Context-Aware Task Execution Using Apprenticeship Learning

Ahmed Faisal Abdelrahman, Alex Mitrevski, and Paul G. Plöger

Introduction

Non-adaptive robot behaviours can be particularly unsatisfactory for human-oriented tasks such as object hand-overs, where factors like posture, approach direction, and individual capability can affect user perceptions and task success [1][2]. We postulate that *contextual adaptivity*: the ability to adapt to the different contexts of a task, can improve the behaviour of service robots.

We present an approach to the acquisition of a context-adaptive robot motor skill, and a user study conducted to validate improvements over a non-adaptive approach. We propose an apprenticeship learning method for the problem of executing contextually adaptive robot-to-human object hand-overs. The strategy enables a robot to learn from an expert's example, encoding motion trajectories in dynamic movement primitives (DMPs), followed by experiential learning of contextualized policies, with which different variants of the demonstrated action are executed according to context.

Three Contexts of a Hand-Over Task: Receiver Posture

We present a model-based version of the C-REPS algorithm [3] for extending the demonstrated, static hand-over policy to a contextually adaptive one, which can learn policies in simulation before transferring to the real system.

Proposed Apprenticeship Learning Approach

We combine **learning from demonstration** and **reinforcement learning**, such that a robot learns to replicate a demonstrator's execution of the task, and then to produce different executions that adapt to context, from experience.

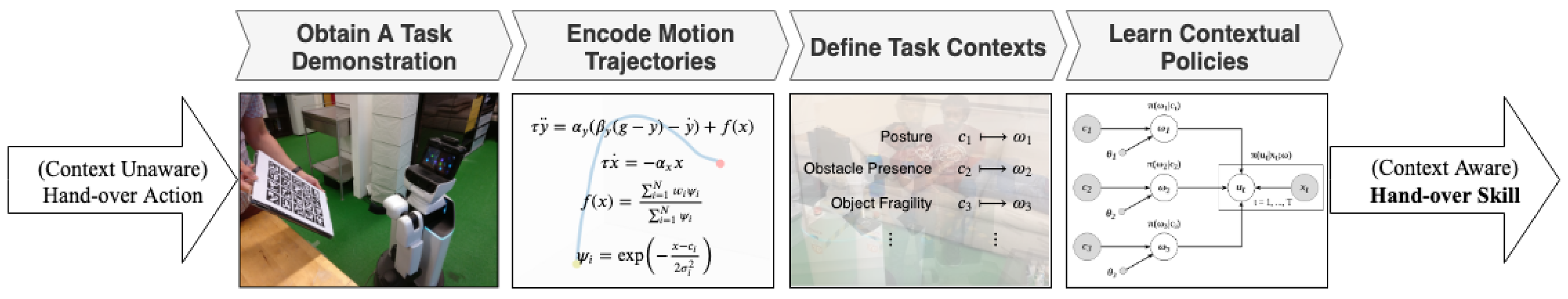
The robot first observes and records the characteristics of a demonstrated hand-over trajectory using a motion-capture method [4].

The trajectory is then encoded in DMPs, by extracting the parameters: initial and final position, $\mathbf{y}(0)$ and g, trajectory shape weights, W, and time constant, τ.

Context variables, $\mathcal{C} = \{c_1, c_2, c_3, ...\}$, are then defined and mapped to these learnable parameters, $\Omega = \{g, W, \tau, ...\}$, such that executions can be made adaptive to them. Examples of hand-over contexts include:

1. Receiver Posture: $c_1 \in \{standing, seated, lying\_down\} \longmapsto g$
2. Obstacle Presence: $c_2 \in \{True, False\} \longmapsto W$
3. Object Fragility: $c_3 \in \{True, False\} \longmapsto \tau$

Finally, we use the C-REPS algorithm to learn these mappings in upper-level policies: $\pi(\omega|c)$, whose sampled parameters alter the behaviour of the previously static, lower-level DMP policy: $\pi(u|x, \Omega)$. The result is hand-over executions that are adaptive to multiple dimensions of context.



The C-REPS Algorithm

Contextual Relative Entropy Policy Search (C-REPS) enables learning contextual policies that optimize behaviours for agent state as well as current context. It relies on a hierarchical policy decomposition consisting of:

- Upper-level policy, $\pi(\omega|c)$: a linear-Gaussian model
- Lower-level policy, $\pi(u|x, \omega)$: DMPs

The algorithm aims to maximize return, $R_{c\omega}$, for the distribution of possible trajectories, $p(c, \omega)$, while bounding the relative entropy between successive policy updates by ϵ:

$$\max_{p} \int_c \int_\omega p(c,\omega) R_{c\omega} dc d\omega$$

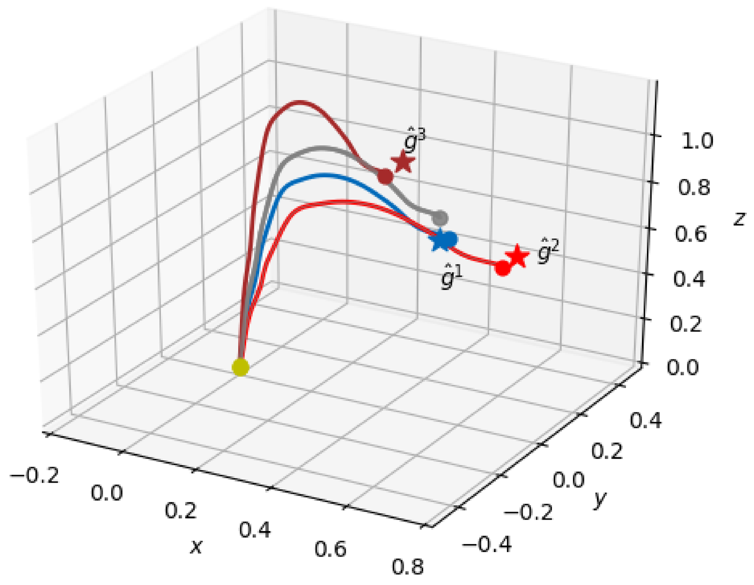
$$\text{s.t.:} \quad \int_c \int_\omega p(c,\omega) \log \frac{p(c,\omega)}{q(c,\omega)} dc d\omega \leq \epsilon$$

Learning Context-Dependent Hand-Over Positions

We focus on learning policy $\pi(g|c_1)$, which samples optimal hand-over position, g, given the posture of the receiver, c_1. The reward function is modelled with knowledge of theoretically ideal positions for each context, $\hat{g}$:

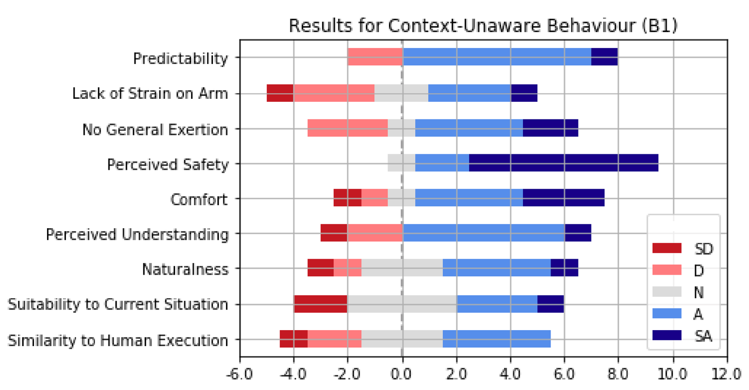
$$\mathcal{R}_g(g, c_1) = \begin{cases} \frac{1}{||g-\hat{g}^1||_2}, & \text{if } c_1 = \text{standing} \\ \frac{1}{||g-\hat{g}^2||_2}, & \text{if } c_1 = \text{seated} \\ \frac{1}{||g-\hat{g}^3||_2}, & \text{if } c_1 = \text{lying down} \end{cases}$$

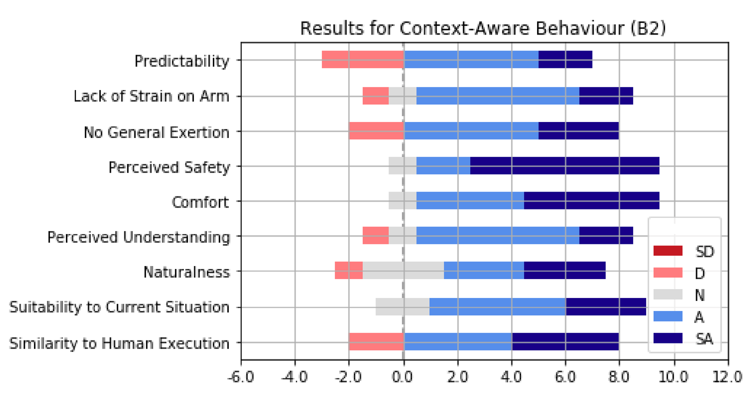
The policy is learned by running roll-outs of simulated hand-over trajectories, and re-estimating parameters $\theta = \{a, A, \Sigma\}$. The sampled positions eventually converge to the desired positions, resulting in adapted trajectories, as shown on the right.



Experimental Results: A User Study

- A simple user study was conducted to validate the benefits of acquired contextual adaptivity to perceived robot behaviour.
- An HSR hands an object over to participants in three contexts (see figure above), under two behaviour modes.
- In the context-unaware mode, the robot reproduces the demonstrated hand-over trajectory in all contexts. In the context-aware mode, it identifies the receiver's posture and uses the learned policy, $\pi(g|c_1)$, to select a hand-over position, and accordingly adapt the trajectory.
- Users fill out a questionnaire that estimates their perceptions of both behaviours across different aspects.
- Results of quantitative (visualized on the right) and qualitative analyses confirms context-adaptive behaviour is instinctively preferable to users.





Conclusions

Our model-based version of C-REPS succeeds in learning contextual hand-over position policies using *mental rehearsal*. Other DMP parameters, W and τ, can be learned to simultaneously adapt executions to multiple contextual dimensions. The user study results validate the advantages of context-aware robot behaviour. In the future, approximating reward functions from multiple demonstrations can enable a life-long learning procedure, while non-parametric variants of C-REPS, and Gaussian Process policy and forward models can be explored.

References

[1] Kheng Lee Koay, Emrah Akin Sisbot, Dag Sverre Syrdal, Mick L Walters, Kerstin Dautenhahn, and Rachid Alami. Exploratory study of a robot approaching a person in the context of handing over an object. In *AAAI spring symposium: multidisciplinary collaboration for socially assistive robotics*, pages 18–24, 2007.

[2] Aaron Bestick, Ravi Pandya, Ruzena Bajcsy, and Anca D Dragan. Learning human ergonomic preferences for handovers. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–9. IEEE, 2018.

[3] Jan Peters, Katharina Mulling, and Yasemin Altun. Relative entropy policy search. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.

[4] Alex Mitrevski, Abhishek Padalkar, Minh Nguyen, and Paul G. Plöger. "Lucy, Take the Noodle Box!": Domestic Object Manipulation Using Movement Primitives and Whole Body Motion. In *Proceedings of the 23rd RoboCup International Symposium*, Sydney, Australia, 2019.

Acknowledgement

We gratefully acknowledge the support by the b-it International Center for Information Technology.

Contact

Ahmed Faisal Abdelrahman, Alex Mitrevski
Autonomous Systems Group, Hochschule Bonn-Rhein-Sieg
Email: ahmed.abdelrahman@inf.h-brs.de,
aleksandar.mitrevski@h-brs.de

b-it Bonn-Aachen International Center for Information Technology

ICRA 2020 31 May • 4 June 2020 Palais des Congrès de Paris Paris • FRANCE

Hochschule Bonn-Rhein-Sieg University of Applied Sciences