Intuitive Social Interaction: Belief-Desire Attribution with Communication

Gabriel B. Margolis
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA
gmargo@mit.edu

Kelly Zhang
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA
kellyz@mit.edu

Abstract

We present an extension of the Bayesian Theory of Mind (BToM) framework to a setting where agents are able to communicate information to one another at a distance. Like the BToM presented by Baker, Saxe, and Tenenbaum (2011), our algorithm jointly infers an agent's desires and beliefs conditioned on observations of the agent's actions. In an extension to this prior work, we allow that the agent's beliefs may be influenced by secondhand information communicated from another agent. We demonstrate a computational model which correctly infers the presence of communication from observed actions. Through human experiments, we demonstrate that subjects can correctly explain agents' irrational behavior by inferring that their beliefs were updates via communication.

Keywords: Theory of Mind; Social Cognition; Action understanding; Bayesian inference

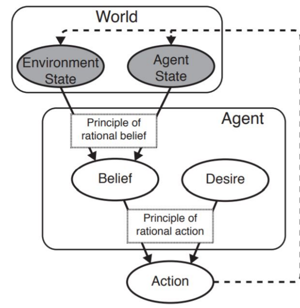


Figure 1: Causal structure of Theory of Mind (Baker et al., 2011).

Introduction

Humans excel at inferring others' complex goals from few observations of their behavior. This ability is known as Intuitive Psychology. Recent work in computational cognitive science has sought to construct models for Intuitive Psychology which enable computers to make inferences on human behavior. Such models have leveraged upon the Bayesian Theory of Mind (BToM) model to achieve human-like predictions in environments of full and partial information (Baker et al., 2011; Baker, Tenenbaum, & Saxe, 2007; Ullman et al., 2010).

A remarkable ability of humans is their ability to adapt their inferences to implicit or explicit rules about their environment. Given new contextual information about the possible mental states of an agent, humans can quickly come up with improved explanations for the agent's observed behavior. This is particularly potent when an agent appears to behave irrationally under the observer's current model; humans are capable of considering what would happen if the agent had various alternative beliefs, and selecting the possibility which most accurately predicts reality.

We present an extension to the work of Baker et al. (2011), which addressed joint goal and belief inference for a single agent seeking rewards in a partially observed environment. Our extension places multiple agents in the same environment and models the possibility that they may have communicated their beliefs to one another with some nonzero probability. This introduces uncertainty into the belief state of each agent.

We develop an inference algorithm based on inverse planning which predicts the goals, beliefs, and probability of communication between agents given a set of observed agent actions. We compare the predictions of our algorithm to the predictions of humans, and evaluate the hypothesis that humans adapt their predictions given context information (i.e. probability of communication) with results similar to explicit Bayesian model-building.

Previous Work

Intuitive Psychology

Intuitive Psychology is the ability of humans and intelligent animals to make predictions about the behaviors and goals of others. In their famous experiment, Warneken and Tomasello (2006) demonstrated that human children as young as 18 months old are able to understand others' goals and aid in their achievement. They also demonstrated that young chimpanzees have similar capacity. These results suggest that Intuitive Psychology is either hardwired into the human brain or learned very at a very young age.

Bayesian Theory of Mind

Humans and animals generally understand each other's actions through the lens of a *Theory of Mind*. Theory of Mind is the ability to conceive of intelligent agents as having internal mental states that motivate their observed behaviors. For example, a squirrel's act of climbing a tree might be attributed to its desire for nuts - an internal mental state that is not directly accessible to the observer. Such an attribution could be made using a simple logical Theory of Mind, with the logical constraint that a squirrel climbs a tree only if it desires nuts.

In addition to desires, beliefs may be thought of as mental states in a Theory of Mind. We may wonder why our squirrel climbs acorn trees and not pine trees. This action can be jointly attributed to the squirrel's desire for nuts and its logical belief that acorn trees contain nuts and pine trees do not.

It turns out that working in purely logical statements is a brittle way to model Theory of Mind. Imagine we observed a squirrel climbing a pine tree: Would we think that this squirrel desired pine needles, or that it believed the pine tree contained nuts? The answer seems uncertain, so we might suppose that our theory of the squirrel's mind should place some nonzero chance on each being the case. This motivates a Bayesian Theory of Mind, in which mental states are modeled within a Bayesian framework.

Our work builds upon the work of Baker et al. (2011), which presents a computational model for Bayesian Theory of Mind and applies this model to jointly infer a simulated agent's beliefs and desires from its behavior in a partially observed environment.

Inverse Planning

With the goal of developing a computational model for thought, and having accepted the premise that Intuitive Psychology is a fundamental building block of human cognition, it is desirable to identify a computational mechanism that can infer an agent's goals from its observed actions. Inverse Planning is one such mechanism that has been proposed in previous literature.

Before we discuss Inverse Planning, we will recall the properties of a rational or optimal planning algorithm. The input of a planning algorithm is the state of an agent's environment and the set of rewards or penalties associated with different actions in that environment. The output of a planning algorithm is an optimal or rational plan, a sequence of actions that maximizes an agent's cumulative expected reward.

The premise of Inverse Planning is that an agent can be approximately assumed to act according to a rational plan, given its beliefs, desires, and current state. Thus, observations of the agent's actions can be viewed as the output of an optimal planning algorithm, and by inverting this optimal planning algorithm we can infer its input: the beliefs and desires that motivated the sequence of observed actions.

Computational models for inverse planning have been proposed in previous work. Baker et al. (2007) demonstrated Inverse Planning in a fully observed environment as a mechanism for goal inference, and found that its predictions were able to closely model the predictions of human subjects. Baker et al. (2011) applied the same formalism to inference in a partially observed environment by inverting a POMDP model.

Communication and Collaboration

Computational modeling of communication and collaboration is an area of active research in computational cognitive science. (Ullman et al., 2010) demonstrated a model for social goal inference in fully observed environments, based on implicit communication. Here, the agents inferred each other's goals and collaborated to bolster each other's reward. Each agent's actions influenced the other agent's desires, but beliefs were not shared or communicated.

Modeling the communication of beliefs and desires is of significant interest to researchers in the robotics community, particularly those concerned with human-computer interaction (HCI). HCI research is concerned with designing ma-

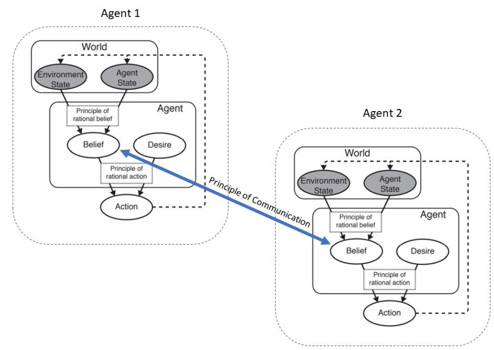


Figure 2: Causal structure of Theory of Mind for two agents with communication.

chines that collaborate with humans to complete tasks and solve problems. By understanding how the human mental state evolves over time in collaborative scenarios, HCI researchers stand to advance the performance of robots collaborating with humans in the context of production lines, surgeries, semi-autonomous driving, and more.

Technical Preliminaries

Path Planning

A path planning algorithm produces a sequence of actions or waypoints which guide an agent from a starting configuration to a goal configuration. For simplicity, we consider grid-based path planning algorithms which operate in discretized state and action space.

Breadth-First Search (BFS) is a simple grid-based path planning algorithm. In effect, BFS considers all paths of length 1, then all paths of length 2, and so on until it identifies a path to the goal. By definition, this must be the shortest path and constitutes an optimal plan. However, the complexity of BFS makes its application to a large grid intractable.

A* Search is a heuristic search algorithm which can perform grid-based path planning much more efficiently than BFS in practice. A* search works by prioritizing the exploration of paths which score better by some heuristic function. Although the worst-case performance of A* is no better than BFS, a good heuristic function can cause it to be much faster in practice. In path planning, the Cartesian distance to the goal is a commonly used and effective heuristic.

In practice, the most popular path planning algorithms in robotics are sample-based algorithms such as Rapidly-exploring Random Tree (RRT). These algorithms generate paths in continuous space by sampling waypoints randomly from the problem domain. Although RRT is fast, it may produce paths that are suboptimal which means they cannot be effectively used for goal inference via Inverse Planning. However, the RRT* variant on this algorithm asymptotically produces the optimal path in continuous space, and therefore we claim that the results of this work could be extended to the continuous domain via the application of RRT*.

Probabilistic Inference

In many interesting classes of problems, we make observations of a few variables and would like to make probabilistic predictions about other, related variables. This is known as probabilistic inference, and the model that describes how variables relate to one another can be represented as a probabilistic program.

At the heart of all probabilistic inference is Bayes' Rule:

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}$$

Bayes' Rule describes the relationship between the posterior ($P(B|A)$) and prior ($P(B)$) distributions of some variable B. Using this Bayes' Rule, we can condition our belief about the distribution of B on an observation of related variable A.

Method

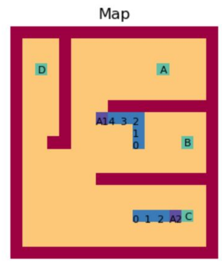
Scenario of Interest

We consider a scenario in which two agents rationally search for their reward-maximizing goal. Four goal locations contain four objects with known reward, but agents initially do not know which object lies at which goal location. Agents can obtain information either by visual observation or by communication:

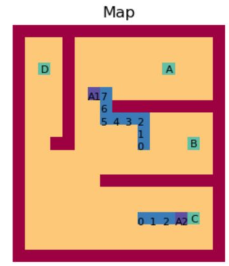
- **Visual observation** occurs whenever an obstacle-free line of sight exists from an agent's location to a goal location.
- **Communication** occurs with probability p at each timestep; each agent receives the full information gathered by the other.

Figure 3 shows a sample plan in this scenario with $p = 0$: no communication between agents. Agent $A2$ immediately travels to goal location C, observing that its most desired reward is available there. Agent $A1$ moves away from goal location B, the only goal it can observe at the start, because this is not its most desired goal. Agent $A1$ first explores the closest unobserved goal at location A, then the next at location D, and finally observes and reaches its most desired goal which turns out to be located in location C.

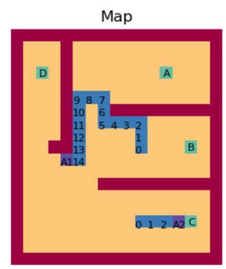
Figure 4 shows another sample plan in the same map configuration, but with $p = 0.2$: communication between agents occurs at the sixth timestep. As before, agent $A2$ immediately travels to goal location C, observing that its most desired reward is available there. Agent $A1$ initially travels toward goal location A as before. However, at the sixth step, $A1$ receives communication from $A2$ and changes course towards C without observing the goal at D. Note that in order to use vision to gain information blocked by an obstacle, an agent must completely pass the obstacle. For example, at 6 steps in Figure 4, agent $A1$ is still unable to see the goal at goal location A because the agent is at the boundary but has not passed the boundary. If $A1$ takes one more step upwards, then it is able to see the goal at goal location A.



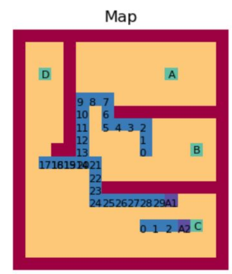
(a) 5 steps



(b) 8 steps



(c) 15 steps



(d) 30 steps

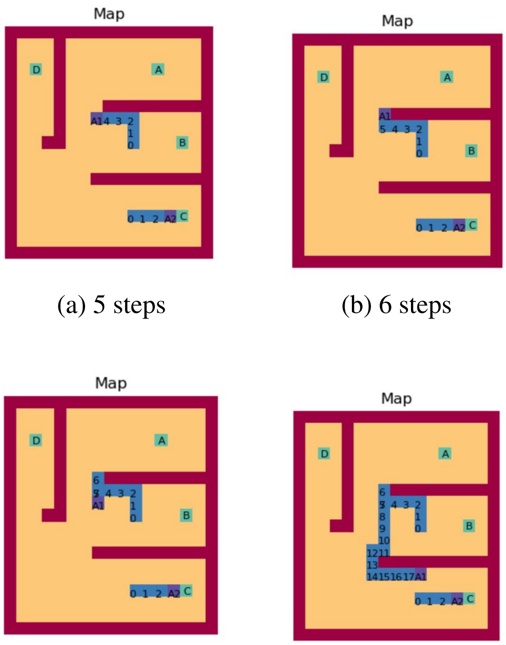
Figure 3: Sample path, no communication

Plan Generation

Plans are generated for each agent by determining the optimal move at each time step. Each agent's optimal known goal is determined and compared to the average value of the unknown goals. If the value of the optimal known goal is greater than that of the unknown goals, then the agent moves towards the optimal known goal. Otherwise, the agent moves toward the closest unknown goal. More specifically, the algorithm iterates through the time steps and updates each agent's plan as follows.

Determining the Optimal Known Goal First, the agent's optimal known goal is determined. The agent's knowledge is updated to account for any newly visible goals in it's current line of sight. The agent can then choose to communicate with another agent. In our implementation, we model communication as an event occurring with some probability p. Our scenarios use $p = 0$ and $p = 0.15$ such that no communication occurs in the former scenario and communication occurs with a probability of 0.15 in the latter scenario. If the agent decides to communicate with another agent, then the agent acquires the other agent's knowledge. The optimal known goal is determined by selecting the goal with the highest reward from the agent's updated knowledge base.

Determining the Next Location Given the optimal known goal, we can now determine the next optimal step. If the reward of the optimal known goal is greater than the average reward value of the unknown goals, then the optimal known goal is the overall optimal goal, so the agent moves toward the optimal known goal. However, if the average reward value of



(a) 5 steps (b) 6 steps

(c) 8 steps (d) 18 steps

Figure 4: Sample path, communication at step 6

the unknown goals is greater than that of the optimal known goal, then there is potentially a goal with greater reward at another goal location. Thus, the agent moves toward the closest goal location of an unknown goal. The location of the next optimal step is determined by the A* search path planning algorithm detailed previously in the technical preliminaries.

Psuedocode for the plan generation algorithm is outlined in Algorithm 1.

Probabilistic Model

An inference model was implemented to predict an agent's goal and communication at each time step.

Goal Inference The purpose of goal inference is to determine the probabilities of the agent going to each goal given the current path. Our inference model is based on the given agent's previous and current locations and inverse planning. First, we determine the agent's location at time t and $t-1$. Then, we use inverse planning to get the actions the agent would have taken at the previous time step given the different goals and determine the posterior probabilities. Psuedocode for the goal inference algorithm is presented in Algorithm 2.

Communication Inference The purpose of communication inference is to determine whether the agent has communicated with another agent at a previous time step given the current path. We implement a simple inference model based on the agent's goals and knowledge. We infer the agent's knowledge at the previous timestep, in the case that communication occurred and in the case it did not. Then, we compute the set of rational plans in each case and compare them to the action that was actually taken. If the action taken can only be rationally explained by communication, the probability of communication is high. Pseudocode for the communication inference algorithm is presented in Algorithm 3.

Algorithm 1 Plan Generation

Load env, agents
t=0
while t $<$ t_max **do**
  for all agents i **do**
    loc[i] $\rightarrow$ optimalStep(loc[i$-$1], knowledge[i$-$1])
    knowledge[i] $\rightarrow$ knowledge[i] + env.visible(loc[i])
    if Unif(0, 1) $>$ p **then**
      knowledge[i] $\rightarrow$ sum(knowledge[:])
    end if
    t = t + 1
  end for
end while

Algorithm 2 Goal Inference

for all goals **do**
  rationalStep = rationalPlans(prevLoc, goal)[0]
  if prevStep = rationalStep **then**
    rationalGoals $\rightarrow$ rationalGoals + goal
  end if
end for
return rationalGoals

Algorithm 3 Communication Inference

$\text{knowledge}_{t-1}^{nocomm}$ = visibleGoals(loc_{t-1}) + knowledge_{t-2}
$\text{knowledge}_{t-1}^{comm}$ = $\text{knowledge}_{t-1}^{nocomm}$ + allAgentsBeliefs
loc_{t}^{nocomm} = rationalPlans(loc_{t-1}, $\text{knowledge}_{t-1}^{nocomm}$)
loc_{t}^{comm} = rationalPlans(loc_{t-1}, $\text{knowledge}_{t-1}^{comm}$)
if loc_{t}^{comm} == loc_{t}^{nocomm} == loc_t **then**
  return 0.3
else if loc_{t}^{comm} == loc_t **then**
  return 0.9
else
  return 0.15
end if

Human Experiments

Human subjects were shown a sequence of images like those in Figures 3 and 4. Subjects were given the following contextual information at the beginning of the experiment:

- There are two agents: A1 and A2.
- Each agent knows the locations of 4 parking spots (A, B, C, and D).
- However, each agent does not know which of the five food trucks (1, 2, 3, 4, or 5) are parked at the four parking spots

until the parking spot comes within the agent's line of sight.

- Each agent has its own level of desire for each food truck. Agents wish to buy from the food truck they most desire as soon as possible.
- Please answer the following questions about agent A1.

Subjects were then shown a sequence of images, each one paired with the following survey questions:

- How much does A1 desire each food truck? (Please assign each truck a value 1-7, with 1 being the least and 7 being the most.)
- What is A1's current goal (the goal location it plans to visit next)? (Please assign each location a probability such that all probabilities sum to 1.)

In scenarios where communication was possible, an additional piece of context was provided at the experiment's start:

- At each timestep, agents communicate their knowledge of food truck locations to one another with probability p.

Additionally, an additional survey question was asked in scenarios with communication:

- Has A2 communicated with A1? (Please provide a probability between 0 and 1 that communication has occurred.)

Following the above guidelines, we surveyed a total of seven subjects.

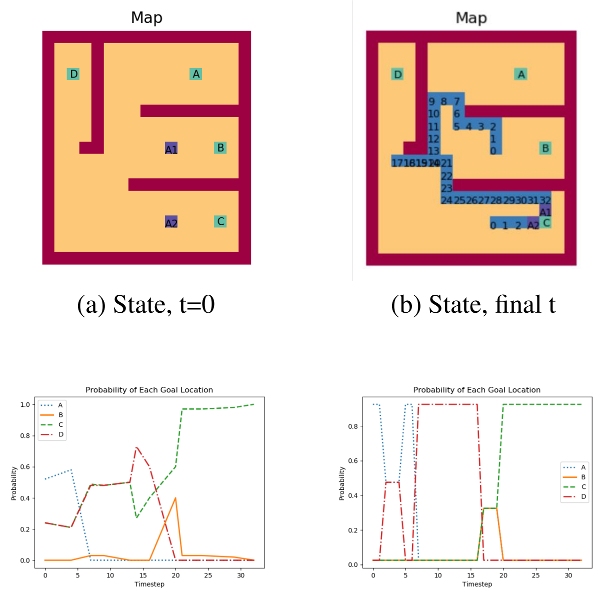
Results

Probabilistic Model

The following section details the path generation and inference results from three scenarios we implemented and tested with our probabilistic model. We analyze the paths generated and the resulting goal and communication probabilities for agent 1 predicted by our model.

Scenario 1: No Communication The first scenario removed the potential for communication by setting the probability of communication to $p = 0$. Figure 5 displays the paths for agent 1 and agent 2 at $t = 0$ and $t = 33$, the final time step. From the complete path in Figure 5b, we can observe that agent 1 checks every goal location while agent 2 goes directly to goal location C.

Figure 5d displays the probability of goal location x being agent 1's current goal (the location it plans to visit next) for all $x \in \{A, B, C, D\}$. From Figure 5d, we observe that our model alters probability predictions primarily when the agent shifts direction. At $t = 0$, the probability of A being agent 1's current goal is high because agent 1 is moving in the direction of goal A, and the probabilities of goals B and C being agent 1's current goal are low since agent 1 is moving away from those goals. However, after agent 1 sees goal A and turns away, the



(a) State, t=0 (b) State, final t

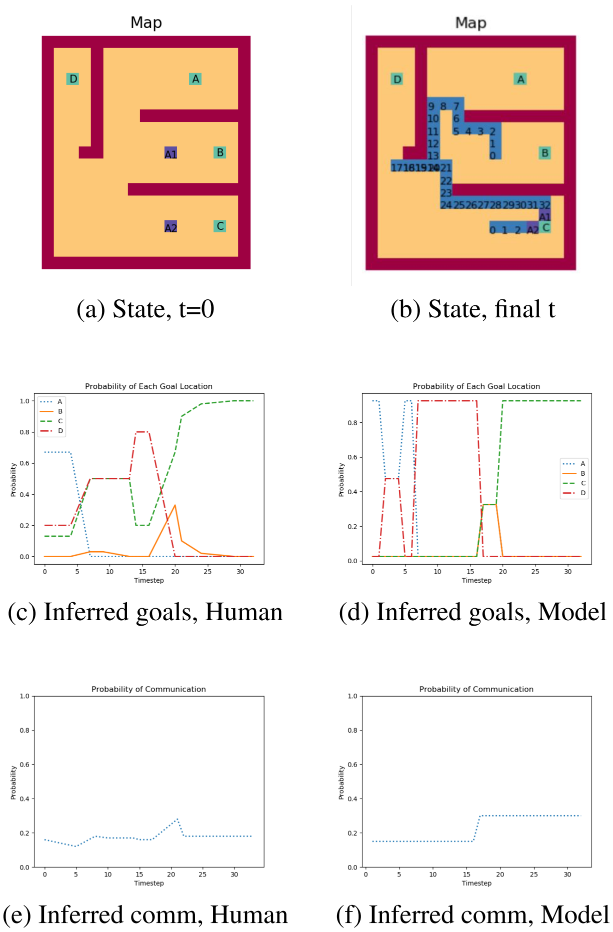
(c) Inferred goals, Human (d) Inferred goals, Model

Figure 5: Scenario 1, $p = 0$

model predicts the probability of goal A being agent 1's current goal as 0 since agent 1 clearly was not interested in what it saw. At this point, the probability of goal D being agent 1's current goal increases since agent 1 moves in the direction of goal D. Once agent 1 sees goal D, it turns around and moves toward goal C, so the probability of goal C being the current goal increases . This probability pattern is expected in a situation with no communication since the agent will want to check every possible goal location for it's optimal goal while minimizing distance traveled. If the agent sees goal x and turns away, then goal x is likely not the agent's optimal goal. The agent will continuously travel to the next closest goal until it finds its optimal goal. Thus, the goal that the agent lands on at the last time step is it's optimal goal. In this scenario, agent 1's optimal goal is goal C.

Scenario 2: Unrealized Communication The second scenario stated that communication had the possibility of occurring with probability $p = 0.15$, but presented a sample where communication was not actually realized. Figure 6 displays the paths for agent 1 and agent 2 at $t = 0$ and $t = 33$, the final time step. From the complete path in Figure 6b, we observe that this is the same path as the first scenario represented in Figure 5b.

Figure 6d displays the model's predicted probability of goal location x being agent 1's current goal for all $x \in \{A, B, C, D\}$. From Figure 6d, we observe that our model makes the same predictions as the first scenario depicted in Figure 5. The same probability distribution is expected since the goal probabilities are purely dependent on the agent's path and not influenced by communication. Although communication may affect the path, the goal probability distribution will

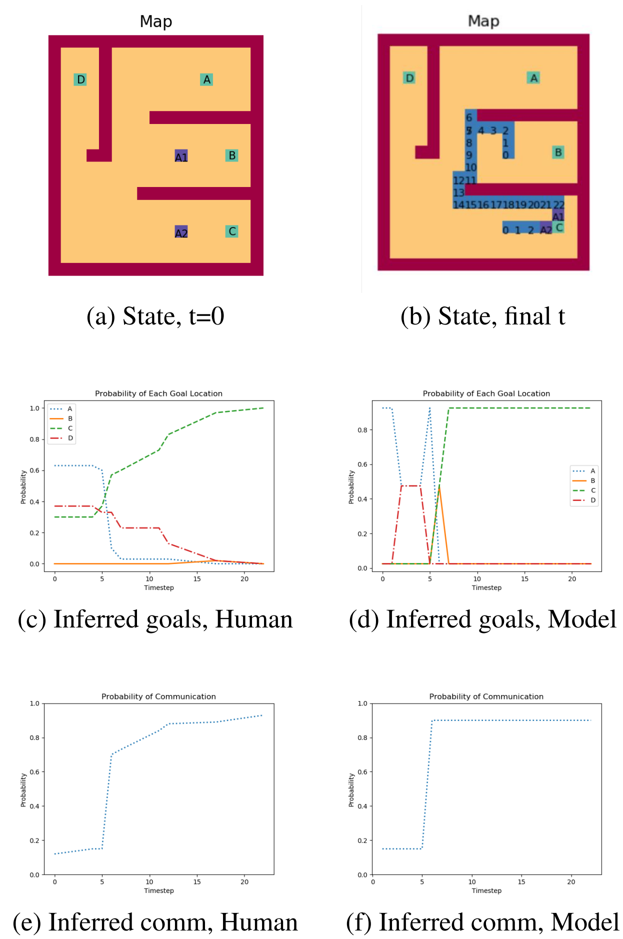


(a) State, t=0 (b) State, final t

(c) Inferred goals, Human (d) Inferred goals, Model

(e) Inferred comm, Human (f) Inferred comm, Model

Figure 6: Scenario 2, $p = 0.15$



(a) State, t=0 (b) State, final t

(c) Inferred goals, Human (d) Inferred goals, Model

(e) Inferred comm, Human (f) Inferred comm, Model

Figure 7: Scenario 3, $p = 0.15$

only be influenced if communication is manifested through a change in path.

Figure 6f displays the model's predicted probability that the agents have communicated at any previous time step given the observed path and environment. In this scenario, the model correctly predicts a low probability of communication throughout.

Scenario 3: Communication The third and final scenario mimicked the second scenario in that agents were able to communication to each other with a probability of $p = 0.15$. Despite the same parameters, the path generated for agent 1 differed from scenario 2. Figure 7b displays the difference in the path generated. We observe that agent 1 does not take the loop to view goals A and D. Instead, agent 1 turns around at the boundary to goal A and moves directly toward goal C.

Figure 7d displays the probabilities of goal location x being agent 1's current goal for all $x \in \{A, B, C, D\}$. As a result of agent 1's different path, we observe in Figure 7d that our model makes different predictions from the first two scenarios. From $t = 0$ to $t = 5$, the probability distributions look the same since agent 1 takes the same path. However, at $t = 7$, the path in scenario 3 diverges from the path in scenarios 1 and 2, resulting in a different probability distribution. Instead of increasing at $t = 6$, the probability of goal D being the current goal remains at 0 for the remainder of the scenario. This change in probability is reflective of agent 1's change in path. Along the same lines, the probability of goal C being the current goal reaches it's peak and remains high at $t = 7$ rather than $t = 20$. This change is also reflective of agent 1's change in path. These results are as expected since a more direct path to goal C will increase the probability that goal C is the current goal at an earlier time step. In this scenario, agent 1 lands on goal C at the final time step. Thus, goal C is still the optimal goal.

Figure 7f displays the model's predicted probability that the agents have communicated at any previous time step given the observed path and environment. In this scenario, the model correctly predicts a high probability that the agents have communicated after step 6 - since the agent takes a step at time 6 which is irrational given the information it has individually gathered (it has not yet observed goal A), the model is able to discern that communication must have occurred with high probability at this time.

Human Experiments

The following section details the goal and communication inferences made by human subjects on the three scenarios constructed with our path generation algorithm. For simplicity, we selected a range of 8 to 12 time steps to show human subjects. We tried to select time steps where we hypothesized a significant change in belief might occur.

Scenario 1: No Communication Figure 5c displays the probability of goal location x being agent 1's current goal for all $x \in \{A,B,C,D\}$ in the scenario depicted in Figure 5. We observe subject uncertainty in current goal determination toward the beginning of the simulation since the probabilities are generally under 0.5 until $t = 15$. Despite subject uncertainty, goal A is predicted as the current goal with the most certainty until $t = 6$. This is expected since agent 1 is taking steps toward goal A until it turns around at $t = 7$. The probability of goal B as the current goal starts at 0 since agent 1 is moving away from goal B. There is a spike in the probability when the agent moves toward goal B after $t = 17$. At $t = 15$, the subject predictions begin to rise above probabilities of 0.5 The probability of goal D being the current goal spikes to reflect the agent altering it's path toward goal D. At $t = 17$, the probability of goal C being the current goal beings to increase as the respective probability for goal D decreases to reflect the agent altering it's path to move toward goal C. At $t = 21$, the probability of goal C being the current goal is about 1 and the probability of goal B decreases to 0 once again as human subjects are certain that goal C is the current goal.

Scenario 2: Unrealized Communication Figure 6c displays the probability of goal location x being agent 1's current goal for all $x \in \{A,B,C,D\}$ in the scenario depicted in Figure 6. Although the exact probabilities are not the same for each time step, the qualitative trends for each goal reflect the same patterns across both scenarios, i.e. probabilities increase or decrease at the same time steps. The slight differences in probability at a given time step can be attributed to the non-deterministic nature of the human mind. The probability distribution similarities between scenario 1 and scenario 2 are expected since the goal probabilities are purely dependent on the agent's path. From Figures 5 and 6, we can see that agent 1 took the same path in both scenarios. Thus, the current goal probability distribution should be the same across both scenarios. Although communication may affect the path, the goal probability distribution is only influenced if communication is manifested through a change in path.

Figure 6e displays the human's predicted probability that the agents have communicated at any previous time step given the observed path and environment. In this scenario, the human subjects predict a low probability of communication throughout, consistent with our model. This makes sense because no strong evidence for communication is presented through either agent's actions in this scenario.

Scenario 3: Communication Figure 7c displays the probability of goal location x being agent 1's current goal for all $x \in \{A,B,C,D\}$ in the scenario depicted in Figure 7. The probability distribution of this scenario differs significantly from those of scenarios 1 and 2. Until $t = 5$, the probability distribution displays the same qualitative trends as in Figures 5c and 6c with goal A having the highest probability, goals C and D having about the same probability less than the probability of goal A, and goal B having a probability of 0. After $t = 5$, the probability distribution diverges. The probability of goal C being the current goal increases, converging at 1, while the probability of goal D being the current goal converges toward 0. Additionally, there is no spike in the probability of goal B being the current goal at around $t = 20$. It is important to note that the probability of goal A being the current goal displays the same qualitative trend since agent 1 still takes steps toward observing goal A. The differences in probability distribution can be attributed to the path taken by agent 1 that was altered by communication. We can observe in Figure 7 that agent 1 does not take steps past the boundary of goal A and toward goal D. Instead, the agent turns around before it can see goal A and moves directly toward goal C. This behavior is most likely explained by communication and manifested in the probability distribution for current goals. The probability of goal C being the current goal increases at an earlier time step, and the probability of goal D being the current goal decreases at an earlier time step.

Figure 7e displays the human's predicted probability that the agents have communicated at any previous time step given the observed path and environment. In this scenario, the human subjects predict a high probability that communication has occurred after time t=6. This makes sense because strong evidence for communication was introduced through A1's otherwise-irrational action at this time.

Analysis

Goal Inference

The inferences made by human subjects mirrored the inferences made by our model, especially in a qualitative manner. The probabilities generally shifted in the same directions at the same time steps. However, there were more significant quantitative differences as our model predicted current goals with higher confidence than humans, especially during earlier time steps. This can be most clearly observed by tracking the probabilities of goal A being the current goal and goal D being the current goal in scenarios 1 and 2. From $t = 0$ to $t = 5$, our model predicts goal A being the current goal with probabilities ranging from 0.5 to 1 and goal D being the current goal with probabilities ranging from 0 to 0.5. On the other hand, human subjects predicted goal A being the current goal with probabilities ranging from 0.3 to 0.6 and goal D being the current goal with probabilities ranging from 0.2 to 0.4. This can be attributed to human uncertainty and more deterministic modeling. It is interesting to note that our model captured the spike that appears for goal B in scenarios 1 and 2, suggesting that our model is able to capture nuances in human inference.

Communication Inference

Our model for communication inference captured the predictions of humans well, both in the case where communication occurred and in the case it did not occur. The scenario in which communication occurred and the time at which it occurred were easily identified with high confidence by both human subjects and our model. Our result supports the hypothesis that humans have a strong intuition for a theory of mind incorporating multiple interacting agents, including the relationship between beliefs, desires, and actions. In particular, our results support the hypothesis that humans have a natural understanding of rational behavior in a complex environment with partial information, and are able to explain irrational behavior by considering the new information an agent may have acquired.

Future Work

Increase Sample Size for Human Surveys

With seven subjects for our human experiments, we were able to demonstrate that humans have an intuitive understanding of rational behavior in environments with limited information. However, to provide a more robust analysis, we aim to collect more data by surveying more subjects.

Planning to Communicate

Our experiments demonstrated that humans can make inferences about communication between agents based on their observed behavior, and that similar inferences can be made by a computational model based. While we modeled communication as a random phenomena, and instructed our subjects to treat it as such, we hypothesize that humans could apply principles of inverse planning to predict communication if it were planned by the agents in a non-random way to help them achieve their goals. In the future, we believe that this is a promising direction of research.

Planning to Observe

Although our agents interacted with a partially observable environment, we generally simplified their plans to ignore uncertainty. This was technically able to lead to suboptimal behavior, which some of our human subjects spotted. For example, our agent would bypass a long hallway because the goal at the end was far away, even if in just a few steps the agent could step into the hallway and observe the goal at a distance; our planner as implemented does not account for the value of information, and therefore does not plan to observe. A more rigorous treatment of the planning problem that accounted for the environment's partial observability would fix this problem.

Unrealized Goals

We restricted our scenarios to domains of two agents and four goals, with all goals present on the map. In Baker et al. (2011), only a subset of the existing goals were present on the map, and subjects were asked to infer the agent's desire for goals that were not present. A scenario with only a subset of goals present could produce interesting behaviors with communication - for example, an agent turning back towards an already-observed goal once it is communicated that its most-desired goal is not present anywhere. A scenario of more than two agents might pose an interesting inference test, where subjects are asked to predict which pair of agents were in communication.

Improved Inference Models

Based on our analysis and comparison of model and human predictions, it would be interesting to improve our inference models to account for human uncertainty when limited information is available during earlier time steps. Our communication inference model in particular is specialized to scenarios where communication can be somewhat definitively determined from action; it accepts the consequences of its observations with very high confidence. In simple settings, this approximates human behavior well. An improved model can account for observational uncertainty or inherent randomness in agent paths. For example, we can assign each agent some probability of randomly changing its goal or taking a random step regardless of communication. Then, the improved model could infer the probability of communication depending on the action or newly inferred goal.

Conclusion

In doing this project, we had the opportunity to apply many of the topics we learned in class this semester to a novel research question. Core ideas such as Bayesian modeling, Intuitive Psychology, and the experimental evaluation of human cognitive behavior were central to our approach.

We are proud of our final product, but it is important for us to reflect on the several challenges we faced and important lessons we learned along the way:

- **Collecting Human Data.** In our initial attempts to collect data, ambiguities in our problem context caused confusion among test subjects. We learned that conducting such trials effectively requires principled design and fine-tuning of instructions to ensure that humans are interpreting the scenario and questions the same way as our computational model.

- **Writing a Clear Project Proposal.** Our initial project proposal was not specific about the problem we intended to solve or the plan we had to solve it. In rewriting the proposal, we were forced to examine what research question our work would actually address, how it would be of importance or relevance to the community, and whether our proposed plan to solve it was feasible. This helped us improve our skills of research ideation, planning, and proposal.

- **Choosing an Interesting, Feasible Project.** Selecting a good project idea is an invaluable skill in research - the project is ideally realistic and somewhat incremental so as not to result in failure, yet interesting and novel enough that success is meaningful. We worked through several brainstorming sessions to identify a project that fulfilled these ideals within the scope of a semester course.

Contributions

Both team members collaborated equally on all major aspects of the project: ideation and experimental design, computational modeling, human data collection, and the write-up. Gabe's specific contributions to the project included design and implementation of inference algorithms, path visualization, and introduction / review of previous work. Kelly's specific contributions to the project included design and implementation of planning algorithms and presentation of results.

Acknowledgments

The authors would like to thank the 6.804 TAs for their hard work and mentorship; in particular, Tristan Thrush provided valuable feedback on our project proposal. The authors would also like to thank Josh Tenenbaum for inspiring this project through the stimulating lectures he delivered this semester.

References

Baker, C. L., Saxe, R. R., & Tenenbaum, J. B. (2011). Bayesian theory of mind: Modeling joint belief-desire attribution. *In Proceedings of the Third Annual Conference of the Cognitive Science Society*.

Baker, C. L., Tenenbaum, J. B., & Saxe, R. R. (2007). Goal inference as inverse planning. *In Proceedings of the Annual Meeting of the Cognitive Science Society*.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2016). Bulilding machines that learn and think like people. *Behavioral and Brain Sciences*.

Ullman, T. D., Baker, C. L., Macindoe, O., Evans, O., Goodman, N. D., & Tenenbaum, J. B. (2010). Help or hinder: Bayesian models of social goal inference. *Neural Information Processing Systems*.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*.

Appendix

The code repository for our plan generation and inference models can be found at: `https://github.com/gmargo11/intuitive-interaction.git`.