Overloaded Communication as Paternalistic Helping

Stephanie Stacy[1] stephaniestacy@g.ucla.edu
Aishni Parab[3] aishni@cs.ucla.edu
Max Kleiman-Weiner[4] maxhkw@gmail.com
Tao Gao[1,2] tao.gao@stat.ucla.edu
[1] Department of Statistics, UCLA [2] Department of Communication, UCLA
[3] Department of Computer Science, UCLA [4] Brain and Cognitive Sciences, MIT

Abstract

Even simple, ambiguous signals can have a rich interpretation when viewed in the context of an interaction in a shared environment. We create a model called Paternalistic Communication by combining an existing model of overloaded language – Rational Speech Acts (RSA) – with a full agent model of Theory of Mind (ToM). This integration allows signals to be processed in conjunction with common ground in a principled manner dependent on task-dependent action utilities. This modeling perspective treats communication as a way to coordinate diverging perspectives in a cooperative setting. Under Paternalistic Communication, a speaker decides what to say by predicting their partner's reaction based on the information in common ground and then evaluates those reactions using their own mind which may contain additional information. We demonstrate the flexibility and performance of Paternalistic Communication in a case study with ambiguous signaling through a set of simulations.

Keywords: communication; common ground; Theory of Mind; Bayesian inference; pragmatics

Introduction

You're walking with a friend in winter when your friend yells "careful!" You look down and observe a patch of black ice underfoot. Without context, "careful!" can mean countless things; however, in context, this sets off a rich inferential process: What is your friend referring to? (the ice) How should this knowledge change your beliefs? (the ground will be slippery) How should this change your actions? (falling hurts, so tread carefully). Traditional models of communication assume that words and their meaning have a one-to-one mapping predefined outside of the current exchange (Shannon, 1948; Valiant, 1984) which would fail at explaining this example. However, instead of an encoding and decoding process, human communication is highly dependent on understanding what is relevant in the current context (Sperber & Wilson, 1986) allowing us to be incredibly successful at expressing rich meaning using sparse, overloaded signals. In this work, we propose a model of signaling that targets how the context of the situation can help solve signal ambiguity.

The example above highlights communication as a cooperative tool for helping, yet communication is a unique type of helping for two reasons. First, it is not the same as instrumental helping because instead of taking actions that change the world, communicators send signals to change the mind. Second, communication requires coordination of minds. Communicators simultaneously track what is shared in the common ground and what is private (Heller, Parisien, & Stevenson, 2016), which requires agents to coordinate their divergent minds. To achieve this, we turn to a previously studied phenomenon: paternalistic helping (Martin, Lin, & Olson, 2016). In the following sections, we introduce a set of components that allow us to build a flexible model of communication using the principle of paternalistic helping.

Overloaded communication in a visual scene has been empirically studied in psychology (J. B. Misyak, Melkonyan, Zeitoun, & Chater, 2014). We model the task in one notable study in which cooperators use and understand overloaded signals in the form of tokens, which can either mean "open" or "avoid", to collect bananas (rewards) and avoid scorpions (punishments) hidden in boxes (J. Misyak, Noguchi, & Chater, 2016). We show that partners who have never interacted with each other before can successfully use ambiguous signals by forming instantaneous conventions that change flexibly, depending on context.

Background

Common Ground

Common ground is mutually shared, public knowledge assumed between communicators. While common ground is theoretically established through infinite recursion (Lewis, 1969), in practice communicators likely assume some common knowledge (Clark & Marshall, 1981). Communication can be viewed as a mechanism to add information to this common ground. In turn this narrows the scope of reasonable signal interpretations (Clark & Brennan, 1991; Clark & Marshall, 1981) making communication more efficient: brief, indirect, and instantaneous. Even pre-linguistic infants use common ground to resolve ambiguity in communication, namely through pointing (Liebal, Behne, Carpenter, & Tomasello, 2009; Liszkowski, Schäfer, Carpenter, & Tomasello, 2009). Knowledge already in the common ground does not need to be discussed, allowing brevity and increased clarity. Thus, a simple "careful," achieves the same effect as a much longer signal. Finally, common ground allows for instantaneous interpretation without requiring a history of interaction because it can be derived from the environment itself (Clark, 1996; Tomasello, 2010). We focus on this type of instantaneously formed common ground which builds on an intuitive understanding of others' minds (Wellman, 1992).

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

Flexible Linguistic Pragmatics

Using pragmatics to consider a signal's context is key to understanding what someone means. The Rational Speech Act framework (RSA) models how to understand a signal in the context of what else a speaker could have said but chose not to (Frank & Goodman, 2012; Goodman & Frank, 2016). Here we describe an extension of this framework that considers additional speaker context and has been proposed in order to capture signaler types or affects (Goodman & Frank, 2016; Kao, Wu, Bergen, & Goodman, 2014).

First, a pragmatic speaker *ps* chooses a *signal* from a set of possible signals and signaling type c to describe a target referent or world state w. Signals are treated as a type of rational action, subject to a utility soft-maximization, where $\beta \in [0, \infty)$ represents the degree of rationality (Luce, 1959). Here, the utility of a signal can be calculated by reasoning how a pragmatic listener (*pl*) will interpret that signal:

$$P_{ps}(signal, c|w) \propto e^{\beta P_{pl}(w,c|signal)} \tag{1}$$

The pragmatic listener models signal interpretation using Bayesian inference, which requires a simple generative speaker model likelihood. A literal speaker *ls* provides an entering point to the recursive reasoning that speakers and listeners can engage in *ad infinitum* when communicating. The literal speaker is defined by uniformly sending true signals according to an indicator function of whether a signal is consistent with the referent state w given the speaker type c. The prior term is defined over both speaker type and state, which are assumed to be independent.

$$P_{pl}(w, c|signal) \propto P_{ls}(signal|w, c)P(w)P(c) \tag{2}$$

RSA is grounded in cooperative logic from linguistic theory which proposes that communicators should choose maximally efficient and straightforward signals (Grice, 1975). While some recent work has begun to develop in the direction of adding action context (Sumers, Hawkins, Ho, & Griffiths, 2021) or grounding pragmatic signals within a utility-driven task (McCarthy, Hawkins, Wang, Holdaway, & Fan, 2021), RSA has primarily been used in purely linguistic settings. In these cases, speakers have the communicative goal of describing a referent by reasoning about how different signals are expected change the listener's beliefs, but in reality communication can be several steps more indirect than this; we communicate about not only referent states (What?) but also social motivations (Why?) and interactions in the shared physical environment that can achieve those motivations (How?).

Tying Signals to Actions: Bayesian Theory of Mind

Theory of Mind (ToM) posits that when deciding how to act, one should rationally take actions that achieve desirable utilities with respect to their underlying mind which contains beliefs, desires, and intentions:

$$P(action|mind) \propto e^{\beta \mathbb{E}[U(action, mind)]} \tag{3}$$

ToM has been successful in a variety of action interpretation tasks (Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; Kleiman-Weiner, Ho, Austerweil, Littman, & Tenenbaum, 2016) where an observer uses Bayesian inverse planning to infer the likely mental states that generate observed actions:

$$P(mind|action) \propto P(action|mind)P(mind) \tag{4}$$

ToM and RSA are both models of rational decisions with respect to a utility maximization, but where do these utilities come from and why do they matter? We use ToM's formulation of agency to connect a signal back to its action utility under the task since, unlike instrumental actions, signals do not directly change the world. We follow the tradition of casting communication as a planing problem driven by a task-based utility maximization, often seen in artificial intelligence (AI) works (Russell, 2019). AI modeling work has shown that grounding communicative interactions in action consequences can tie the value of a signal to the value of expected outcome actions (Gmytrasiewicz & Durfee, 2001; Gmytrasiewicz & Doshi, 2005). These approaches are promising in formalizing signal utility but assume a fixed one-to-one mapping between the signal and meaning. To move beyond codebook mode of communication, we integrate the RSA linguistic pragmatics with a task-oriented definition of signal utility.

A Paternalistic Perspective on Communication

To understand how communication serves to coordinate minds, it is useful to view cooperative communication as a type of paternalistic helping. Here a speaker understands and predicts a listener's actions according to shared knowledge but evaluates them according to private knowledge. For example, parents often make decisions for their children "for their own good," regardless of the child's preferences. A paternalistic perspective has been successful in modeling how to interpret helpful pointing under ambiguity (Jiang et al., 2021).

Paternalistic helping acts as a binding agent between common ground, RSA, ToM, and signal utilities derived from actions. A pragmatic paternalistic signaler chooses what to say by evaluating the utility of different signals, equivalent to the pragmatic RSA speaker (Equation 1). However, instead of deriving utility directly from the listener's beliefs $P_{pl}(w, c|signal)$, we replace this with a more general utility function grounded in task-specified actions stated below. The speaker creates an expectation of how good a signal is by predicting how a receiver will *act* upon hearing the signal $P(a|signal)$ using public, common ground information $mind_{cg}$ and evaluates how good that action is $U(a, mind)$ using private knowledge within their own mind:

$$\mathbb{E}[U(signal, mind)] = \mathbb{E}_{P(a|signal,c)}[U(a, mind)] \tag{5}$$

There are two terms connecting signals to actions. First, $P(mind_{cg}|signal, c)$ can be derived from inverse planning in ToM where signals are treated as a type of rational action

(Equation 4) and is similar to modeling a RSA listener (Equation 2), but is capable of reasoning more generically over other components of the mind. Second, $P(a|mind_{cg})$, which can be derived from ToM rational action planning (Equation 3).

$$P(a|signal,c) = \sum_{mind_{cg}} P(mind_{cg}|signal,c)P(a|mind_{cg}) \quad (6)$$

Here, the signal and speaker type are assumed to be independent from other components of the mind. The integration of common ground, ToM and RSA under the paradigm of paternalistic helping gives a flexible, context driven approach to overloaded communication.

Case Study Modeling

We demonstrate the power of Paternalistic Communication (PaCo) by modeling a case study with impromptu, overloaded signaling: Misyak, Noguchi, Chater (2016). Through a non-linguistic cooperative communication task, the authors empirically demonstrate that humans coordinate to form instantaneous conventions using contextual cues from the common ground, even when a signal can mean opposite things ("go to" or "avoid" a location). We provide a computational account of these behaviors as a special case of ambiguous communication captured by PaCo and compare it to a baseline model: the version of RSA pragmatics adapted for speaker type. Context in PaCo includes both the world features and how an agent can act based on the knowledge of those features; whereas, RSA is only able to consider world features.

Task

During each trial, participants saw three boxes, each containing either a banana (a reward) or a scorpion (punishment). The goal of this task was to open as many boxes with bananas as possible while avoiding boxes with scorpions. The signaler had full information about the contents in the boxes but could not open them. The receiver has no information about the contents in the box but could use axes to open the boxes. At each trial, the signaler had some number of tokens to mark boxes with to provide information to the receiver, and the receiver had some number of axes to open boxes. Both individuals knew how many tokens and axes were available (see Figure 1). Extra information about the total number of bananas and scorpions was either shown in the common ground or occluded by a wall to hide that information (not shown in figure). Four key conditions highlighted how humans flexibly convey meaning across context: Two Token, Inversion, One Ax, and Wall, summarized in table 1. Our analysis focuses on a version of this game with one-shot interactions between partners that represent instantaneously formed conventions without learning and rapport building over a sequence of plays (see Experiment 2 (J. Misyak et al., 2016) for details).

This experiment emphasizes the importance of common ground as context to solve ambiguous communication. It

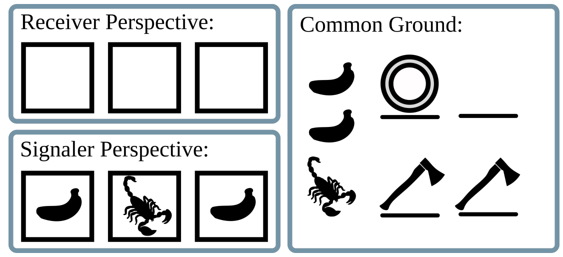

Figure 1: Schematic of Inversion condition setup for experiment. Information is split into agent specific knowledge and shared common ground.

Condition	Tokens	Axes	Wall Present
Two Token	2	2	False
Inversion	1	2	False
One Ax	1	1	False
Wall	1	2	True

Table 1: Experimental conditions in (J. Misyak et al., 2016)

fully supports a ToM model with beliefs about the possible contents of boxes and desires to collect reward and avoid punishments for utility maximization which compose $mind_{cg}$, and axes that can be used to open boxes which define possible actions $a \in A$. In this context, $mind_{cg}$ is effectively equivalent to w in Equation 1, but can naturally generalize to include uncertainty in joint desires. The number of tokens defines the space of all possible signals, while the number of axes, rewards, and signals define a prior over the shared content of the common ground. Communication in the task is always fully overloaded because placing a token on a box can have two opposite interpretations: "go there" or "avoid that," depending on the speaker type c. Thus, disambiguation occurs on a trial-by-trial basis as receivers flexibly and jointly infer the tokens' meaning and, as a direct consequence, the boxes' contents.

Simulation 1: Capturing Human-like Use of the Same Signal for Opposite Meanings

While RSA can use signals flexibly to maximally resolve beliefs about the world state this may not always be the optimal communication strategy: a fact which humans are sensitive to. For example, in the Inverse and One Ax conditions the world state and available signals remain the same. However, in the One Ax condition, because participants can get at most one reward, extra information about the second reward is extraneous. Humans use tokens to denote punishments in the Inverse condition (providing maximal information about the world) but in the One Ax condition where they could provide maximal information by marking punishments, they tend to use their token to mark a reward (providing an action directive). Signaling in opposite ways in these two conditions relies on the signaler's expectation formed through ToM action prediction that the receiver will act differently based on the rational integration of beliefs and the available actions. We predict that PaCo will robustly capture human-like flexible use of tokens in these conditions as well as the other key conditions tested in the original study.

Methods

Task Specification To translate the task's goal into an explicit utility calculation, we assign a positive value (+1) for each banana and a negative cost (-1) for a scorpion. Unlike traditional RSA, this cost ratio could natural vary using PaCo; however, this is not a factor considered in the original behavioral experiment, thus we choose a fixed constant where the benefit of choosing a banana is equivalent to the cost of choosing the scorpion. There is no explicit cost of using more tokens, if available, following the original study. PaCo and RSA can both be characterized by two free parameters: β and $P(open)$. β offers an estimation of how rational an agent is; we assume partners are equally rational. $P(open)$ represents the prior distribution over signaler type. We focus on the two types primarily employed by humans: $c \in \{avoid, open\}$. An open-type signaler may only place tokens on bananas while an avoid-type signaler may only place tokens on scorpions. The prior over beliefs $p(mind_{cg})$ is uniformly split across all possible assignments of bananas and scorpions; when there is no wall, all assignments inconsistent with the common ground beliefs are given 0 probability. To test the robustness of the models, we compare model predictions of how the signaler will act under a wide range of parameter combinations (β = [1, 2, ..., 17], $P(open)$ = [.4, .425, ..., .675, .7]). For each combination, we let the two models play the same task as seen by humans in the original experiment.

Descriptive Statistics An averaged root-mean-squared-error ($\overline{RMSE}$) quantifies how closely the model approximates human signal generation, where a smaller $\overline{RMSE}$ indicates better agreement between human and model. For a particular condition, we first categorize behavior into the two strategies a signaler could employ and take the RMSE between model x and human $x*$ distribution. Then, across the four conditions, these RMSEs are averaged to get obtain $\overline{RMSE}$:

$$\overline{RMSE} = \frac{1}{4} \sum_{m \in Condition} \left(\sqrt{\frac{1}{2} \sum_{x \in open, avoid} (x_m - x_m^*)^2} \right) \quad (7)$$

Results

To understand how robust each model is to changes in hyperparameters, we calculate the $\overline{RMSE}$ across the grid of β and meaning priors for each model, summarized in Figure 2.To compare overall tolerance to parameter changes between the two models, we conducted a one-sided Wilcoxon signed-rank test for matched-pairs. Under equivalent conditions, the median error under PaCo is significantly smaller than RSA (W = 630, $p = 1.1 \times 10^{-34}$). This supports PaCo's robustness across a wide range of parameters, and suggests that these properties are not the product of over-fitting human data, but rather, a specific example of a general class of phenomena a paternalistic perspective is capable of handling.

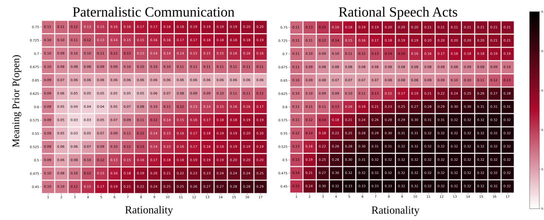

Figure 2: PaCo and RSA heatmaps of $\overline{RMSE}$ for key trials: The $\overline{RMSE}$ for each model and parameter combination is represented as a color intensity in the heatmap. Lighter colors represent better agreement between human and model.

Beyond the overall fit, we look at specific strategies employed in the four key conditions, paying specific attention to difference between Inversion and One Ax, where humans tend to change their strategy between conditions. To do this, we select the parameter set that best approximates human strategies in terms of error minimization for each model (PaCo: β = 3, P(open) = .575 results in $\overline{RMSE} = 2.94 \times 10^{-2}$, RSA: β = 5, $P(open)$ = .65 results in $\overline{RMSE} = 7.05 \times 10^{-2}$).

Like humans, PaCo is sensitive to the common ground: how many signals and axes were available and the presence/absence of the wall, and instantaneously changes which strategy is dominant between the Inversion and One Ax conditions (Human $P(open)$: Inv =.42, One Ax = .63; PaCo: Inv = .47, One Ax =.57). In contrast, RSA fails to make this strategy switch or even distinguish between these conditions (Inversion = One Ax = .58). Next, we explore these phenomena through simulation results beyond the original study.

Simulation 2: Understanding the Effects of Action Driven Utility

PaCo and RSA behave differently at capturing human signaling flexibility: PaCo derives its utility from how desirable action outcomes under the task are expected to be whereas RSA focuses on minimizing the uncertainty in a listener's beliefs. We explain these differences by dividing the context into two separate sources of uncertainty within the common ground: world space and action space. We expect RSA to be sensitive only to the world space knowledge, whereas PaCo's performance should depend on whether considering the receiver's action space can act as a constraint on signaling. Specifically, we expect PaCo's action-based reasoning to become more important in cases where the world state is highly uncertain.

Methods

Task Specification We use the same task utility structure as before, adding a small cost (-.1) per token used to encourage shorter signals. In addition, to look at how performance varies across scaled-up environments, we expand the world to have five boxes. Token meaning priors are set at the optimal ones that match human performance in Simulation 1, and the models are set to high rationality ($\beta = 20$) to emphasize theoretical performance. The number of axes are manipulated (1, 2, 3, 4), with and without a wall. The number of tokens are set to be high (3, 4) which ensures that a signaler has the means to send a longer signal if desired. Similarly, the number of rewards are set to be high (3, 4) which ensures the possibility of achieving a high utility. We sample N=250 environments for each combination of wall and number of axes.

Descriptive Statistics To test how PaCo and RSA communicated using different strategies, we used the Kullback-Leibler (KL) divergence between P, the true belief that the signaler privately knows and Q, the receiver's belief posterior about the box contents $mind_{cg}$ to describe the uncertainty over the set of possible beliefs M. Because the receiver's posterior is highly dependent on which signal they observe, the expectation accounts for the signaler's probability of sending each signal given the true world:

$$\mathbb{E}[KL(P||Q)] = \mathbb{E}[\sum_{mind_{cg} \in M} P(mind_{cg}) \log \frac{P(mind_{cg})}{Q(mind_{cg})}] \quad (8)$$

A larger KL divergence occurs when the receiver is uncertain about the true state of the world, here, contents of boxes.

Results

When there is no wall, both models achieve the upper bound of possible performance. Consistent with our hypothesis, these models make different predictions when there is higher uncertainty in the world from adding a wall. When the wall is added, performance drops for both models; however, multiple comparison tests show that PaCo outperforms RSA at each level of ax (all $p_{adj} < .05$ under Tukey's HSD) except when there are four axes ($p_{adj} = .074$) (see Fig. 3). When the receiver has four axes, there are no constraints on the action space and thus, considering actions is not able to restrict signaling behavior. Because PaCo cooperators take into account the receiver's action space, a less capable agent requires less information to do its best making PaCo predict that it is sometimes better to tell their partner exactly how to act.

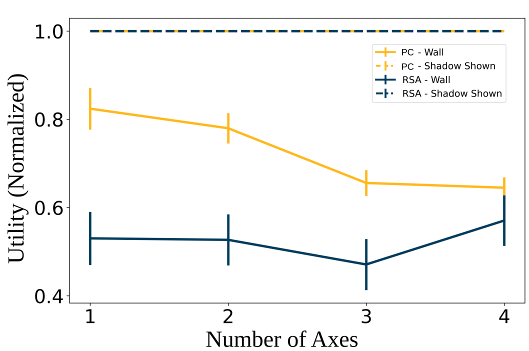

Figure 3: Utility achieved as a function of axes for RSA and PaCo with 95% CI. Dashed lines represent when there is no wall and solid lines represent cases when there is a wall.

Even more striking, PaCo uses fewer tokens than RSA to achieve a higher task utility under high uncertainty (Fig. 4). Constraints of the action space help reduce the space of reasonable signals; however, when the wall is absent, PaCo uses more tokens than RSA, seemingly *over*-informing. When PaCo judges their partner as capable, it prefers a longer, more cautious signal to ensure clarity even when a shorter signal can be understood with high probability.

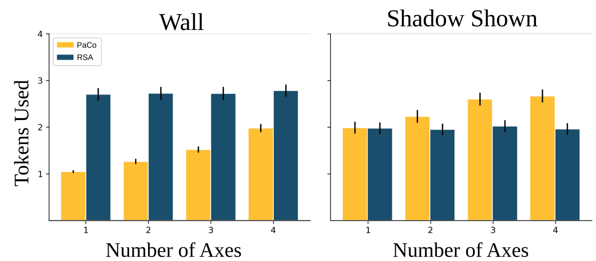

Figure 4: Proportion of tokens used by PaCo and RSA given the available receiver actions for cases with a wall (left) and when the shadow is present (right).

By definition, RSA always aims to provide the most informative message, whereas PaCo's action driven utility sends a task outcome oriented one. From Fig. 5, we see this clearly in the breakdown of model KL divergences. When the shadow is shown, both models always have virtually 0 divergence, indicating that the signal can fully resolve the state of the world. However, in the wall condition, higher uncertainty leads to a different pattern of results. RSA achieves a much smaller KL divergence than PaCo, indicating that RSA agents are likely to have a better understanding of the true world state, but that this alone is not enough to succeed at the task.

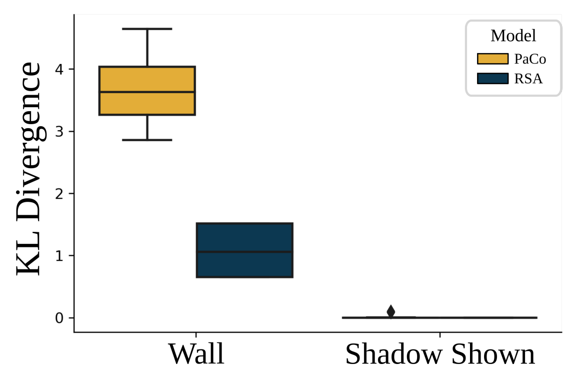

Figure 5: Distribution of KL divergence between receiver belief posterior and true world state for each sampled environment. True world belief distribution is adjusted to give incorrect world states a negligible, non-zero (10^{-6}) weight.

Simulation 3: Generalized Performance

Coordinating minds and maximizing a utility calculus are two modeling pillars in PaCo. Given this, we investigate to what extent adding recursion improves performance for PaCo and RSA. We also use this task to examine whether PaCo's advantage generalizes beyond the specific conditions from the previous simulations to cases on a larger scale and with few constraints in the environment. We expect PaCo's coordination of minds will reduce the need for deep recursion and allow generalization across a larger variety of settings.

Methods

Task Specification For both PaCo and RSA, we include a simple partner where the signaler simulates a receiver, but the receiver does not model the signaler (naive). We compare this to models with an added level of recursion to the receiver to see how having a partner model can change performance (pragmatic). To scale up reasoning in the environment, we look at environments with three to six boxes and remove all free parameters from the models, focusing on how well the models can perform in general settings without any prior biases. To measure the best possible performance under uncertainty, partners greedily select the action or signal with the maximum expected utility. We put a uniform prior over a token's meaning and remove all signaling costs. We then uniformly sample from the space of possible worlds all possible worlds with three to six boxes which have at least one scorpion and one banana. The number of axes and tokens are sampled independently such that there is at least one and at most $n-1$ for each, given a world with n boxes. The presence of a wall is also sampled as a binary variable. Each model at each reasoning level has a total of 3000 simulated trials.

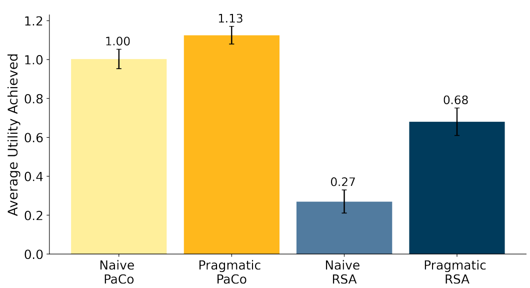

Figure 6: Average achieved utility by model (PaCo/RSA) and pragmatics (yes/no). Labels denote means.

Results

A factorial ANOVA (2 x 2) indicates a significant interaction between receiver pragmatics (Yes/No) and model: that is, the contribution of pragmatics on utility is different depending on whether you're using PaCo or RSA ($p < .05$). Moreover, we see that even the version of PaCo without pragmatics is consistently outperforming RSA with pragmatics. Post Hoc analysis using Tukey's HSD Test for multiple comparisons finds this difference in performance to be significant (p = 0.001, 95% C.I. = [-0.4295, -0.2165]). That is, the simple version of PaCo is still able to outperform RSA.

When comparing PaCo and RSA with and without recursion, we see that even the version of PaCo without recursion consistently outperforms RSA with pragmatics (Fig. 6). These results indicate that PaCo's success does not necessarily rely heavily on deep recursion. Instead sensitivity to other task-related information may shift some of the burden off complex reasoning. Here, PaCo's flexibility in conveying information about actions and not just beliefs about the environment allow it to outperform RSA, especially in the absence of common ground information.

Both RSA and PaCo receiver models benefit from adding pragmatics; however, this benefit seems especially large for RSA. Even without receiver pragmatics, PaCo is able to outperform the equivalent RSA model and the more complex RSA model. This suggests that in place of complex pragmatics, more flexible processing of the mind and actions seen in PaCo can stand in without a cost to performance.

Discussion

PaCo builds upon RSA's pragmatic reasoning framework by integrating infrastructure from cognitive science to take task-driven action context into account. This provides a holistic view of the interplay between common ground, the mind, and the shared environment which allows communicators to reason beyond beliefs. PaCo also uses predicted actions to determine the value of a signal, allowing us to argue for communication as a way to align cooperators' minds. Through modeling a case study, we highlighted (1) the importance of treating common ground as a multi-faceted constraint to signaling, which requires treating partners as rational and capable of achieving things in the world and (2) the benefit of framing communication as a means to coordinate perspectives, which highlights how different components of cooperators' minds interact to reduce reliance on deep social recursion.

In this task, restrictions on world beliefs and available actions led to different human behaviors contributing uncertainty to the common ground. While both models switch between signaling strategies, RSA selects a signal based on informativeness whereas PaCo considers action consequences and underlying beliefs in conjunction. Using only expected outcome utility, PaCo naturally switched between sending signals that were maximally informative and signals that were imperative. This behavior was supported in Simulation 2 through token usage and KL divergence in the wall condition. These results can motivate future behavioral study exploring this phenomenon in humans.

Moreover, achieved task utility in Simulation 2 established the theoretical improvement of PaCo's action-driven model which generalized in this task, as demonstrated by Simulation 3. PaCo reached a higher asymptotic performance under maximal rationality across different sized environments without relying on informative signal meaning priors or costs. In addition, while increased recursion could improve performance within a model, even a shallow PaCo model outperformed a recursive RSA one as seen in Simulation 3. Principled use of common ground information ultimately allowed PaCo signalers to use the same signal in opposite ways in Simulation 1 to flexibly and robustly capture human behavioral data and additionally shifts some of the inferential burden of deep recursion to other heuristics such as utility.

Acknowledgments

We thank our anonymous reviewers for their thoughtful feedback. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650604 to SS and by DARPA PA 19-03-01 and ONR MURI project N00014-16-1-2007 to TG.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*, 1–10.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10 – 63). Cambridge: Cambridge University Press.

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

Gmytrasiewicz, P. J., & Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *Journal of Artificial Intelligence Research*, *24*, 49–79.

Gmytrasiewicz, P. J., & Durfee, E. H. (2001). Rational communication in multi-agent environments. *Autonomous Agents and Multi-Agent Systems*, *4*(3), 233–272.

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 & M. J (Eds.), *Syntax and sematics* (Vol. 3, p. 41 - 58). New York: Academic Press.

Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition*, *149*, 104–120.

Jiang, K., Stacy, S., Chan, A., Wei, C., Rossano, F., Zhu, Y., & Gao, T. (2021). Individual vs. joint perception: A pragmatic model of pointing as smithian helping. In *Proceedings of the 43rd annual conference of the cognitive science society.*

Kao, J. T., 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.

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 *Proceedings of the 38th annual conference of the cognitive science society.*

Lewis, D. K. (1969). Convention: A philosophical study.

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, *12*(2), 264–271.

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science*, *20*(5), 654–660.

Luce, R. D. (1959). On the possible psychophysical laws. *Psychological Review*, *66*, 81–95.

Martin, A., Lin, K., & Olson, K. R. (2016). What you want versus what's good for you: Paternalistic motivation in children's helping behavior. *Child Development*, *87*(6), 1739–1746.

McCarthy, W. P., Hawkins, R., Wang, H., Holdaway, C., & Fan, J. E. (2021). Learning to communicate about shared procedural abstractions. In *Proceedings of the 43rd annual meeting of the cognitive science society.*

Misyak, J., Noguchi, T., & Chater, N. (2016). Instantaneous conventions: The emergence of flexible communicative signals. *Psychological Science*, *27*(12), 1550–1561.

Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: Virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, *18*(10), 512–519.

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Penguin.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (Vol. 142). Cambridge: Harvard University Press.

Sumers, T., Hawkins, R., Ho, M., & Griffiths, T. (2021). Extending rational models of communication from beliefs to actions. In *Proceedings of the 43rd annual conference of the cognitive science society.*

Tomasello, M. (2010). *Origins of human communication*. MIT Press.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, *27*(11), 1134–1142.

Wellman, H. M. (1992). *The child's theory of mind*. The MIT Press.