# Hierarchical Prediction and Planning for Human-Robot Collaboration

An T. Le[1], Philipp Kratzer[1], Simon Hagenmayer[1], Marc Toussaint[2,3] and Jim Mainprice[1,2]
firstname.lastname@ipvs.uni-stuttgart.de
[1]Machine Learning and Robotics Lab, University of Stuttgart, Germany
[2]Max Planck Institute for Intelligent Systems ; IS-MPI ; Tübingen/Stuttgart, Germany
[3]Technische Universität Berlin ; TUB ; Germany

***Abstract*— In this paper, we tackle the problem of human-robot coordination in sequences of manipulation tasks. Our approach integrates hierarchical human motion prediction with Task and Motion Planning (TAMP).**

**We first devise a hierarchical motion prediction approach by combining Inverse Reinforcement Learning and short-term motion prediction using a Recurrent Neural Network. In a second step, we propose a dynamic version of the TAMP algorithm Logic-Geometric Programming (LGP) [1]. Our version of Dynamic LGP, replans periodically to handle the mismatch between the human motion prediction and the actual human behavior. We assess the efficacy of the approach by training the prediction algorithms and testing the framework on the publicly available MoGaze dataset [2].**

## I. INTRODUCTION

As robots become more capable, they will increasingly share space with humans. Consider the case of tidying a kitchen, where the human is interested in having maximal support from the robot while requiring a minimal amount of interference with its task objectives. Humans do this naturally when they collaborate. For instance, one could put food back into the fridge while the other collects and cleans the dishes. The case falls into Human-Robot Collaboration (HRC) category, which focuses on robotic systems able to perform joint actions with humans [3], [4].

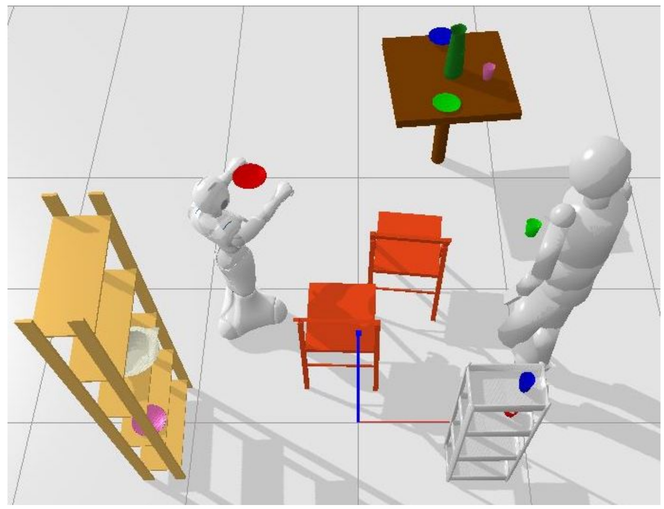

Fig. 1: Pepper carries a plate while the human from the MoGaze dataset is carrying a green cup. The supporting motion plan for setting the table resulting with from Dynamic LGP minimally interferes with the human task.

## II. HIERARCHICAL MOTION PREDICTION & PLANNING

For motion planning, we introduce Dynamic LGP, which is a variant of LGP, to solve the minimal interference Human-Robot tasks. The basic idea of LGP is to decompose the task with two levels of abstraction. At the highest level we consider a discrete set of actions $A = \{a_i\}_N^{i=1}$, for instance move, pick and place. We call a *skeleton*, a sequence of symbolic actions $a_{1:K}$. A fully instantiated plan is then a skeleton, together with a motion trajectory $x : [0, T] \to \chi$, where $\chi = \mathcal{C} \times \mathcal{H} \times \mathcal{O}$, the Cartesian product of the robot, human and movable object configuration spaces respectively.

In our experiments, we consider a cost function $c : (q_t, \dot{q}_t, \ddot{q}_t, s) \mapsto c_t \in \mathbb{R}$, is a combination of differentiable maps, penalizing velocities and accelerations of the robot.

Obstacle avoidance and goal manifold are enforced using equality and inequality constraints $h_p,\ g_p$ in the phase $k(t) \in [t/T]$ conditioned on a discrete symbolic state $s_k \in \mathbb{S}$.

To impose transition conditions between phases, the switch functions $h_{sw},\ g_{sw}$ define equalities and inequalities constraints conditioned on the transition action $a_k$. We assume that the equality and inequality functions are differentiable.

The task is then to find a global path $x : t \mapsto x_t$, which minimizes the following LGP:

$$\min_{x, a_{1:K}, s_{1:K}} \int_0^{KT} c(x(t), \dot{x}(t), \ddot{x}(t), s_{k(t)}) \mathrm{dt}$$

$$\begin{aligned}
\text{s.t.}\quad & \\
& x(0) = x_0,\ h_{\text{goal}}(x(KT)) = 0,\ g_{\text{goal}}(x(KT)) \leq 0 \\
& \forall t \in [0, KT] : h_p(x(t), \dot{x}(t), s_{k(t)}) = 0, \\
& \qquad\qquad g_p(x(t), \dot{x}(t), s_{k(t)}) \leq 0 \\
& \forall k \in \{1, ..., K\} : h_{sw}(x(t), \dot{x}(t), a_k) = 0 \\
& s_k \in \text{exec}_{a_k}(s_{k-1}) \\
& s_K \in \mathbb{S}_{\text{goal}}
\end{aligned}$$

where the path is global continuous $x$ and contains $K \in N$ phases, each has fixed duration $T > 0$.

For a given symbolic goal set, we rely on action feasibility and state transition checks operations, which allow us to instantiate a search process using any tree search algorithm (e.g., depth first, breadth first). If a feasible skeleton $a_{t:K}$ leading to symbolic goal state $s^g \in \mathbb{S}_{\text{goal}}$ is found, a Non-Linear "trajectory optimization" Program (NLP) is defined. The NLP considers geometric switches in the system kinematics with long-term dependencies. In our implementation, we use an interior point method [5], [6] to optimize this NLP.

| | |
|---|---|
| (on X Y) | check if exists a stable 3D $xy\phi$ joint from X to Y |
| (at X Y) | check if $\|x_X - x_Y\|_2 \leq r \| r \in \mathbb{R}$ |
| (carry X Y) | check if exists a stable free joint (6D) from X to Y |

TABLE I: Predicate inference

| | |
|---|---|
| Start State | (0, 4, 0, 1, 0, 3, 1, 0, 1, 2) |
| Actions | Go to white shelf<br>Pick up cup<br>Go to table<br>Place |
| End State | (1, 3, 0, 1, 0, 3, 1, 0, 1, 0) |

TABLE II: Example high-level trajectory

*1) Dynamic planning:* As the actual human behavior may deviate from the prediction, the motion trajectory or the skeleton $a_{1:K}$ may become sub-optimal or even unfeasible. Hence a crucial component for dynamic LGP is to infer the current symbolic state from the current environmental condition.

For example, the predicate (on X Y) is inferred by checking in the system kinematic tree if there is a stable 3D $xy\phi$ joint from X to Y. Table I describes our setup symbolic inference for the predicates using the system kinematics. Specifically, querying (human-carry, ?x - object) or (agent-carry, ?x - object) predicates can be done using (carry X Y) check. This symbolic query is the primary mechanism encoding the human intention into the domain design. For example, in the set-table task, the proposition describes the object carried by the human (human-carry, object) defined to be in the goal set, assuming that the human intentions are always to cooperate to complete the task. Then the robot can plan the remaining actions to reach the goal.

### A. Long-Term Motion Prediction using Hierarchies

The motion prediction layer infers likely symbolic and geometric changes in the workspace, given an initial symbolic state $s_t$ and a geometric state $x_t$. Recall that the geometric state $x_t = (q_t, h_t, o_t)$, concatenates the robot $q_t$, the human $h_t$, and the movable object $o_t$ configurations.

At the top level, our hierarchical motion prediction uses Maximum Entropy Inverse Reinforcement Learning (MaxEnt IRL) [7] and a low-level which performs full-body motion prediction conditioned on sequence of sub-goals induced by sequence of high-level actions given by the top-level.

*1) Goal-Conditioning:* To be able to use motion prediction as a sub-policy, we need a predictive function $h_{t+1:T} = f(h_{0:t}, g^*)$ that computes a trajectory of future human states $h_{t+1:T}$ given previous observed states $h_{0:t}$ and a goal $g^*$. We use VRED, a recurrent neural network-based model for predicting motion [8] and make it goal-conditioned by adding a three-dimensional position $g_t$ to the input of the network at every timestep. Note that we could also use other planning-based predictors together with MaxEnt IRL. However, we use VRED due to the scaling property of deep models that learns high-dimensional configuration trajectory of human-motion captures.

| | Single planning | Dynamic planning |
|---|---|---|
| Success rate | 91.2% | 100% |
| Symbolic plan time | $0.0005 \pm 0.0001(sec)$ | $0.0006 \pm 0.0002(sec)$ |
| Task time reduction | $0.298 \pm 0.078$ | $0.300 \pm 0.100$ |
| Path ratio | 1.000 | $0.626 \pm 0.155$ |
| LGP replan count | - | $3.0 \pm 0.87$ |

TABLE III: Dynamic LGP with Human Prediction

*2) Network State Representation:* We learn a policy $\pi$ that can schedule discrete goals using tabular MaxEnt IRL algorithms based on state frequency calculations. For the MoGaze dataset the discretized state is given by the number of objects on a location and the human position as follows: (cups-table, cups-shelf1, cups-shelf2, plates-table, plates-shelf1, jugbowl-table, jugbowl-shelf1, jugbowl-shelf2, humanPos). The action space is discretized similarly. An example skeleton can be seen in Table II.

We use heuristics for interfering with the exact goal for the human hand or pelvis, for example, by computing the closest point on the table to the human which is not occupied. The heuristics could be further improved by the use of human intention prediction as in [9].

## III. Results & Conclusions

### A. Long-Term Motion Prediction using Hierarchies

We first compare the original VRED implementation with the VRED conditioned on goal inputs on the MoGaze dataset. Results show that the goal-conditioned prediction network achieves both a better angular loss of 7.99 instead of 10.14 and a significantly better position loss of 3.84 instead of 12.56, than the network without goal-conditioning. This is expected because the goal-conditioned network uses oracle information of the goal position.

To test the accuracy of the high-level policy, we extracted the task of setting up the table for one person from the dataset. We then run tabular MaxEnt IRL, showed that the learned policy solved the task in 80% of the cases. However, a perfect imitation was achieved solely in 16% of the test runs of the cross-validation. This is because the algorithm is limited by our symbolic state and action representation. Including more complex state features, e.g., from the 3d environment, could further improve the algorithm.

### B. Dynamic LGP with Long-Term Prediction

In this experiment, we choose 8 data segments from MoGaze, and produce the Long-Term Prediction outputs described in Section II-A. We run 5 task instances for each segment to capture the human motion prediction statistics due to its stochasticity. The overall task IoU between the robot and the human objects is $0.34 \pm 0.13$. Table III reports task statistics for this experiment. It shows that Dynamic LGP has higher success rates and produces shorter paths and needs slightly more time to complete than single planning. These plans also reduce the total time to execute the task by a factor approaching 2, which is what one would expected when two agents collaborate at a task.

ACKNOWLEDGMENT

This work is partially funded by the research alliance "System Mensch". The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Philipp Kratzer.

## REFERENCES

[1] M. Toussaint, "Logic-geometric programming: An optimization-based approach to combined task and motion planning." in *IJCAI*, 2015, pp. 1930–1936.

[2] P. Kratzer, S. Bihlmaier, N. B. Midlagajni, R. Prakash, M. Toussaint, and J. Mainprice, "Mogaze: A dataset of full-body motions that includes workspace geometry and eye-gaze," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 367–373, 2020.

[3] A. M. Bauer, D. Wollherr, and M. Buss, "Human-Robot Collaboration - a Survey." *I. J. Humanoid Robotics*, vol. 05, no. 01, pp. 47–66, 2008. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/S0219843608001303

[4] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, "Progress and prospects of the human–robot collaboration," *Autonomous Robots*, vol. 42, no. 5, pp. 957–975, Nov. 2017. [Online]. Available: "http://dx.doi.org/10.1007/s10514-017-9677-2

[5] J. Mainprice, N. Ratliff, M. Toussaint, and S. Schaal, "An interior point method solving motion planning problems with narrow passages," in *IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, 2020.

[6] J. Mainprice, "bewego: Differentiable kinetics optimization library," 2019–2021. [Online]. Available: https://github.com/humans-to-robots-motion/bewego

[7] B. D. Ziebart, A. L. Maas, J. A. Bagnell, and A. K. Dey, "Maximum entropy inverse reinforcement learning." in *Aaai*, vol. 8, 2008, pp. 1433–1438.

[8] H. Wang and J. Feng, "Vred: A position-velocity recurrent encoder-decoder for human motion prediction," *arXiv preprint arXiv:1906.06514*, 2019.

[9] P. Kratzer, N. B. Midlagajni, M. Toussaint, and J. Mainprice, "Anticipating human intention for full-body motion prediction in object grasping and placing tasks," in *IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, 2020.