UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
What Is the point? a Theory of Mind Model of Relevance

Permalink
https://escholarship.org/uc/item/7hz7p7f3

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Jiang, Kaiwen
Stacy, Stephanie
Dahmani, Annya L.
et al.

Publication Date
2022

Peer reviewed

eScholarship.org
Powered by the California Digital Library
University of California

What Is the Point? A Theory of Mind Model of Relevance

Kaiwen Jiang[1] kaiwenj@g.ucla.edu
Annya Dahmani[3] adahmani@g.ucla.edu
Stephanie Stacy[1] stephaniestacy@g.ucla.edu
Boxuan Jiang[1] bxjiang@g.ucla.edu
Federico Rossano[4] frossano@ucsd.edu
Yixin Zhu[1] yixin.zhu@ucla.edu
Tao Gao[1,2] tao.gao@stat.ucla.edu

[1] Department of Statistics, UCLA [2] Department of Communication, UCLA
[3] Department of Psychology, UCLA [4] Department of Cognitive Science, UCSD

Abstract

Although pointing is sparse, overloaded, and indirect, it allows humans to effectively decode shared information, (ex)change their minds, and plan accordingly. Pointing is an invitation to jointly attend to an object, which triggers the mutual inference between agents of each other's mind. Relevance is a fundamental assumption underlying all human communication, including pointing. We define relevance as how much a signaler's belief can make a positive difference to its receiver's well being. We build a Theory of Mind (ToM) model to test our definition of relevance and use pointing as a case study. In two experiments, we test our relevance model in a classic artificial intelligence (AI) task, the Wumpus world, with the key difference that there is a guide that points to help a hunter. Agents with our relevance model gain significantly higher rewards than agents who ignore signals from the guide. Agents with our model also achieve better performance than agents who receive an additional observation of the environment. The results show that the power of pointing comes from the ToM inference of relevance, rather than providing more precise individual perception.

Keywords: pointing; relevance; joint attention; theory of mind

Introduction

Pointing: condensed but powerful communication

Imagine two hunters hunting in a forest. The young hunter sees a broken branch on the ground. Assuming that the branch was broken by the wind, he gets ready to continue his search. At this moment, his partner, the experienced hunter points the broken branch to him. The young hunter suddenly realizes that the branch was broken by their prey. He holds his breath and prepares to hunt.

In this pointing example, one's attention to an observation is not a spotlight to enhance individual sensory accuracy and more than an action label commonly adopted in the computer vision community. Instead, it involves rich cognitive inference and demonstrates properties of human unique communication. First, pointing is **sparse**; the semantics of such a succinct act has complex meanings, much beyond the gesture itself: "Take a look at this broken branch caused by a prey." Such rich information is condensed spatially and temporally into one extension of the index finger, which lasts no more than a few seconds. Second, pointing is **overloaded**. As Wittgenstein and Anscombe (1953/2001) stated: "Point to a piece of paper. Now to its color, to its shape." Multiple features may coexist in the location where a pointing signal directs, and each of these features may be the referent of the pointing. Third, pointing is **indirect**. The meaning of the pointing can go far beyond the referent visual feature. In the hunting example, the experienced hunter simply points to the broken branch, but she does not mean the brokenness of the branch. Instead, she means they should get ready to hunt. Throughout the paper, we use female pronouns to represent the signaler and male pronouns to represent the receiver.

Fortunately, humans can interpret gestures (Kendon, 2004; Lascarides & Stone, 2009) such as pointing with decent accuracy despite their key properties of being sparse, overloaded, and indirect. As a rich form of communication, pointing is effective in changing its receiver's mind and actions.

Relevance: key assumption of communication

We use pointing as an example to highlight the jointness of communication. Pointing leads to a **joint attention**, qualitatively different from attending to an object individually. When an agent attends to his surroundings *individually*, he owns his observations; it is his own job to evaluate the relevance of perceived objects and filter out irrelevant information (Wilson & Sperber, 2002). In contrast, when a signaler points an object to a receiver, she invites the receiver to become a "guest" to the observation. The receiver can safely assume that the information given by the "host" must be relevant to the shared task. In the hunting example, no matter how clearly the young hunter sees the broken branch individually, he is not likely to change his explanation of the shape of the branch. However, from a point, he can make much richer inferences and revise his plans more drastically. Both the signaler and the receiver of the pointing are aware of the effect of this joint attention, so they reserve it to convey relevant information.

Pointing is built on the mutually acknowledged assumption that human communication must be relevant (Sperber & Wilson, 1986). Infants as young as 12 months prefer to share information only when it matches their partner's current goal: when infants watch an adult misplace an object and search for it, they point to the exact object more often than to other objects not needed by the adult (Liszkowski, Carpenter, Striano, & Tomasello, 2006). Meanwhile, the receiver has to also resonate with the relevance of what is being pointed to. Infants point significantly more to responsive adults than to ignorant ones; when the adult expresses disinterest, children no longer repeat the gesture (Carpenter & Liebal, 2011).

Relevance allows for an agent to decide which meaning (m) among many compatible meanings is the most likely. Mathematically, this can be formulated by evaluating the relevance of each possible meaning $m \in M$, and the most relevant meaning is the signaler's intended meaning m^*,

$$m^* = \arg\max_{m \in M} Relevance(m). \tag{1}$$

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*. ©2022 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

In the hunting example, the young hunter can list many meanings: the prey broke the branch, the wind broke the branch, a storm broke the branch, and so on. He must choose the meaning that is most relevant to the hunt, which is the prey.

Now that the importance of relevance has been acknowledged, the real challenge is how to define it. Classic information theory ignores relevance in its models of communication. They focus on retrieving the accurate signal from a noisy channel and then interpret it with a codebook (MacKay, 2003). In the field of machine learning, one mainstream approach is to define relevance as the associations between variables. It is usually trained in a data-driven fashion from a large dataset. As an exemplar of this model family, deep-learning based models in natural language processing and computer vision analyze the mapping from syntax or visual stimuli to meaning (Collobert et al., 2011; Vaswani et al., 2017; Kenton & Toutanova, 2019; Dosovitskiy et al., 2020). Like models in information theory, these models do not consider overloaded signals. They rely on one-to-one mappings between signal and meaning, which is not how people understand pointing.

These association models of relevance also fail to capture that human communication is causally transparent (Pearl & Mackenzie, 2018). When interpreting pointing, the receiver should not rely on a massive training on interpretation, but an understanding of the underlying model of how pointing is generated. To understand the pointing act, the young hunter must ask the key questions: "Why was the pointing signal sent? What would I have done if I did not receive the point?" To address these questions, we devise a causal model based on *agency* and *utility calculus*. More specifically, the causal process of signal generation and interpretation is modeled by how the agents' actions, both instrumental and communicative, are generated and driven by their mental states.

Our model of relevance can be considered as the utility function in a special type of rational speech act (RSA) (Frank & Goodman, 2012). However, this utility definition is not trivial. In fact, the majority of this paper is focused on deriving a definition of relevance based on utility theory and coordination of minds. The reason this is challenging is because we must go beyond a language game. We must have a full model of an agent that can change the physical world. This leads to a different focus in the context of pragmatics. In a language game, the focus is on the vocabulary: an agent can think what else the speaker could have said. Here, our context is focused on action: an agent thinks what else it could do.

We start from the causal interpretation of relevance by defining relevance as to what degree a receiver's well being can be improved by a signaler sharing her belief. In communication, signalers tend to be altruistic (Tomasello, 2010). There is no point in a signaler telling a receiver information that does not make a difference to the receiver's well being. Relevance must serve as an intervention to change the receiver's well being.

Under this assumption, the causal model of agency and utility theory are important for defining relevance. A signaler must know their receiver's mental state and predict their actions and the consequences of those actions, as these are key to evaluating a receiver's well being.

Our definition of relevance is not limited to pointing, but can be extended to other forms of communication. We use pointing as a case study to describe relevance because pointing can be mapped to any meaning. Pointing serves an extreme test of overloaded signals for agents to interpret by assuming relevance.

Theory of mind: stage for relevance

The capacity to infer others' mental states and predict their future actions relies on Theory of Mind (ToM) abilities. ToM allows agents to infer others' beliefs as the informative state, desires as the motivational state, and intentions as the deliberative states of the mind (Bratman, 1987). Communication is used as an effort for agents to increase their mutual benefits through synchronizing their minds. With beliefs and desires as components of the mind, agents are capable of decision-making, perspective-taking (Barnes-Holmes, McHugh, & Barnes-Holmes, 2004), inferring other agents' minds based on their actions, changing beliefs according to their sensory input, and revising their plans.

Work in developmental psychology has shown that ToM is a social commonsense that develops in early infancy (Gergely & Csibra, 2003; Woodward, 1998; Wellman, 2014). ToM has also been successfully formulated as Bayesian inference and inverse planning that explains both infant (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Liu, Ullman, Tenenbaum, & Spelke, 2017) and adult (C. L. Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017) cognition. Bayesian ToM is successful in modeling physical and social goal inference (C. Baker, Saxe, & Tenenbaum, 2011; Ullman et al., 2009). Subjects in these models observe the environment or the social interaction from the outside, not as a part of the social interaction. While these ToM models have succeeded in modeling individual behavior, the full potential in modeling transparent communication through the coordination of minds has not been fully developed. Our goal here is to introduce a causal model of relevance based on how agents synchronize their minds using ToM inferences.

Preliminaries: modeling Theory of Mind

With ToM, an agent can make sense of other agents' actions a from their belief b and desire d. We leave out intention in our planning model because there are no competing desires in communication. Using Bayesian ToM, we can infer an agent's belief b and desire d from its action a,

$$P(b,d|a) \propto P(b,d)P(a|b,d). \tag{2}$$

The planning model $P(a|b,d)$ in Eq. (2) describes how an agent plans its actions based on its belief and desire. In utility theory, desire can be modeled as a utility function. Here we use the belief-action value function $Q(b,a)$ to represent an agent's evaluation of the desirability of an action a based on its belief b. A rational agent chooses the actions that they believe to be the most desirable,

$$a^* = \arg\max_a Q(b,a). \tag{3}$$

To capture a certain degree of stochasticity and irrationality in human nature, a softmax function can be used to calculate the probability

$$P(a|b) \propto e^{\alpha Q(b,a)}, \tag{4}$$

where α is a parameter of the agent's rationality. Desire is not a part of the inference because a fixed desire of accomplishing the task is known by the agents in communication. It is represented as the Q function in Eq. (4).

One way of evaluating $Q(b,a)$ is to model the environment as an Markov decision process (MDP) (Sutton & Barto, 2018) or a partially observable Markov decision process (POMDP) (Kaelbling, Littman, & Cassandra, 1998). In these models, the environment has certain states s. An agent's belief is represented as a probabilistic distribution over all possible states,

$$b(s) = P(s|b). \tag{5}$$

In our study, beliefs are defined over states in the physical world, not including other agents' minds. When an agent takes an action a in a state s, it gains a reward $r(s,a)$. The agent's desire can be represented as the state-action value function $Q(s,a)$. It represents the optimal discounted cumulative reward that an agent can receive by taking the action a starting from the state s. When the agent maintains a belief, it can evaluate each action a with its belief-action function $Q(b,a)$. This process can be simply calculated by taking the expectation of the state-action value function over the belief,

$$Q(b,a) = \mathbb{E}_b(Q(s,a)) = \sum_s P(s|b)Q(s,a), \tag{6}$$

where $Q(b,a)$ can be obtained by using algorithms called solvers in POMDP.

Our model must utilize the ToM framework and treat communication as a rational action. The generative process of these actions requires both agents to coordinate their two minds, one from the self, the other from the partner. This is more complicated than the equations in this section and is the key problem in building a computational model that formally captures the essence of pointing. In addition, communicative actions do not change the physical world, but only change one's mind. The utility function of the communicative actions should be defined by the agents' minds, which was discussed by Frank and Goodman (2012).

A ToM Model of Relevance

A utility-based definition of relevance

Based on the mental states outlined by ToM, we start to define relevance of the signaler's belief in communication.

Relevance is evaluated in multiple steps. First, the signaler predicts the receiver's actions based on the receiver's belief. In the hunting example, the young hunter believes that the wind broke the branch. Knowing this, the experienced hunter predicts that he will walk away. We can write down this action prediction as

$$a_{Rec} = \arg\max_a Q(b_{Rec},a), \tag{7}$$

where the subscript *Rec* represents the receiver.

Next, the signaler evaluates the predicted action with her own belief. The experienced hunter knows that the prey is around, so she knows that the hunt will be ruined if the young hunter walks away. In this case, the signaler's evaluation of the receiver's action $Q(b_{Sig},a_{Rec})$ is low, where b_{Sig} is the signaler's belief. Here, we can see the discrepancy of the evaluation of the same action based on different beliefs; a_{Rec} is the best choice of actions to the receiver, while it may be undesirable to the signaler.

But that is not the end. If the receiver knows what the signaler knows, his utility and action can be improved based on the signaler's evaluation. Receiving the signaler's mind, the rational receiver will take the action that maximizes $Q(b_{Sig},a)$, which will improve his well being as evaluated by the signaler. In the hunting example, if the young hunter knows that the prey is around, he will stay silent and prepare to hunt. To the experienced hunter, this action is much better than the young hunter's original plan. In this case, the signaler's belief is relevant.

With the evaluation of the receiver's plans before and after receiving the pointing, we can calculate the utility of the pointing. Formally, we define the relevance of the signaler's mind as the gap between the best thing a receiver can do with the signaler's belief and the outcome that the receiver is actually going to get evaluated by the signaler.

$$Relevance(b_{Sig},b_{Rec}) = \max_a Q(b_{Sig},a) - Q(b_{Sig},a_{Rec}). \tag{8}$$

Sharing information must make a difference. For example, imagine an adult points a plane to a child. One possible interpretation of this pointing is that the child should fly. However, this interpretation is not relevant. It makes no difference to the child's well being because flying is not possible for the child. On the other hand, if the receiver of this pointing is a pilot, the "you should fly" interpretation may be relevant, as flying may increase the pilot's well being.

The evaluation of beliefs seems simple, but it is a more advanced type of ToM than a false belief task (Wimmer & Perner, 1983). In a false belief task, people only need to take others' perspective and predict their actions with ToM. However, in communication, when evaluating the receiver's utility, the signaler a) uses the receiver's belief to *predict* his action and b) uses her own belief to *evaluate* the receiver's action. This type of crossing of minds has not been looked at in communication, but has been shown in altruistic behavior (Tomasello, 2010). The closest concept in developmental psychology to this is paternalistic helping. A child can offer help in a way that she believes to be helpful to others, not what they believe to be helpful to themselves (Martin, Lin, & Olson, 2016). The signaler's evaluation of the receiver's mind is also often taken by parents. Parents improve a child's well being not based on what the child wants, but on what is best for them from the parents' perspective. Therefore, we call this **paternalistic evaluation of beliefs**. Our definition of relevance between two beliefs is shown in Fig. 1.

We propose that this crossing of minds goes beyond the classic AI definition of the value of evidence. The value of a piece of evidence is defined to be "the difference in expected value between best actions before and after information is obtained " (Russell, Norvig, & Davis, 2010). If we use Q function to represent the value, then $V(e) = \max_a Q(b,a|e) - \max_a Q(b,a)$. If the value of evidence is adopted as the definition of relevance, this will lead to a bad news paradox. Under this assumption, if the goal of the communication is to improve the listeners expected utility, then bad news should never be told. Providing bad news as a piece of evidence will drop the receiver's expected utility. This is because value of evidence only uses the receiver's perspective without any crossing of minds. The bad news paradox can be illustrated in the context of a hunting example: if a young hunter believes a prey is in front of him, he would have a high expected utility in shooting. However, based on an experienced hunter's knowledge this is a bad action because she knows the prey is not nearby. Using value of evidence as a relevance definition, she should not tell this piece of evidence at all because it would decrease the young hunter's expected utility. He would no longer expect to kill the prey. But if she uses our paternalistic evaluation of beliefs, this evidence carries positive value because from the experienced hunter's perspective it prevents the young hunter from shooting and wasting an arrow.

In practice, due to the limitation of communication, the receiver may not recover the signaler's exact belief. Instead, he may interpret the signaler's belief as b'. Formally, we define the utility of pointing as the utility change before and after communication, evaluated based on the signaler's belief. For a pointing signal u, the utility is

$$U(u) = Q(b_{Sig}, a'_{Rec}) - Q(b_{Sig}, a_{Rec}), \tag{9}$$

$$b'_{Rec} = P(s|b_{Rec}, u), \tag{10}$$

$$a'_{Rec} = \arg\max_a Q(b'_{Rec}, a). \tag{11}$$

Relevance in Eq. (8) can be used as the entry-level utility. With the utility of pointing clearly defined, we can model it as a rational action. Relevance is directly connected to the instrumental utility change caused by the signal. Therefore, it can be used as the utility in Eq. (4). Then the probability that the signaler takes the action of pointing u is

$$P_{Sig}(u|b_{Sig}) \propto e^{\alpha U(u)}. \tag{12}$$

The receiver's interpretation of the pointing signal can be modeled with Eq. (2),

$$P_{Rec}(b_{Sig}|u) \propto P_{Rec}(b_{Sig}) P_{Sig}(u|b_{Sig}). \tag{13}$$

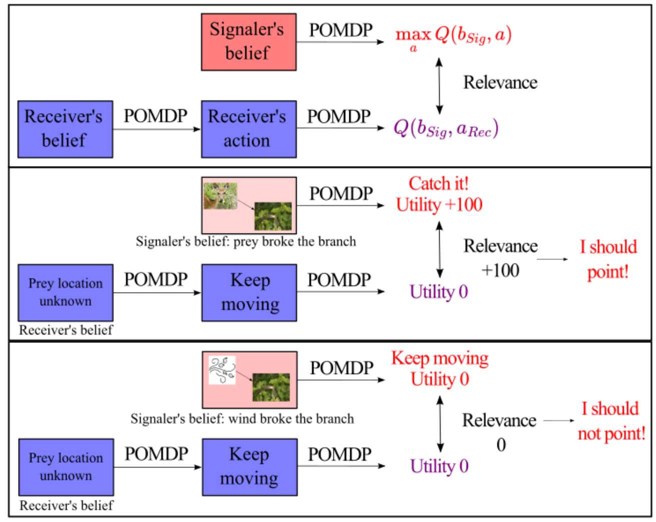

Figure 1: **Relevance calculation. Top:** Relevance is evaluated by coordinating the signaler's belief (red) and the receiver's belief (blue). Purple represents the crossing of beliefs. **Middle:** In the hunting example, the belief *prey broke the branch* is relevant as it increases the signaler's evaluation of the receiver's utility. **Bottom:** the belief *wind broke the branch* is irrelevant.

When the signaler has full knowledge of the world, b_{Sig} is the same as a single state s, which is the case in our experiments.

Experiments

We test the relevance model of pointing with an augmented version of the classic AI task, the Wumpus world (Russell et al., 2010), which is partially observable. In the Wumpus world, a hunter tries to kill a monster called Wumpus. However, he cannot see the location of the Wumpus and can only infer its location by his observation of the stench it emits. To simulate communication, we add another agent, the guide, who observes everything about the environment and the hunter. However, the only way she can communicate to the hunter is to point to an observation the hunter has already observed. The hunter needs to infer the meaning of the pointing and act accordingly. We call this game the *guided Wumpus hunting*. It is inspired by the hunting example, with highly sparse, overloaded, and indirect communication.

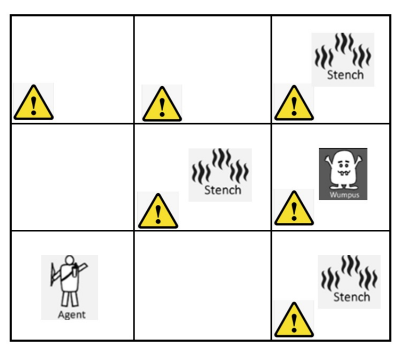

Figure 2: **Environment for Experiment 1.**

We conduct two simulation experiments with the *guided Wumpus hunting*. We start by testing our model of relevance with one observation. Of note, there is no uncertainty about which observation the point is referring to. The second experiment raises this challenge by adding another observation, incorporating overloadedness into the experiment. In both experiments, we compare the performance of agents with our relevance model and agents who use a single agent model as the baseline. In both models, the belief-action value function $Q(b,a)$ is calculated by a POMDP solver, the PERSEUS algorithm (Spaan & Vlassis, 2005). In addition, to distinguish the relevance model from enhancing individual perception, we add a control condition called "double observation." In this condition, the agent uses the single agent model, but he receives a second observation from the environment as if he observed the world twice.

Experiment 1

Task The environment of the experiment is shown in Fig. 2. The Wumpus is located in one of the six tiles with a warning sign. It does not move.

The hunter starts from the bottom-left corner tile. He can move or shoot in four directions: up, down, left, and right. A moving action will move the hunter one tile in the selected direction. If the action moves the hunter outside of the map, the hunter will not move. A shooting action will shoot an arrow to the adjacent tile in the selected direction.

The hunter can move unlimited steps in the map, but moving each step has an action cost of 5. The game ends when the hunter shoots or enters the tile of the Wumpus. If the hunter moves to the tile of the Wumpus, he gains -100. If he shoots and hits the Wumpus, he will gain a reward of 100. However, if he misses the shot, he will get -100.

The hunter cannot see the Wumpus, but he can infer the location of the Wumpus by observing its stench. There are two possible observations in the environment, stench or nothing. If the hunter is in a tile next to the Wumpus, he will have a high probability $p > 0.5$ of observing the stench. If the hunter is not next to the Wumpus, he will have a low probability $1 - p$ of observing the stench. The observation accuracy p is manipulated as an experiment condition. In the classic Wumpus world, $p = 1$. We add more stochasticity to increase the task difficulty and the need for pointing. Although we only have one possible referent to point to, which is the stench, the pointing is still overloaded. It can mean that the Wumpus is in one of all possible tiles (see Fig. 3) or that the hunter should take one of all possible actions.

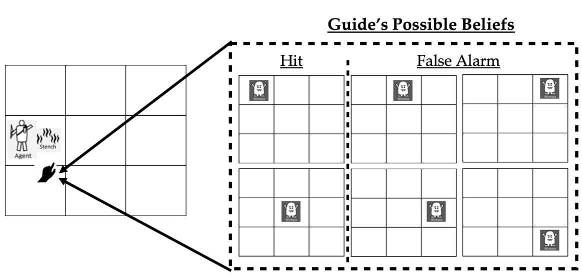

Figure 3: **Overloadedness of the pointing act in Experiment 1.** The pointing to the stench can mean an accurate observation or a false alarm, each leading to multiple possible world states.

Conditions The experiment is a 3×7 design; there are three models and seven observation accuracies.

The baseline model is the single agent model. With this model, the hunter uses a POMDP model to hunt the Wumpus. He ignores all the pointing signals from the guide, so that he engages in a single agent task. The second model is the relevance model of pointing. In this condition, the guide and the hunter use the relevance model to generate and interpret pointing signals. The third model is the double observation model. We test all three models with seven different levels of observation accuracies, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, and 1.

We run 100 game simulations for each condition and record the reward gained by the hunter. We predict that a) the hunters who use the relevance model will gain more rewards than the hunters who use the other two models, and b) as the observation accuracy decreases, the performance of the relevance model does not decrease because the power of our relevance model comes from ToM inference instead of observation accuracy.

Results The average reward across trials for each model under various observation accuracies is depicted in Fig. 4. Overall, agents who use the proposed relevance model achieve a higher average reward than agents who use the single agent POMDP model or the double observation model. The main effect of model type is significant ($F(2,2079) = 141.926$, $p < 0.001$), and the main effect of observation accuracy is also significant ($F(6,2079) = 58.370$, $p < 0.001$). The interaction between models and observation accuracy is significant ($F(12,2079) = 16.303$, $p < 0.001$). A post-hoc test with Bonferroni correction shows that agents who use the relevance model of pointing gain a higher reward than agents who use the double observation model ($F(1,1398) = 109.882$, $p < 0.001$). Our results show the power of relevance-based pointing, especially when the observation accuracy is low. It is more effective in helping its receiver than providing more accurate observations to his individual attention.

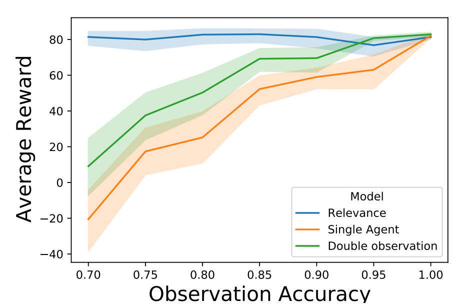

Figure 4: **Results of Experiment 1.** Shaded areas represent 95% bootstrap confidence interval.

Experiment 2

In Experiment 1, we have only one type of feature in the observation: the stench. Although the pointing signal may still have multiple interpretations as shown in Fig. 3, it has only one referent. This lacks the overloadedness discussed by Wittgenstein and Anscombe (1953/2001). To capture this overloadedness, we added another observation of glitter to the environment. Here, the glitter and stench coexist in the same grid. This setting offers a more complete evaluation of the Wumpus world.

Task The setup of Experiment 2 is identical to Experiment 1 except for a few aspects. We reduced the size of the environment because including an additional observation exponentially decreased the speed of our POMDP solver. The environment of Experiment 2 is shown in Fig. 5. A gold bar is added to the game as a source of the glitter. Picking up the gold bar gives the hunter a reward. The Wumpus and the gold bar are located in two different tiles of the three tiles with a warning sign. They do not move. They are invisible to the hunter but visible to the guide.

The hunter starts from the bottom-left corner tile. His moving and shooting actions have the same effect as in Experiment 1. In addition, the hunter has another action of picking up the gold bar. This action removes the gold bar if the hunter is in the same tile with the gold bar. Otherwise, it does not change the environment. The action of picking up the gold bar has a cost of 5 if he misses the gold bar. The hunter will gain a reward of 100 if he successfully picks up the gold bar.

The gold bar spreads glitter to its nearby tiles. The observations of glitter and stench of the Wumpus have the same probability model as the stench in Experiment 1. The observations of the stench and the glitter are independent, resulting in four possible observations in the environment. For example, if the hunter is in a tile that is adjacent to the Wumpus but not the gold, he may observe a) both glitter and stench with probability $p(1-p)$, b) single glitter with probability $(1-p)^2$, c) single stench with probability p^2, or d) nothing with probability $p(1-p)$.

Conditions The design and conditions are the same as Experiment 1. We predict that there will be an advantage in rewards with the relevance model compared to a single agent model and double observation model.

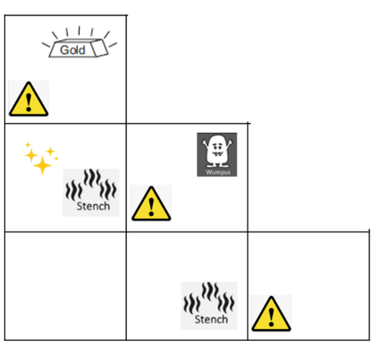

Figure 5: **Environment for experiment 2.**

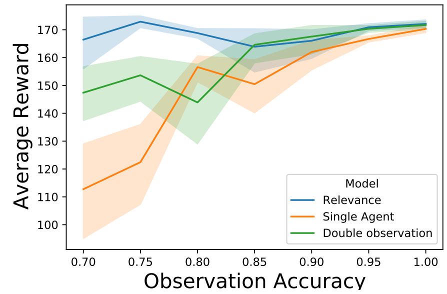

Figure 6: **Results for Experiment 2.** Shaded areas represent 95% bootstrap confidence interval.

Results The average reward across trials for each model under various observation accuracies is depicted in Fig. 6. Similar to Experiment 1, agents who use the relevance model achieve a higher reward on average than agents who use the single agent POMDP model or the double observation model. The main effect of model type is significant ($F(2,2079) = 41.732$, $p < 0.001$), and the main effect of observation accuracy is also significant ($F(6,2079) = 20.049$, $p < 0.001$). The interaction between models and observation accuracy is significant ($F(12,2079) = 9.130$, $p < 0.001$). A post-hoc test with Bonferroni correction shows that agents who use the relevance model of pointing gain higher reward than agents who use the double observation model ($F(1,1398) = 21.163$, $p < 0.001$). Our results are consistent with the results in Experiment 1. The relevance model of pointing still achieves high performance in highly overloaded communication.

Discussion

Our relevance model was successful in capturing the essence of transparent communication. In both Experiment 1 and Experiment 2, the agents who used the relevance model for communication achieved better performance than agents who ignored the communication. In addition, agents who used the relevance model achieved better performance than agents with more precise individual perception by having two samples from the environment. The high performance of the relevance model was robust over all observation accuracies.

Our results showed that by leveraging ToM and utility theory, agents can achieve overloaded communication without a predefined codebook. In the experiments, the guide and the hunter do not have a codebook that regulates the relevance between signals and meanings. Crucially, they never learned the relevance through massive training. They promptly calculate the relevance based on their context of the cooperative task. Our results also showed that the power of pointing comes from the ToM inference which supports the relevance calculation. The power of pointing does not come from enhancing individual perception, which is supported by the robust performance of the relevance model across all observation accuracies.

Acknowledgments

This research was supported by DARPA PA 19-03-01 and ONR MURI project N00014-16-1-2007 to TG. This material is also based upon work supported by NSF-GRFP Grant No. DGE-1650604 to SS.

References

Baker, C., Saxe, R., & Tenenbaum, J. (2011). Bayesian theory of mind: Modeling joint belief-desire attribution. In *Proceedings of the annual meeting of the cognitive science society.*

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*(4), 1–10.

Barnes-Holmes, Y., McHugh, L., & Barnes-Holmes, D. (2004). Perspective-taking and theory of mind: A relational frame account. *The Behavior Analyst Today*, *5*(1), 15.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press.

Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. In *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 159–181).

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, *12*, 2493–2537.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... others (2020). An image is worth 16x16 words: Transformers for image recognition at scale. In *International conference on learning representations (iclr).*

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

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, *101*(1-2), 99–134.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kenton, J. D. M.-W. C., & Toutanova, L. K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the conference of the north american chapter of the association for computational linguistics (naacl).*

Lascarides, A., & Stone, M. (2009). A formal semantic analysis of gesture. *Journal of Semantics*, *26*(4), 393–449.

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12-and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, *7*(2), 173–187.

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041.

MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

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.

Pearl, J., & Mackenzie, D. (2018). *The book of why: the new science of cause and effect*. Basic books.

Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial intelligence: a modern approach*. Prentice Hall.

Spaan, M. T., & Vlassis, N. (2005). Perseus: Randomized point-based value iteration for pomdps. *Journal of Artificial Intelligence Research*, *24*, 195–220.

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

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

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

Ullman, T. D., Baker, C. L., Macindoe, O., Evans, O., Goodman, N. D., & Tenenbaum, J. B. (2009). Help or hinder: Bayesian models of social goal inference. In *Proceedings of advances in neural information processing systems.*

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. In *Proceedings of advances in neural information processing systems.*

Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.

Wilson, D., & Sperber, D. (2002). *Relevance theory.* Blackwell.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Wittgenstein, L., & Anscombe, G. (2001). *Philosophical investigations: the german text, with a revised english translation* (3rd ed.). Blackwell. (Original work published 1953)

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34.