Modeling Communication to Coordinate Perspectives in Cooperation

Stephanie Stacy[1], Chenfei Li[1], Minglu Zhao[1], Yiling Yun[1], Max Kleiman-Weiner[3,4], Tao Gao[1,2]

[1]Department of Statistics, UCLA, [2]Department of Communication, UCLA, [3]Department of Brain and Cognitive Sciences, MIT,
[4]Department of Psychology, Harvard University

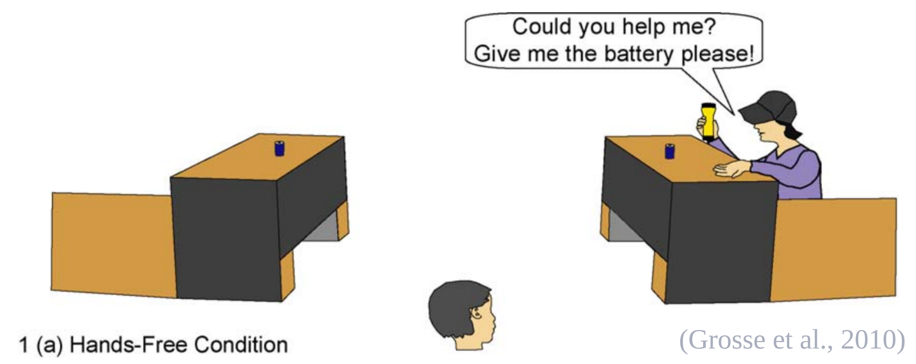
Motivation

Emergence of communication — an adaptation for human cooperation?

Still, communication is highly overloaded

"Point to a piece of paper. And now point to its shape — now to its color — now to its number ... How did you do it?"
-Wittgenstein, 1953

Despite this, even young children are good at interpreting cooperative requests



(Grosse et al., 2010)

How can we model cooperative communication under this ambiguity?

Model Components

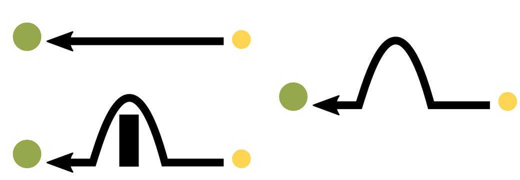
Building Blocks

Intuitive Action Utility from Physical Commonsense

$$Utility = U(a, mind)$$

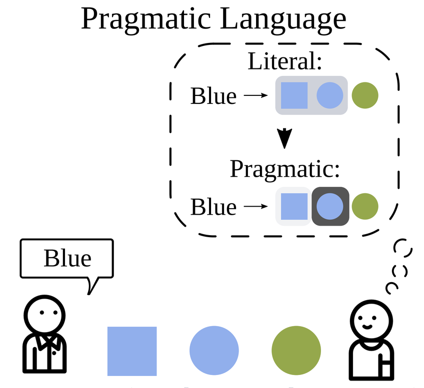
$$P(a|mind) \propto e^{\beta \mathbb{E}[U(a,mind)]}$$

Not Surprising: Surprising:



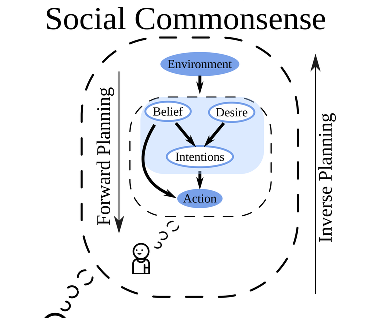
(Gergley et al., 1995; Jara-Ettinger et al., 2019)

Rational Speech Acts for Pragmatic Language



(Frank & Goodman, 2012)

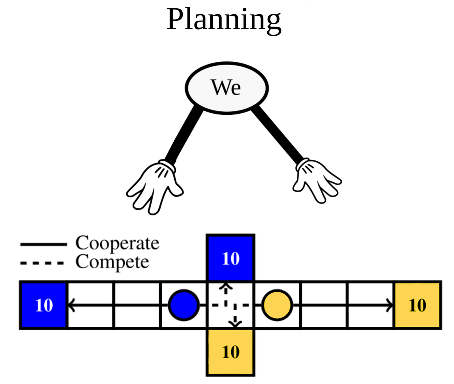
Theory of Mind (ToM) as Social Commonsense



$$P(mind|a, env) \propto P(a, mind|env)P(mind|env)$$

(Baker et al., 2009)

"We Mode" for Cooperative Planning



(Kleiman-Weiner et al., 2016)

Challenges / Solutions

1. **Challenge:** ToM is designed to handle actions not communication. **Solution:** Treat signals as rational actions. Subject them to the same modeling framework as actions

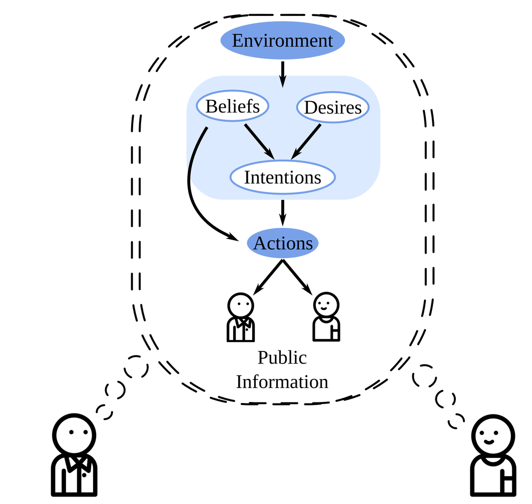
2. **Challenge:** The utility of a signal is not straightforward to define. **Solution:** Signals change the mind. The mind determines rational actions. Actions have well defined utilities

3. **Challenge:** Rationality is defined in terms of self interests. **Solution:** Selfishness creates ambiguity. Cooperative logic as a joint utility can resolve the ambiguity

Imagined We for Communication

Rational Speech Acts Style Pragmatics → Signal

Shared Agency ToM → Joint Mind



Joint Utility Calculus → Joint Action

Computational Model

Given private knowledge, a signaler **chooses a signal**

$$P(signal|mind_t) \propto e^{\beta \mathbb{E}[U(signal, mind_t)]}$$

According to **how good that signal is** expected to be

$$\mathbb{E}[U(signal, mind_t)] = \mathbb{E}_{P(a|signal)}[U(a, mind_t)]$$

Which is measured by how likely cooperators **take an action** upon **hearing that signal** (through updating the joint public mind)

$$P(a|signal) = \sum_{mind_j} P(mind_j|signal)P(a|mind_j)$$

Bayesian inference formalizes how signals affect the joint public mind

$$P(mind_j|signal) \propto P(signal|mind_j)P(mind_j)$$

And rational planning tells us which actions are rational given a particular mind

$$P(a|mind_j) \propto e^{\beta U(a, mind_j)}$$

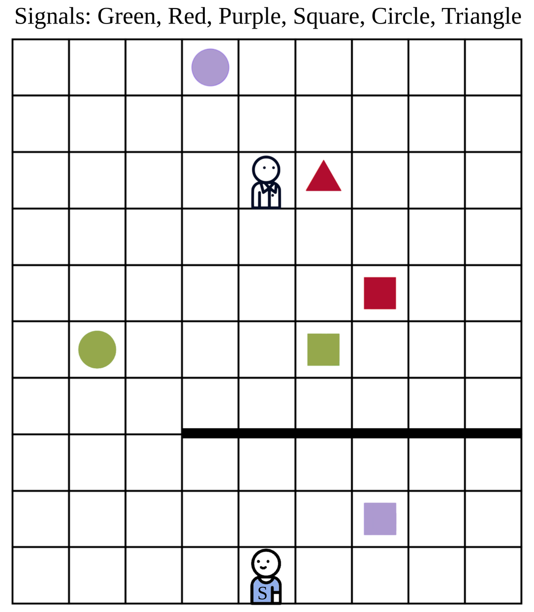
And **how good those actions are** according to the private knowledge in the signaler's mind

$$P(a|mind_t) \propto e^{\beta U(a, mind_t)}$$

Signaling Task

When and how should I ask for help?

Signals: Green, Red, Purple, Square, Circle, Triangle

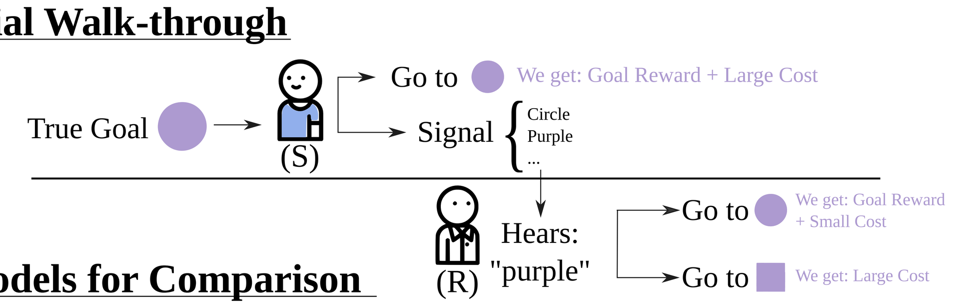


(Private Speaker Knowledge) Target:

Sampled:
- Item #, location, and identity
- Target item

Fixed:
- Grid size
- Agent locations
- Item feature set (3 colors; 3 shapes)

Trial Walk-through



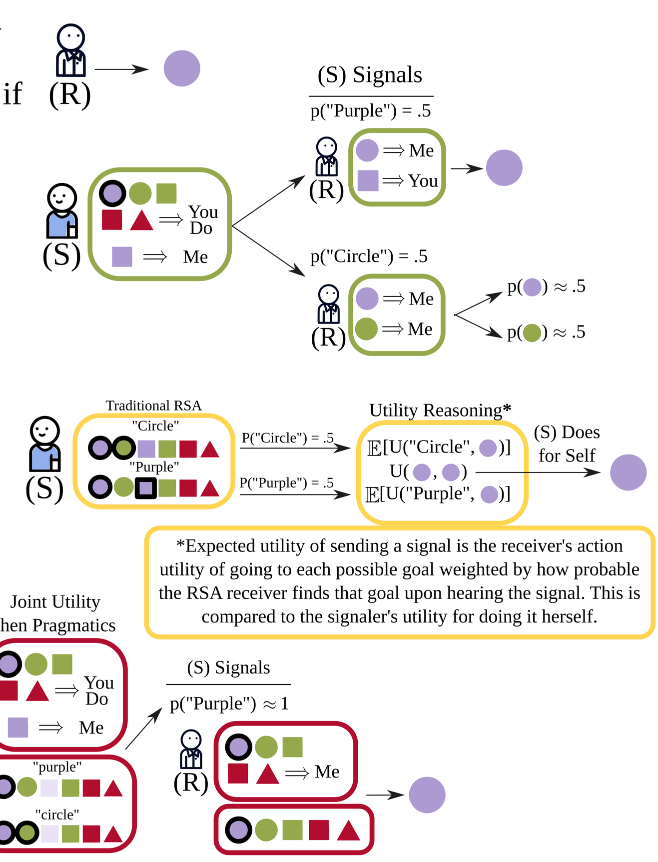
Models for Comparison

Optimal Central Control (CC)
Best action by joint utility if goal were known to all

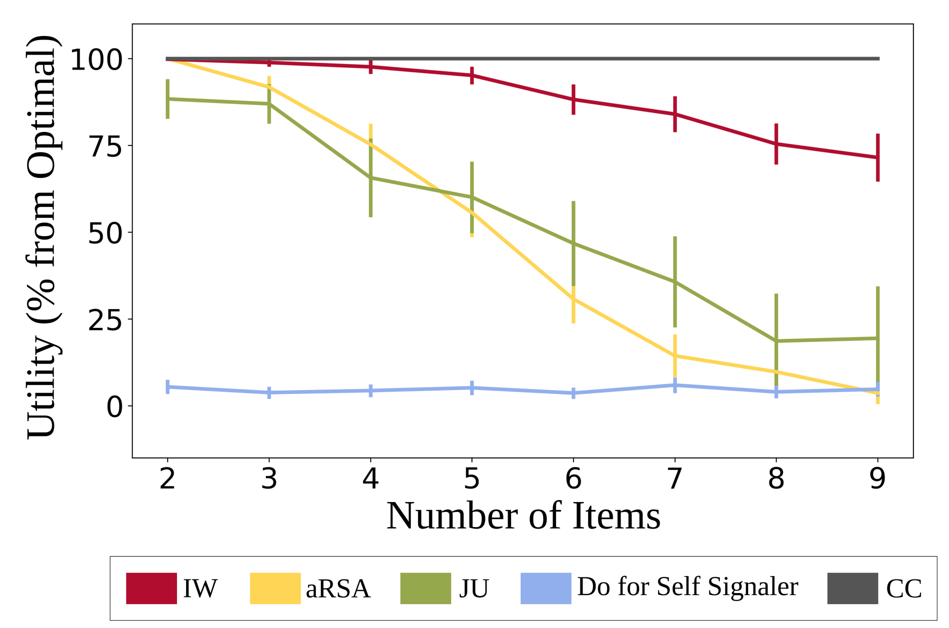
Joint Utility (JU)
Signaler uses joint utility to decide who should do it; if receiver should, signaler sends a truthful signal uniformly

Pragmatics (aRSA)
Traditional RSA for language augmented with individual utility reasoning

Imagined We (IW)
First joint utility reasoning then joint pragmatics

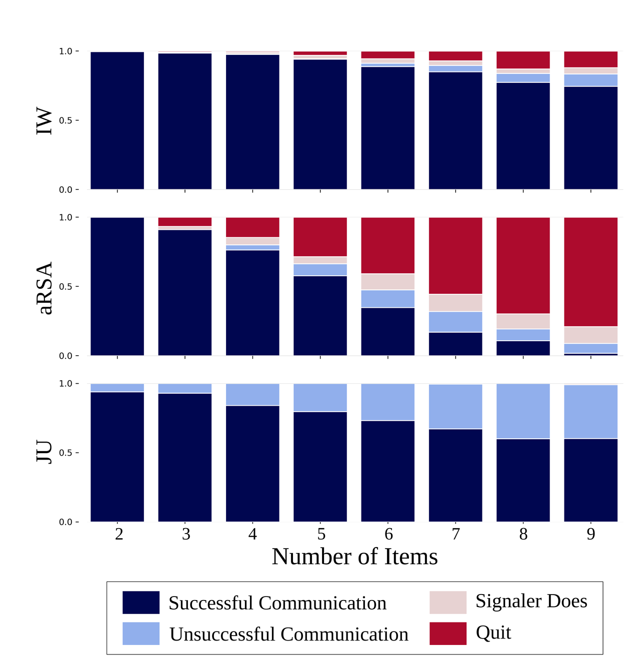


Signaling Joint Intentions



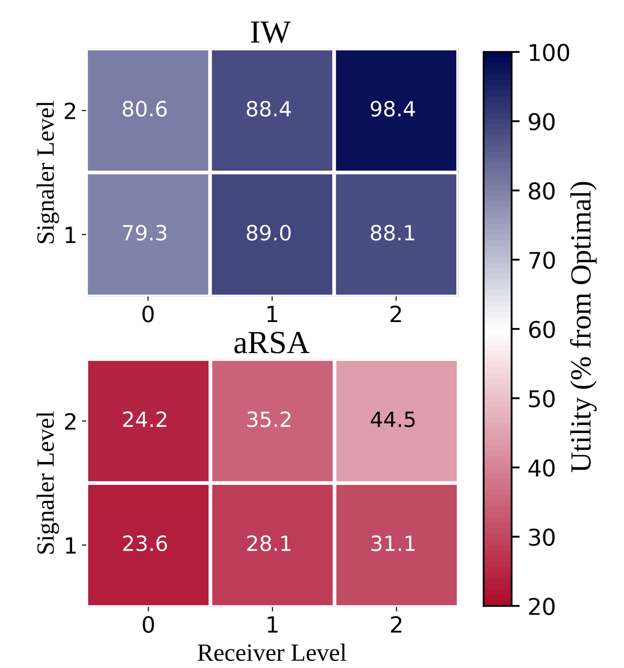
We examine how different models perform in the signaling task under increasing ambiguity (# items).

1. Communication improves performance, except when overloading is at its maximum
2. IW cooperators are able to successfully communicate under high ambiguity; whereas performance in JU and aRSA decay substantially



Achieved utilities are broken down to demonstrate the ways in which models behave differently.

1. JU always signals, but the receiver is unable to understand
2. aRSA predicts negative utilities and quits instead of communicating



IW	Receiver 0	Receiver 1	Receiver 2
Signaler 2	80.6	88.4	98.4
Signaler 1	79.3	89.0	88.1

aRSA	Receiver 0	Receiver 1	Receiver 2
Signaler 2	24.2	35.2	44.5
Signaler 1	23.6	28.1	31.1

When communicators of different reasoning complexities play each other, how does this affect utility?

1. Adding deeper reasoning to communicators improves performance
2. The least complex IW outperforms even the most complex aRSA, indicating deep reasoning may not be needed

Conclusions

1. Leveraging Multiple Types of Context
   Both joint utility and joint pragmatics add to disambiguating overloaded signals
2. Robust Inferences
   The Imagined We framework outperforms competing models with many items and when highly overloaded
3. Less Reliance on Recursion
   Although recursion helps task performance, the IW does not need deep recursion to perform well because the burden of inference is spread to joint utility and cooperative assumptions

References: Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. Cognition, 113(3), 329-349.
Frank, M., and Goodman, N. (2012). Predicting pragmatic reasoning in language games. Science 336.6084, 998-998.
Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. Cognition, 56(2), 165-193.
Grosse, G., Moll, H., & Tomasello, M. (2010). 21-month-olds understand the cooperative logic of requests. Journal of Pragmatics, 42, 3377–3383.
Jara-Ettinger, J., Floyd, S., Huey, H., Tenenbaum, J. B., & Schulz, L. E. (2019). Social Pragmatics: Preschoolers Rely on Commonsense Psychology to Resolve Referential Underspecification. Child development.
Kleiman-Weiner, M., Ho, M. K., Austerweil, J. L., Littman, M. L., & Tenenbaum, J. B. (2016). Coordinate to cooperate or compete: abstract goals and joint intentions in social interaction. In CogSci.
Misyak, J., Noguchi, T., & Chater, N. (2016). Instantaneous conventions: The emergence of flexible communicative signals. Psychological science, 27, 1550–1561.
Wittgenstein, L. (1953). Philosophical investigations.

Acknowledgments: This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650604 to SS and DARPA PA 19-03-01 and ONR MURI project N00014-16-1-2007 to TG.