Knowledge Transfer from a Human Perspective

Ramya Ramakrishnan
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139
ramyaram@mit.edu

Julie A. Shah
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139
julie_a_shah@csail.mit.edu

ABSTRACT

Transfer in reinforcement learning (TiRL) is a challenging research problem. Agents are still preprogrammed for specific tasks or can only transfer knowledge under limited circumstances. To quicken the learning process, people, who are extremely adaptable, may be able to provide feedback to guide the agent. This requires an interpretable medium for transfer that both the human and agent can understand. In this work, we propose a transfer medium based on object mappings between tasks. We conduct human subject experiments to test whether *people* are able to effectively use these mappings in the form of advice to play video games. Preliminary results show that good mappings improve people's transfer performance on some games, but can hurt people when they misunderstand. The potential interpretability benefits of using an object-based representation for TiRL can guide the development of more interpretable transfer learning algorithms for agents.

1. INTRODUCTION

Robots acting in the real world must adapt to new situations and environments. The process of generalizing previously learned knowledge from one or more source tasks to a new target task is known as transfer learning. Transfer in reinforcement learning (TiRL) [1, 2] is challenging for robots, but people are very adaptable and can easily detect similarities between various situations. If a robot could use human feedback to help the transfer process, it could learn new tasks much faster. This, however, requires the robot to express its knowledge in an interpretable way so that humans can understand and accordingly help.

Prior works in cognitive science [3, 4] have indicated the importance of objects in human cognition and have shown that even infants are able to understand the concept of objects. This has inspired much work on object-based representations for transfer in RL [5, 6, 7] to help speed up learning in a task. Based on these works, we hypothesize that objects can be an interpretable medium to relate two similar tasks. We define "interpretability" as communication of an agent's reasoning for decision-making to improve human understanding of the learning process. We assess interpretability by measuring subjective perception of the information provided as well as objective task performance when a person uses this information to perform a task.

In this work, we propose a transfer framework that provides mappings between objects in a target task and objects in a source task. Our hypothesis is that by providing relationships between objects, the framework can communicate interpretable information useful for transfer. We conduct human experiments to assess whether this framework is interpretable to people. We present preliminary results and insights from these experiments, which will guide future work in developing more interpretable RL transfer algorithms.

Appears in: *Proceedings of the 1st Workshop on Transfer in Reinforcement Learning (TiRL) at the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), A. Costa, D. Precup, M. Veloso, M. Taylor (chairs), May 8–9, 2017, São Paulo, Brazil.*

2. RELATED WORK

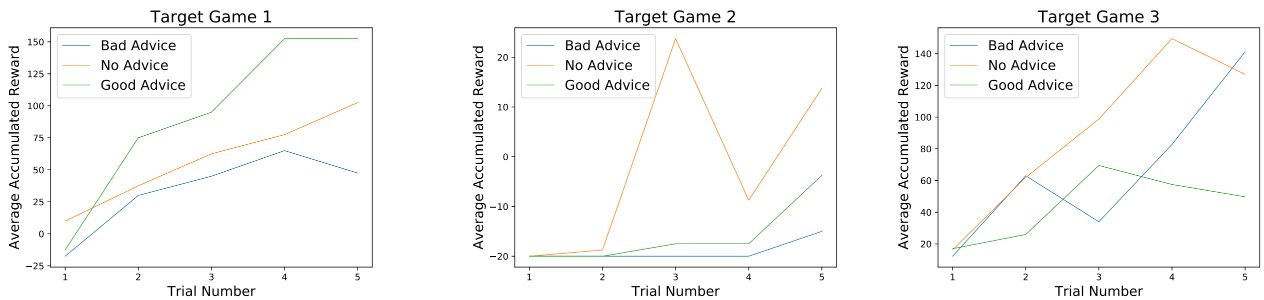
Many works [8, 9, 10, 11] have studied human feedback in agent learning. Similar to our goal, these methods aim to bridge the gap between people and agents, but they do not focus on the transfer process. On the other hand, TiRL [2] has been explored in many applications such as games [7, 12] and robotics [13], but there is limited work on assessing the interpretability of these approaches.

One transfer approach [14] uses task-level descriptors to perform zero-shot transfer in a target task without any training data. Another work [15] develops an algorithm to use advice from multiple teachers to aid an agent in transfer. Results show that theoretically, good teachers help, while bad teachers hinder performance. Some works [16, 17] have assessed the benefits and drawbacks of transferring knowledge from a learned neural network to a range of target tasks. Another method [18] trains a large policy network to learn many tasks through the use of several expert networks. This single network is able to learn new games more quickly with no other guidance. However, these deep learning approaches are limited in that they do not explicitly give information about the learning in the hidden layers of the network. Thus, they cannot be directly used in human-in-the-loop systems.

A recent work [19] aims to combine the advantages of both deep networks and symbolic reasoning. They propose an end-to-end framework that generates symbols from raw game pixels, finds relationships between objects, and learns a Q-network for each pair of object types. They evaluate the transfer capability of their approach but do not explicitly address interpretability of the transferred knowledge.

3. TRANSFER EXPERIMENT

We propose a medium for transfer across tasks based on object mappings that we hypothesize will be interpretable to people. We consider a transfer scenario that has a source task T_S with a set of object classes $C_S = \{c_1, ..., c_m\}$ and



a target task T_T with object classes $C_T = \{c_1, ..., c_n\}$. An object class represents objects with similar behavior (e.g., bullets). We define an object mapping between the tasks as $C_T \rightarrow C_S \cup \{c_{new}\}$, where c_{new} is a new object class. A "good" mapping would specify the most similar source object class for each object class in the target task, or if none were similar, the target object class would be mapped to c_{new}.

We conducted a human subject experiment with 12 participants to assess whether our object mapping medium, as a form of advice, would help people transfer knowledge across tasks. Each participant played three pairs of games, where each pair was a transfer scenario with a source game and a target game. The games were chosen based on pilot experiments such that they have some object similarities but have enough differences to make the transfer task hard for people. For each game pair, participants played the source game up to 10 times, received one of three advice conditions (G-good, B-bad, or N-no mapping), and then played the target game up to 10 times. Each participant saw all advice conditions. We used a Graeco-Latin square design [20] to balance the ordering of advice and game. For example, participant 1 received the ordering B1, G2, N3, where letters specify the advice condition, and numbers specify the game condition; and participant 2 received G3, N1, B2. Participants filled out a questionnaire after the source game, after receiving advice, and after the target game. The questionnaires were used to assess subjective perception of task performance and the usefulness of the mappings. The objective metric was performance on each target game, and the subjective measures were Likert-scale responses from the questionnaires.

Good advice was an object mapping relating objects with similar behavior, and bad advice was an object mapping intended to confuse the participant in playing the new game. For example, good advice would map two objects that you should avoid, while bad advice would map an object that you should retrieve with an object that you should avoid.

4. PRELIMINARY RESULTS

We first analyzed our subjective measures. We ran the Mann-Whitney U test comparing good vs. bad advice, good vs. no advice, and bad vs. no advice. We found that people who were presented with good advice agreed statistically significantly stronger that the advice helped them play the new game ($p = 0.004$) than people who got bad advice. Participants also agreed that they learned the target game more quickly with good advice than with bad ($p = 0.034$). Another very interesting trend was that participants rated the pair of games to be more similar when given good advice than when given bad ($p = 0.084$). This means that by presenting a mapping that highlights the similarities between two tasks, people *perceived* the tasks to be more similar than when presented with an intentionally bad mapping.

While we saw intuitive results in the subjective measures, we observed different trends in the objective metric (i.e., game scores). We considered each game pair separately as the performance across games varied. The figure above shows the learning curve for each game, where each line averages 4 participants' learning curves over the first 5 trials.

In the first game pair, participants learned more quickly with good advice than with bad or no advice. This game pair had "good" objects that needed to be retrieved and "bad" objects that needed to be avoided. This is a simple rule, so a mapping to either of these objects would clearly indicate what action to take. In the second game pair, we find that no advice does better than good and bad. Rather than having simple reactive rules, these games involved "protecting", a vague concept. For example, the source game involved hitting (a short-distance shooting action that requires you to go one cell away from the object and shoot) flames to protect cities, and the target game involved the same hitting action towards friends to protect them from an enemy. While an agent would interpret the two hitting objects as similar, a person may assess similarity on a higher-level, such as which objects should be protected. For the final game pair, both no advice and bad advice did better than good. Similar to the previous pair, the objects in these games involved more complex rules, such as avoiding and shooting the same object. Since the mappings did not indicate what features were similar between the objects, people could relate the wrong set of features and perform worse.

5. DISCUSSION

Results from these experiments showed that people perceived good advice to be better than bad based on subjective ratings. An interesting finding was that we can shape people's perception of task similarity. This insight can help us develop interpretable transfer learning algorithms that present information strategically to enable human understanding of the tasks in a way similar to the agent. From objective measures, we observed that object mappings can improve performance on transfer scenarios that involve simple rules. However, on tasks that involve complicated object interactions, mappings can hinder performance because they do not indicate *how* or *why* the objects map and can lead to misunderstanding. This implies the need for more informative explanations of the transfer process, rather than simple object mappings. These experiments show initial promise in using object-based representations for greater interpretability in transfer, which can inspire the development of TiRL algorithms with more interpretable repesentations.

REFERENCES

[1] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[2] Matthew E Taylor and Peter Stone. Transfer learning for reinforcement learning domains: A survey. *JMLR*, 10:1633–1685, 2009.

[3] Nelson Cowan. The magical mystery four how is working memory capacity limited, and why? *Current directions in psychological science*, 19(1):51–57, 2010.

[4] Elizabeth S Spelke and Katherine D Kinzler. Core knowledge. *Developmental science*, 10(1):89–96, 2007.

[5] Carlos Diuk, Andre Cohen, and Michael L Littman. An object-oriented representation for efficient reinforcement learning. In *Proceedings of the 25th international conference on Machine learning*, pages 240–247. ACM, 2008.

[6] Luis C Cobo, Charles L Isbell, and Andrea L Thomaz. Object focused q-learning for autonomous agents. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 1061–1068. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[7] Alexander Braylan and Risto Miikkulainen. Object-model transfer in the general video game domain. 2016.

[8] Bei Peng, James MacGlashan, Robert Loftin, Michael L Littman, David L Roberts, and Matthew E Taylor. An empirical study of non-expert curriculum design for machine learners. 2016.

[9] Jeffery A Clouse and Paul E Utgoff. A teaching method for reinforcement learning. In *Proceedings of the Ninth International Conference on Machine Learning (ICML?92)*, pages 92–101, 1992.

[10] Richard Maclin and Jude W Shavlik. Creating advice-taking reinforcement learners. *Machine Learning*, 22(1-3):251–281, 1996.

[11] Ofra Amir, Ece Kamar, Andrey Kolobov, and Barbara J Grosz. Interactive teaching strategies for agent training. International Joint Conferences on Artificial Intelligence, 2016.

[12] Andrei A Rusu, Neil C Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. Progressive neural networks. *arXiv preprint arXiv:1606.04671*, 2016.

[13] Coline Devin, Abhishek Gupta, Trevor Darrell, Pieter Abbeel, and Sergey Levine. Learning modular neural network policies for multi-task and multi-robot transfer. *arXiv preprint arXiv:1609.07088*, 2016.

[14] David Isele, Mohammad Rostami, and Eric Eaton. Using task features for zero-shot knowledge transfer in lifelong learning. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-16)*, July 2016.

[15] Yusen Zhan, Haitham Bou Ammar, and Matthew E. Taylor. Theoretically-Grounded Policy Advice from Multiple Teachers in Reinforcement Learning Settings with Applications to Negative Transfer. In *Proceedings of the 25th International Conference on Artificial Intelligence (IJCAI)*, July 2016. 25

[16] Yunshu Du, Gabriel V. de la Cruz, Jr., James Irwin, and Matthew E. Taylor. Initial Progress in Transfer for Deep Reinforcement Learning Algorithms. In *Proceedings of Deep Reinforcement Learning: Frontiers and Challenges workshop (at IJCAI)*, New York City, NY, USA, July 2016.

[17] Ruben Glatt, Felipe Leno da Silva, and Anna Helena Reali Costa. Towards knowledge transfer in deep reinforcement learning. In *Intelligent Systems (BRACIS), 2016 5th Brazilian Conference on*, pages 91–96. IEEE, 2016.

[18] Emilio Parisotto, Jimmy Lei Ba, and Ruslan Salakhutdinov. Actor-mimic: Deep multitask and transfer reinforcement learning. *Proceedings of the 4th International Conference on Learning Representations (ICLR)*, 2016.

[19] Marta Garnelo, Kai Arulkumaran, and Murray Shanahan. Towards deep symbolic reinforcement learning. *arXiv preprint arXiv:1609.05518*, 2016.

[20] Henry B Mann. The construction of orthogonal latin squares. *The Annals of Mathematical Statistics*, 13(4):418–423, 1942.