Scalable pragmatic communication via self-supervision

Jennifer Hu (jennhu@mit.edu)[1], Roger Levy[1], Noga Zaslavsky[1,2]
[1]Department of Brain and Cognitive Sciences, MIT; [2]Center for Brains Minds and Machines, MIT

Introduction

- Pragmatic reasoning is an integral part of communication
- Rational Speech Act framework (RSA; [1]) has successfully modeled pragmatics in small settings
- Large-scale applications of RSA have relied on imitating human behavior in contextually grounded datasets
- **We propose a new approach to scalable pragmatics using self-supervised learning**, building upon results that characterize pragmatic reasoning in terms of general information-theoretic principles [2]

Background

Rational Speech Act model (RSA; [1])

Literal listener	$l_0(m\|u) \propto \mathcal{L}(u,m)P(m)$
Literal speaker	$s_0(u\|m) \propto \mathcal{L}(u,m)\exp(-\kappa(u))$
Pragmatic speaker	$s_t(u\|m) \propto \exp(\alpha(\log l_{t-1}(m\|u) - \kappa(u)))$
Pragmatic listener	$l_t(m\|u) \propto s_t(u\|m)P(m)$

- $\mathcal{L}(u,m)$: lexicon function (learned in our models)
- $\kappa(u)$: utterance cost function (estimated from Google Books n-gram frequencies)
- $\alpha \geq 0$: "rationality" parameter (tuned to fit data)
- $P(m)$: prior over meanings (uniform in our models)

New understanding of RSA [2]

- RSA implements an alternating-maximization (AM) algorithm for optimizing

$$\mathcal{G}_\alpha[s,l] = H_s(U|M) - \alpha\mathbb{E}_s[\log l(M|U) - \kappa(U)]$$

- Maximizing $\mathcal{G}_\alpha \approx$ least effort principle ⇒ call this **LE-RSA**
- With small adjustment, RSA can be grounded in Rate-Distortion (RD) theory
- Suggests that RSA is only one instance of more general model class, varying along two axes (Figure 1):

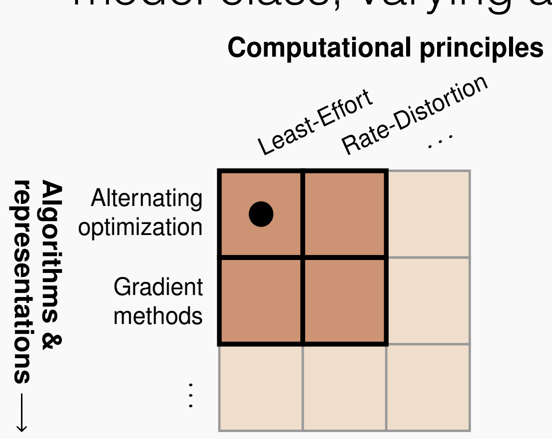

Figure 1

1. **Algorithms & representations:** we compare AM to gradient descent (GD), as GD is scalable and may enable generalization across domains
2. **Computational principles:** we focus on LE, but our models could easily be adapted for RD

Dataset & task

- Existing corpus of color reference games [3]
- Speaker and listener see context of 3 colors; speaker describes privately assigned target and listener clicks on inferred target
- Simplified the dataset to perform exact AM (18K training rounds)
  - Only kept rounds with single speaker message and took top 100 messages as space of utterances

	Context	Human	Base	SSL-AM	SSL-GD	SL
C_1	m^*	gray	purple	blue	gray	purple
C_2	m^*	purple	purple	red	purple	purple

Table 1: target m* nearly identical in both contexts, showing that human descriptions are sensitive to context

Models

Self-supervised learning (SSL) approach

Learn θ here: maximize likelihood of **decontextualized** data

θ → **Lexicon** $\mathcal{L}(u,m;\theta) \in (0,1)$ - - → Non-pragmatic

Base — Literal RSA agents, taking each context to be space of meanings

Pragmatic

SSL-AM
- Enrich base agents with RSA iterations
- Implements AM for optimizing intrinsic LE objective

SSL-GD
- Agents' pragmatic parameters updated for each context via GD on LE objective
- Lexicon kept frozen
- Pragmatic parameters can continually adapt

Achieve pragmatics with no additional data!

Supervised learning (SL) approach

θ → **Lexicon** $\mathcal{L}(u,m;\theta) \in (0,1)$ → **Base** → Pragmatic

SL
- AM form of RSA iterations baked into architecture, as in [4]
- Same base agents as in SSL approach

Learn θ here: maximize likelihood of **contextualized** data

Results

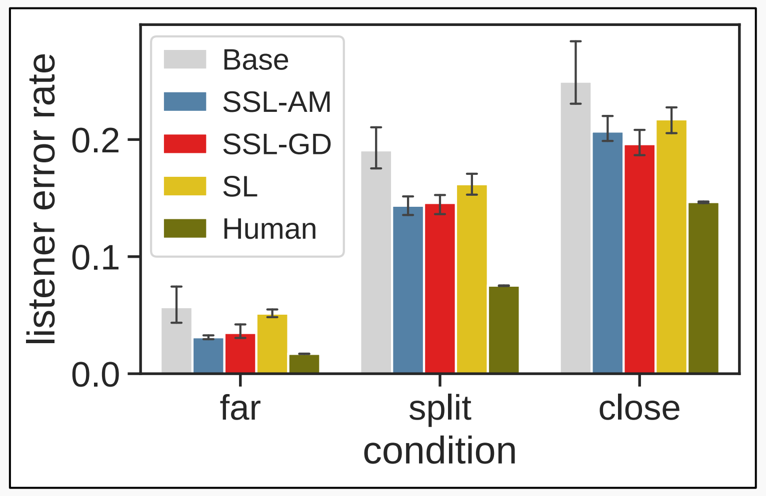

Figure 2: prop. of test set rounds where listener fails to select ground-truth target color, given human speaker utterance

- Pragmatic listeners improve upon base listener and achieve accuracy comparable to SOTA [3,4]
- No sig. difference between SSL-AM and SSL-GD
- SSL comparable to SL (possibly slightly better) on this simplified task, while never accessing contextualized data

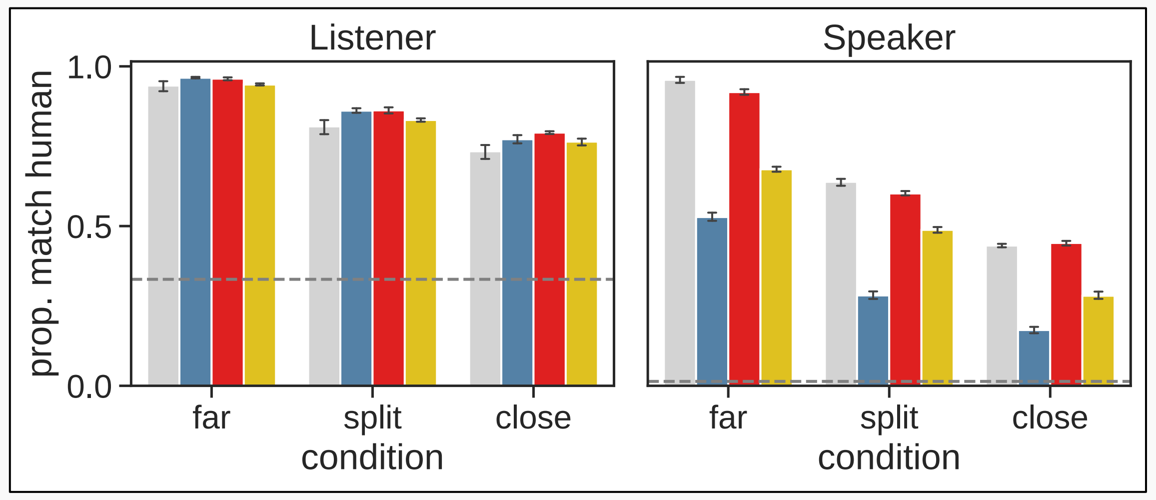

Figure 3: prop. of test set rounds where models match human behaviors

- Among pragmatic speakers, SSL-GD = best fit to human speaker
- Pragmatic speakers sig. decrease fit compared to base speaker
- SSL-AM may be exploiting gradedness of neural lexicon, resulting in pragmatic drift (see Table 1)

Discussion

- We proposed scalable self-supervised approach: learn pragmatic policies by optimizing agent-intrinsic objective instead of imitating human behavior
- **SSL-GD more data efficient than SL and more scalable than SSL-AM, while achieving similar performance**
- Future research:
  - Use SSL-GD to study how pragmatic knowledge might be shared across contexts and domains, by allowing agents' parameters to continuously adapt
  - Use non-contextualized datasets for training
  - Test more complex domains and architectures
- Our models execute a form of algorithmic computation [5] grounded in pragmatic theory and information theory

References. [1] Frank, M. & Goodman, N. (2012). Predicting pragmatic reasoning in language games. Science. [2] Zaslavsky, N., Hu, J., & Levy, R. (2020). A Rate-Distortion view of human pragmatic reasoning. [3] Monroe, W., Hawkins, R.X., Goodman, N. & Potts, C. (2017). Colors in context: A pragmatic neural model for grounded language understanding. TACL. [4] McDowell, B. & Goodman, N. (2019). Learning from Omission. ACL. [5] Veličković, P. & Blundell, C. (2021). Neural algorithmic reasoning. Patterns. **Funding.** JH: NSF GRFP. RL: NSF BCS-1551866 & BCS-1456081, Google Faculty Research Award, Elemental Cognition, MIT Quest for Intelligence. NZ: BCS Fellowship in Computation.