A Theory of Mind Approach to Nonverbal Communication

From the Pespective of Utility

Zhancun Mu
YuanPei College
Peking University
yhbylch@stu.pku.edu.cn

Abstract

Nonverbal communication, including pointing, gazing, and gesturing, is an essential part of human interaction. Unlike verbal communication, nonverbal cues rely more heavily on pragmatic reasoning. The Rational Speech Act (RSA) framework is one influential approach to modeling pragmatics, but on its own it is insufficient for nonverbal communication. In this essay, we propose integrating Theory of Mind (ToM) into the RSA framework to enable rational agents for nonverbal communication.[1]

1 Introduction

Nonverbal communication plays a major role in human interaction. Humans express and infer mental states through expressions, gestures, and other nonverbal cues. Effectively communicating nonverbally requires a deep understanding of social dynamics, which remains challenging for current AI systems.

To begin, we must identify the key requirements for nonverbal communication. A major distinction from verbal communication is that nonverbal cues rely more heavily on situational and contextual information. To elaborate, semantics refers to the literal meaning of words, while pragmatics involves interpreting meaning in context [6]. Because nonverbal signals lack rich semantics, understanding them requires leveraging situational cues and pragmatic reasoning to effectively convey or comprehend nonverbal messages.

The Rational Speech Act (RSA) framework is an influential approach for making quantitative predictions about pragmatic inference [3]. It models communication as a recursive reasoning process between speakers and listeners. A core assumption is that humans are rational agents who try to maximize informativeness. This essay will first introduce RSA for verbal communication. However, RSA alone is insufficient for nonverbal cues. We propose integrating Theory of Mind (ToM) [1] while retaining the concept of utility maximization. This extended framework may better capture nonverbal communication. We will demonstrate it on a problem requiring utility reasoning, serving as a benchmark for testing agents' nonverbal communication skills.

2 Rational Speech Act Model

2.1 Introduction to RSA

Although this essay focuses on nonverbal communication, we will first introduce a basic RSA framework for verbal communication using the simple example in Fig. 1. The set of possible world states is $S = \{\text{blue-square}, \text{blue-ball}, \text{green-square}\}$ and the set of possible utterances is

[1]This essay serves as an early draft to outline my idea, which requires further research.

$U = \{\text{square}, \text{ball}, \text{blue}, \text{green}\}$. In this reference game, the speaker aims to communicate the world state (the object) to the listener.

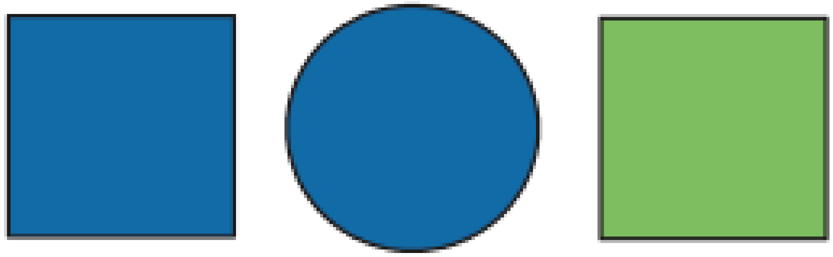

Figure 1: A classic referential task from [3].

The RSA framework comprises three levels: a pragmatic speaker, a pragmatic listener, and a literal listener. The pragmatic speaker selects the optimal utterance for the literal listener, who interprets the utterance literally and identifies the most compatible object. The pragmatic listener reasons about the pragmatic speaker's reasoning process to interpret appropriately. Formally:

$$\begin{cases} P_{L_1}(s|u) \propto P_{S_1}(u|s)P(s) & \text{pragmatic listener} \\ P_{S_1}(u|s) \propto \exp(\alpha U_{S_1}(u;s)) & \text{pragmatic listener} \\ P_{L_0}(s|u) \propto u(s)P(s) & \text{literal listener} \end{cases}$$

- **The literal listener** L_0 naively interprets an utterance according to its meaning which means computing $P_{L_0}(s|u)$ directly by combining the semantic and prior.
- **The pragmatic speaker** S_1 is (almost by α) rational, which means he chooses the utterance by its utility.
  - The speaker chooses the utterances to better communicate with the (hypothesized) literal listener.
  - Frank and Goodman [3] choose the utility function to be both efficient and informative, *i.e.*
    $$U_{S_1}(u;s) = \log P_{L_0}(s|u) - C(u),$$
    We will have a discussion on it later.
- **The pragmatic listener** L_1 does Bayesian inference on the speaker's reasoning.

2.2 Problems of RSA

Although [3] characterize RSA as recursive reasoning, it is more accurately depicted as a hierarchical model. Specifically, the pragmatic speaker can only model the literal listener, not the full pragmatic listener. Additionally, real communication involves bidirectional reasoning, whereas RSA only captures unidirectional processes. Finally, for nonverbal cues, the literal listener is a poor model because nonverbal semantics are underspecified. To address these limitations, an effective communication policy must capture bidirectional belief modeling, relying on Theory of Mind (ToM).

However, the utility maximization concept remains relevant. For instance, when pointing, a naive agent may continue pointing redundantly to achieve a goal, whereas humans exhibit more discerning behavior. Prior work adds terms like "social costs" to penalize such actions [5]. However, directly optimizing informativeness through utility functions is more intuitive than imposing social costs, since effective communication should convey information.

In the following section, we will extend the RSA framework with ToM, demonstrated on a problem requiring utility reasoning in communication. This will serve as a useful benchmark for evaluating agents' understanding of utility for communication.

3 Modelling Theory of Mind in communication

We illustrate the problem below and an example can be seen in Fig. 2.

Problem. *A and B want to go to a restaurant. There exists two restaurant X,Y. A knows which restaurant (goal g) it is going to while B does not. The goal of A is to lead B to the right restaurant. They take turns to move. The quality of their communication policy can be evaluated by the result and their walking distance.*

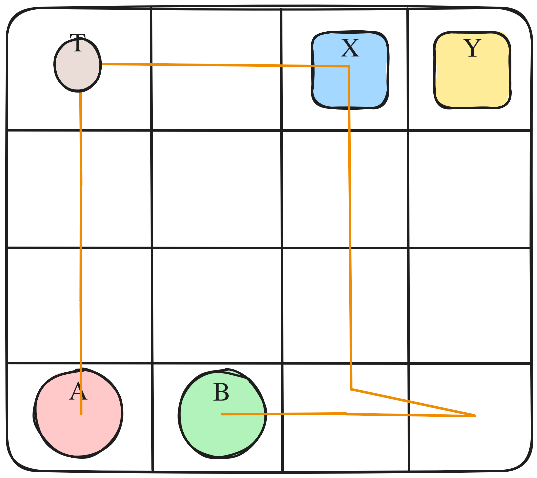

Figure 2: An example of the problem. The orange line is an example of the communication result between A and B. After turn 3, B figures out the goal and directly goes to X while A knows B knows the goal.

This problem is complex because sender A must model receiver B's beliefs about the goal, while B must infer the real goal and signal confusion if off track. Effective communication demands balancing efficacy and efficiency. Overly effective behavior can reduce efficiency. Thus, well-performing agents need both ToM capabilities to model others' perspectives and utility optimization to weigh each action's informativeness. In the following part, we will develop a policy to address this problem.

We can model the problem as a Markov Decision Process (MDP) [7], in which s denotes the state, a is the action, g is the real goal, $b_A = P_A(g_B)$ is A's belief of B's goal and $b_B = P_B(g_A)$ is B's belief of the real goal (A's goal). For a rational agent, $g_B = \arg\max P_B(g_A)$. Inspired by the five minds theory by Fan et al. [2] ,we also introduce $b_{AB} = P_A(b_B)$ and $b_{BA} = P_B(b_A)$ as the second-order estimation.

Using Bayesian ToM (BTOM), we can infer belief by:

$$P(b|h) \propto P(b)P(h|b),$$

where h is the past history, h_t represents $\{s_1, a_A^1, s_2, a_B^2, \cdots\}$.

We denotes the Q-value function as $Q(s, g_A, g_B)$, then:

$$Q(b_A, s, a) = \sum_{g_B} P_A(g_B)Q(s, g, g_B),$$

and

$$Q(b_B, s, a) = \sum_{g_A} P_B(g_A)Q(s, g_A, g_A).$$

Adapted from RSA, the policy of (almost) rational agents is:

$$P(a|s, b) \propto \exp(\alpha U(b, s, a)).$$

We estimate utility in a counterfactual way: the utility function can be defined as :

$$U_i^{t-1}(a) = \sum_{g_{-i}} [P_{i,-i}(g_{-i}|h_t) - P_i(g_{-i})]Q(s, [g_i, g_{-i}]).$$

$P_{i,-i}(g_{-i}|h_t)$ is "how i's belief will change, predicting $-i$'s belief will change due to observation of h". We can estimate it using the second-order belief:

$$P_{i,-i}(g_{-i}|h_t) = \sum_{b_{-i} \sim b_{i,-i}} P_i(b_{-i}|b_{i,-i}) \sum_{\hat{g}_{-i}} P_{-i}(\hat{g}_{-i}) \sum_{a_{-i}} P(a_{-i}|\hat{g}_{-i}) P_i(g_{-i}|[h_t, a_{-i}]).$$

4 Conclusion

In conclusion, we propose modifying the RSA framework by incorporating Theory of Mind (ToM). This extended framework can enable developing policies for nonverbal communication across varying contexts. Related work by Lee et al. [5] develops a policy for senders, but their receiver's policy does not generalize to settings with changing communicative goals. Jiang et al. [4] build a model considering relevance, but without accounting for second-order belief modeling. Our framework aims to provide a strong foundation for designing policies that can flexibly and effectively communicate nonverbally across diverse settings. Further research should focus on implementing this framework in realistic multi-agent environments and comparing its performance to prior methods.

Acknowledgement

We would like to thank Dr. Yuxi Ma in BIGAI for her early experiments similar to Fig. 2.

References

[1] Michael Bratman. Intention, plans, and practical reason. 1987. 1

[2] Lifeng Fan, Shuwen Qiu, Zilong Zheng, Tao Gao, Song-Chun Zhu, and Yixin Zhu. Learning triadic belief dynamics in nonverbal communication from videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7312–7321, 2021. 3

[3] Michael C Frank and Noah D Goodman. Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998, 2012. 1, 2

[4] Kaiwen Jiang, Stephanie Stacy, Annya L Dahmani, Boxuan Jiang, Federico Rossano, Yixin Zhu, and Tao Gao. What is the point? a theory of mind model of relevance. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 44, 2022. 4

[5] Jin Joo Lee, Fei Sha, and Cynthia Breazeal. A bayesian theory of mind approach to nonverbal communication. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 487–496. IEEE, 2019. 2, 4

[6] Dan Sperber and Deirdre Wilson. *Relevance: Communication and cognition*, volume 142. Citeseer, 1986. 1

[7] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018. 3