

# CSE 598 Project Proposal: Imitation Learning with Baxter Robot using Hi-Fives

Michael Drolet, Frankie Liu, Evan Lam

**Abstract**—The goal of this project is to successfully learn hi-fives for human-robot interaction. We will be using an Imitation Learning approach by incorporating Bayesian Interaction Primitives [1]. Through expert-guided demonstrations, we train the robot to learn relationships between human and robot trajectories. We demonstrate that the robot is able to complete the interaction with a human and successfully issue a hi-five.

## I. INTRODUCTION

Teaching robots how to interact with humans is an interesting challenge in the field of Human Robot Interaction. The athletic and social components which make humans successful in daily interactions are often difficult to capture. Imitation Learning is a useful tool for these situations, where we are able to capture the ideal actions performed by a human and partner. However, teaching robots how to learn new tasks and interact with humans can be done using a variety of methods. Imitation learning, which uses a human expert to guide the interaction, is one popular approach that we seek to use for this experiment. Reinforcement Learning is another popular option for learning tasks that humans are experienced at. We hypothesize that a Reinforcement Learning approach over a continuous and high-dimensional state space will require a significant number of demonstrations in order for the robot to learn the hi-five interaction, compared to the Bayesian Interaction approach. As a result, for a simple interaction that is able to be generalized across different robots using limited samples, we found the Bayesian Interaction approach to be suitable for our project.

Throughout the semester in CSE598, we were introduced to a number of concepts in Imitation Learning. Dynamic Motor Primitives, which are a powerful tool for Imitation Learning, are compelling due to their biologically-inspired approach. We saw in the case of the frog example, how linear dynamical systems can be created to model the trajectory of an object. However, since many systems have nonlinear behaviour, there is a need to add a forcing function. The forcing function makes the dynamical system nonlinear, and as a drawback, things can quickly get unstable. We hypothesize that this may not be an issue with a more simplistic hi-five interaction, so the DMP approach is also attractive. Ultimately, the uncertainty of how much control theory and differential equations are needed to use DMP's led us to consider other options. On the other hand, we are confident that we could model the forcing by learning weights for the basis function decompositions of a trajectory, as this is what we are doing in our current approach.

## II. RELATED WORK

### A. Bayesian Interaction Primitives

The Bayesian Interaction Primitive (BIP) framework is a novel and powerful approach that is being used by the Interactive Robotics Lab at Arizona State University. This SLAM inspired algorithm is useful for encoding spatiotemporal information about the interaction into a state space. BIP approximates each dimension using a weighted linear combination of time-dependent basis functions. In this case, we are using Gaussian basis functions. Included in a state vector are the weights for every basis function, the phase, and phase velocity. As a result, BIP allows us to infer the current state, given previous states (demonstrations) and current observations. Formatting this probabilistically allows the inference to be done using Bayes Filters (in particular, a Kalman Filter) during testing to guide the robot. Since the robot's degrees of freedom are no longer being observed during testing, the BIP framework is based on obtaining a partial observation of the current state to generate the positions/variables of the whole state space. This approach is effective for learning on different types of robots, as there is no domain specific knowledge encoded into the algorithm.

## III. PROBLEM STATEMENT

The goal of this project is for the Baxter robot to successfully issue hi-fives. In our judgment (and speaking with Dr. Amor), the criteria for success consists of a quick hi-five with a high success rate. We initially wanted to do most of our training and demonstrations on just the robot, and not using VREP. This would allow us to focus more on the results and methods, rather than having to deploy onto two different platforms. Additionally, all members of our group have never used VREP before, so it would be a bit of a learning curve. However, with the coronavirus present, our plans had to shift to using the VREP simulation tool for testing our algorithms. Overall, we have decided to use a hybrid approach- that is, using simulation and in-person examples- to present our demonstrations. This includes capturing data in real world through the Optitrack motion capture system. The training can be easily verified in-person; however, we would like to demonstrate the ability to train and test in simulation. This would require additional components, such as OpenPose and Kinect, considering we will not have access to an Optitrack system at home.

### A. Simulation

The simulation tool we are using is VREP. This is a nice tool since its interface is user-friendly and there are useful

resources to begin developing with it. On the other hand, we were considering using Gazebo due to its ability to integrate tightly with ROS. However, Gazebo appeared to be heavy compared to other simulators, so we ultimately decided on VREP. One of the challenges we faced while using VREP was that its interface is different than the way we currently interact with the Baxter robot. That is, the simulator has commands to set joint angles, joint velocities, and other options by using the built-in commands. However, we are using ROS topics within the IntPrim framework to control the Baxter robot. Consequently, switching from simulation to real-life would require a large number of code changes and a need to maintain two different types of controllers. This lead us to looking into using ROS topics for the VREP simulator, and using PyRep to handle the callbacks from receiving a message. PyRep can call internal LuaScript functions in the simulator, and has greater speed than the Blue Origin (simx). This means that the high frequency topics would be handled better using this PyRep/ROS integration, compared to vanilla VREP, and also gives us the ability to only maintain one ROS controller for both the simulation and real Baxter robot. The next challenge is to figure out which components should be done in-person (which is difficult due to the virus), and see which components can be done by using the simulator.

#### *B. In person*

We predict that we will be using the simulator to handle a greater portion of the project if the virus prevents us from meeting up; however, we will see if it is practical to also use the real robot during this time. One benefit of conducting in-person experiments is that the physics is perfect (ground truth) and we do not have to deal with any physics engine computation delays which would alter the results of perception and motion planning. A huge advantage of using the in-person Baxter robot is that we are able to find joint angle key frames for the training pose easily. For example, we put the robot into puppet mode and then recorded its key positions for a hi-five. Then, all that is required is to interpolate between these points to obtain a motion sequence. This method allowed us to quickly create different hi-five poses and allows us to come up with new poses if we decide to improve our hi-five motion. Lastly, the biggest benefit of using the Baxter in-person is that we have access to a controlled environment- with an Optitrack system, high speed network, and device drivers all configured- so that development and deployment is seamless.

### IV. EXPERIMENTS

Our group conducted approximately 75 demonstrations (hi-fives) on the Baxter robot. Most of these test hi-fives were done using Frankie as the HRI partner, however we did not want our model to overfit the trajectories by using only one person. Therefore, we decided to also use Evan as a HRI partner during training. The hi-five was generated using key frames, as discussed in the section above. Since we learned how to conduct the experiments repeatedly and generate a model for testing, we are now focusing on creating

a better hi-five to be used. This will require retraining, since the robot is now learning a new interaction. We plan to use the Extended Kalman Filter (EKF) to perform the filtering in the experiment, as this filter is suitable for smaller samples sizes (compared to the Extended Kalman Filter) and offers good performance in practice.

#### *A. Data Generation*

During the experiments, we collected subject data by using the Optitrack motion capture system, where 3D poses were collected at 120 Hz. Subjects wore a headband and two gloves that are equipped with motion capture markers. By using only three degrees of freedom, we avoid performing inference in a space that is too high-dimensional. After all, hands and head position are likely the only relevant information needed to perform a successful hi-five. We are using ROS as our communication protocol and to handle messages. Topic names that come with the Baxter robot are used, but we will need to replicated/subscribed inside of VREP in order for the deployment to work on both methods.

#### *B. Biomechanics*

It was suggested by Dr. Amor that we do a biomechanical analysis on the hi-five motion. This would be a nice complement to our work, since we can compare the mechanics of a human-human hi-five to a human-robot hi-five. However, at first glance, we were unable to find research on the biomechanics of a hi-five. We decided that- instead- we can perform our own biomechanical analysis on the human-human motion and compare that to the human-robot motion. This would include using the Optitrack motion capture system for two people, and recording the motion of the hi-fives. We can break up the interaction into one phase, and analyze each part of the phase, similar to how a gait cycle can be evaluated. This will give us a useful benchmark to compare against.

### V. PARAMETER TUNING

Some of the parameters that need to be tuned are the number of basis functions per degree of freedom and the scaling factor for each. This is done through a grid-search that tests various combinations during testing and reports the configuration that results in the lowest Mean Absolute Error (MAE) from the ideal trajectory.

### VI. GOING FORWARD

Since we are in the beginning stages of the project, it would inappropriate to include a Discussion/Analysis and Conclusion in this report. Hence, we will discuss our plans going forward. In the future, we plan to finish writing the ROS subscribers and PyRep interfaces in VREP. This will allow us to do offline experiments (simulations, per say) and when/if things subside with the virus, we can test the results in person. Additionally, we plan to meet and use the Optitrack system for the biomechanical analysis, as this will provide us with the most accurate results.

## REFERENCES

- [1] J. Campbell, A. Hitzmann, S. Stepputtis, S. Ikemoto, K. Hosoda, and H. Ben Amor. Learning Interactive Behaviors for Musculoskeletal Robots Using Bayesian Interaction Primitives. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, November 2019.