)}{\sum_{s' \in S} Pr_{speaker}(u|s')}$

‘Pragmatic’ listener			
“hat”	0	0	1
“glasses”	0	0.69	0.31
“smiling”	0.64	0.25	0.11

Enriching the model: additional parameters

- Listener prior over meanings: 'what do I think my interlocutor is trying to say, before observing an utterance'?

'Literal' listener: $Pr_{lit}(s|u) = \frac{[\![u]\!]^s * P_{prior}(s)}{\sum_{s' \in S}[\![u]\!]^{s'} * P_{prior}(s')}$

- 'Cost' term on utterances:
  - Grice's maxim of manner: be brief (avoid unnecessary prolixity).
  - Speaker utility formalized as a tradeoff of informativity and cost.
- Rationality term α 'penalizes' low-utility utterances.

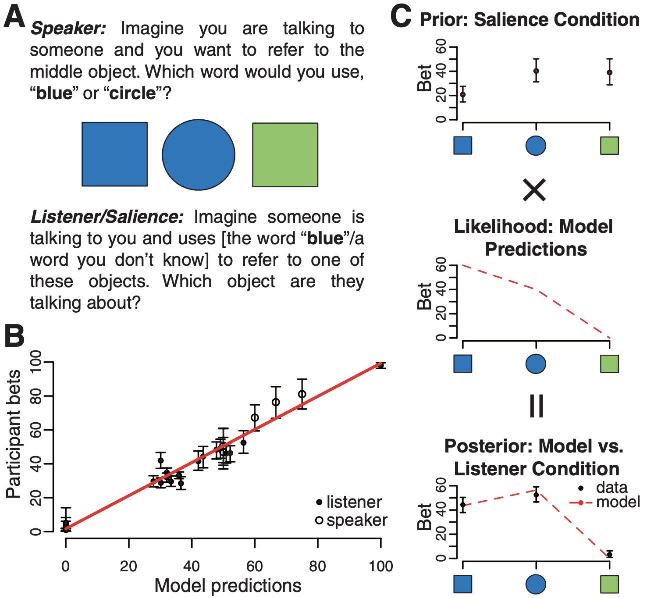
'Pragmatic' speaker: $Pr_{speaker}(u|s) \propto e^{\alpha[ln(Pr_{lit}(s|u)) - cost(u)]}$

Resources

- `http://webppl.org/`: online interpreter for WebPPL, a probabilistic programming language (Goodman & Stuhlmueller, 2014)
  Javascript syntax; functional programming properties (e.g., no for-loops)
- `http://forestdb.org/`: a repository with several WebPPL implementations of published RSA models
- `https://www.problang.org/` (Scontras, Tessler, & Franke, online): an intro to RSA taught in WebPPL.

Implementation of the 'smiley' model:

- Model code: `https://tinyurl.com/rsa-model`
- Copy and paste into WebPPL interpreter `http://webppl.org/`



Frank and Goodman (2012)

Enriching the model: multiple speaker ‘goals’

In ‘vanilla’ model, speakers have one communicative goal (e.g., listener identifies intended referent from )

Enriching the model: multiple speaker 'goals'

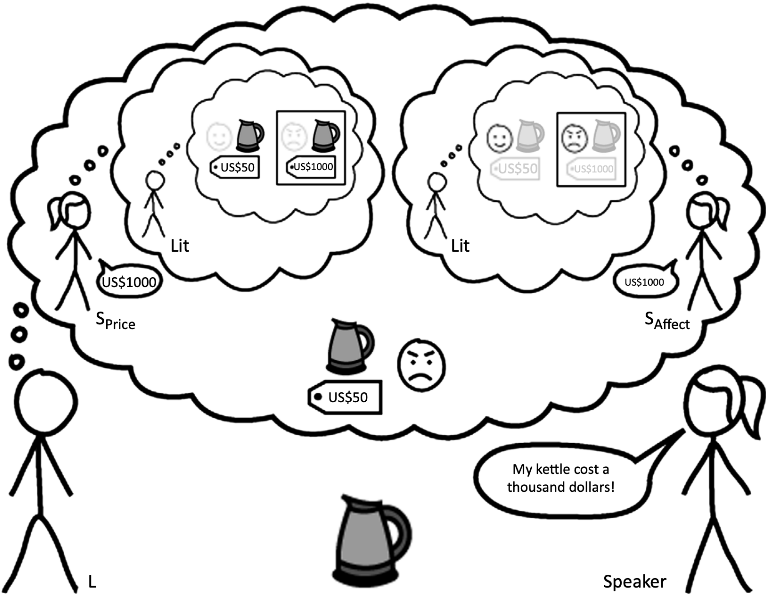
In 'vanilla' model, speakers have one communicative goal (e.g., listener identifies intended referent from )

In practice, goals change across contexts, and utterances are themselves indicative of speaker goals.

S: "The kettle cost me a million dollars"

'Price' goal: *How much did S pay?*
'Affect' goal: *How does S feel about the price?*

Kao et al. (2014)



Goodman and Frank (2016)

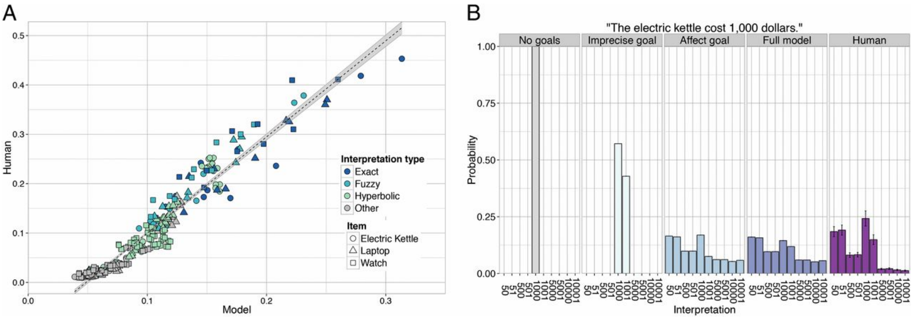
Jointly inferring multiple variables

The kettle cost me a thousand dollars!

$S = \{50, 51, \cdots, 10001\}$ $G = \{\texttt{price}, \texttt{affect}\}$

$$P_L(s, g|u) \propto P_{speaker}(u|s, g) * P_{\text{state-prior}}(s) * P_{\text{goal-prior}}(g)$$

- If goal were `price`, "\$1000!" would be communicatively optimal [$P_{speaker}$("\$1000!"$|1000, \texttt{price})$ is high].
- But kettles are unlikely to cost that much [$P(1000)$ is low].
- $\rightarrow$ Listener explains utterance via `affect` goal.


A
Human
Model
Interpretation type
Exact
Fuzzy
Hyperbolic
Other
Item
Electric Kettle
Laptop
Watch
B
"The electric kettle cost 1,000 dollars."
No goals
Imprecise goal
Affect goal
Full model
Human
Probability
Interpretation

Assessing RSA model fit

Recall: models frequently include free parameters (α optimality parameter, cost terms on utterances)

Free parameter values often estimated using Bayesian Data Analysis techniques

Bayes' rule: for a hypothesis H and data d:

$$Pr(H|d) \propto Pr(d|H) * Pr(H)$$

In BDA, *H*ypotheses are model parameterizations, *d*ata are empirical (often experimental) observations.

- Step 1: 'condition' on experimental data to compute posterior distributions over model parameter values.
- Step 2: sample from posterior to 'run the model forward'
- Step 3: assess predictions against observations

Some other domains of application

- Scalar inference (the 'drosophila' of pragmatics):
  John ate some of the cookies $\rightarrow$ 'John didn't eat all of the cookies'.
  (Goodman & Frank, 2016; Potts, Lassiter, Levy, & Frank, 2015; Waldon & Degen, 2020, inter alia)
- Linguistic vagueness & imprecision (e.g., *John is tall*, *The townspeople are asleep*):
  Listeners jointly infer speaker meanings and contextual variables (thresholds of predication, precision standards)
  (Lassiter & Goodman, 2013; Schuster & Degen, 2020; Waldon, 2022, inter alia)
- Social meaning: listeners jointly infer speaker meanings and social goals (e.g., appear 'relaxed'/'professional')
  (Burnett, 2019)

NLP application 1: contrastive image captioning

"[H]igh-quality captions are not merely *true*, but also *pragmatically informative* in the sense that they highlight salient properties and help distinguish their inputs from similar images."

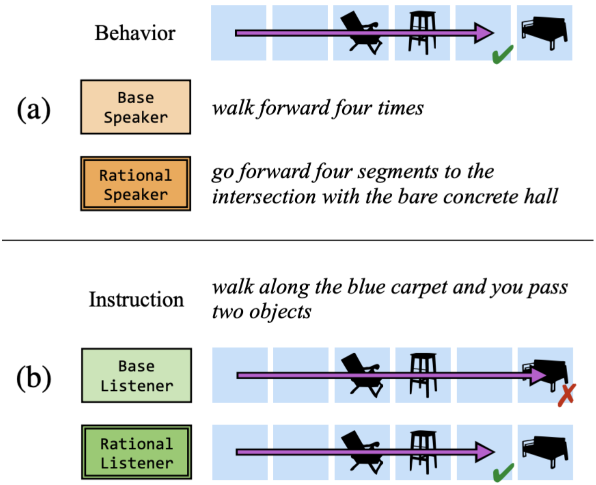
(Cohn-Gordon, Goodman, & Potts, 2018, "Pragmatically Informative Image Captioning with Character-Level Inference")

- $S_0(caption|image) \propto$ sequence probability from a neural image captioning model (trained on individual images).
- $L_{lit}(image|caption) \propto S_0(caption|image)$
- Pragmatic $S_1(caption|image) \propto L_{lit}(image|caption)$

S_0 caption: a double decker bus
S_1 caption: a red double decker bus

- Implements an incremental, character-level S_1 (more tractable utterance space!)

NLP application 2: directing agents in navigable space



(Fried, Andreas, & Klein, 2018, "Unified Pragmatic Models for Generating and Following Instructions")

References I

Burnett, H. (2019). Signalling games, sociolinguistic variation and the construction of style. *Linguistics and Philosophy*, *42*(5), 419–450.

Cohn-Gordon, R., Goodman, N., & Potts, C. (2018). Pragmatically informative image captioning with character-level inference. In *Proceedings of the 2018 conference of the north american chapter of the association for computational linguistics: Human language technologies, volume 2 (short papers)* (pp. 439–443).

Degen, J. (2023). The rational speech act framework. *Annual Review of Linguistics*, *9*(1), 519–540.

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

References II

Fried, D., Andreas, J., & Klein, D. (2018). Unified pragmatic models for generating and following instructions. In *Proceedings of the 2018 conference of the north american chapter of the association for computational linguistics: Human language technologies, volume 1 (long papers)* (pp. 1951–1963).

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

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3: Speech acts* (p. 41-58). New York: Academic Press.

Kao, J., Wu, J. Y., Bergen, L., & Goodman, N. D. (2014). Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences*, *111*(33), 12002–12007.

References III

Lassiter, D., & Goodman, N. D. (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. In *Semantics and linguistic theory* (Vol. 23, pp. 587–610).

Potts, C., Lassiter, D., Levy, R., & Frank, M. C. (2015). Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics*, *33*(4), 755–802.

Schuster, S., & Degen, J. (2020). I know what you're probably going to say: Listener adaptation to variable use of uncertainty expressions. *Cognition*, *203*, 104285.

Waldon, B. (2022). A novel probabilistic approach to linguistic imprecision. In *Measurements, numerals and scales: Essays in honour of stephanie solt* (pp. 307–327). Springer.

Waldon, B., & Degen, J. (2020). Modeling behavior in truth value judgment task experiments. *Proceedings of the Society for Computation in Linguistics*, *3*(1), 10–19.