Enabling Customer-Driven Learning and Customisation Processes for ML-Based Domestic Robots

Fabrice Matulic, Yuta Kikuchi, Jason Naradowsky
Preferred Networks Inc.
Tokyo, Japan
{fmatulic,kikuchi,narad}@preferred.jp

ABSTRACT

Smart domestic robots are poised to revolutionise the way household chores and everyday tasks are carried out in the home of the future. At the heart of the "intelligence" and behaviour of these robots will be complex machine learning (ML) systems that, in addition to extensive training at the manufacturing stage, will most likely require further on-site adjustments to adapt to customers and their environments. Drawing from the robotics literature on Learning from Demonstration and Human Robot Interaction, we review relevant techniques which we hypothesise customers could realistically use to perform these adaptation and customisation steps as smoothly and effortlessly as possible.

KEYWORDS

Human Robot Interaction, Learning from Demonstration

INTRODUCTION

Recent years have seen an increasing number of personal robots providing support for various domestic tasks, ranging from everyday household chores to assisting disabled and elderly people. The market is expected to grow annually at double-digit rates in the next decade, reaching $35 million by 2022, according to P&S Market Research [1]. This growth is fuelled by advances in manufacturing, electronics, and robotics, but lately increasingly also by the rapid progress made in “artificial intelligence”. The recent boom of deep learning is gradually enabling the transition from dumb machines executing simple repetitive tasks or requiring full manual control, to smart devices that can autonomously service humans and more naturally interact with them. Machine learning is involved

HCML Perspectives Workshop at CHI 2019, May 4 2019, Glasgow, UK

at several stages within that sophisticated process: in the algorithms that analyse the environment and the task context, as well as in the modules that enable efficient communication with users using human-based interactions such as speech, gestures, etc.

The more sophisticated the processes and the more diverse the target objects to manipulate, the more machine learning models require examples for training. Considering the wide variety of households and people, it is unlikely that factory-trained models will be sufficient for complex tasks involving many unknown elements, e.g. tidying up a room or minding a baby. To tackle these kinds of highly context-dependent challenges, the robot will have to learn or hone their skills directly from the customers, i.e. from non-technical users with limited time and patience. This means that the learning and customisation processes need to be as simple and short as possible.

A relevant line of research addressing such issues is Learning from Demonstration (LfD) also known as imitation learning[5, 7]. In LfD a robot is trained to execute tasks via interaction with humans, potentially even non-experts. There is considerable work dealing with LfD in the robotics literature with many of the explored techniques addressing tasks relevant to personal robots, e.g. picking and placing, finding motion paths in a room with obstacles, recognising common objects, responding to natural speech input etc. In this short paper, we outline the machine-learning needs of future smart domestic robots and examine what we believe are relevant LfD approaches that can be leveraged so that customers will be able to efficiently teach and customise their products.

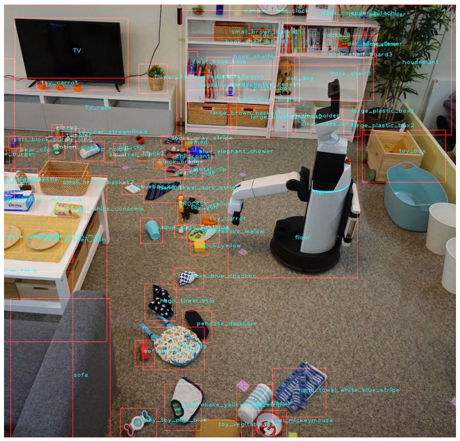
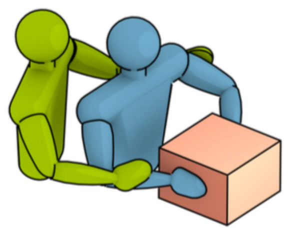


Figure 1: Prototype of our autonomous room-tidying robot presented at CEATEC 2018 [25]. The robot was trained to recognise and classify hundreds of different household objects, and to move these objects to their desired places. This destination can vary depending on object type, e.g. rubbish in one bin, toys in another, and slippers arranged next to each other. Users can specify areas to tidy up and indicate destination containers using speech input and pointing gestures, but the robot is not able to recognise new objects and integrate them in the tidying process. Nor can the user provide assistance to the robot, to improve its understanding of the task or to teach it entirely new tasks.

MACHINE LEARNING-DRIVEN DOMESTIC ROBOTS

There are a variety of domestic tasks that can potentially be accomplished by robots (some of which already are). We describe the challenges of customer-driven robot teaching in the context of tidying up a cluttered room. This complex task involves a number of technical hurdles and subproblems that can be overcome with the help of (user-guided) machine learning. Based on our recent experience of building a prototype of an autonomous room-tidying robot (see Figure 1 and [25]), we identify the following subtasks to be performed by the machine: moving about in the room (while avoiding obstacles), recognising objects to put away, determining how to pick them up (grasping), deciding where they belong, and how to place them at their destination or dispose of them. While the robot's machine-learning models will have been bootstrapped with extensive prior knowledge of typical household items and room environments, unavoidably the customer will have to instruct the robot what to do with these items as well as help it recognise previously unseen objects. Further support from the user to ensure smooth and safe navigation of the robot will likely also be required. We believe that this additional context-specific information can be provided by the user with minimum overhead using adequate techniques from LfD and related learning methods.

Direct Teaching



Observational Learning

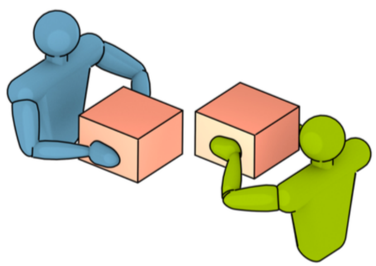


Figure 2: Two main categories of LfD techniques (illustrations from Calinon[7]).

Learning Behaviour from Demonstration

In the machine learning field, a major distinction of LfD approaches is between action-driven and goal-driven learning. In **action-driven learning** (often referred to as behaviour cloning), the demonstration is treated as supervision and learning takes place as the robot tries to mimic the basic movements of the demonstrator. Despite the low-level view of the problem, it is not a simple task as the physical abilities of the robot do not correspond directly to the human's, and the robot often has a different viewpoint and orientation. A second approach to LfD is **goal-driven learning** (sometimes referred to as inverse-reinforcement learning), where the robot attempts to infer what the goal of the task is, and then determines which policy (sequence of its actions) would best achieve it. In doing so, the robot abstracts away from trying to find a direct correspondence to each movement, therefore becoming more robust to physical differences between robot and demonstrator.

The field of robotics typically divides LfD along a different axis, depending on what and how the demonstration is provided (See left column)[5]. In **direct teaching**, (a.k.a. kinaesthetic teaching), the teacher physically manipulates the robot, directly guiding its movements to perform the task. The robot's sensors can record task information directly, allowing it to perform well in situations close to the original demonstration where there is little need to generalise[8, 22]. In **observational learning** a human demonstrates the task with their own body, and the robot must learn how these actions translate to its own physical constraints. Typically, the observation is captured with visual sensors (camera), and can utilise recent advances in vision[26], object detection [10] and pose estimation [24].

The conditions of personal robotics tend to favour goal-driven observational learning over the alternatives as the demands on a non-expert user are comparatively low [2]. For instance, in the case of direct teaching, it may be difficult for users to guide robotic arms to pick up and manipulate household items. Furthermore, the fast-changing nature of the household environment means that robots will likely have to make strong generalisations as items are constantly found in new locations and orientations. Many tasks are also hard to decompose into movements, for instance, in cooking eggs there is a natural variance to the materials that warrants focusing on understanding the goal, rather than the actual demonstrated movements leading to the result.

Unlike the academic setting, a personal robot exists in an extended interaction with the users in the household, and thus continual learning becomes an important concern. How should the behaviour of the robot be corrected? While direct teaching is a demanding form of interaction if it is the basis for all learning, it may be useful to correct or fine-tune the robot's behaviour. Such hybrid approaches are rarely studied in the academic setting, but may be an effective strategy. Similarly, behaviour cloning may be of use to bootstrap a specific action that can then be followed by more abstract goal-based teaching and possibly even reinforcement learning techniques, where the robot could autonomously perfect its execution of a task. An example of the latter approach has been proposed for towel folding tasks [6]. In these contexts, to keep user intervention at a low level, one-shot learning strategies might also be employed (see left column).

The task demonstration itself is not the only way human interaction can improve such systems. For example, a user can also communicate positive or negative feedback, helping to refine the robot's understanding of the desired behaviour[9, 21].

Learning New Words and Concepts

LfD enables robots to learn sophisticated behaviours from human interaction, but a personal robot must also learn new concepts. For instance, how would a robot understand what is meant by "Sarah's jacket", which room is "the office", or that a certain toy has a name and is subject to different considerations? For a robot to generalise to new environments, it must be able to acquire this knowledge of previously unseen objects.

This is a challenge for domestic robotics. Both object detection and language grounding typically rely on pre-trained models, which are often trained on images and vocabulary consisting of tens of thousands, even millions of items. Standard datasets are manually created by ML researchers, requiring technical knowledge and preprocessing which is difficult to recreate within the home with non-experts.

We thus argue for the need to learn new concepts and new vocabulary in the context of learning the behaviours they support. Recent research in visual question answering [18, 19] has shown that current deep models are capable of inducing such concepts from pairs of images and text instructions of the sort, "*Is there a red cube to the left of the green sphere?*". Such methods induce a simple syntax, allowing them to isolate which sections of the image are likely to refer to which words. Analogously, if a demonstrator accompanies the action of moving a coat with the statement "*I'm putting Sarah's jacket in the closet*." it could provide the necessary cues for inducing the concept in a similar way. This has been previously demonstrated for navigation tasks [16].

Scaling this to the personal home domain has several challenges. Humans interact with each other not in strict, unambiguous text commands, but with fast, informal speech, thus necessitating the development of more targeted and accurate speech recognition. Still, the contents of each speech utterance are difficult to decode. Having a rich understanding of the world (and other humans) allows us to exclude information that might be inferable from the context. Modelling such pragmatic reasoning in the context of interaction is an exciting and necessary direction of research for natural verbal interaction [15]. Yet another challenge is the development of less synthetic data for bootstrapping command-based learning, or finding less data-hungry methods [4].

One-shot Learning

Machine learning algorithms typically require exposure to thousands (or millions) of examples to learn sophisticated behaviour. This need for big data renders many methods impractical for personal robot applications. In contrast, there is a growing interest in learning from just a few, or even just a single example. Such *one-shot* approaches are a promising strategy for learning from end users, who would otherwise likely grow frustrated from repeatedly demonstrating a task.

One-shot learning differs from traditional supervised learning in its emphasis on knowledge *transfer*, helping to abstract away from inconsequential aspects of the task, and thus allowing the model to generalise quickly to new situations. Precisely what and how to transfer remains an active research topic. Existing work has shown success in one-shot learning of object classes [12, 13], and for task learning from demonstration [11]. Common approaches utilise shared parameters or representations, or meta-learning [14].

However, there are many challenges in extending one-shot learning to real-world settings. Consider a demonstration of placing slippers together. The robot must determine "*Did the user place the object (slipper) there because that spot (x-y-z coordinate) is ideal, or because that spot is next to another similar object (the other slipper)?*" Humans reason with a broader prior knowledge of the consequences of these decisions (e.g. wanting slippers to be together when putting them on). Here a deeper interaction with users (for instance, pairing instructions with gesture or language) may aid models in overcoming such ambiguity.

Customisation after Deployment

How to adapt a trained model to new situations is an active topic of academic research. As models learn new tasks, existing model parameters may be adjusted, and the performance on previously-mastered tasks may suffer (this process is known as *catastrophic forgetting*). Current research aims to maintain the model's ability to perform old tasks during additional training, sometimes by holding parameters fixed, and sometimes by increasing model capacity as more tasks are added[17, 20, 23].

However, customisation in a real product poses its own unique set of challenges and prompts a discussion of what degree user customisation in ML products should be supported. A looming concern is user safety: A robot trained to chop vegetables, whose parameters for that task are altered, may potentially behave in dangerous ways. Malicious users may also intentionally teach robots to perform dangerous activities. It is therefore important to consider the extent to which users can customise such systems.

One standard way to limit customisation while supporting novel behaviour is to provide only an API to pre-trained behaviours. Behind such a wall, users can create new skills by combining existing ones, or register newly-taught skills in a manner where they can be reviewed prior to allowing them to be deployed in the robot. Further risk assessment is warranted to determine potential safety concerns and an appropriate set of control measures.

CONCLUSION

While in academia there has been considerable research efforts looking at teaching robots tasks via user-friendly techniques such as learning from demonstration, there are still few applications of those concepts in the consumer market. There is currently no personal robot that can autonomously accomplish complex domestic tasks that require am in-depth understanding of the household environment and/or that involves numerous different (including previously unknown) objects. In this paper, we have identified promising approaches from the literature that we believe can be leveraged to tackle some of those challenges. One aspect that has not been extensively tested in academic validation experiments is scalability and it is unclear how tasks like tidying a room and cooking a meal can be reliably taught so that the robot can perform them well in a large variety of contexts. Moreover, there are a host of other issues that need to be addressed before such robots can be deployed in the homes of regular consumers, including safety and security matters, which are often neglected in academic research (see left column and [3]), as well as human robot interaction questions surrounding communication with the robot when it cannot make a decision on its own. We are in the process of developing LfD and interactive machine learning approaches for these concrete real-world contexts with non-expert customers as users. We are hopeful that many techniques will prove useful and will help smart robots become an integral part of the home.

While the all-round robot butler that can autonomously perform multiple household chores is still far away, we believe that we are at the dawn of a new era in which smart domestic robots powered by advanced machine learning systems will be able to efficiently assist people for specific household tasks. Successful solutions to train, refine and customise those ML systems will likely combine several machine- and user-centric methods from traditional programming and teaching to different flavours of learning from demonstration techniques.

REFERENCES

[1] 2017. *Global Personal Robots Market Size, Share, Development, Growth and Demand Forecast to 2022*. Technical Report. P&S Market Research.

[2] Saleema Amershi, Maya Cakmak, William Bradley Knox, and Todd Kulesza. 2014. Power to the people: The role of humans in interactive machine learning. *AI Magazine* 35, 4 (2014), 105–120.

[3] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul F. Christiano, John Schulman, and Dan Mané. 2016. Concrete Problems in AI Safety. *arXiv* abs/1606.06565 (2016).

[4] Jacob Andreas, Dan Klein, and Sergey Levine. 2018. Learning with Latent Language. In *NAACL-HLT*. 2166–2179.

[5] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. 2009. A Survey of Robot Learning from Demonstration. *Robot. Auton. Syst.* (2009), 469–483.

[6] Benjamin Balaguer and Stefano Carpin. 2011. Combining imitation and reinforcement learning to fold deformable planar objects. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 1405–1412. https://doi.org/10.1109/IROS.2011.6094992

[7] Sylvain Calinon. 2018. *Learning from Demonstration (Programming by Demonstration)*. Berlin, Heidelberg.

[8] Sylvain Calinon and Aube Billard. 2007. Incremental learning of gestures by imitation in a humanoid robot. In *HRI*. 255–262.

[9] Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep Reinforcement Learning from Human Preferences. In *NeurIPS*. 4299–4307.

[10] Coline Devin, Pieter Abbeel, Trevor Darrell, and Sergey Levine. 2018. Deep Object-Centric Representations for Generalizable Robot Learning. In *ICRA*. 7111–7118.

[11] Yan Duan, Marcin Andrychowicz, Bradly Stadie, OpenAI Jonathan Ho, Jonas Schneider, Ilya Sutskever, Pieter Abbeel, and Wojciech Zaremba. 2017. One-Shot Imitation Learning. In *Advances in Neural Information Processing Systems 30*. 1087–1098.

[12] Li Fei-fei. 2006. Knowledge transfer in learning to recognize visual object classes. In *ICDL*.

[13] Michael Fink. 2005. Object Classification from a Single Example Utilizing Class Relevance Metrics. In *NeurIPS*, L. K. Saul, Y. Weiss, and L. Bottou (Eds.). 449–456.

[14] Chelsea Finn, Tianhe Yu, Tianhao Zhang, Pieter Abbeel, and Sergey Levine. [n. d.]. One-Shot Visual Imitation Learning via Meta-Learning. In *CoRL*.

[15] Daniel Fried, Jacob Andreas, and Dan Klein. 2018. Unified Pragmatic Models for Generating and Following Instructions. In *NAACL-HLT*. 1951–1963.

[16] Daniel Fried, Ronghang Hu, Volkan Cirik, Anna Rohrbach, Jacob Andreas, Louis-Philippe Morency, Taylor Berg-Kirkpatrick, Kate Saenko, Dan Klein, and Trevor Darrell. 2018. Speaker-Follower Models for Vision-and-Language Navigation. In *NeurIPS*. 3318–3329.

[17] Spyros Gidaris and Nikos Komodakis. 2018. Dynamic Few-Shot Visual Learning Without Forgetting. In *CVPR*. 4367–4375.

[18] Ronghang Hu, Jacob Andreas, Marcus Rohrbach, Trevor Darrell, and Kate Saenko. 2017. Learning to Reason: End-to-End Module Networks for Visual Question Answering. In *ICCV*. 804–813.

[19] Justin Johnson, Bharath Hariharan, Laurens van der Maaten, Li Fei-Fei, C. Lawrence Zitnick, and Ross Girshick. 2017. CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning. In *CVPR*.

[20] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences* 114, 13 (2017), 3521–3526.

[21] W. Bradley Knox and Peter Stone. 2010. Combining Manual Feedback with Subsequent MDP Reward Signals for Reinforcement Learning. In *AAMAS*. 5–12.

[22] D. Lee, C. Ott, Y. Nakamura, and G. Hirzinger. 2011. Physical human robot interaction in imitation learning. In *ICRA*. 3439–3440.

[23] German Ignacio Parisi, Ronald Kemker, Jose L. Part, Christopher Kanan, and Stefan Wermter. 2018. Continual Lifelong Learning with Neural Networks: A Review. *arXiv* abs/1802.07569 (2018).

[24] Xue Bin Peng, Angjoo Kanazawa, Jitendra Malik, Pieter Abbeel, and Sergey Levine. 2018. SFV: Reinforcement Learning of Physical Skills from Videos. *ACM Trans. Graph.* 37, 6 (2018).

[25] Preferred Networks Inc. 2018. Autonomous Tidying-up Robot System. https://projects.preferred.jp/tidying-up-robot/en/.

[26] Tianhe Yu, Chelsea Finn, Sudeep Dasari, Annie Xie, Tianhao Zhang, Pieter Abbeel, and Sergey Levine. 2018. One-Shot Imitation from Observing Humans via Domain-Adaptive Meta-Learning. In *RSS*.