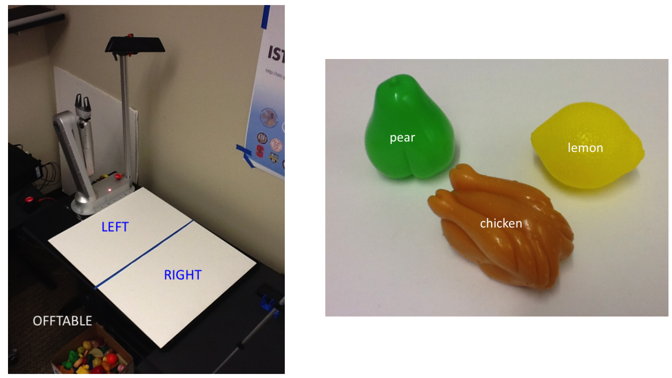
A Developmental Approach to Goal-Based Imitation Learning in Robots

Michael Jae-Yoon Chung, Jinna Lei, Ankit Gupta, Dieter Fox, Andrew N. Meltzoff, and Rajesh P. N. Rao
Computer Science & Engineering
University of Washington
Seattle, Washington 98195-2350
Email: {mjyc,jinna,ankit,fox,rao}@cs.washington.edu,meltzoff@uw.edu

***Abstract*—We propose a new developmental approach to goal-based imitation learning that allows a robot to: (1) learn probabilistic models of actions through self-discovery and experience, (2) utilize these learned models for inferring the goals of human demonstrations, and (3) perform goal-based imitation for human-robot collaboration. Our approach is based on Meltzoff's "Like-me" hypothesis in developmental science, which states that children use self-experience to bootstrap the process of intention recognition and imitation. Such an approach allows a robot to leverage its increasing repertoire of learned behaviors to interpret increasingly complex human actions, even when the robot has very different actuators from humans. We present preliminary results illustrating our approach using a simple robotic tabletop organization task. We show that the robot can learn a probabilistic model of its actions on a small set of objects, and use this model for both goal inference and goal-based imitation of human actions. We also present results demonstrating that the robot can use its learned probabilistic model to seek human assistance whenever it recognizes that its inferred actions are too uncertain, risky, or impossible to perform.**



(a) (b)

Fig. 1. **Robotic tabletop task setup.** (a): The robot is located left side of work area and the Kinect® looks down from left side in the robot perspective. The three predefined areas that distinguish object states are notated. (b): Toy tabletop objects.

I. Introduction

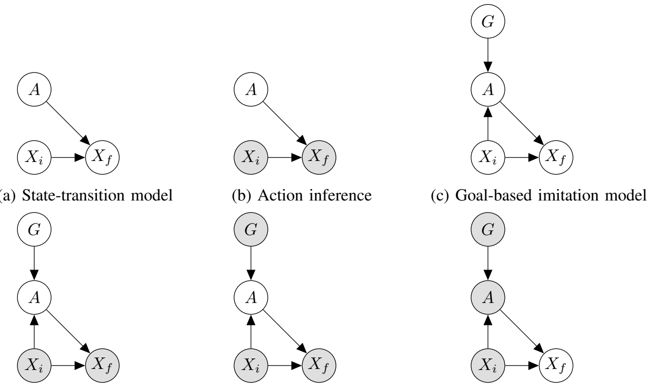
Recent advances in imitation learning (also referred to as *learning by demonstration* and *apprenticeship learning*) have led to a number of impressive demonstrations of robotic skill learning from humans [1]–[8]. Most of these results involve following action trajectories given by a human in the same space as the robot's actuator space. For example, demonstrations in [1], [5], [6] were collected by a human manually moving the robot's arm while in [2], demonstrations were obtained by joystick control of a helicopter. Even in the case of imitation learning in humanoids, e.g., [4], [7], [8], where manual or joystick control is not possible, a motion-capture system is typically used to provide a trajectory of tracked human poses for the humanoid robot to imitate.

In many instances of imitation learning, it is not the trajectory that is important but the goal of the action. For example, if the goal is to open a box, it does not matter which hand is used to hold the box and which fingers are used to lift the lid. Thus, if opening the box is part of a complex task being demonstrated by a human, the robot needs to recognize the goal of the human action and then employ its own actuators to achieve the same goal. Such an approach, which we call *goal-based imitation* [9], acknowledges the fact that robots often have different actuators than humans but may still be able to achieve the same goals as a human demonstrator, albeit using different means. Goal-based imitation is thus a solution to the problem of *heterogeneous imitation learning* [10], where the human and robot have different actuators (in contrast, the traditional case of *homogeneous imitation learning* assumes the robot and human have the same actuator space).

In this paper, we present a new *developmental* approach to goal-based imitation learning that builds on the "Like-me" hypothesis [11] regarding human development. The hypothesis, which is supported by infant gaze following and imitation studies (see [11], [12], [13]), states that children utilize internal models learned through self-experience to interpret the acts and goals of others. Our framework demonstrates how the "Like-me" hypothesis can realized in robots using graphical models for probabilistic reasoning and learning under uncertainty.

Our approach utilizes three components: (1) a self-discovery module that allows the robot to learn probabilistic action models through self-exploration, (2) a goal inference module that allows the robot to infer the intention of a human demonstrator using its learned probabilistic models, and (3) a module for goal-based imitation that uses the inferred goal and the learned probabilistic model to infer the action most likely to achieve the goal.

An important distinction between our model and the many previous approaches to human action understanding and ac-



(a) State-transition model (b) Action inference (c) Goal-based imitation model

(d) Goal Inference (e) Goal-based action inference (f) State prediction

Fig. 2. **Graphical Models**. (a) through (f) illustrate the use of graphical models for learning state-transitions, action inference, goal inference, goal-based imitation, and state prediction. Shaded nodes denote evidence (known values).

tivity recognition that have been proposed (e.g., [14]) is that our approach leverages *robotic self-discovery*: the robot *learns* probabilistic models of state changes in the world through its own actions and subsequently uses these learned models for interpreting human actions. Thus, rather than relying on pre-wired models of human actions or labeled human data, our method allows the robot to learn increasingly sophisticated models of action based on its ongoing interactions with the world and use these models for interpreting the actions of others. Such an approach to action understanding is consistent with recent cognitive theories of human intention recognition [11].

To illustrate our approach, we present preliminary results from a simple human-robot collaborative learning task in which a robot learns to organize objects of different shapes on a table. The robotic system consists of an arm-and-gripper for manipulation and an RGBD camera for visual feedback. Due to the physical limitations of the arm-and-gripper system, some objects may be difficult or impossible to grasp but easier to push, while others may be reliably picked up. In such a setting, attempting to imitate the actual trajectories followed by the human hand and fingers will likely result in failure but a goal-based approach may succeed. We demonstrate that (1) the robot can learn probabilistic models of its actions on objects through self-exploration, and (2) use these models for both goal-inference and goal-based imitation of human actions on objects. Our results additionally demonstrate that the robot can leverage its probabilistic models to seek human assistance whenever it finds that its inferred actions are too uncertain, risky, or impossible to perform.

The paper is organized as follows. We describe the proposed framework for probabilistic goal inference and imitation learning in Section II. Section III describes the robotic hardware and experimental set-up. We present results from our goal inference and imitation experiments in Section IV and conclude in Section V by discussing some current limitations of our approach and future work.

II. Learning through Graphical Models

In this section, we describe how probabilistic graphical models can be used for learning transitions of object state, inferring goals, and performing goal-based imitation. We describe the approach in the context of a simple task involving manipulating a small set of objects on a table but the approach itself is general and could be applied to other scenarios.

A. Overview of the task

The environment for the task used in this paper consists of a set of objects on a tabletop which can be moved around by a human or a robotic arm. The position and in-plane rotation of the object defines its *state*. We define *goals* based on whether the object has reached a particular state. The robotic arm can manipulate the state of any object using a set of *actions*. The robotic arm needs to learn probability models over the states, actions, and goals.

B. Notations

Let Ω^X be the set of states in the environment, and let Ω^A be the set of all possible actions available to the robot (these can be different from the possible actions of the human demonstrator). Let Ω^G be the set of possible goals. We assume all three sets are finite. Each goal G represents an abstract task, which can be achieved using one or more actions in Ω^A. The basic template for the probabilistic graphical model we use is the Markov state-transition model in Fig. 2(a). Initially, the robot is in state X_i; when it executes action A, it stochastically enters a state X_f as determined by the transition probability $P(X_f|X_i, A)$.

C. Learning through self-experience

We assume that the transition probability distribution is initially unknown to the robot and must be learned through exploration, similar to the manner in which infants explore the consequences of their actions through exploration and "body babbling" [15]. The robot collects training data tuples $\{x,\ a,\ x'\}$ for the graphical model in Fig. 2(a), where $x, x' \in \Omega^X, a \in \Omega^A$. Given the training data, simple maximum likelihood parameter estimation is used to learn the transition probability distribution.

D. Goal-based graphical models and goal inference

After the transition parameters are learned, goal-based graphical models for achieving specific goals can be learned as follows. We consider the case where "goals" are labels for abstract tasks, such as moving an object from location A to location B. The object could be picked and placed at B, pushed to B, or transported using some other action, but the important point is that goal remains the same. The robot learns goal-based models as follows. To achieve a goal (or task) g, the robot identifies the initial state and desired final state (X_i and X_f) and computes the marginal distribution $Pr(A|X_i, X_f)$ using Bayesian inference in the graphical model shown in Fig. 2(b). Note that although the present implementation uses

Algorithm 1 Learning Through Self Experience(Ω^X, Ω^A)

for all $x \in \Omega^X$ and $a \in \Omega^A$ pairs **do**
    for $k = 1 \rightarrow n$ **do**
        Execute action a and record observed state x'.
    end for
    Compute $Pr(X_f|X_i = x, A = a)$ based on observed x's.
end for
for all X_i and X_f pairs **do**
    Compute $Pr(A|X_i, X_f)$.
    Determine g using X_i, X_f.
    Set $Pr(A|X_i, G = g) = Pr(A|X_i, X_f)$
end for
Construct graphical model, $\mathcal{G}$, from $Pr(X_f|X_i, A)$ and $Pr(A|X_i, G = g)$.

Algorithm 2 Goal Inference and Action Selection($\mathcal{G}$,x_i,x_f,τ)

Compute $Pr(G|X_i, X_f)$ using $\mathcal{G}$ and junction tree algorithm.
$g_{MAP} \leftarrow \texttt{max}(Pr(G|X = x_i, X_f = x_f))$
$a_{MAP} \leftarrow \texttt{max}(Pr(A|X_i = x_i, X_f = x_f, G = g_{MAP}))$
if $Pr(X_f = x_f|X_i = x_i, A = a_{MAP}, G = g_{MAP}) < \tau$ **then**
    $a_f = ask\_human$
else
    $a_f = a_{MAP}$
end if
return a_f

1-step inference, the approach generalizes readily to multi-step planning, i.e., inference of a sequence of actions.

Once a distribution over actions is inferred, a goal-based model is created by augmenting the initial model in Fig. 2(a) with a new node G as shown in Fig. 2(c). This provides a compact way of representing and reasoning about abstract goals. For each specific goal or task g, we set the conditional $Pr(A|X_i, G = g) = Pr(A|X_i, X_f)$ where X_i and X_f are the initial and desired states for goal g.

For goal inference, the robot observes object states x_i and x_f from human demonstration, e.g., change in location (or orientation) A to B. The robot then computes the posterior distribution over goals G given x_i, and x_f, as depicted in Fig. 2(d) (note that the variable A is marginalized out during inference).

E. Goal-based imitation and action selection

Goal-based imitation is implemented as a two-stage process: first, we infer the likely goal of the human demonstration using the goal-based graphical model described above, and second, we either execute the action most likely to achieve this goal or seek human assistance.

In more detail, the robot senses the current state x using its sensors and infers the human's goal g by taking the mode of the posterior distribution of G from the goal-inference step. It then computes the posterior over actions A as shown in Fig. 2(e) and selects the maximum a posteriori action a_{MAP}. Since a_{MAP} is not guaranteed to succeed, we predict the probability of reaching the desired final state x' using a_{MAP} by computing the posterior probability of X_f as shown in 2(f). If this probability of reaching the desired state is above a prespecified threshold, τ, the robot executes a_{MAP}, otherwise it executes the "Ask human" action to request human assistance.

Algorithm 1 and Algorithm 2 summarize the complete sequence of steps as described above.

III. Robotic Learning Task

A. Hardware

For our experiments, we use the Gambit robot arm-and-gripper designed at Intel Labs Seattle. The Gambit is well-suited to tabletop manipulation tasks and has previously been shown to perform well in tasks with humans in the loop [16]. It has seven controllable degrees of freedom (DoFs): base rotation, shoulder joint, elbow joint, forearm rotation, two wrist joints, and a parallel jaw gripper. The three revolute DoFs (base, shoulder, and forearm) provide position control while the other DoFs (forearm and two wrist joints) provide orientation control via a roll-pitch-roll spherical wrist. For our task, the tabletop working area is a circle with a radius of $\approx$ 60 cm. The low-level gambit driver software is built on top of ROS (Robot Operating System) [17] which runs on a dedicated Intel Atom net-top PC. Application programs running on a separate computer commands the arm using high-level drivers.

For sensing the current state of objects on the table, we use the Kinect® camera, which provides a stream of registered color and depth images, and has a working range of approximately 0.5 m to 5 m. The camera is mounted on the base frame of Gambit and looks down on the table surface as shown in Figure 1. The location of the camera was chosen to maximize the view of the robot's work area while minimizing physical interference with the robot arm movements.

For the preliminary experiments described in this paper, we defined three discrete areas that are used for defining state of the objects as shown in Fig. 1. These fall within the area where the robot's gripper can reach without kinematic failures.

The robot takes as input the stream of color and depth images from the Kinect camera and first segments out the background and the human's hands. The remaining pixels correspond to the objects which are then used to determine the state of objects on the table. Once the object state is identified, the robot can perform a fixed set of actions on the object. We now describe these components in detail.

B. Segmenting the Image Stream

For every color-depth image, we use background subtraction technique on the depth channel to remove the background. We do this by noting the background depth at each pixel when the

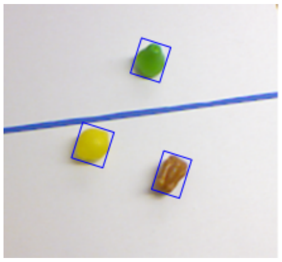


Fig. 3. **Object bounding boxes:** We use a simple rotated bounding box detection algorithm provided by OpenCV. The detected bounding boxes are later used for grasping.

system starts. For each next frame, we segment out the pixels for which the depth is greater than or equal to the background depth at that pixel. To make our system robust to noise in the Kinect's depth channel, any depths within 15mm of the background depth are also segmented out.

The remaining foreground pixels consist of human's hands and objects. We learn a color model for the skin in the HSV color space and classify the foreground pixels as skin or not. We perform connected components analysis on skin pixels and remove the components which connect to the table edge. This is based on the justifiable assumption that the human's arms will come from outside the table region and hence the hand pixels will be connected to the table edges.

After segmenting out the hand pixels, the remaining pixels correspond to the objects. Performing a connected component analysis on these pixels gives us the pixel components corresponding to each of the objects. For our experiments, we assume that: (1) the maximum number of objects on the table is known ahead of time, and (2) two objects cannot be created or removed in a single image frame. These reasonable assumptions allow object identification without requiring a complex object recognition engine.

C. Object State Estimation

We use discrete states to characterize the current state of an object. The experiments in this paper assumed three discrete states for objects based on their location: (1) "LEFT" signifying that the object is on the left side of the blueline on the table as shown in Figure 1; (2) "RIGHT" denoting that the object is on the right side of the blueline; and (3) "OFFTABLE" signifying that the object is not in the tabletop work area. For each object identified on the table by the method in the previous section, we determine its state based on the location of its centroid position. We also fit a bounding box to each identified object. The centroid position and rotation angle of the bounding box are stored with the object state information for robotic manipulation.

D. Robot Actions

The Gambit robot was provided with a fixed set of six high-level actions for manipulating objects: place LEFT (PlaceL), place RIGHT (PlaceR), place OFFTABLE (PlaceOT), push to LEFT (pushL), push to RIGHT (pushR), and push OFFTABLE (pushOT). For computing these actions, we use a custom

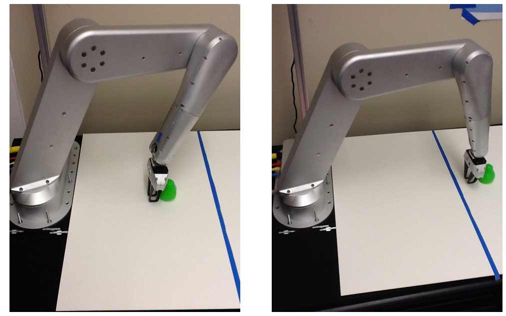


Fig. 4. **The push to right action, before and after:** The push to left action works similarily. To push off table, the arm goes down to the lower side of table. For the place actions, we use almost identical actions to the picking up chess pieces used in [16].

inverse kinematics solver used successfully in a previous project involving the Gambit [16].

For the "place" actions, the robot first attempts to pick up the object by moving its end effector above the centroid of the object and rotating the gripper to align itself perpendicular to the major axis of the object (determined by the bounding box). If the robot successfully picked up the object, it places the object down at the location (LEFT, RIGHT, or OFFTABLE) indicated by the place command.

For the "push" actions, the robot first positions its end effector behind the object based on its centroid and direction of the push. For pushL and pushR, the gripper yaw angle is rotated perpendicular to the major axis of the table, while for pushOT, it is rotated parallel to the major axis of the table. This ensures that object contact area is maximized to reduce the chance of the object slipping while pushing. The robot pushes the object until it changes state (or the inverse kinematic solver fails to find a possible solution).

IV. Experiments and Results

We illustrate the proposed approach using a tabletop organization task involving plastic objects of three different shapes: a pear, a lemon, and a miniature broiled chicken. These objects could be in one of three different states: LEFT, RIGHT, and OFFTABLE, as defined in the previous section. The robot's aim is to learn and imitate the actions of humans on these objects as they organize these objects on the table.

The objects were chosen to demonstrate the different types of challenges inherent in the two types of actions available to the robot: place and push. The pear-shaped objects are pointed and slippery at the top, which makes them almost impossible for the robot to successfully pick up for a place action. On the other hand, the wide bodies of these objects make them easy to push. The lemon-shaped objects are hard to manipulate using either place or push actions because their spherical shapes makes them both hard to pick up (because they can slip through the grippers) and push (because they can roll away from the gripper). For humans, none of these objects poses a challenge for either picking or pushing.

Fig. 5 shows the learned transition probabilities based on the robot's self-discovery phase. The transition models are learned by performing 10 trials for each initial state and action pair

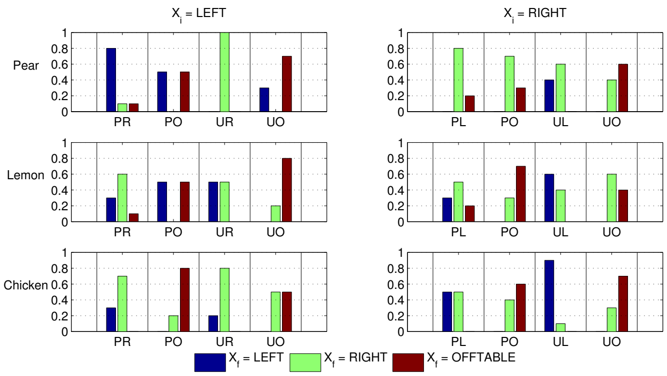


Fig. 5. **Learned transition model.** Each row represents different types of objects, and each column represents the different initial states X_i. The colors of bars represent different final states X_f. The y-axis represents the range of probabilities and the x-axis represents the six different manipulation actions available to the robot (PL = place LEFT, PR = place RIGHT, PO = place OFFTABLE, UL = pUsh LEFT, UR = pUsh RIGHT, UO = pUsh OFFTABLE). We do not show actions that cause self-transitions given an initial state.

(X_i, A) for each object type. We deliberately used a small number of trials to test whether the method could cope with less training data and more uncertainty in the transition model.

Since the state space is small, we are able to enumerate and test all of the interesting possible human demonstrations. By interesting, we mean that the state changes after action execution (for example from RIGHT to OFFTABLE) and the initial state is not OFFTABLE.[1]. There are total four interesting state changes for each object that can be demonstrated by a human: 1) LEFT to RIGHT, LEFT to OFFTABLE, RIGHT to LEFT, and RIGHT to OFFTABLE.

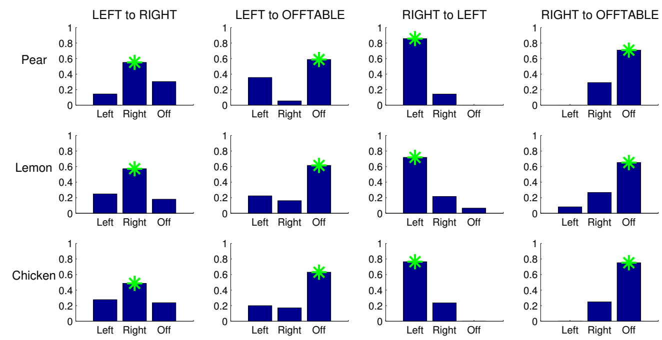
Fig. 6(a) shows the inferred goals given all possible interesting initial and final state transitions, using the graphical model in Fig. 2(a). For all cases, our model correctly infers the intended goal state of the human. Fig. 6(b) shows the MAP action for a given initial state and goal. Our model correctly identifies whether a "place" or a "push" action is better, given the dynamics of the object as encoded in the learned probabilistic model for the object. Note that the two most preferred actions are always push or place actions in the correct direction.

Finally, Fig. 6(c) shows the predicted state distribution given an action and an initial state. The robot calculates the posterior probability of getting to the desired state, and executes the action if this probability is above a predetermined threshold. Otherwise, it asks the human collaborator for help.[2]

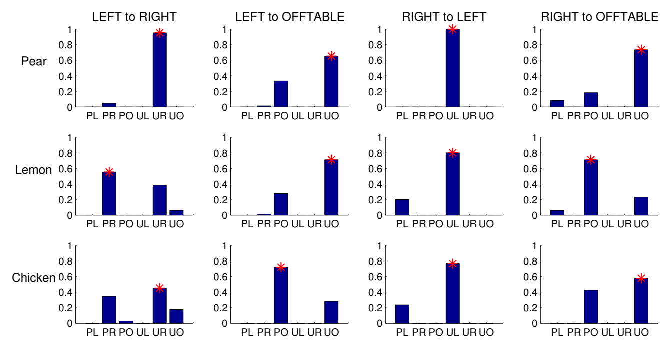
Table I compares trajectory-based imitation with our proposed goal-based approach. The trajectory-based approach simply mimics the human action without considering the goal or uncertainty, i.e., it executes a place action if the human executes a place, and a push action if the human executes a push. The goal-based approach on the other recognizes the goal and uses the best action it has available to achieve the goal.

[1]The current implementation does not allow the robot to pick up objects that are located OFFTABLE

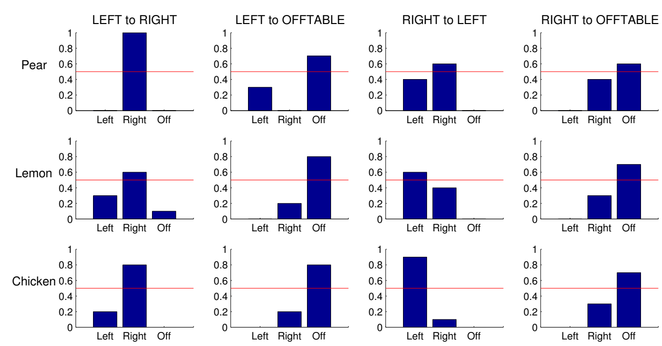
[2]In future implementations, we hope to select the threshold automatically based on a reward function within a POMDP framework.



(a) Goal Inference



(b) Action Inference



(c) Final State Prediction

Fig. 6. (a) **Most likely goals**. Initial and final states are at the top of each column. The height of the bar represents the posterior probability of each goal state, with the true goal state marked by an asterisk. (b) **Inferring actions**. For each initial and desired final state, the plots show the posterior probability of each of the six actions, with the MAP action indicated by an asterisk. (c) **Predicting final state**. The plots show the posterior probability of reaching the desired final state, given the initial state. The red bar marks 0.5, the threshold below which the robot asks for human help in the Interactive Goal-Based mode.

Using our computed transition probabilities, we can calculate the hypothetical success rate of a purely trajectory-based approach. For our goal-based approach, we use the posterior distribution shown in Fig. 6(b). Finally, the "Interactive Goal-Based" mode assumes that the robot may ask a human for help, with a 100% success rate when the human executes the requested action. The third column in Table I shows what the performance would be if we require the robot to be 50% sure

TABLE I

COMPARISON OF APPROACHES: THE SUCCESS RATES OF THREE DIFFERENT APPROACHES TO IMITATION ARE SHOWN.

	Trajectory	Goal-Based	Interactive Goal-Based
Pick & Place Demonstration			
Pear	0.2250	0.6750	**0.8250**
Lemon	0.5250	**0.6750**	**0.6750**
Chicken	0.6500	**0.8000**	**0.8000**
Push Demonstration			
Pear	0.6750	0.6750	**0.7250**
Lemon	0.5750	**0.6750**	**0.6750**
Chicken	0.7250	**0.8000**	**0.8000**

of reaching a desired state.[3] These results demonstrate the advantage of a goal-based approach over purely trajectory-based imitation.

V. Summary and Conclusion

We have proposed a new developmental approach to robotic imitation learning based on inference of goals using graphical models that are learned by the robot through self-experience. Our approach follows Meltzoff's "Like-me" hypothesis of human development and intention recognition. We demonstrated, using a simple tabletop organization task, that a robot can: (1) learn probabilistic models of the consequences of its actions on objects through self-exploration, (2) use these learned models to infer the goal of a human action on these objects, and (3) perform goal-based imitation.

Our results show how a goal-based approach can allow imitation learning even when the robot has different actuators and motor capabilities from a human. Our results also illustrate the usefulness of probabilistic models for human-robot interaction: the robot use its estimates of uncertainty to decide when to execute a computed action and when to seek human assistance to avoid accidents while maximizing human-robot collaborative throughput.

Goal-based imitation and probabilistic models have attracted attention in the robotics community [18]–[20]. Some of these approaches propose imitation based on high-level goals [18], [19] but do not rely on robotic self-discovery of probabilistic models, a central tenet of the developmental approach proposed here for bootstrapping goal-based imitation. Other approaches have focused on attempting to model continuous low-level goals [20]. None of these approaches emphasize the utility of probabilistic models for human-robot interaction tasks, which is a major focus of our work.

Our results point to a number of interesting open issues. For the preliminary studies in this paper, we used a simple exhaustive exploration strategy, but for larger state spaces, a more sophisticated approach based on reward functions (e.g., [21]) could be employed. Additionally, following the example of human infants, some form of directed self-exploration based on observing human teachers (e.g., [4], [8]) may be desirable. Better generalization could also be achieved using continuous state representations and nonparametric Bayesian models such as Gaussian processes, e.g., [22].

Finally, the approach we have presented lends itself naturally to generalization based on relational probabilistic models [23], [24] and hierarchical Bayesian representations [25]. Such models have the potential to significantly increase the scalability and applicability of our suggested approach to large-scale scenarios, besides facilitating transfer of learned skills across tasks and domains. We intend to explore such relational models in future work.

[3]We do not see perfect imitation results on the third column because we do not ask the human for help in every case. In some cases, the probability of success will surpass the confidence threshold, but not make the state transition successfully.

Acknowledgments

This research was supported by the ONR Cognitive Science program grant no. N000140910097 and by the Intel Science and Technology Center (ISTC), Seattle. The authors would like to thank Cynthia Matuszek and Kendall Lowrey for their advice and help during the initial stages of this project.

References

[1] P. Kormushev, S. Calinon, and D. Caldwell, "Robot motor skill coordination with em-based reinforcement learning," in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2010, pp. 3232–3237.

[2] P. Abbeel, A. Coates, and A. Ng, "Autonomous helicopter aerobatics through apprenticeship learning," *The International Journal of Robotics Research*, 2010.

[3] R. Chalodhorn and R. P. Rao, "Learning to imitate human actions through eigenposes," in *From Motor Learning to Interaction Learning in Robots*. Springer, 2010, pp. 357–381.

[4] D. B. Grimes and R. P. Rao, "Learning actions through imitation and exploration: Towards humanoid robots that learn from humans," in *Creating Brain-Like Intelligence*. Springer, 2009, pp. 103–138.

[5] J. Kober and J. Peters, "Policy search for motor primitives in robotics," *Advances in Neural information processing systems*, vol. 21, pp. 849–856, 2009.

[6] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 2, pp. 286–298, 2007.

[7] R. Chalodhorn, D. B. Grimes, K. Grochow, and R. P. Rao, "Learning to walk through imitation," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2007.

[8] D. Grimes, R. Chalodhorn, and R. Rao, "Dynamic imitation in a humanoid robot through nonparametric probabilistic inference," in *Proceedings of Robotics: Science and Systems*, 2006.

[9] D. Verma and R. Rao, "Goal-based imitation as probabilistic inference over graphical models," *Advances in Neural information processing systems*, vol. 18, p. 1393, 2006.

[10] B. Price and C. Boutilier, "Imitation and reinforcement learning in agents with heterogeneous actions," in *Seventeenth International Conference on Machine Learning ICML2000*, 2000.

[11] A. Meltzoff, "Like me: a foundation for social cognition," *Developmental science*, vol. 10, no. 1, pp. 126–134, 2007.

[12] A. N. Meltzoff and R. Brooks, "Self-experience as a mechanism for learning about others: A training study in social cognition," *Developmental psychology*, vol. 44, no. 5, p. 1257, 2008.

[13] A. N. Meltzoff, P. K. Kuhl, J. Movellan, and T. J. Sejnowski, "Foundations for a new science of learning," *Science*, vol. 325, no. 5938, pp. 284–288, 2009.

[14] L. Liao, D. Patterson, D. Fox, and H. Kautz, "Learning and inferring transportation routines," *Artificial Intelligence*, vol. 171, no. 5, pp. 311–331, 2007.

[15] R. Rao, A. Shon, and A. Meltzoff, "A Bayesian model of imitation in infants and robots," *Imitation and social learning in robots, humans, and animals*, pp. 217–248, 2004.

[16] C. Matuszek, B. Mayton, R. Aimi, M. Deisenroth, L. Bo, R. Chu, M. Kung, L. LeGrand, J. Smith, and D. Fox, "Gambit: An autonomous chess-playing robotic system," in *2011 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 4291–4297.

[17] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Ng, "ROS: an open-source robot operating system," in *ICRA Workshop on Open Source Software*, vol. 3, no. 3.2, 2009.

[18] D. Song, K. Huebner, V. Kyrki, and D. Kragic, "Learning task constraints for robot grasping using graphical models," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2010, pp. 1579–1585.

[19] C. H. Ek, D. Song, K. Huebner, and D. Kragic, "Task modeling in imitation learning using latent variable models," in *10th IEEE-RAS International Conference on Humanoid Robots (Humanoids), 2010*. IEEE, 2010, pp. 548–553.

[20] Z. Wang, M. Deisenroth, H. Amor, D. Vogt, B. Schölkopf, and J. Peters, "Probabilistic modeling of human movements for intention inference," in *Proceedings of Robotics: Science and Systems*, 2012.

[21] M. Deisenroth and C. Rasmussen, "PILCO: A model-based and data-efficient approach to policy search," 2011.

[22] A. Friesen and R. Rao, "Gaze following as goal inference: A Bayesian model," in *Carlson L. and Hölscher C. and Shipley T. (Eds.) Proceedings of the 33rd Annual Conference of the Cognitive Science Society, Boston, MA*, 2011.

[23] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. The MIT Press, 2009.

[24] M. Richardson and P. Domingos, "Markov logic networks," *Machine learning*, vol. 62, no. 1, pp. 107–136, 2006.

[25] J. Tenenbaum, C. Kemp, T. Griffiths, and N. Goodman, "How to grow a mind: Statistics, structure, and abstraction," *Science*, vol. 331, no. 6022, p. 1279, 2011.