Towards Learning Task Intentions from Human Demonstrations

Tim Welschehold  Nichola Abdo  Christian Dornhege  Wolfram Burgard

Abstract—Learning from demonstrations is an active research area and many methods to instruct robots have been developed recently. However, learning tasks directly from a human teacher is still an unsolved problem. In this paper we present an approach that learns action models for joint motions of the robot base and gripper as well as a representation of the overall task both from observing demonstrations carried out by a human teacher. To this end we adapt RGBD observations of the human teacher to the capabilities of the robot for action learning. Furthermore, we introduce a framework called teach-and-improvise to imitate tasks without explicitly specifying their goals. We demonstrate the effectiveness of both action and task learning in complex real world experiments.

I. INTRODUCTION

To pave the way for complex mobile robots into common households the ability to easily instruct a robot is essential. Users need to be able to teach their robot to achieve a custom task in their own home environments without any explicit knowledge of how robots learn such tasks. In this work we address the problem of jointly learning both action models and task goals from human teacher demonstrations. We consider the teacher to demonstrate to the robot how to manipulate the objects in a scene in order to achieve a goal. The envisioned teaching process should be intuitive and manageable for non-expert users. Thus, for learning the individual actions we do not rely on kinesthetic teaching like [1], or programming skills. The challenge in the context of human action demonstrations arises from the fact that these cannot be directly reproduced by the robot due to differing grasping capabilities and kinematic constraints. To overcome this issue the demonstrations must be adapted to the robot. Grasp planners [2] while able to generate high-quality grasps lack the desired connection to the specific task demonstration and are thus not suited to always choose the appropriate grasp. In our work we introduce a system that is able to adapt the observed human grasp to the robot capabilities, while also considering the observed underlying grasping and manipulation motions. We generate both robot gripper and base motions based on the demonstrated human hand and torso motions.

Most current work on learning the embedding of the action models in tasks builds on existing predicates to instantiate or extract symbolic or logical rules from teacher demonstrations like Höfer and Brock [3]. In this work we do not rely on an expert to program those predicates, but instead do not assume to have prior semantic knowledge about tasks and actions and moreover rely only on a small number of teacher demonstrations to learn them. When addressing the task learning itself, in this work we focus on inferring the desired goal states from few demonstrations. A major challenge in this context is the potential ambiguity in the demonstrations. The teacher might demonstrate different valid object configurations, compare Fig. 1. Without prior knowledge about the task, we aim at generalizing these configurations and reproducing the task from unseen initial states without explicitly specifying a goal state. Traditional task and motion planning techniques [4] expect a concrete goal state description as input. In this work we propose a teach and improve approach to generate the goals by formulating an optimization problem, in which we maximize the likelihood that the final goal state aligns with the intention of the teacher given the demonstrations.

II. PROBLEM STATEMENT

In our work we aim to enable a robot to learn a task from RGBD observations of human teacher demonstrations. The challenge lies in learning both the representation of the individual actions as well as the intention and composition of the task itself. We consider a task to consist of the individual manipulations of all objects involved in it. We assume that each object manipulation can be segmented into three parts, the reaching or grasping, the manipulation itself and the releasing or retreat part. For learning action models a demonstration is given by the tracks of 6-dof trajectories of the human hand and torso as well as the handled objects. Hand and object trajectories describe how the object should be handled, while the torso motion gives insight about the general positioning for the action. As these trajectories are not directly executable for the robot, the action learning
problem is to adapt these to the capabilities of the robot, while retaining the intended demonstrated movements. For task learning a demonstration consists of the manipulation actions performed for the task and the intermediate states of the world between them. A state is defined by all object poses and the final state is considered to be an instance of the task goal. Here, different demonstrations of the same task are allowed to lead to ambiguous goals and the object manipulations may be performed in different orders. The task learning problem is therefore to produce action sequences for previously unseen initial states of the same task that imitate the intention of the user without explicitly formulating a goal state.

III. ACTION LEARNING

We propose an approach in which the robot motion is designed to follow the demonstrations as close as possible while deviating as necessary to fulfill the constraints posed by its geometry. We account for the robot’s grasping skills and kinematics as well as for collisions with the environment. We furthermore can assume for most actions that the grasp on the object is fixed during manipulation and all trajectories should be smooth in the sense that consecutive poses should be near each other. Incorporating this constraints and assumptions we formulate a graph optimization problem to adapt the recorded human motions in a way that they are suited for the robot to learn a motion model. For details on the graph structure and the implementation we refer to our recent work [5], [6].

IV. TASK LEARNING

Task learning computes a sequence of manipulation actions $a_t$ and their corresponding intermediate goals $s_t$ that represent the intention of the demonstrations for a given initial state $s_0$. This is done without explicitly specifying any goal state. Based on pairwise object relations, we define an intention likelihood $\Psi(s_T)$, which holds the information of how well a state $s_T$ fits as the intended task goal.

We further introduce templates $\lambda_t$ which serve as underlaying frames in which the actions can be interpreted and executed. Executing the intention likelihood we formulate a search-based optimization with

$$
\max_{\lambda_0:T-1, s_0:T-1, s_1:T} \Psi(s_T) - \sum_{t=0}^{T-1} \text{cost}(a_t)
$$

(1)

where $s_0:T-1, s_1:T$ and $\lambda_0:T-1$ are the desired manipulations, their corresponding goals and the interpreting reference frames. We use Monte Carlo Tree Search to search task solutions and reason about interpretations of actions and their goals.

A. Tree Structure

Classical Monte Carlo Tree Search iteratively grows a search tree to approximate the returns of states and actions using Monte Carlo Simulations. Typically there are two types of nodes: decision and chance nodes. Decision nodes select actions according to the node’s associated state. The chance nodes reason about potential outcomes of the action usually reflecting the stochasticity of the addressed problem. In our case we also use decision nodes to decide how to proceed from a given state. For the chance nodes we also consider a stochasticity which results from the ambiguity in the intended goal of the task in contrast to unexpected or random events. Fig. 2 shows the structure of the proposed tree. For further insight and a general overview of MCTS-based algorithms and applications we refer the reader to Browne et al. [7].

B. Teach-and-Improvise

Our algorithm consists of four main steps. First we select a promising leaf in the tree that we expand. Then the selected leaf is expanded by appending template-selection nodes, goal-selection nodes and action-selection nodes. Using the intention likelihood we simulate scores for the added nodes and in a last step do a backpropagation to compute the best solutions. We repeat this for a predefined number of iterations or until no node can be extended further. The branch with the highest score gives us the plan for the task.

V. EXPERIMENTAL EVALUATION

A. Action Learning

We attached markers to hand and torso to record the human teacher poses. Using the adapted data of human demonstrations from our approach, we build action models using a Gaussian mixture model representation. Fig. 3 shows exemplary trajectories generated for the actions of grasping a door handle and opening the door. A detailed experimental evaluation of the action learning can be found in our current work [5], [6].

B. Task Learning

1) Improvising Task Goals: Using our proposed approach we conducted experiments to reproduce demonstrated object arrangements from random starting states. Fig. 4 shows how the system improvises solutions for the task demonstrated in Fig. 1 for different starting states. In the third row in Fig. 4
Fig. 3: In the top row the figure illustrates exemplary generated trajectories to grasp the handle (left) and then open the door (right). The bottom row shows an image of the action demonstration (left) and the PR2 during task execution (middle and right).

Fig. 4: Improvising solutions for the task of arranging the objects on the table. In the left column the respective initial states are visualized. The right column shows the solutions generated by our approach. The third row shows an improvised solution in clutter of other objects. In this case as many relations as possible are satisfied while others, i.e., the relations considering the location on the table cannot be fulfilled.

Fig. 5: On the top row we show two final states of demonstrations for the task of arranging the objects on the table. In the bottom row we see a random initial state with the same objects (left) and the improvised arrangement achieved by the robot (right).

On the other hand we demonstrated how demonstrations of tasks can be used to infer solutions that satisfy the user intention. Experiments show that we are able to reproduce complex actions with constrained trajectories as well as improvise solutions for tasks involving multiple objects.

So far the described modules are implemented independently on the robot. Current work aims to fuse the approach of learning individual actions into our task learning approach. Our experiments have shown that on the one hand the representation of actions as a joint model for base and arm motions is well suited for concatenation as part of complex tasks. It allows an easy transition between actions as it is suited to reposition the robot base for the next action’s demands. On the other hand the task learning module provides the information about which object to handle and the desired goal pose. So far these were manually given to the robot when reproducing demonstrated actions. Experiments with the task learning framework only considered point to point actions that do not require a certain trajectory. The combination of both approaches will be able to include complex actions as part of the planning as, for example, opening a shelf door to fetch something from inside.

VI. CONCLUSION AND OUTLOOK

We have shown that it is possible for a robot to learn task and action representations directly from a human by observing demonstrations with an RGBD camera. On the one hand we presented an approach to learn mobile manipulation actions relying only on human demonstrations and knowledge about the robot’s kinematic and grasping capabilities.

we see the proposed solution for the same task in a clutter of other, irrelevant objects.

2) Robot Experiments: We implemented our approach on a PR2 robot. Fig. 5 shows the improvised accomplished solution for the object arrangement task, where individual manipulation actions were carried out by an existing motion planner.

REFERENCES