

Skill Generalization via Inference-based Planning

Muhammad Asif Rana, Mustafa Mukadam, S. Reza Ahmadzadeh, Sonia Chernova and Byron Boots
Institute for Robotics & Intelligent Machines, Georgia Institute of Technology, Atlanta, GA, USA
{asif.rana, mmukadam3, reza.ahmadzadeh, chernova, bboots}@gatech.edu

Abstract—We present a novel approach which unifies conventional learning from demonstration (LfD) and motion planning using probabilistic inference, for generalizable skill reproduction. As a part of this approach, we also present a new probabilistic skill model that requires minimal parameter tuning, and is more suited for encoding skill constraints and performing inference in an efficient manner. Preliminary experimental results on a manipulation skill are also provided.

I. INTRODUCTION

As robots assume more collaborative roles alongside humans in dynamic environments, they must have the ability to learn new skills and adapt them to novel scenarios (including, changes in start/goal states, or the environment). Toward this end, probabilistic trajectory-based LfD enables robots to learn skills from multiple human demonstrations.

Purely probabilistic approaches [1, 2] however, attract reproduced trajectories towards an average form of the demonstrated motions, without regard to the initial robot state. Relatively more generalizable probabilistic approaches [3–5], need extensive parameter tuning to avoid under-fitting or over-fitting. Non-parametric Gaussian process (GP) based approaches [6, 7] overcome the extensive manual parameter tuning problem by learning the GP hyper-parameters from demonstrations via maximum likelihood, but fail to include the underlying system dynamics which are crucial for certain skills. Probabilistic movement primitives (ProMPs) [8], include system dynamics in the GP formulation and perform inference to generalize skills over different initial and/or goal states. However, ProMPs require a phase variable to encode time-dependence of system states, making them prone to temporal distortions. A major drawback for all these GP-based methods is that they carry a high computational overhead due to inversion of a dense kernel matrix.

Furthermore, many conventional LfD approaches are not equipped to handle arbitrarily placed obstacles [1, 2]. Of those that do, generated trajectories only preserve reaching an end goal, and not other properties [3, 8, 9]. An exception is the work of Ye and Alterovitz [10], which learns a probability distribution from demonstrations and adapts it for the current scenario using a sampling-based motion planner as an ad-hoc post-processing step. Another exception [11], perhaps more similar to ours, instead carries out trajectory optimization over the learned distribution, first in the presence of obstacles and then for the given via points. However, in both these approaches, there is a certain level of redundancy induced by carrying out skill generalization as a multi-step process and assuming that each step is independent. Conversely, our

approach provides a unifying framework for LfD and motion planning by carrying out generalized skill reproduction as a one-shot process.

Our specific contributions include: (i) a probabilistic skill model (trajectory prior) that extracts the spatial and temporal correlations among the demonstrations and considers the underlying system dynamics while requiring minimal parameter tuning; and (ii) a novel reproduction method that finds a continuous-time trajectory via probabilistic inference, which is *optimal* with respect to the learned skill model while remaining *feasible* when subjected to different scenarios.

II. TRAJECTORY OPTIMIZATION AS PROBABILISTIC INFERENCE

We argue that for generalizable skill reproduction, LfD should also have the same motivation as motion planning, that is, finding trajectories that are optimal and feasible. However, in contrast to motion planning (where optimality is pre-specified and hard coded), LfD would require optimality to be learned from demonstrations. We adopt the probabilistic inference perspective on motion planning [12], where we seek to find the posterior distribution of the trajectory, $p(\theta|\mathbf{e}) \propto p(\theta)p(\mathbf{e}|\theta)$, given some random binary events, \mathbf{e} , for example, collision avoidance, starting from a given location, reaching a desired goal/via-point or a combination thereof. Here,

1) *The Prior*: $p(\theta) \propto \exp\{-\frac{1}{2}\|\theta - \mu\|_{\mathcal{K}}^2\}$, defines optimality and is learned from demonstrations. This can also be interpreted as learning the hyper-parameters (mean μ and covariance \mathcal{K}) of the trajectory distribution. We use structured Gaussian processes to model the prior [13](Section III).

2) *The Likelihood*: $p(\mathbf{e}|\theta) \propto \exp\{-\frac{1}{2}\|\mathbf{h}(\theta, \mathbf{e})\|_{\Sigma}^2\}$, encodes feasibility and defines the probability of the events occurring given the trajectory. We model this as a distribution in the exponential family [12].

We can find the *maximum a posteriori* (MAP) trajectory i.e. the mode of the posterior distribution through inference,

$$\theta^* = \operatorname{argmax}_{\theta} \{p(\theta)p(\mathbf{e}|\theta)\} \quad (1)$$

This gives the desired trajectory that is optimal and feasible. Our key insight is that skill reproduction in any new scenarios is in fact equivalent to performing planning as inference.

Following [12], we use factor graphs [14] to represent distributions and use the duality between inference and optimization to arrive at a fast and efficient approach that solves (1).

III. TRAJECTORY PRIOR AS SKILL MODEL

We model the prior using Gaussian processes (GPs) such that for any collection of times $\mathbf{t} = \{t_0, \dots, t_N\}$, the trajectory parametrized by support states θ_i , follows a Gaussian distribution, $\theta \doteq [\theta_0 \dots \theta_N]^\top \sim \mathcal{N}(\boldsymbol{\mu}, \mathcal{K})$. Unlike earlier mentioned GP based approaches, our choice of structured GPs not only takes care of the higher-order system dynamics (velocities, accelerations etc.), but also allows learning and inference to be computationally efficient.

A. Structured Heteroscedastic GP From LTV-SDE

We consider trajectories as solutions of a linear time-varying stochastic differential equation (LTV-SDE), $\dot{\boldsymbol{\theta}}(t) = \mathbf{A}(t)\boldsymbol{\theta}(t) + \mathbf{u}(t) + \mathbf{F}(t)\mathbf{w}(t)$. Here, $\mathbf{A}(t)$ and $\mathbf{F}(t)$ are time-varying system matrices, $\mathbf{u}(t)$ is a bias term, with an additive white noise process, $\mathbf{w}(t) \sim \mathcal{GP}(\mathbf{0}, \mathbf{Q}_C(t)\delta(t-t'))$. The state $\boldsymbol{\theta}(t)$, comprises of the vectorized positions and any higher-order time derivatives (for all degrees of freedom). Taking the first and second order moment of the solution to the LTV-SDE, yields the desired GP [13]. Learning the GP hyperparameters is thus equivalent estimating underlying LTV-SDE parameters. The inverse covariance matrix of this GP has a sparse block diagonal structure [13] that enables efficient learning and inference.

Usually, only the demonstrated workspace trajectories are relevant for skills. Therefore, we choose to *learn from demonstrations*, a prior distribution $p(\mathbf{x}|\boldsymbol{\theta})$, generated by using the LTV-SDE over the workspace state $\mathbf{x}(t)$, instead of that in configuration space $\boldsymbol{\theta}(t)$. However, since the problem of finding the associated trajectories in configuration space becomes under-constrained for high-degree-of-freedom robots, we add an additional *pre-specified* smoothness constraint (constant-velocity prior). This is given by $p(\boldsymbol{\theta}) \propto \exp\{-\frac{1}{2}\|\boldsymbol{\theta} - \boldsymbol{\mu}^\theta\|_{\mathcal{K}^\theta}^2\}$, analogous to that used in [12, 15]. Eventually for skill reproduction, a combined prior, $p_x(\boldsymbol{\theta}) = p(\boldsymbol{\theta}|\mathbf{x}) \propto p(\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta})$, is utilized for MAP inference in (1).

B. Learned Workspace Prior

The workspace prior distribution mentioned in Section III-A is defined as, $p(\mathbf{x}|\boldsymbol{\theta}) \propto \exp\{-\frac{1}{2}\|\mathbf{C}(\boldsymbol{\theta}) - \boldsymbol{\mu}^x\|_{\mathcal{K}^x}^2\}$. Here, \mathbf{C} imposes the robot kinematic constraints, mapping the configuration space to workspace. For the skills we consider and ease of implementation, a discrete version of the LTV-SDE above, proved sufficient to learn the LTV-SDE parameters. We learn these parameters from demonstration by performing maximum likelihood estimation with ridge regression. We used an end-effector state composed of 3D positions and linear velocities, but our approach can be easily used to encode orientations and angular velocities as well. Note that the learned prior requires no manual parameter tuning and is directly available for inference.

IV. EXPERIMENTS

We validated the proposed method on a *box-opening* skill using a Kinova JACO² 6-DOF arm. We provided 6 demonstrations using kinesthetic teaching with varying initial end-effector states (varying initial position, zero initial velocity),

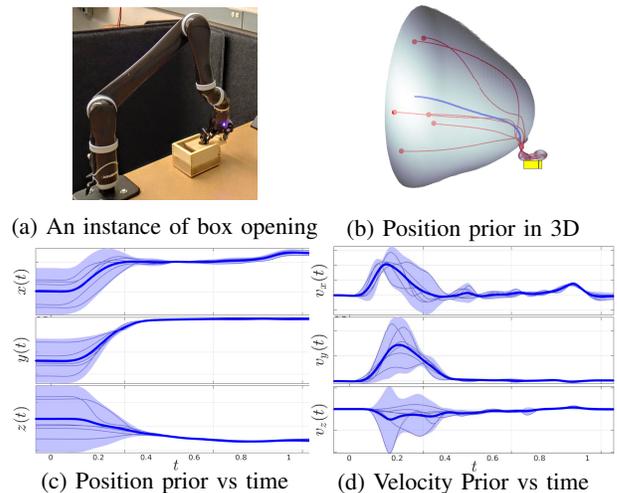


Fig. 1: Prior for *box-opening* skill. The mean is in blue with 95% confidence interval. Demonstrations are overlaid.

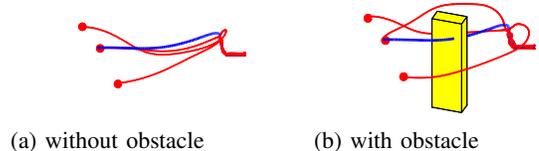


Fig. 2: Reproduced trajectories in red from different initial states. The obstacle is in yellow and the prior position mean is in blue.

and aligned them using dynamic time warping. The end-effector position data with time was recorded at 100Hz, and the corresponding linear velocities were estimated by fitting a cubic smoothing spline and differentiating. The workspace support states were selected by uniformly re-sampling the spline. Fig. 1 shows the learned prior distribution, $p(\mathbf{x}|\boldsymbol{\theta})$.

The skill is composed of two primitive actions, reaching and sliding the lid of the box. The sliding part of the skill is highly constrained compared to the reaching part. Hence, as shown in Fig. 1 (c)-(d), the variance in the state variables (i.e. positions and velocities) becomes much smaller during the sliding part. It should be noted here that the prior also encodes the coupling between the state variables.

Fig. 2 shows the reproduced MAP trajectories $\boldsymbol{\theta}^*$, found after conditioning the combined prior over the likelihood, $p_x(\boldsymbol{\theta})p(\mathbf{e}|\boldsymbol{\theta})$. For reproduction with different initial states, the likelihood contained the observation for the given initial state with a very small Σ , since we are certain about the initial state. For reproduction in a new environment with obstacles, the likelihood also contained the collision-free likelihood from [12], where the obstacle cost is evaluated using a precomputed signed distance field. Here we set the parameter Σ manually such that it enables desired clearance of the robot from the obstacle. In general, Σ depends on the size of the robot, desired clearance and the environment itself. As shown in Fig. 2, the robot is able to carry out the crucial sliding motion from three different initial states and in the presence of a new obstacle in the environment. Since the direction of motion is highly relevant to this skill, encoding velocities in the prior proved particularly beneficial.

REFERENCES

- [1] Sylvain Calinon, Florent Guenter, and Aude Billard. On learning, representing, and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(2):286–298, 2007.
- [2] Benjamin Reiner, Wolfgang Ertel, Heiko Posenauer, and Markus Schneider. Lat: A simple learning from demonstration method. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4436–4441. IEEE, 2014.
- [3] S Mohammad Khansari-Zadeh and Aude Billard. Learning stable nonlinear dynamical systems with gaussian mixture models. *IEEE Transactions on Robotics*, 27(5):943–957, 2011.
- [4] Sylvain Calinon, Irene Sardellitti, and Darwin G Caldwell. Learning-based control strategy for safe human-robot interaction exploiting task and robot redundancies. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 249–254. IEEE, 2010.
- [5] Aleš Ude, Andrej Gams, Tamim Asfour, and Jun Morimoto. Task-specific generalization of discrete and periodic dynamic movement primitives. *IEEE Transactions on Robotics*, 26(5):800–815, 2010.
- [6] David B. Grimes, Rawichote Chalodhorn, and Rajesh PN Rao. Dynamic imitation in a humanoid robot through nonparametric probabilistic inference. In *Robotics: Science and Systems (RSS)*, pages 199–206, 2006.
- [7] Markus Schneider and Wolfgang Ertel. Robot learning by demonstration with local gaussian process regression. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 255–260. IEEE, 2010.
- [8] Alexandros Paraschos, Christian Daniel, Jan R Peters, and Gerhard Neumann. Probabilistic movement primitives. In *Advances in neural information processing systems*, pages 2616–2624, 2013.
- [9] Peter Pastor, Heiko Hoffmann, Tamim Asfour, and Stefan Schaal. Learning and generalization of motor skills by learning from demonstration. In *2009 IEEE International Conference on Robotics and Automation (ICRA)*, pages 763–768. IEEE, 2009.
- [10] Gu Ye and Ron Alterovitz. Demonstration-guided motion planning. In *International symposium on robotics research (ISRR)*, volume 5, 2011.
- [11] Dorothea Koert, Guilherme Maeda, Rudolf Lioutikov, Gerhard Neumann, and Jan Peters. Demonstration based trajectory optimization for generalizable robot motions. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 515–522. IEEE, 2016.
- [12] Jing Dong, Mustafa Mukadam, Frank Dellaert, and Byron Boots. Motion planning as probabilistic inference using Gaussian processes and factor graphs. In *Proceedings of Robotics: Science and Systems (RSS-2016)*, 2016.
- [13] Sean Anderson, Timothy D. Barfoot, Chi Hay Tong, and Simo Särkkä. Batch nonlinear continuous-time trajectory estimation as exactly sparse gaussian process regression. *Autonomous Robots*, 39(3):221–238, 2015.
- [14] Frank R Kschischang, Brendan J Frey, and H-A Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2):498–519, 2001.
- [15] Mustafa Mukadam, Xinyan Yan, and Byron Boots. Gaussian process motion planning. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9–15, 2016.