Goal Inference as Inverse Planning

Chris L. Baker, Joshua B. Tenenbaum & Rebecca R. Saxe
{clbaker,jbt,saxe}@mit.edu
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology

Abstract

Infants and adults are adept at inferring agents' goals from incomplete or ambiguous sequences of behavior. We propose a framework for goal inference based on *inverse planning*, in which observers invert a probabilistic generative model of goal-dependent plans to infer agents' goals. The inverse planning framework encompasses many specific models and representations; we present several specific models and test them in two behavioral experiments on online and retrospective goal inference.

Keywords: theory of mind; action understanding; Bayesian inference; Markov Decision Processes

Introduction

A woman is walking down the street, when suddenly she pauses, turns, and begins running in the opposite direction. Why? Is she crazy? Did she complete an errand unknown to us (perhaps dropping off a letter in a mailbox) and rush off to her next goal? Or did she change her mind about where she was going? These inferences derive from attributing goals to the woman and using them to explain her behavior.

Adults are experts at inferring agents' goals from observations of behavior. Often these observations are ambiguous or incomplete, yet we confidently make goal inferences from such data many times each day. Developmental psychologists have shown that infants also perform simple forms of goal inference. In experiments using live-action stimuli, Woodward found evidence that 6-month old infants attribute goals to human actors, and look longer when subsequent behavior is inconsistent with the old goal (1998). Meltzoff (1995) showed that 18-month olds imitate intended acts of human actors rather than accidental ones, and Csibra and colleagues found evidence that infants infer goals from incomplete trajectories of moving objects in simple two-dimensional animations (Csibra, Biró, Koós, & Gergely, 2003), both suggesting that children infer goals even from incomplete actions.

The apparent ease of goal inference masks a sophisticated probabilistic induction. There are typically many goals logically consistent with an agent's actions in a particular context, and the apparent complexity of others' actions invokes a confusing array of explanations, yet observers' inductive leaps to likely goals occur effortlessly and accurately. How is this feat of induction possible?

A possible solution, proposed by several philosophers and psychologists, is that these inferences are enabled by an intuitive theory of agency that embodies the *principle of rationality*: the assumption that rational agents tend to achieve their desires as optimally as possible, given their beliefs (Dennett, 1987; Gergely, Nádasdy, Csibra, & Biró, 1995). However, many authors (*e.g.* Nichols and Stich (2003); Baker, Tenenbaum, and Saxe (2006)) have argued that the qualitative descriptions of the principle of rationality that have been proposed are insufficient to account for the complexities of human goal inference. Further, the qualitative predictions of noncomputational models lack the resolution for fine-grained comparison with people's judgments.

Here, we propose a computational version of this approach to goal inference, in terms of inverse probabilistic planning. It is often said that "vision is inverse graphics": computational models of visual perception – particularly in the Bayesian tradition – often posit a causal physical process of how images are formed from scenes (*i.e.* "graphics"), and this process must be inverted in perceiving scene structure from images. By analogy, in inverse planning, planning is the process by which intentions cause behavior, and the observer infers an agent's intentions, given observations of an agent's behavior, by inverting a model of the agent's planning process. Like much work in computer vision, the inverse planning framework provides a rational analysis (Anderson, 1990) of goal inference. We hypothesize that people's intuitive theory of goal-dependent planning approximates scientific models of human decision making proposed by economists and psychologists, and that bottom-up information from inverting this theory, given observations of behavior, is integrated with top-down prior knowledge of the space of goals to allow rational Bayesian inference of goals from behavior.

The inverse planning framework includes many specific models that differ in the complexity they assign to the beliefs and desires of agents. Prior knowledge of the space of other agents' goals is necessary for induction, and in this paper, we will present and test several models that differ in their representations of goal structure. Our experimental paradigm tests each model with a wide range of action trajectories in a simple space for which our models make fine-grained predictions. (Our stimuli resemble those of Gergely et al. (1995)). Some of these stimuli display direct paths to salient goals, and have simple intentional interpretations. Other stimuli display more complex behaviors, which may not have simple intentional interpretations. These sorts of trajectories allow us to distinguish between alternative models that differ in their representation of complex goal structure. By varying the length of the trajectories, we measure how subjects' goal inferences change over time, and by eliciting both online and retrospective inferences, we measure how subjects integrate information over time.

To illustrate the space of models we present, consider the introductory example. Each of the three queries raised about the woman's goals correspond to a particular representation of goal structure that we test. The first model (M1) assumes a single invariant goal across a trajectory, and explains any deviation from the optimal behavior as noise, or bounded rationality. The second model (M2) assumes that agents can have subgoals along the way to their final goal, and is able to explain indirect paths. The third model (M3) assumes that agents' goals can change over time, and can also explain indirect paths or changes in direction.

The plan for the paper is as follows. We first describe our framework for inverse planning, and present three specific inverse planning models for goal inference. We then describe two new behavioral experiments designed to distinguish between our specific inverse planning models, and provide quantitative results of our model for each experiment.

Inverse planning framework

Although the definition of rationality has been left informal in prior work on intentional reasoning, formal models of rationality have been well developed in the field of decision theory. Markov Decision Problems (MDPs) are the standard formalism for sequential decision making, or planning, under uncertainty. Solving an MDP entails finding an optimal policy, or rule of action, that leads to the maximum expected discounted reward, given the environment. A rational agent is one that follows an optimal policy.

At its core, the inverse planning framework assumes that human observers represent other agents as rational planners solving MDPs. The causal process by which goals cause behavior is generated by probabilistic planning in MDPs with goal-dependent reward functions. Using Bayesian inference, this causal process can be integrated with prior knowledge of likely goal structures to yield a probability distribution over agents' goals given their behavior. Our framework builds on previous work by Baker et al. (2006) and Verma and Rao (2006), who propose similar inverse planning frameworks. Here we consider a wider range of hypothesis spaces for goal structures, and present the first quantitative tests of this framework as an account of human goal inference.

Let S be the set of agent states, let W be the set of environmental states, let G be the set of goals, and let A be the set of actions. Let $s_t \in S$ be the agent's state at time t, let $w \in W$ be the world state (assumed constant across trials), let $g \in G$ be the agent's goal, and let $a_t \in A$ be the agent's action at time t. Let $P(s_{t+1}|s_t, a_t, w)$ be the state transition distribution, which specifies the probability of of moving to state s_{t+1} from state s_t, as a result of action a_t, in world w. In general, the dynamics of state transitions depend on the environment, but for the stimuli considered in this paper, state transitions are assumed to yield the desired outcome deterministically.

Let $C_{g,w}(a,s)$ be the cost of taking action a in state s for an agent with goal g in world w. In general, cost functions may differ between agents and environments. For our 2D motion scenarios, action costs are assumed to be proportional to the negative length of the resulting movement (staying still incurs a cost as well). The goal state is absorbing and cost-free, meaning that the agent incurs no cost once it reaches the goal and stays there. Thus, rational agents will try to reach the goal state as quickly as possible.

The value function $V^{\pi}_{g,w}(s)$ is defined as the infinite-horizon expected cost to the agent of executing policy π starting from state s (with no discounting):

$$V^{\pi}_{g,w}(s) = E_{\pi}\left[\sum_{t=1}^{\infty}\sum_{a_t} P_{\pi}(a_t|s_t,g,w)C(a_t,s_t)\middle| s_1 = s\right]. \quad (1)$$

$Q^{\pi}_{g,w}(s_t,a_t) = \sum_{s_{t+1}} P(s_{t+1}|s_t,a_t)V^{\pi}_{g,w}(s_{t+1}) + C_{g,w}(a_t,s_t)$ is the state-action value function, which defines the infinite-horizon expected cost of taking action a_t from state s_t, with goal g, in world w, and executing policy π afterwards. The agent's probability distribution over actions associated with policy π is defined as $P_{\pi}(a_t|s_t,g,w) \propto \exp\big(\beta Q^{\pi}_{g,w}(s_t,a_t)\big)$, sometimes called a *Boltzmann policy*. The optimal Boltzmann policy and the value function of this policy can be computed efficiently using value iteration (Bertsekas, 2001). This policy embodies a "soft" principle of rationality, where the parameter β controls how likely the agent is to deviate from the rational path for unexplained reasons. The β parameter plays an important role in each of our models, weighing randomness against high-level goal structure, and we vary its value for each of our models to determine its effect on prediction in our experiments.

Next, we describe three candidate representations for people's prior knowledge about goals in our framework, roughly corresponding to the three kinds of explanations we offered for the woman's anomalous behavior in our introductory example. These candidate models, denoted M1(β), M2(β,κ), and M3(β,γ), are formalized in the subsections below.

Model 1: single underlying goal

Our first candidate model assumes that the agent has one underlying goal that it pursues across all timesteps. We denote this model M1(β). Unlike M2 and M3, this model must explain all deviations from the shortest path to the goal in terms of unlikely choices by the agent, governed by the parameter β. Given a state sequence of length T, the distribution over the agent's goal in this model is obtained using Bayes' rule:

$$P(g|s_{1:T},w) \propto P(s_{1:T}|g,w)P(g|w), \quad (2)$$

where $P(s_{1:T}|g,w) = \prod_{t=1}^{T-1} P(s_{t+1}|s_t,g,w)$. The probability of the next state s_{t+1}, given the current state s_t, the goal g, and the environment w, is computed by marginalizing over actions, which are only partially observable though their effect on the agent's state: $P(s_{t+1}|s_t,g,w) = \sum_{a_t \in A_{s_t}} P(s_{t+1}|s_t,a_t,w)P_{\pi}(a_t|s_t,g,w)$. M1($\beta$) is a special case of both M2(β,κ) and M3(β,γ), with κ and γ equal to 0.

Model 2: complex goals

The next model we consider is based on the complex goal model of Baker et al. (2006). We denote this model M2(β,κ).

In this model, with prior probability κ, the agent picks a complex goal, which includes the constraint that the agent must pass through a particular "via-point" on the way to its end goal. With prior probability $1-\kappa$ the agent picks a simple goal, which is just a single goal point, as in M1. Given the agent's type of goal (simple or complex), the distribution over goals within each type is assumed to be uniform. Inferences about an agent's end goal are obtained by marginalizing over goal types, and within the complex goal type, marginalizing over possible via-points.

The evidence that people represent and reason about complex goals (Baker et al., 2006) has lead us to consider M2 as a hypothesis for explaining people's goal inferences. However, in the stimuli we consider, there may be less evidence for complex goals than in the experiments of Baker et al. (2006). Thus, we vary the parameter κ to assess the effect of greater a priori probability of complex goals. The next model we consider also represents sequences of goals, but in a way that is more generic, and only depends on the agent's tendency to "change its mind".

Model 3: changing goals

Our final model assumes that agents' goals can change over time for reasons unknown to observers. This model takes the form of a Dynamic Bayes net, which we denote as M3(β,γ), where γ is the probability of changing goals. (In this section, we omit the conditional dependence of probability distributions on w for readability). Let k be the number of goals, and let $P(g_1)$ be the prior over initial goals at time t=1. $P(g_{t+1}|g_t)$ is the conditional distribution over changing to goal g_{t+1} at time $t+1$ given the goal g_t at time t:

$$P(g_{t+1}{=}i|g_t{=}j) = \begin{cases} 1-\gamma & \text{if } i = j \\ \gamma/(k-1) & \text{otherwise.} \end{cases} \quad (3)$$

When $\gamma = 0$, this model reduces to M1. When $\gamma = (k-1)/k$, the conditional distribution $P(g_{t+1}|g_t)$ is uniform; in this case the model is equivalent to choosing a new goal at random at each time step. Intermediate values of γ between 0 and $(k-1)/k$ interpolate between these extremes.

To compute the posterior distribution over goals at time t, given a state sequence $s_{1:t+1}$, we recursively define the *forward distribution*:

$$P(g_t|s_{1:t+1}) \propto P(s_{t+1}|g_t,s_t) \sum_{g_{t-1}} P(g_t|g_{t-1})P(g_{t-1}|s_{1:t}), \quad (4)$$

where the recursion is initialized with $P(g_1)$. This allows us to model subjects' online inferences in Experiment 1.

To compute the marginal probability of a goal at time t given $s_{1:T}$, $t{<}T$, we use a variant of the forward-backward algorithm. The forward distribution is defined by Eq. 4 above. The *backward distribution* is recursively defined by:

$$P(s_{t+2:T}|g_t,s_{1:t+1}) = \sum_{g_{t+1}} P(s_{t+2}|g_{t+1},s_{t+1})P(s_{t+3:T}|g_{t+1},s_{t+2})P(g_{t+1}|g_t). \quad (5)$$

The marginal probability of goal g_t given the state sequence $s_{1:T}$ is the product of the forward and backward messages:

$$P(g_t|s_{1:T}) \propto P(g_t|s_{1:t+1})P(s_{t+2:T}|g_t,s_{1:t+1}). \quad (6)$$

This distribution allows us to infer what the agent's goal was at time t, given its past movements from 1:t+1, and future movements from time t+1:T, allowing us to model subjects' retrospective inferences in Experiment 2. The parameter γ plays a key role in how information from the past and future is integrated into the distribution over current goals. When $\gamma = (k-1)/k$, past and future movements carry no information about the current goal. When $\gamma = 0$, changing goals is prohibited, and future information constrains the probability of all past goals to be equal to $P(g_{T-1}|s_{1:T})$. For each experiment, we tested the range of predictions of M3 across this parameter space.

Experiments

As candidate models for our experiments, we tested each of M1, M2, and M3 with β values in $\{0.25,0.5,1.0,1.5,2.0,4.0\}$. For M2 and M3, which each have an additional free parameter, we tested a range of values for these parameters as well. These values are listed in Tables 1 and 2. For M2, we omit the full range of β values from Tables 1 and 2 for readability.

Experiment 1

Our first experiment tested the power of our alternative models to predict people's judgments in a task of inferring agents' goals from observations of partial action sequences.

Participants Participants were 16 members of the MIT community.

Materials and Procedure Subjects were told they would watch 2D videos of intelligent aliens moving around in simple environments with visible obstacles, with goals marked by capital letters.

There were 100 stimuli in total. An illustrative subset is shown in Fig. 1(a). Each stimulus contained 3 goals. There were 4 different goal configurations, and two different obstacle conditions: gap and solid, for a total of 8 different environments. There were 11 different complete paths: two paths headed toward 'A', two paths headed toward 'B', and 7 paths headed toward 'C' (to account for C's varying location). Partial segments of these paths starting from the beginning were shown in each different environment. Because many of the paths were initially identical, and because many of the paths were not possible in certain environments (i.e. collided with walls), the total number of unique stimuli was reduced to 100.

Stimuli were presented shortest lengths first in order to not bias subjects toward particular outcomes. Stimuli of the same length were shown in random order. After each stimulus presentation, subjects were asked to rate which goal they thought was most likely (or if two or more were equally likely, to pick one of the most likely). After this choice, subjects were asked to rate the likelihood of the other goals relative to the most

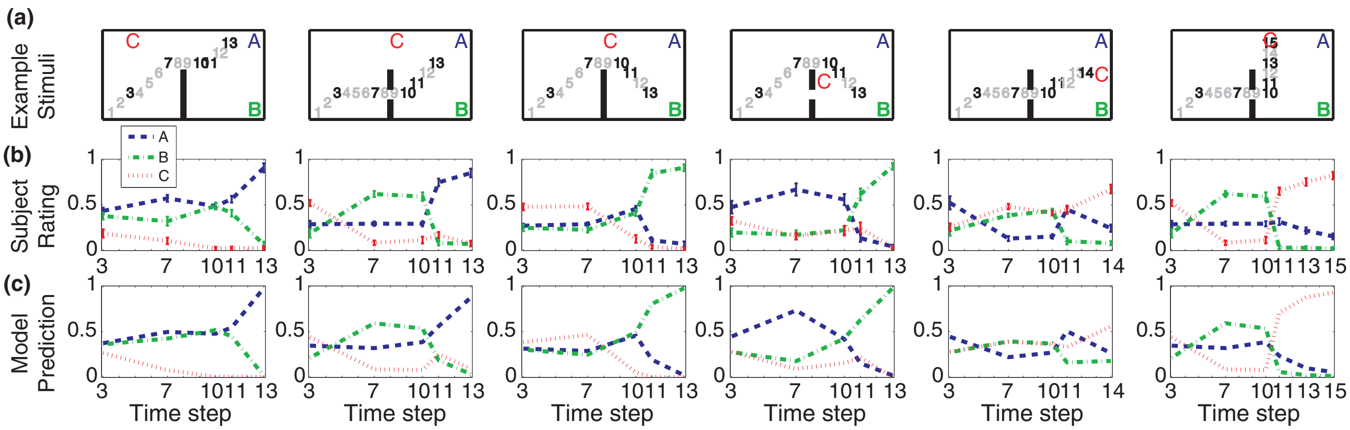

Figure 1: Experiment 1. (a) Example stimuli. Plots show all 4 goal conditions and both obstacle conditions. Both 'A' paths are shown, one of two 'B' paths is shown and 2 of 7 'C' paths are shown. Dark colored numbers indicate displayed lengths. (b) Average subject ratings with standard error bars for above stimuli. (c) Model predictions. Model predictions closely match people's ratings. Displayed model: M3(1.5,.5).

Model													
M1(β)		M2(1.5,κ)		M2(2.0,κ)		M3(β,0.67)		M3(β,0.50)		M3(β,0.25)		M3(β,0.10)	
β	r	κ	r	κ	r	β	r	β	r	β	r	β	r
0.25	0.82	0.10	0.89	0.10	0.89	0.25	0.92	0.25	0.92	0.25	0.92	0.25	0.91
0.5	***0.83***	0.25	0.90	0.25	0.91	0.5	0.93	0.5	0.94	0.5	0.95	0.5	0.94
1.0	0.82	0.50	0.92	0.50	0.93	1.0	0.95	1.0	0.95	1.0	0.96	1.0	0.96
1.5	0.82	0.67	0.92	0.67	0.93	1.5	0.96	1.5	0.96	1.5	0.97	1.5	0.97
2.0	0.81	0.75	0.93	0.75	0.94	2.0	***0.96***	2.0	***0.97***	2.0	***0.98***	2.0	***0.98***
4.0	0.78	0.90	***0.93***	0.90	***0.94***	4.0	0.96	4.0	0.96	4.0	0.96	4.0	0.96

Table 1: Experiment 1 results. M1, M2 and M3 were tested with β values in $\{0.25, 0.5, 1.0, 1.5, 2.0, 4.0\}$. Odd columns contain parameter settings for the various models. Even columns contain r-values of the various models' correlations with people's ratings. For M2 and M3, which each have an additional free parameter, we tested a range of values for these parameters as well. We omit the full range of β values for M2 for readability. The M3(β, 0.67) column corresponds to the condition in which a new goal is sampled at random at each time step.

likely goal, on a 9-point scale from "Equally likely", to "Half as likely", to "Extremely unlikely". Ratings were normalized to sum to 1 for each stimulus, then averaged across all subjects and renormalized to sum to 1. Example subject ratings are plotted with standard error bars in Fig. 1(b).

Each model makes strong predictions about people's ratings in this experiment. If M1 is correct, then people should weigh evidence from old and recent movements equally, and react slowly to new evidence that conflicts with past evidence. Conversely, M2 and M3 predict that people should react quickly to recent movements strongly indicating a particular goal. M2 achieves this by inferring that a subgoal has been reached, and that recent movements reflect the end goal. M3 achieves this by inferring that the agent has changed its goal. Example model predictions from M3$(1.5, 0.5)$ are plotted in Fig. 1(c); these match subjects' ratings very closely.

Results The results of Experiment 1 are summarized in Table 1. All instances of M3 correlate highly with subjects' ratings, indicating that subjects were quick to respond to evidence of a new goal. Because of this, M2 also correlates highly with people's judgments. M1 clearly does a poorer job of predicting people's judgments. Fig. 3(a) shows scatter plots of model predictions versus subject ratings for the model with the highest correlation from each class. Although the predictions of M3 correlate slightly higher with subjects' ratings than the predictions of M2, only M1 is ruled out as a viable model for people's judgments in this experiment.

Experiment 2

Our second experiment sought to provide a context within which predictions of M2 and M3 could be distinguished. We showed subjects long trajectories and asked them to make retrospective judgments about the agent's goal at several earlier points in the action sequence. As explained below, in cases where the early and late stages of a trajectory are locally best explained by different goals, only the changing goal model (M3) predicts that people's retrospective goal inferences will vary accordingly.

Participants Participants were 16 members of the MIT community (distinct from the first group).

Materials and Procedure The procedure of Experiment 2 was similar to that of Experiment 1, except now subjects were told they would see an alien's movement, and that after this movement, an earlier point along the alien's path would be marked. Subjects were told they would then be asked to indicate "which goal the alien had in mind" at the marked point, in light of the entire subsequent path they observed.

There were 95 stimuli in this experiment. Stimuli were taken from Experiment 1 as follows. Each path from each environment was used. However, only paths of maximal length were displayed. The marked points were taken to be evenly

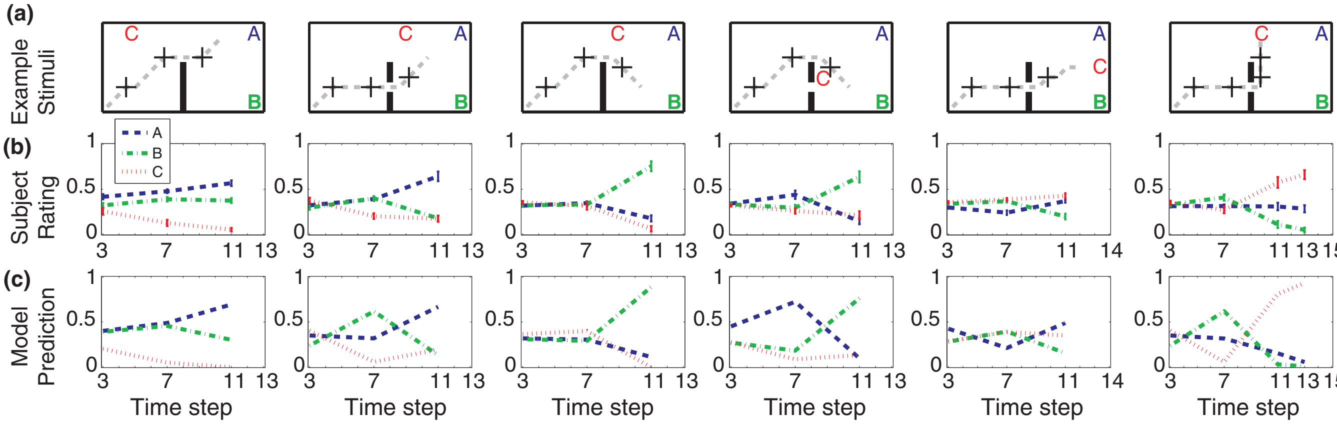

Figure 2: Experiment 2. (a) Example stimuli. Dashed line corresponds to the movement subjects saw prior to rating the likelihood of each goal at the marked point. Black +'s correspond to test points in the stimuli. Compare to corresponding column of Fig. 1. (b) Subjects' ratings: compare to Fig. 1. (c) Model predictions. Displayed model: M3(1.5,0.5).

Model													
M1(β)		M2(1.5,κ)		M2(2.0,κ)		M3(β,0.67)		M3(β,0.50)		M3(β,0.25)		M3(β,0.10)	
β	*r*	κ	*r*	κ	*r*	β	*r*	β	*r*	β	*r*	β	*r*
0.25	***0.60***	0.10	0.78	0.10	0.77	0.25	0.91	0.25	0.95	0.25	***0.94***	0.25	***0.87***
0.5	0.58	0.25	0.78	0.25	0.77	0.5	0.91	0.5	***0.95***	0.5	0.93	0.5	0.86
1.0	0.57	0.50	0.78	0.50	0.77	1.0	***0.92***	1.0	0.94	1.0	0.92	1.0	0.84
1.5	0.56	0.67	0.78	0.67	0.77	1.5	0.92	1.5	0.93	1.5	0.90	1.5	0.83
2.0	0.56	0.75	0.78	0.75	0.77	2.0	0.91	2.0	0.92	2.0	0.88	2.0	0.82
4.0	0.55	0.90	***0.78***	0.90	***0.77***	4.0	0.90	4.0	0.90	4.0	0.86	4.0	0.78

Table 2: Experiment 2 results. M1, M2 and M3 were tested with β values in {0.25,0.5,1.0,1.5,2.0,4.0}. Odd columns contain parameter settings for the various models. Even columns contain *r*-values of the various models' correlations with people's ratings. For M2 and M3, which each have an additional free parameter, we tested a range of values for these parameters as well. We omit the full range of β values for M2 for readability. The M3(β,0.67) column corresponds to the condition in which a new goal is sampled at random at each time step.

spaced points corresponding to shorter path lengths previously displayed in Experiment 1. Thus, each rating in Experiment 2 had a corresponding rating in Experiment 1. Fig. 2(a) shows illustrative stimuli that directly correspond to the stimuli in Fig. 1(a).

In this experiment, M2 and M3 make opposite predictions. M2 (and M1) predict that the strong evidence provided for particular end goals by the long paths shown in Experiment 2 will have a large effect on people's inferences about agents' goals earlier in their paths. This is because both M1 and M2 assume the agent's goal (simple or complex) to be constant throughout its path. Thus, if an agent gives a strong indication of pursuing a particular goal at the end of its path, these models assume that it must have been headed for this goal all along. M3, however, uses Eq. 6 to compute the probability of the agent's goal at the marked point. As discussed earlier, the parameter γ controls how much information the distribution over goals in the future provides about goals in the past. M3 predicts that people's prior assumptions about the likelihood of changing goals will have a large effect on their judgments. If γ is close to $(k-1)/k$, M3 predicts that people's ratings in Experiment 2 will correlate highly with the corresponding ratings from Experiment 1.

Results The results of Experiment 2 are summarized in Table 2. M1 continues to perform poorly. Now, however, M2 is also a relatively poor predictor of people's judgments, while M3 continues to correlate most highly with people's judgments. Interestingly, as predicted earlier, the correlation of people's ratings in Experiment 2 with people's ratings from the corresponding stimuli from Experiment 1 was 0.89; fairly high given the difference in tasks. Fig. 3(b) shows scatter plots of model predictions versus subject ratings for the model with the highest correlation in each class.

Discussion

The high correlations between our models and subjects' predictions from Experiment 1 and Experiment 2 provide strong quantitative evidence in support of the inverse planning framework. These results also provide support for the goal structures of M3 as plausible representations for human goal inference. However, the lower correlations of M2 with subjects' predictions from Experiment 2 do not rule out subgoals as a possible goal structure representation. Subgoals could be useful in some cases, such as in our earlier work, where we showed that people's action predictions are well-explained by M2 when an agent persistently pursues complex goals (Baker et al., 2006). As mentioned previously, the stimuli used in the current paper provide less evidence for subgoals than the stimuli used in Baker et al. (2006). People's use of different models to explain and reason about different data might be captured by a hierarchical Bayesian model that incorporates both M2 and M3 as submodels, as well as many

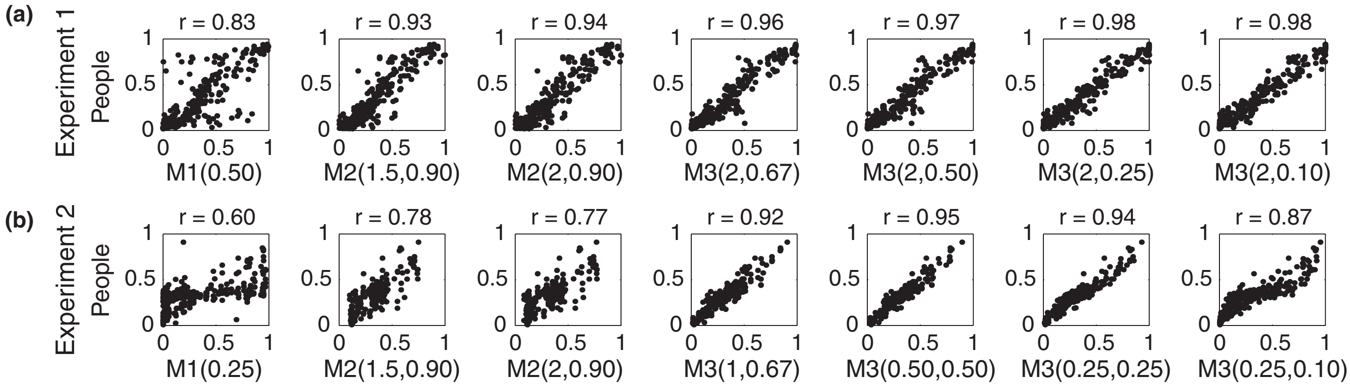

Figure 3: Example scatter plots of model predictions against subject ratings. Plots of model predictions use the parameter settings with the highest correlation from each model column of Tables 1 and 2. (a) Experiment 1 results. (b) Experiment 2 results.

other submodels with different goal representations, since M2 and M3 merely scratch the surface of possible goal structures.

Although our results suggest that people explain apparent deviations from the most direct path by inferring a change in goals, inferring that an agent's goal has changed does not account for all the complexities of behavior – the parameter β also plays an important role in generating accurate predictions of subjects' ratings in our model. Thus, people's goal inferences in our experiments reflect a tradeoff between explaining complex behavior in terms of unlikely deviations from the shortest path and attributing a change in goals.

Two well-known qualitative accounts of action understanding, theory-theory (Gopnik & Meltzoff, 1997) and simulation theory (Goldman, 2006), can be seen as cases of inverse planning. On a theory-theory interpretation, inverse planning consists of inverting a theory of rational action to arrive at a set of goals that could have generated the observed behavior, and inferring individual goals based on prior knowledge of the kinds of goals the observed agent prefers. On a simulation theory account, goal inference is performed by inverting one's own planning process to narrow down the set of goals that could have generated the observed behavior, and inferring individual goals from this set according to their desirability under one's own preferences.

Unlike previous proposals, our computational framework shows precisely how to integrate top-down prior knowledge about goals with bottom-up observations of behavior using Bayesian inference. Similar ideas have been sketched out qualitatively, but our inverse-planning models are the first to quantitatively describe people's probabilistic goal inferences, and to explain rationally how these inductive leaps can be successful given only sparse, incomplete observation sequences.

Conclusion

We presented a computational framework for explaining people's goal inferences based on inverse planning. Within this framework, we presented three specific inverse planning models, each using a different representation of goal structure. We tested these models with two novel experiments designed to distinguish between the different goal representations of the models. Our experiments yielded high resolution data, which provided empirical support for the inverse planning framework, and gave quantitative evidence for our changing goals model as a plausible goal representation. It will be important to test whether inverse planning will generalize to explain human goal inference outside of the laboratory, but we believe that the power of the rationality assumption, combined with rich representations of goal structure, can account for much of people's everyday reasoning about agents' actions and goals.

Acknowledgments: This work was supported by AFOSR MURI contract FA9550-05-1-0321, the James S. McDonnell Foundation Causal Learning Collaborative Initiative, the Department of Homeland Security Graduate Fellowship (CLB) and the Paul E. Newton Chair (JBT).

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Baker, C. L., Tenenbaum, J. B., & Saxe, R. R. (2006). Bayesian models of human action understanding. *Advances in Neural Information Processing Systems*, Vol. 18 (pp. 99–106).

Bertsekas, D. P. (2001). *Dynamic programming and optimal control*. Belmont, MA: Athena Scientific, 2nd edition.

Csibra, G., Biró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*, 111–133.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Gergely, G., Nádasdy, Z., Csibra, G., & Biró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.

Goldman, A. I. (2006). *Simulating minds*. Oxford University Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*(5), 838–850.

Nichols, S., & Stich, S. (2003). *Mindreading*. Oxford University Press.

Verma, D., & Rao, R. (2006). Goal-based imitation as probabilistic inference over graphical models. *Advances in Neural Information Processing Systems*, Vol. 18 (pp. 1393–1400).

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