

Inferring Latent States with the RSA Model

Nicholas Tomlin

13 July 2018

1 Background

The Rational Speech Acts model (Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013) provides a computational analogue to Gricean pragmatics in which agents recursively reason about each other’s communicative intentions. Traditionally, this involves base speakers and listeners (S_0 and L_0) who learn literal semantic mappings from utterances $u \in U$ to characteristic world states $w \in W$. Pragmatic agents then use Bayesian reasoning to produce and interpret utterances probabilistically, according to the following proportions:

$$\begin{aligned} P_{L_0}(w \mid u) &\propto [[u]](w) \cdot P(w) & P_{S_0}(u \mid w) &\propto \exp(\lambda_1(\log([[u]](w)) - C(u))) \\ P_{S_1}(u \mid w) &\propto \exp(\lambda(\log P_{L_0}(w \mid u) - C(u))) & P_{L_1}(w \mid u) &\propto P_{S_0}(u \mid w) \cdot P(w) \\ P_{L_2}(w \mid u) &\propto P_{S_1}(u \mid w) \cdot P(w) & P_{S_2}(u \mid w) &\propto \exp(\lambda_2(\log P_{L_1}(w \mid u) - C(u))) \end{aligned}$$

For a worked example of how this recursive reasoning can produce typical Gricean implicatures, see Monroe and Potts (2015).

Calculating these inverse probabilities proves difficult at scale, however. The earliest renditions of RSA, such as Smith et al. (2013), implement a fixed lexicon so that the truth condition $[[u]](w)$ can be evaluated directly. By contrast, Learned RSA (Monroe and Potts, 2015) allows optimization of the semantic lexicon through backpropagation, but requires hand-built feature representations of utterances and world states. Most recently, neural RSA (Andreas and Klein, 2016; Monroe et al., 2017; Fried et al., 2017) samples utterances from the base models and makes pragmatic inferences accordingly. This is the method we propose to implement in the experiment described below.

2 The Negotiation Game

Lewis et al. (2017) describe a multi-agent bargaining game in which two agents use dialogue to negotiate in a semi-cooperative environment. Agents, each with their own private goal vectors, must reach an agreement to divvy a shared set of objects; failure to cooperate results in zero reward. After collecting game data through Amazon Mechanical Turk, the original experimenters built dialogue agents capable of negotiating with real human players.

In order to produce goal-oriented dialogue, these agents used *dialogue rollouts* as a planning mechanism. Specifically, agents were trained to simulate full conversations and select utterances with the highest expected reward. We suggest that the proper application of pragmatic reasoning may eliminate the need for such dialogue rollouts.

3 Reasoning About Latent States

Since the negotiation game described above is semi-cooperative, we expect the strongest agents may not adhere to the Gricean cooperative principle. For example, it may be in the speaker’s best interest to keep their private goal vector hidden, while still conveying information about desired outcomes. Furthermore, these desired outcomes may change over time; Chen (2017) suggests that the strongest agents behave aggressively at the start of negotiation, and gradually become more agreeable.

In line with the Rational Speech Acts framework, we propose pragmatic agents who reason recursively about each other’s intents and desired outcomes, rather than about the explicit state representation.

References

- Andreas, J. and Klein, D. (2016). Reasoning about pragmatics with neural listeners and speakers. *arXiv preprint arXiv:1604.00562*.
- Chen, D. (2017). Building effective goal-oriented dialogue agents.
- Frank, M. C. and Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.
- Fried, D., Andreas, J., and Klein, D. (2017). Unified pragmatic models for generating and following instructions. *arXiv preprint arXiv:1711.04987*.
- Goodman, N. D. and Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in cognitive science*, 5(1):173–184.
- Lewis, M., Yarats, D., Dauphin, Y. N., Parikh, D., and Batra, D. (2017). Deal or no deal? end-to-end learning for negotiation dialogues. *arXiv preprint arXiv:1706.05125*.
- Monroe, W., Hawkins, R. X., Goodman, N. D., and Potts, C. (2017). Colors in context: A pragmatic neural model for grounded language understanding. *arXiv preprint arXiv:1703.10186*.
- Monroe, W. and Potts, C. (2015). Learning in the rational speech acts model. *arXiv preprint arXiv:1510.06807*.
- Smith, N. J., Goodman, N., and Frank, M. (2013). Learning and using language via recursive pragmatic reasoning about other agents. In *Advances in neural information processing systems*, pages 3039–3047.