Curriculum Learning in Reinforcement Learning (Doctoral Consortium)

Sanmit Narvekar
Department of Computer Science
University of Texas at Austin
Austin, TX 78712 USA
sanmit@cs.utexas.edu

ABSTRACT

Transfer learning in reinforcement learning is an area of research that seeks to speed up or improve learning of a complex target task, by leveraging knowledge from one or more source tasks. This thesis will extend the concept of transfer learning to *curriculum learning*, where the goal is to design a sequence of source tasks for an agent to train on, such that final performance or learning speed is improved. We discuss completed work on this topic, including an approach for modeling transferability between tasks, and methods for semi-automatically generating source tasks tailored to an agent and the characteristics of a target domain. Finally, we also present ideas for future work.

General Terms

Algorithms; Performance; Experimentation

Keywords

Reinforcement Learning; Transfer Learning; Curriculum Learning

1. INTRODUCTION

As autonomous agents are called upon to perform increasingly difficult tasks, new techniques will be needed to make learning such tasks tractable. Transfer learning [3, 9] is a recent area of research that has been shown to speed up learning on a complex task by transferring knowledge from one or more easier *source tasks*. Most existing transfer learning methods treat this transfer of knowledge as a one-step process, where knowledge from all the sources are directly transferred to the target. However, for complex tasks, it may be more beneficial (and even necessary) to gradually acquire skills over multiple tasks *in sequence*, where each subsequent task requires and builds upon knowledge gained in a previous task. This is especially important, for example, if the knowledge or experience needed from a source requires prerequisites to obtain.

The goal of this thesis work is to extend transfer learning to the problem of *curriculum learning*. As a motivating example, consider the game of Quick Chess[1] (Figure 1). Quick Chess is a game designed to introduce players to the full game of chess, by using a sequence of progressively more difficult "subgames." For example, the first subgame is a 5x5 board with only pawns, where the player learns how pawns move and about promotions. The second subgame is a small board with pawns and a king, which introduces a new objective: keeping the king alive. In each successive subgame, new elements are introduced (such as new pieces, a larger board, or different configurations) that require learning new skills and building upon knowledge learned in previous games. The final game is the full game of chess.

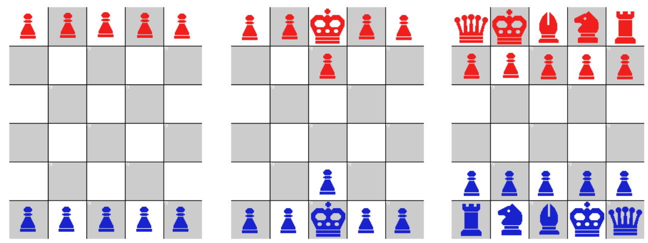

Figure 1: Different subgames in Quick Chess

The question that motivates this line of work is: can we find an optimal sequence of tasks (i.e. a curriculum) for an agent to play that will make it possible to learn some target task (such as chess) fastest, or at a performance level better than learning from scratch?

Designing an effective curriculum is a complex problem that ties task creation, sequencing, and transfer learning. A good set of source tasks is crucial for having positive transfer. A curriculum designer must be able to suggest source tasks using knowledge of the target domain, and that are tailored to the current abilities of the agent. Second, the tasks must be sequenced in a way that allows new knowledge to be accumulated by the agent along each step. During this process, the agent also needs to know how long it should spend in a source in order to acquire the necessary skills, the kind of knowledge it should obtain for transfer, and how to transfer it to the next stage of the curriculum. In the following sections, we discuss progress made towards some of these objectives, and plans for future work in this direction.

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

[1] http://www.intplay.com/uploadedFiles/Game_Rules/P20051-QuickChess-Rules.pdf

2. COMPLETED WORK

There are two pieces of completed work that relate to this thesis. Last year [7], we examined the problem of learning to select which tasks could serve as suitable source tasks for a given target task. Unlike previous approaches, which typically relied on a model of the two task MDPs, or experience (such as sample trajectories) from the two tasks, our approach used a set of meta-level feature descriptors describing each task. We trained a regression model over these feature descriptors to predict the benefit of transfer from one task to another, as measured by the jumpstart metric [9], and showed using a large scale experiment that these descriptors could be used to predict transferability.

This method serves as one way of identifying which task to select as part of a curriculum. However, it assumed that a fixed and suitable set of tasks was provided, which was created using knowledge of the domain. It also only showed successful transfer for a 1-stage curriculum.

At AAMAS 2016, we will present the overall problem of curriculum learning in the context of reinforcement learning [6], and show concrete multi-stage transfer. Specifically, as a first step towards creating a curriculum, we consider how a space of useful source tasks can be generated, using a parameterized model of the domain and observed trajectories of an agent on the target task. We claim that a good source task is one that leverages knowledge about the domain to simplify the problem, and that promotes learning new skills tailored to the abilities of the agent in question. Thus, we propose a series of functions to semi-automatically create useful, agent-specific source tasks, and show how they can be used to form components of a curriculum in two challenging multiagent domains. Particularly, we show that a curriculum is useful and can provide *strong transfer* [9].

The methods we propose are also useful for creating tasks in the classic transfer learning paradigm (1 step curriculum). Past work in transfer learning has typically assumed a fixed set of source tasks are provided, using a static analysis of the domain. This work allows new source tasks to be suggested from a dynamic analysis of the agent's performance.

3. DIRECTIONS FOR FUTURE WORK

In our most recent work [6], we showed how one form of transfer – value function transfer – could be used to transfer knowledge via a curriculum. However, transfer learning literature has shown that alternate forms of transfer are also possible, using options [8], policies [2], models [1], or even samples [5, 4]. Thus, an interesting extension is characterizing what type of knowledge is useful to extract from a task, and whether we can combine multiple forms of transfer for use in a curriculum. For example, it may be that some tasks are more suitable for transferring options, while others are more suitable for transferring samples. Thus, when designing a curriculum, the designer would ask both which task to transfer from, and the type of knowledge to transfer.

Another key question to address is how long to spend on a source task. In our previous work [6, 7], agents trained on source tasks until convergence. However, this is suboptimal for two reasons. First, spending too much time training in a source task can result in overfitting to the source. In curriculum learning, we only care about performance in the target task, and not whether we also achieve optimality in a source. Thus, this excess training can be detrimental. Second, from a practical standpoint, if the necessary skills can be acquired from a source task before reaching convergence, there is no benefit to keep training on it. Doing so would make it harder to show benefits such as strong transfer.

Finally, there is the overall goal of automatically sequencing tasks in the curriculum. Our current work [6] proposed a way of creating a space of tasks, but assumed a separate process (e.g., a human expert) chose which tasks to use. Automating this process completes the loop. Once we have a curriculum for an agent, we can also examine related questions such as how to individualize that curriculum for a new set of agents that have different representations or abilities.

Acknowledgements

This work was done in collaboration with Jivko Sinapov, Matteo Leonetti, and Peter Stone as part of the Learning Agents Research Group (LARG) at the University of Texas at Austin. LARG research is supported in part by NSF (CNS-1330072, CNS-1305287), ONR (21C184-01), AFRL (FA8750-14-1-0070), & AFOSR (FA9550-14-1-0087).

REFERENCES

[1] A. Fachantidis, I. Partalas, G. Tsoumakas, and I. Vlahavas. Transferring task models in reinforcement learning agents. *Neurocomputing*, 107:23–32, 2013.

[2] F. Fernández, J. García, and M. Veloso. Probabilistic policy reuse for inter-task transfer learning. *Robotics and Autonomous Systems*, 58(7):866 – 871, 2010. Advances in Autonomous Robots for Service and Entertainment.

[3] A. Lazaric. Transfer in reinforcement learning: a framework and a survey. In M. Wiering and M. van Otterlo, editors, *Reinforcement Learning: State of the Art*. Springer, 2011.

[4] A. Lazaric and M. Restelli. Transfer from multiple MDPs. In *Proceedings of the Twenty-Fifth Annual Conference on Neural Information Processing Systems (NIPS'11)*, 2011.

[5] A. Lazaric, M. Restelli, and A. Bonarini. Transfer of samples in batch reinforcement learning. In A. McCallum and S. Roweis, editors, *Proceedings of the Twenty-Fifth Annual International Conference on Machine Learning (ICML-2008)*, pages 544–551, Helsinki, Finalnd, July 2008.

[6] S. Narvekar, J. Sinapov, M. Leonetti, and P. Stone. Source task creation for curriculum learning. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, Singapore, May 2016.

[7] J. Sinapov, S. Narvekar, M. Leonetti, and P. Stone. Learning inter-task transferability in the absence of target task samples. In *Proceedings of the 2015 ACM Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. ACM, 2015.

[8] V. Soni and S. Singh. Using homomorphisms to transfer options across continuous reinforcement learning domains. In *In Proceedings of American Association for Artificial Intelligence (AAAI)*, 2006.

[9] M. E. Taylor and P. Stone. Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10(1):1633–1685, 2009.