Principles for an Alternative Design of Movement Primitives that Uses Probabilistic Inference in Learned Graphical Models

Elmar Rückert[1], David Kappel[1], Gerhard Neumann[2], Marc Toussaint[3] and Wolfgang Maass[1]

***Abstract*— For motor skill learning in high-dimensional continuous action spaces we propose to endow movement representations with an intrinsic probabilistic planning system, integrating the power of stochastic optimal control methods within a movement primitive. The parametrization of the primitive is a graphical model that represents the dynamics and constraints, encoded by an intrinsic cost function, such that inference in this graphical model yields the control policy. We parametrize the intrinsic cost function using task-relevant features, such as the importance of passing through certain via-points. The system dynamics as well as intrinsic cost function parameters are learned in a reinforcement learning setting. This alternative movement representation complies with salient features of biological movement generation, i.e. its modular organization in elementary movements, its characteristics of stochastic optimality under perturbations, and its efficiency in terms of learning [1].**

I. INTRODUCTION

Stochastic optimal control or probabilistic inference methods [2], [3] are promising approaches for motor control of high-dimensional redundant systems. Movement primitives based on probabilistic inference in learned graphical models have interesting features, i.e. they can model motor variability often observed in human movement generation, they can adapt to changing task constraints without re-learning, and they can compete with classical control approaches in terms of learning speed and robustness to noise [1].

Movement planning can be formulated as graphical model inference problem [2], where states, controls and constraints (e.g. desired goals, abstract features or obstacles) are implemented as random variables. An example of a graphical model for T time-steps is shown in Figure 1 (A). Here states are denoted by $\mathbf{x}_{1:T}$, controls by $\mathbf{u}_{1:T-1}$ and constraints by $z_{1:T}$. The stochastic process is defined by:

$$P(\mathbf{x}_{1:T}, \mathbf{u}_{1:T-1}, z_{1:T}) = P(\mathbf{x}_1) \prod_{t=1}^{T-1} P(\mathbf{u}_t|\mathbf{x}_t) \times \prod_{t=1}^{T} P(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_{t-1}) \prod_{t=1}^{T-1} P(z_t|\mathbf{x}_t, \mathbf{u}_t) \ ,$$

where $P(\mathbf{u}_t|\mathbf{x}_t)$ denotes the state dependent prior for the controls, the distribution $P(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_{t-1})$ the state transition model and $P(\mathbf{x}_1)$ the initial state distribution.

[1]Institute for Theoretical Computer Science, Graz University of Technology, Austria {rueckert,kappel, maass}@igi.tugraz.at

[2]Intelligent Autonomous Systems, Technische Universität Darmstadt, Germany neumann@ias.tu-darmstadt.de

[3]Department of Computer Science, Universität Stuttgart, Germany marc.toussaint@informatik.uni-stuttgart.de

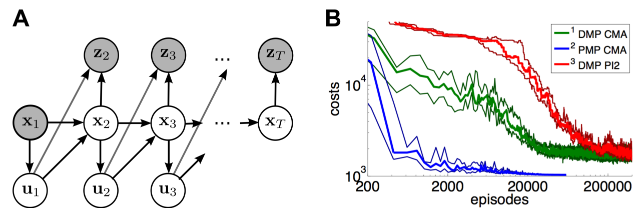

Fig. 1. (**A**) Movement generation as graphical model inference problem. (**B**) Via-point task learning curves of the the proposed movement representation (PMP) and a standard dynamic movement primitive (DMP) approach [1].

For inferring the controls, we compute the posterior $P(\mathbf{x}_{1:T}, \mathbf{u}_{1:T-1}|z_{1:T} = 1)$ over trajectories, conditioned on fulfilling the constraints ($z_t = 1$) at each time-step t.

The advantage of this approach is that there is no distinction between sensor and motor, perception and action. We can include a multitude of random variables, some of which might represent features of the state, some might represent goals or constraints in the future. Moreover, many sophisticated inference techniques can be used, e.g. message passing, belief propagation, variational inference methods or Markov Chain Monte Carlo sampling approaches.

Also for learning the parameters of these models, a multitude of policy search methods are known [4]. In [1] it was shown that reinforcement learning can be used to learn the system dynamics and the intrinsic cost function. This method significantly outperformed a standard dynamic movement primitive approach in terms of convergence speed and noise robustness, which is illustrated for a noisy via-point reaching task in Figure 1 (B).

Recently probabilistic inference has also been proposed as a promising brain computing principle that may explain neural activities [5]. The dynamics of spiking neural networks can be characterized by the same graphical model, which provides an interesting paradigm towards understanding biological motor control. In first results we reproduced important properties of cursor reaching tasks of monkeys.

REFERENCES

[1] E. Rückert, G. Neumann, M. Toussaint, and W. Maass, “Learned graphical models for probabilistic planning provide a new class of movement primitives,” *Frontiers in Comp. Neuro.*, vol. 6, no. 97, 2013.

[2] M. Toussaint and C. Goerick, “A bayesian view on motor control and planning,” in *From Motor Learning to Interaction Learning in Robots*, ser. Studies in Computational Intelligence, 2010, vol. 264, pp. 227–252.

[3] H. Kappen, V. Gómez, and M. Opper, “Optimal control as a graphical model inference problem,” *Mach. Learn. 87(2)*, pp. 159–182, 2012.

[4] J. Kober and J. Peters, “Policy Search for Motor Primitives in Robotics,” *Machine Learning*, vol. 84, no. 1-2, pp. 171–203, 2011.

[5] M. Botvinick and M. Toussaint, “Planning as inference,” *Trends in Cognitive Sciences*, vol. 16, no. 10, pp. 485 – 488, 2012.