REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
06-07-2017	Final Report	28-Sep-2015 - 27-Mar-2017

4. TITLE AND SUBTITLE
Final Report: Modeling Social Common Sense for Seamless Human-Machine Teaming: Inverting the "Intuitive Game Engine" with Probabilistic Programming"

5a. CONTRACT NUMBER
W911NF-15-1-0639

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS
Chris Baker, Tao Gao, Vikash Kumar, Emanuel Todorov, Joshua Tenenbaum

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Massachusetts Institute of Technology (MIT
77 Massachusetts Ave.
NE18-901
Cambridge, MA 02139 -4307

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
68239-LS-DRP.4

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Humans use commonsense knowledge of agents and the physical world to solve problems interactively. We model people's knowledge of agents and the world using rich "game-engine"-based models of 3D physical scenes with realistic physical dynamics and agents autonomously acting and interacting with others, based on individual mental states and shared tasks. Cast as a probabilistic program, this intuitive game engine can be inverted to support inferences about other's beliefs, desires, goals, and tasks, which are vital for successful social interaction. We show the promise of this approach in application to modeling human inferences of the targets of the observed reaching

15. SUBJECT TERMS
cognitive science, artificial intelligence, robotics, human-robot teaming, social cognition, social intelligence, theory of mind, coordination, cooperation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Joshua Tenenbaum
UU	UU	UU	UU		19b. TELEPHONE NUMBER 617-452-2010

Report Title

Final Report: Modeling Social Common Sense for Seamless Human-Machine Teaming: Inverting the "Intuitive Game Engine" with Probabilistic Programming"

ABSTRACT

Humans use commonsense knowledge of agents and the physical world to solve problems interactively. We model people's knowledge of agents and the world using rich "game-engine"-based models of 3D physical scenes with realistic physical dynamics and agents autonomously acting and interacting with others, based on individual mental states and shared tasks. Cast as a probabilistic program, this intuitive game engine can be inverted to support inferences about other's beliefs, desires, goals, and tasks, which are vital for successful social interaction. We show the promise of this approach in application to modeling human inferences of the targets of the observed reaching actions of others.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received		Paper
07/05/2017	1	Chris L. Baker, Julian Jara-Ettinger, Rebecca Saxe, Joshua B. Tenenbaum. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing, Nature Human Behaviour, ( ): 0064. doi:
		1,049,852.00
TOTAL:	**1**	

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received		Paper
TOTAL:		

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper
07/06/2017	3 Max Kleiman-Weiner, Mark K. Ho, Joseph L. Austerweil, Michael L. Littman, Joshua B. Tenenbaum. Coordinate to cooperate or compete: abstract goals and joint intentions in social interaction, 38th Annual Conference of the Cognitive Science Society. 10-AUG-16, Philadelphia, PA. : ,
07/06/2017	2 Vikash Kumar, Emanuel Todorov, Sergey Levine. Optimal Control with Learned Local Models: Application to Dexterous Manipulation, EEE International Conference on Robotics and Automation. 16-MAY-16, Stockholm, Sweden. : ,
TOTAL:	**2**

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper
TOTAL:	

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received	Paper
TOTAL:	

Number of Manuscripts:

Books

Received Book

TOTAL:

Received Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED	DISCIPLINE
Vikash Kumar	0	
FTE Equivalent:	**0.00**	
Total Number:	**1**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
Chris Baker	0.00
FTE Equivalent:	**0.00**
Total Number:	**1**

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Joshua Tenenbaum	0.00	
Emanuel Todorov	0.00	
FTE Equivalent:	**0.00**	
Total Number:	**2**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:..... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):..... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: 0.00

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME
Total Number:

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

See attachment.

Technology Transfer

Modeling Social Common Sense for Seamless Human Machine Teaming:
Inverting the "Intuitive Game Engine" with Probabilistic Programming
Final Report

Introduction

Humans need rich knowledge of the world and other people to succeed in everyday life. Our work is based on the premise that this knowledge can be modeled as an intuitive version of modern computer game engines, capable of representing 3D physical scenes with realistic physical dynamics and characters autonomously acting and interacting with others, based on individual mental states and shared tasks. Cast as a probabilistic program, this intuitive game engine can be inverted to support inferences about other's beliefs, desires, goals, and tasks, which are vital for successful social interaction. Endowing machines with probabilistic programs for inverting the intuitive game engine will enable robot teammates that are sensitive to social context, and that can coordinate with humans on shared goals and tasks.

Modern computer game engines have many components, but our focus here is on two key modules: a physics simulator, and a mechanism for rational planning. Our physics models allow simulation of mass and forces to capture the dynamics of the world. Our models of planning are based on the assumption that other people plan efficient actions given their beliefs, desires, goals and the physical laws of the world. Combining models of physics and planning allows us to capture notions of ability, effort, and cost which differ greatly between humans and robots, and are thus essential for facilitating human-robot teamwork.

Our research program has followed several thrusts, each detailed below.

Inferring the intent of embodied motion

The aim of this thrust is to infer the goal of a human actor from video (RGB-D data from a Microsoft Kinect) of their actions and environment. This will enable a robot to predict a human teammate's future movements, and to intervene in ways the human finds helpful.

Technical approach

We use the Mujoco (Multiple Joint Control) physics engine of Emo Todorov and the UW Movement Control Laboratory to build an approximate physical model of a human agent and the structure of the task they are performing.

The model of planning we use is based on non-linear trajectory optimization. We use an implementation of the iterative Linear-Quadratic Gaussian control algorithm by Professor Todorov's group to synthesize trajectories given a physical model and a task-specific cost function defined in MuJoCo.

We embed the intuitive game engine consisting of the physics and planning models within a probabilistic program. The probabilistic program assumes that physics and planning are the processes that generate a human actor's observed actions, conditioned on the physical model of the world, and a physical model of the human actor. We assume that the actions are sampled from a Gaussian distribution surrounding the optimized trajectory, resulting in a distribution over

trajectories, which can be used to score observations. We use Bayes' rule to invert this probabilistic program to infer the actor's goal:

P(Goal|Trajectory,World) = P(Trajectory|Goal,World) P(Goal|World) / P(Trajectory|World). (1)

Fig. 1 shows this probabilistic program, specialized for the goal-directed reaching setting we describe below.

In order to seamlessly interact with humans in productive teams, robots must be able to make graded inferences about the goals that guide human actions. To capture this, we build probabilistic programs that can weigh the evidence for and against different possible goals using Bayesian inference. We consider inferences about the targets of human reaching actions. Our probabilistic programs use models analogous to an "intuitive game engine", based on physically-realistic simulations of reaching actions of humanoid robots. We use the MuJoCo simulation engine, developed by Emanuel Todorov's lab at the University of Washington to simulate reaching trajectories. Given a goal and an initial state, the probability distribution over a reaching trajectory of length *t* specified by the model is:

$$P(x_{1:t}, u_{1:t-1} | x_0, g, \varepsilon).$$

To compare our simulated trajectories with noisy, real-world observations requires us to model stochastic observations of the state sequence using an extended Kalman filter (EKF). We assume that at time *t*, an observation y_t is generated conditioned on x_t such that $y_t = h(x_t) + m_t$, where $m_t \sim \mathcal{N}(0, \Sigma_y)$ is normally distributed. Furthermore, we assume that the state transitions are Markovian and normally distributed: $x_{t+1} = f(x_t, u_t) + n_t$, where $n_t \sim \mathcal{N}(0, \Sigma_y)$ is normally distributed. Because the dynamics of our complex robotic model are nonlinear, we linearize the model around x_t, u_t to perform EKF estimates, such that H = $\partial h(x_t)/\partial x_t$ and F = $\partial f(x_t, u_t)/\partial x_t$. We can then treat these as the observation and state transition matrices of a linear Kalman filter at each time step. The MuJoCo physics engine makes computing these linearizations simple and efficient, by providing functions to inexpensively compute Jacobians of the dynamics around any site on the model.

Estimating the probability of a particular goal, given an observation sequence, uses Bayes' rule:

$$P(g|y_{0:t}, u_{1:t-1}, \varepsilon) \propto P(y_{0:t}, x_{1:t}, u_{1:t-1} | x_0, g, \varepsilon) P(g)$$
$$= P(y_{0:t}|x_{0:t}) P(x_{1:t}, u_{1:t-1} | x_0, g, \varepsilon) P(g).$$

We can compute these probabilities using the EKF and exploiting the conjugacy properties of the normal distribution.

We use Microsoft Kinect to process the world and observed human actions into a format compatible with MuJoCo, iLQG, and the probabilistic program. First, we extract the structure of the task – the physical dimensions of the human actor, and the physical configuration of the environment, including a tabletop with an array of objects placed on top. These can be used as input to directly define a MuJoCo model. For each object in the array, we define a different cost function, with that object as the goal. Currently, this process includes several manual steps, but it could be automated using computer vision algorithms for object and scene recognition. Next, we use the dynamic estimate of the human actor's skeleton given by Kinect as the observation of the action. This can be directly compared with the trajectory output by MuJoCo and iLQG. The probabilistic program for inference of the intent of human reaching actions generates hypothetical

plans and trajectory distributions for each potential intended goal. Observed human reaching trajectories are compared against each hypothesis, yielding scores which are integrated with the prior to obtain a posterior distribution over goals (see Fig. 1(a) and Equation 1).

Results

Fig. 1(a) shows the probabilistic program and MuJoCo simulation of our experimental setting. The height of the tabletop and the location of 16 target objects are initialized from the Kinect RGB-D stream. On observing a trajectory, the probabilistic program is inverted to infer the goal that generated the actions.

Fig. 1(b) shows several example model inferences. These examples use synthetic data, generated by conditioning the model on a particular target and synthesizing MuJoCo reaching trajectories for that target, with Gaussian noise added to the reaching controls. These synthetic observations are used as input to the model, which infers the goal of each noisy trajectory as it progresses. The model infers the correct goal in each example, and these inferences illustrate the ambiguity inherent in the task of inferring goals from behavior. In each case, the observed actions are initially consistent with many goals, but the model predicts that the goal is harder to infer for some trajectories than others. The model infers Red_3 after the fewest frames, and Blue_3 after the most frames. It infers that Blue_4 will be confused with the correct goal Red_3 about midway through trajectory 3, and that White_3, then Blue_2, then White_4 will be confused with the correct goal Blue_3 for trajectory 2.

Fig. 1(c) shows a further set of examples, from a simulated "car parking" domain. These results use a different context and environment, with a different kind of "body", but the structure of the probabilistic program, and the planning algorithm for generating hypothesized plans and trajectories, iLQG, are identical. This shows that our approach is not specific to humanoid robots, but can be naturally applied to other settings by changing the model of physics, the contextual input, and the observations to the model. The model infers the correct goal of the observed actions for each example, and these inferences have interesting structure – the model predicts that the intended parking spot is initially ambiguous, but eventually resolves at the specific point in each trajectory when the path is no longer consistent with alternative goals.

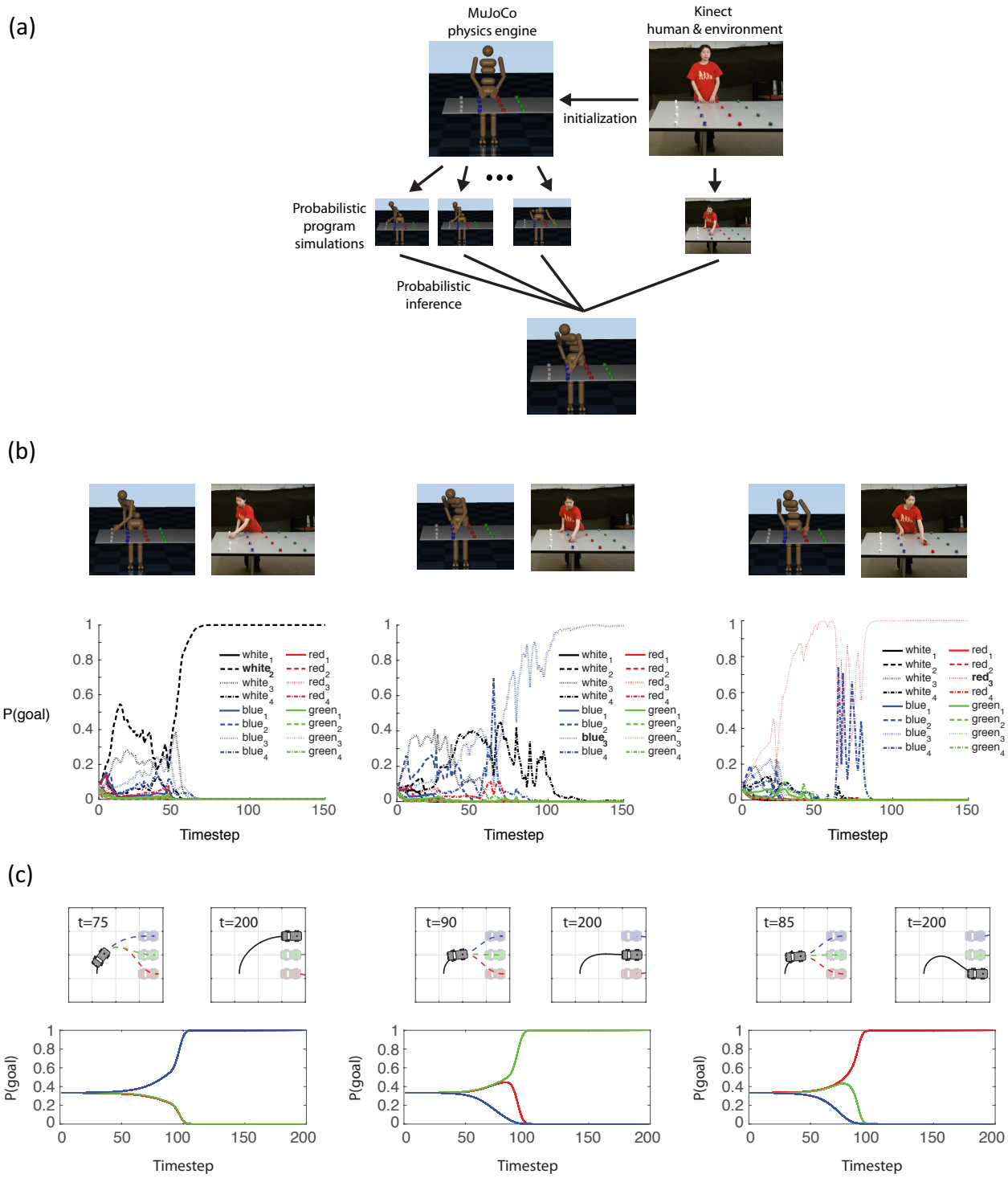


Figure 1: Inferring the intent of motion. **(a)** Structure of our probabilistic programming framework for intent inference. **(b)** Example results of inferring the intent of bodily motion. The model infers the correct goal in each case. **(c)** Example results of inferring a car's intent space in a parking simulation.

Next, we systematically explore the predictions of the model in a range of reaching scenarios. This experiment used a 6 degree-of-freedom robot model, reaching for target objects in a 4x4 array on a tabletop, shown in Fig. 2. Five stochastic repetitions of reaches to each target were generated to characterize the variation across trials.

The model predicts considerable variation across different reaches to the same target for some, but not all targets. For example, Figure 3 compares model inferences across the 5 example reaches for the targets in rows 1 and 4 of column 2 (targets 5 and 8), respectively. The model inferences for row 1 are much quicker, and also less variable, than those for row 4. This pattern is intuitive – evidence for nearer targets accumulates more quickly than for targets that are further away. During the extension of the arm toward a far-away target, multiple different judgments are possible. Nevertheless, inferences in some trials resolve toward the correct goal fairly quickly, while others take nearly the entire trial to converge to the correct target.

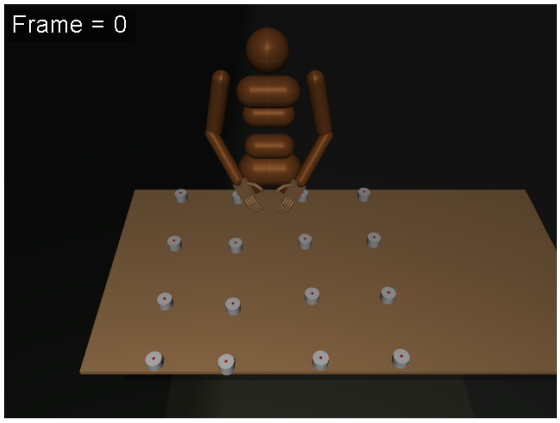


Figure 2. The humanoid robot model and physical context of our experiment. Sixteen targets objects are arrayed in a 4x4 grid. The robot reaches for 1 of the 16 objects with its left hand. Each reaching trial lasts 120 frames.

Figure 4 focuses on a single reach to the target in row 4 of column 2. Heatmaps of the distribution of model inferences are shown at three frames of the video: frames 25, 35 and 45. A different inference is made at each point: first, the target in row 3 of column 1, then the target in row 4 of column 1, then the correct target. In this example and others, the targets inferred in error by the model tend to be close or adjacent to the correct targets.

To assess our earlier observation that closer targets are inferred faster, we performed a two-way ANOVA on the speed with which the model inferred the correct target. The first factor was the row of the target, and the second factor was the "side" of the target: "right" or "left". We hypothesized that reaches to the right side would produce slower inferences, because the reach was performed with the left hand, and thus reaches to the right side went across the body. We found that the effect of row was highly significant ($F(3,79)=12.53$, $p<0.00001$). We also found that the effect of side was significant ($F(1,79)=4.44$, $p<0.05$), with reaches to the left side producing faster inferences.

Figure 5 shows inferences from single reaches to all 16 targets. The strong effect that closer targets are inferred faster is readily apparent (with some exceptions due to random variation in the reaching trajectories). The difference in timing between reaches to left side and reaches to the right side (across the body) is more variable.

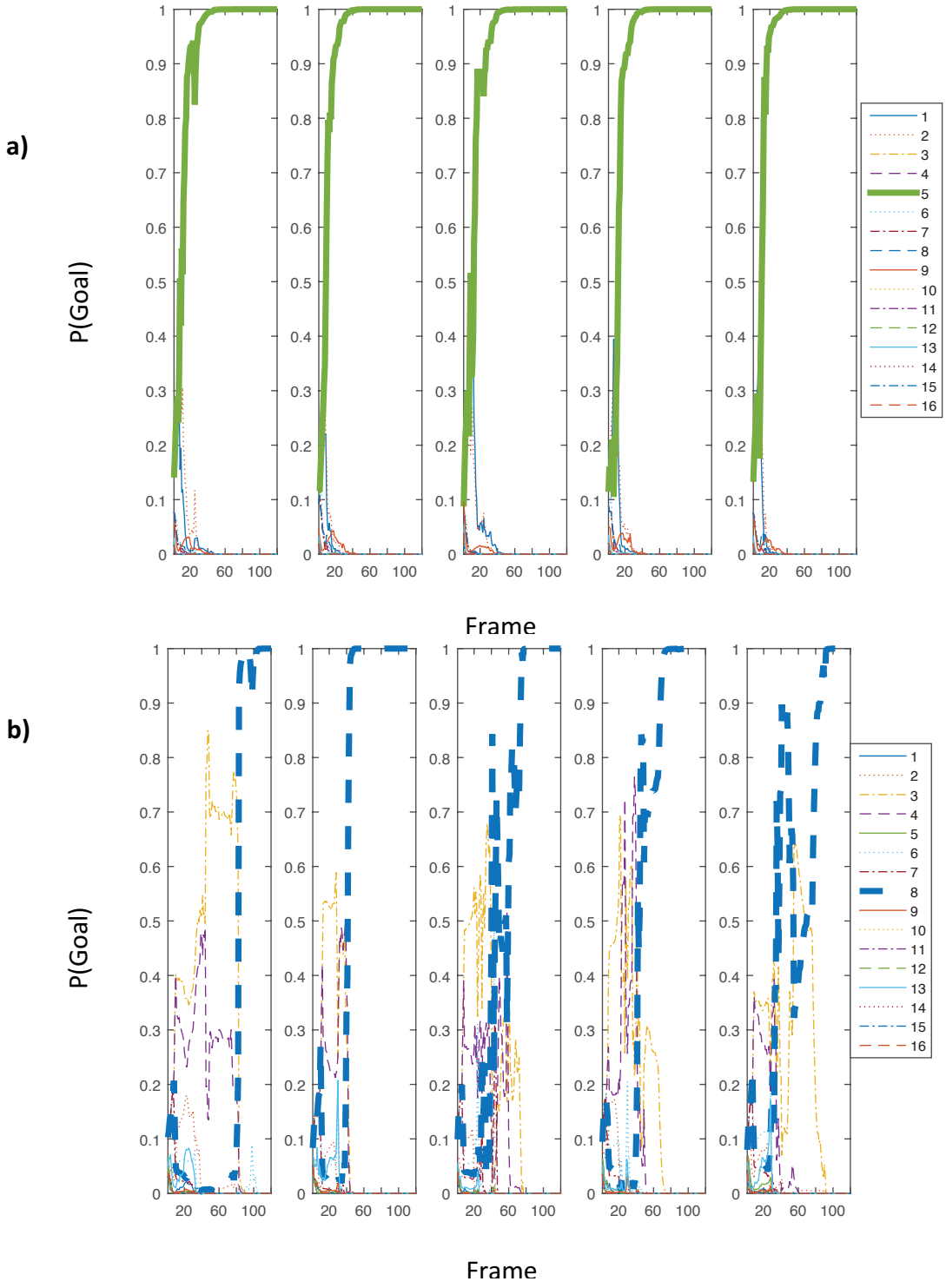


Figure 3. Model predictions across 5 reaches to the targets in rows 1 and 4 of column 2 (targets 5 and 8). The probability of the correct inference is plotted with a bold line. The inferences given the reaches to the target in row 1 of column 2 (a) converge to the correct target faster, and are less variable overall, than those given the reaches to the target in row 4 of column 2 (b).

Taken together, these model inferences provide a rich set of qualitative and quantitative predictions that can be compared with human judgments. At the individual trial level, the fine-grained dynamics of the target inferences of the model can be compared directly with human data. At the level of different targets, the qualitative effects that closer targets are inferred faster, and that reaching across the body produces slower goal inferences can both be tested experimentally.

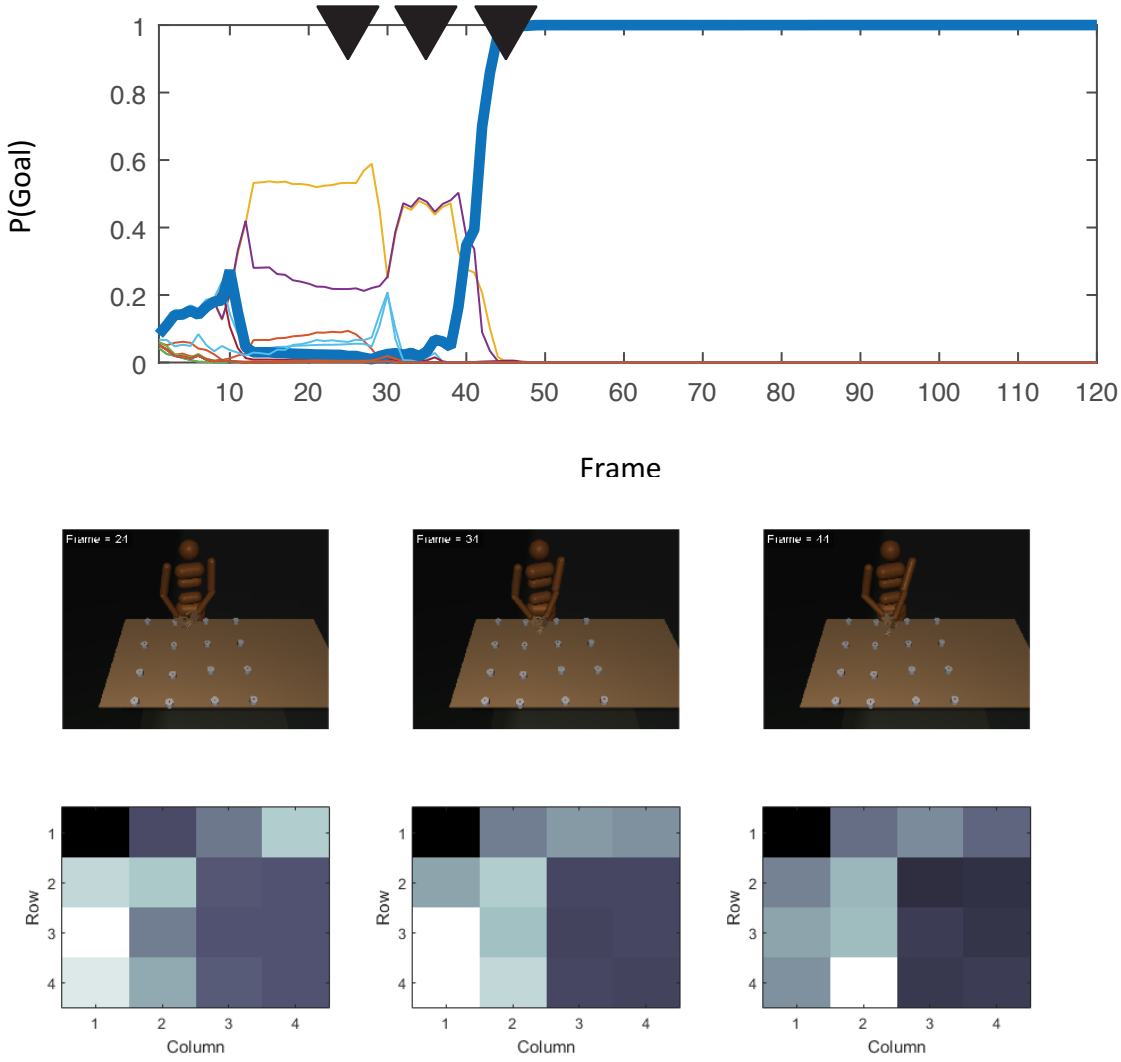


Figure 4. Model predictions at three points during one reach to the target in row 4, column 2.

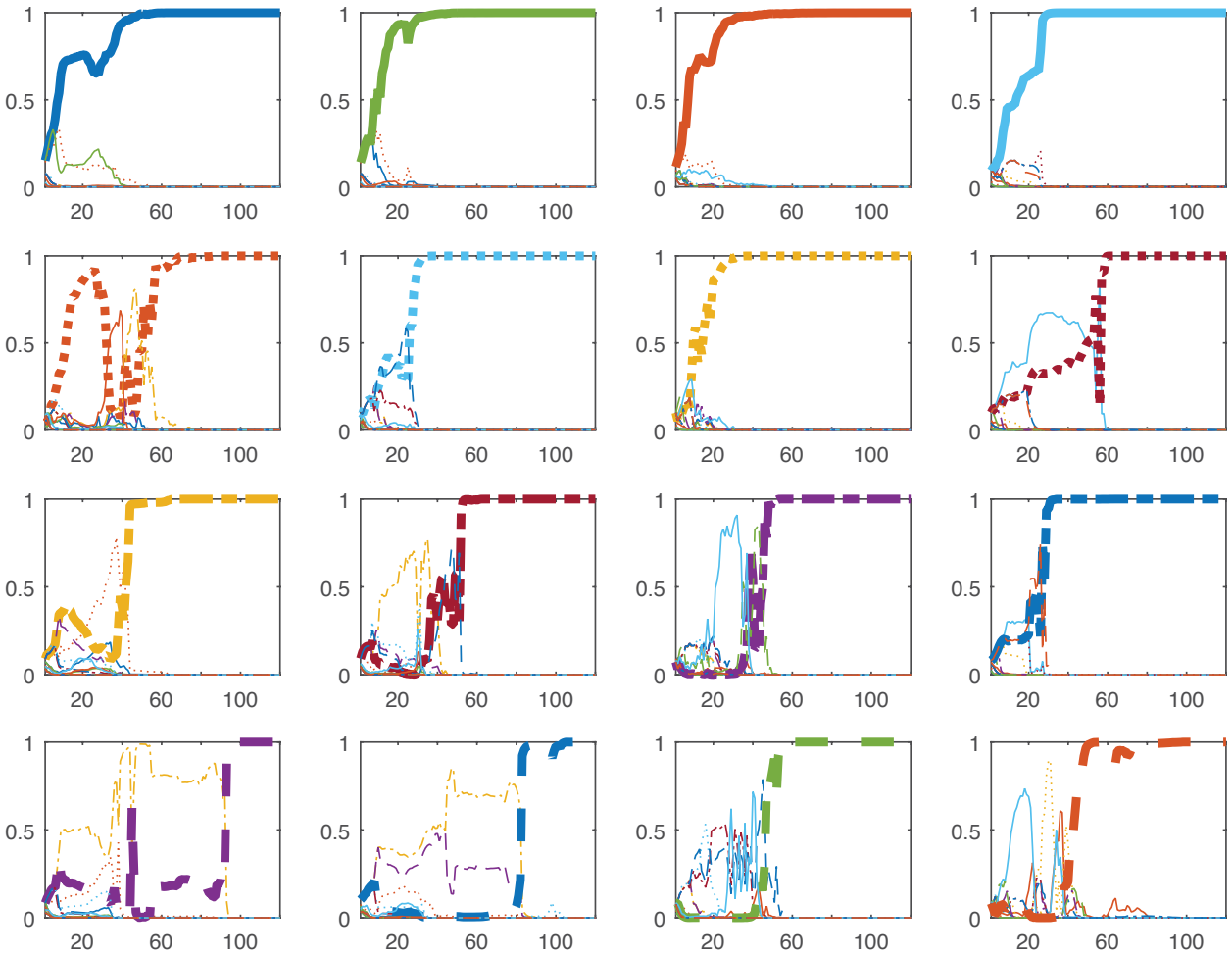


Figure 5. Model predictions for single reaches to each of the 16 targets. In each plot, the probability of the correct inference is plotted with a bold line. The location of each subplot (i.e., row, column) corresponds to the target locations in the grid shown in Fig. 2.

Evaluation

The stimuli we use to evaluate our approach consist of Microsoft Kinect RGB-D recordings of the reaching actions of 4 human participants in the setting shown in Fig. 1. Each participant performed reaches to 16 target objects. The collection of this behavioral data was performed prior to the present DARPA grant. The dataset of recorded reaches is broken into 64 clips (4 actors times 16 targets), and we present these clips to the model to test its ability to infer intentions. In future work, we will also present these clips to a second group of human judges, who will infer the goal of the actor's reach at various frames in each clip. This will produce behavioral data in a format comparable to the model predictions in Fig. 1(b), with a distribution over potential goals at various timesteps.

Our hypothesis is that the probabilistic program described above will be able to infer the intended goal of observed reaches in a way that matches the accuracy and response time of naive human observers. Fig. 1(b) illustrates the predictions the model makes about the judgments of human

participants. At each frame of the video, the model predicts a distribution over the goal that participants will infer, including the pattern of errors, and the point at which performance will exceed chance. We expect that the input provided by the Kinect sensor will be noisier, and thus the model predictions will be more variable when applied to the human reaching data.

Modeling human reasoning about beliefs and desires

In addition to reasoning about other people's goals, the ability to reason about others' beliefs and desires is critical for successful teamwork. For example, recognizing that a teammate is searching for a tool that you borrowed requires reasoning jointly about their desire to obtain the tool, and about their initially false belief that the tool was in the location where they left it. Only by these inferences can one know what the appropriate helping action is in this case: to return the tool. In this thrust, we have designed algorithms for jointly inferring the beliefs and desires of agents from observations of their actions in a partially observable world. We performed two human behavioral experiments, each with a large number of stimulus conditions and multiple judgments per condition, and quantitatively compared our model predictions to human judgments. All human behavioral studies in this publication were conducted prior to the present DARPA grant. The resulting modeling of these data were described in the following journal paper.

Publications

Human mentalizing supports rational attribution of beliefs, desires, and percepts (2017). Chris L. Baker, Julian Jara-Ettinger, Rebecca Saxe, & Joshua B. Tenenbaum. *Nature Human Behavior*, 1(0064).

Technical approach

The model uses a probabilistic program for inverting a model of agents' belief- and desire-dependent planning. The model of planning is given by partially observable Markov decision processes (POMDPs): an agent-based framework for rational planning and state estimation, inspired by the classical theory of decision-making by maximizing expected utility, but generalized to agents planning sequential actions that unfold over space and time with uncertainty due to incomplete information.

POMDPs capture three central causal processes of core mentalizing highlighted by Fig. 6(a): A rational agent (1) forms percepts that are a rational function of the world state, their own state, and the nature of their perceptual apparatus -- for a visually guided agent, anything in their line of sight should register in their world model; (2) forms beliefs that are rational inferences based on the combination of their percepts and their prior knowledge -- a process at least roughly analogous to Bayesian belief updating on the agent's part; and (3) plans rational sequences of actions -- actions that, given their beliefs, can be expected to achieve their desires efficiently and reliably.

These representations provide the key likelihood terms for the Bayesian model: $P(Observation|State)$, $P(Belief|Observation,Belief_0)$, and $P(Action|Belief,Desire)$, respectively. We invert this probabilistic program by integrating the likelihoods with the prior $P(Belief_0,Desire,State)$ using Bayes' rule (abbreviating variables by their first letters):

$$P(B,D,O,S|A) = P(A|B,D)\, P(B|O,B\_0)\, P(O|S)\, P(B\_0,D,S) / P(A).$$

Results

Fig. 6(b) shows example results from one trial of a behavioral experiment on jointly inferring beliefs and desires (see caption for details). The model predicts human judgments with high accuracy across all 73 trials of this experiment; the correlation between the model and human desire judgments is r=0.91, and the correlation between the model and human belief judgments is r=0.78. The collection of this human behavioral data was performed prior to, and not supported by, the present DARPA grant.

Fig. 6(c) shows example results from one trial of a second behavioral experiment on jointly inferring beliefs along with the partially observed state of the world (see caption for details). Again, the model accurately predicts human judgments, with a correlation of r=0.91 across 19 trials. The collection of this human behavioral data was performed prior to, and not supported by, the present DARPA grant.

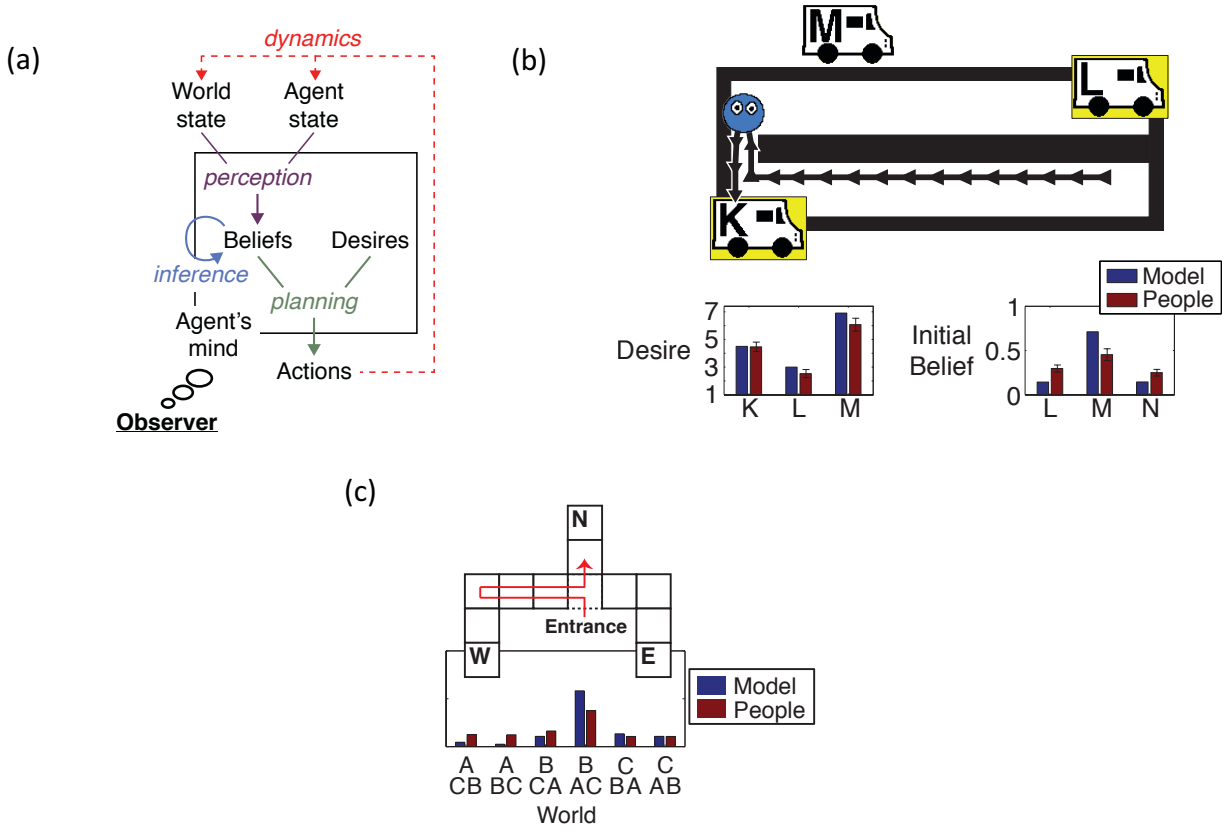


Figure 6: Modeling human reasoning about beliefs and desires. **(a)** Causal schema of the Bayesian probabilistic program. **(b)** In a scenario with 3 food-trucks (Korean (K), Lebanese (L), and Mexican (M)) but only two parking spots, participants attributed a strong desire for the Mexican food-truck (Desire: M > K > L), and a (false) initial belief that this truck was present (Belief: M > L, N (nothing)), after observing an agent pass the Korean truck to check the second spot, then see the Lebanese truck, and then return to the Korean truck. The Bayesian model captures these predictions. **(c)** In a scenario with 3 food-carts (Afghan (A), Burmese (B), and Colombian (C)), participants were told the agent liked A most and C least, and carts could be open or closed; on observing the trajectory shown they inferred that A, B, and C were most likely to be in the West, North and East spots, respectively.

Optimal control for dexterous manipulation with learned local models

Manipulating objects with the hands is a foundational skill for humans, but it is computationally difficult for robotic control algorithms. This thrust aims to develop algorithms for dexterous manipulation of objects with a robotic hand. The method learns a model of the local dynamics of the system, along with local control policies to perform a variety of manipulation tasks. Good models of manipulation planning will be key for building probabilistic programs that can understand the goals of human object manipulation, and robots that can assist humans by manipulating objects.

Publications

Optimal control with learned local models: Application to dexterous manipulation (2016). Vikash Kumar, Emanuel Todorov, & Sergey Levine. IEEE International Conference on Robotics and Automation (ICRA). Best Manipulation Paper Award.

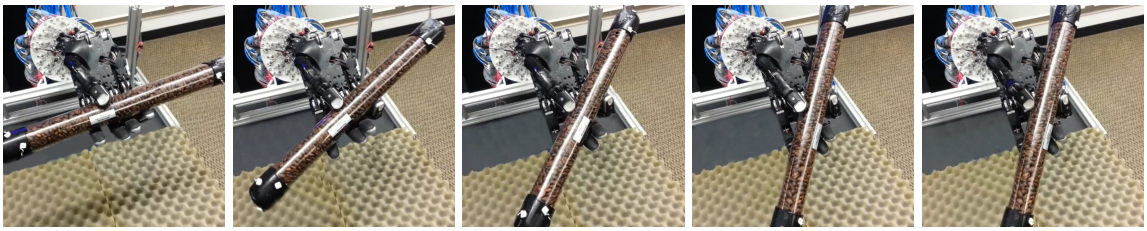
Technical approach

The approach is based on model-based reinforcement learning, which simultaneously learns a model of the world (i.e. the dynamics of the hand and the object being manipulated) along with optimal control laws for performing the task. Currently only a local model is learned, valid around the set of previously executed trajectories. Making these models generalize by using the data to train some neural network is an important next step. The control problem uses a 100-dimensional state space, so efficient algorithms for learning and control are essential. The algorithm is tested in simulation using the MuJoCo physics engine, and in hardware with the Adroit platform, a pneumatically actuated robotic hand.

Results

Fig. 7(a) shows an example manipulation behavior performed by the algorithm on the Adroit robotic hand, involving twirling a cylindrical object in a clockwise direction. The learning curves in Fig. 7(b) show that the algorithm is able to quickly learn stable hand poses both in simulation and in hardware, and Fig. 7(c) shows that various twirling behaviors are learned quickly as well.

(a)



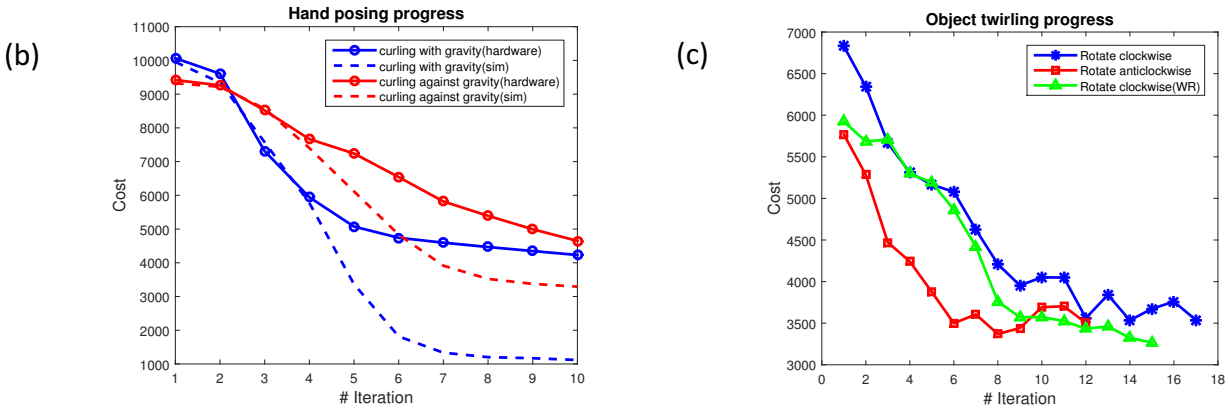


Figure 7: Dexterous manipulation with learned local models. **(a)** Example "twirling" behavior learned by the algorithm. **(b)** Learning curves for various hand poses in simulation and hardware. **(c)** Learning curves for various twirling behaviors in hardware.

Coordinated action

The aim of this thrust is to build probabilistic programs that can infer teammates' intent to coordinate, and respond appropriately, with actions that others expect and find helpful. For example, when performing a task such as cleaning up a table, human teammates will coordinate on which objects to pick up based on their proximity to different objects, on the mass of the objects relative to their respective strengths, and on the relative status relations between participants.

Publications

Max Kleiman-Weiner, Mark K. Ho, Joseph L. Austerweil, Michael L. Littman, & Joshua B. Tenenbaum. (2016). Coordinate to cooperate or compete: Abstract goals and joint intentions in social interaction. In Proceedings of the 38th Annual Conference of the Cognitive Science Society.

Technical approach

Our approach is based on the assumption that for a given cooperative task, each participant considers what the optimal joint action for all participants would be. Each participant then executes their portion of the joint task, while monitoring the execution of the other player, inferring the correct underlying task being executed, and replanning their own execution based on the actions of the other player, and on the nature of the underlying task inferred.

We consider two formulations of this basic framework. The first uses joint planning in MuJoCo to generate dynamic humanoid trajectories. The second uses discrete grid-games, which although less physically realistic, allow us to investigate more sophisticated game-theoretic concepts, e.g., modeling the decision to cooperate or compete with other players in various interactions.

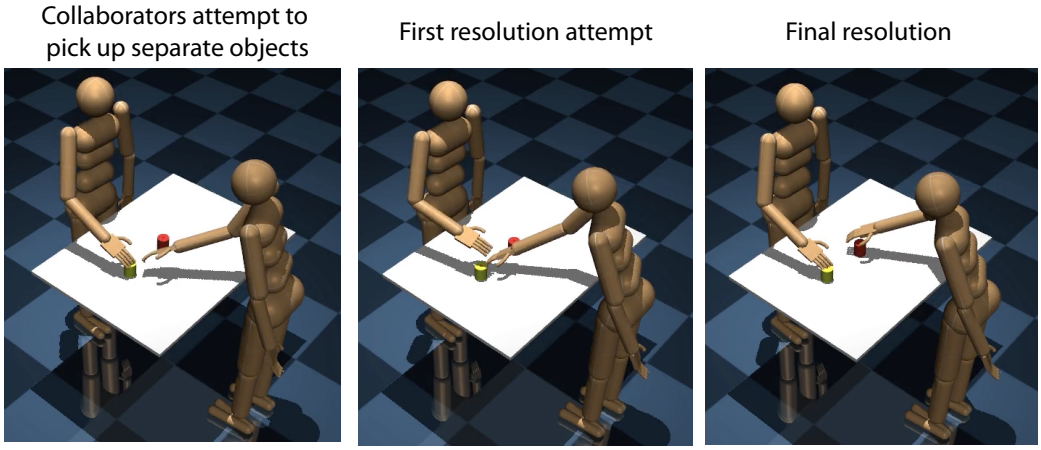
Results

Fig. 8(a) shows example results of a coordinated reaching task. In this task, the agents must coordinate on the goal to reach for separate objects, without knowing which object the other will reach for. In the first frame, the far agent attempts to coordinate on each reaching for the closer object, while the near agent attempts to do the opposite. In the second frame, the far agent attempts to resolve the situation by formulating a new plan to reach for the red object. In the third frame,

the near agent has reacted by switching to the red object, and the far agent has returned to the preferred, nearer yellow object.

Fig. 8(b) shows the results of an experiment on modeling human coordination and competition in grid games. In each of the examples, two players must move in the grid from their starting location to the location of various goals (marked by the amount of reward they provide), while coordinating with the other player to avoid collisions. The model captures many features of the human data, including the proportion of participants who cooperate within each experiment. The collection of this human behavioral data was not supported by the present DARPA grant.

(a)



(b)

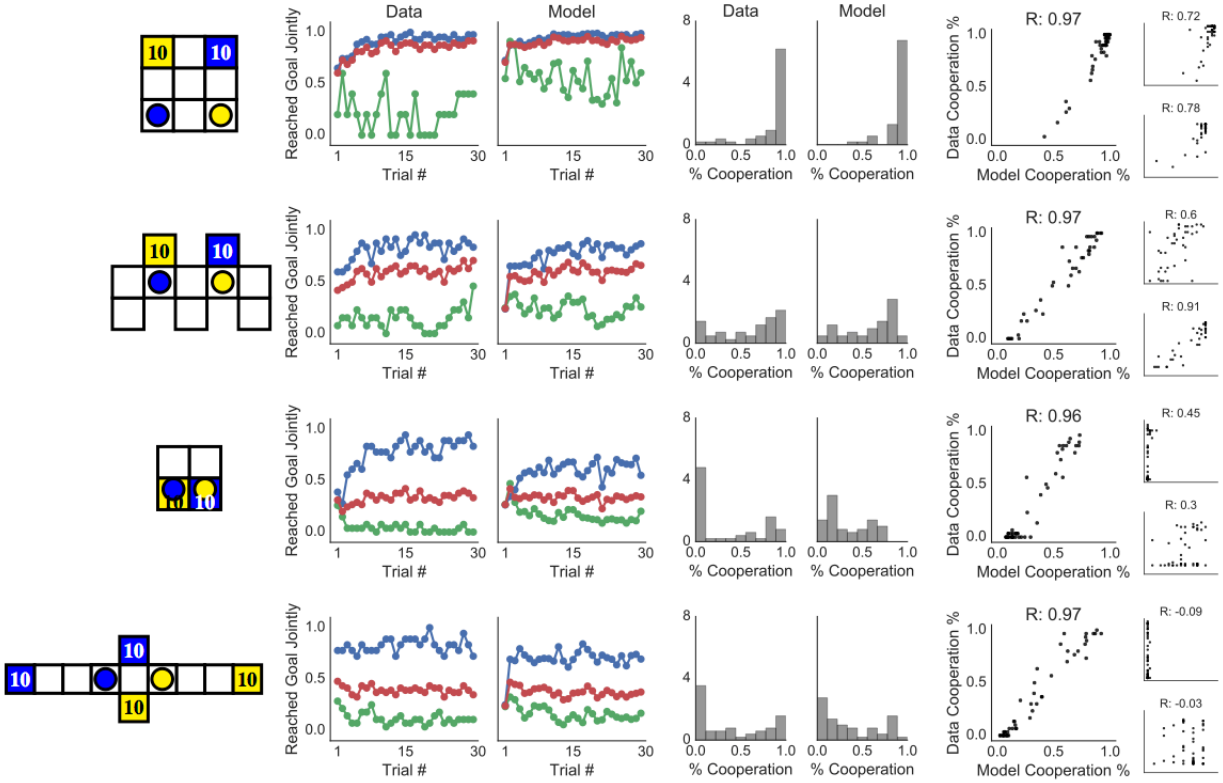


Figure 8: Coordinated action. **(a)** Example results of our coordination framework using simulated humanoids in MuJoCo. **(b)** Participant data and model predictions in column 1 which was repeated 30 times. Rows 1 and 2 are coordination games and rows 3 and 4 are social dilemmas. Column 2 shows the average rate of cooperation for each round of play averaged over the high-cooperating cluster of participants (blue), low-cooperating cluster of participants (green) and all participants (red). Column 3 are histograms of the proportion of cooperation for all pairs of participants. Column 4 quantifies the model predictions where each point represents the frequency of cooperation for a given dyad observed in the data and as predicted by the model. The inset shows correlations of the two lesioned models with the same human data: (top) only compete (bottom) only cooperate.