Leveraging Human Knowledge for Machine Learning Curriculum Design

Matthew E. Taylor
Department of Computer Science
The University of Southern California
taylorm@usc.edu

Abstract

Transfer learning is a successful technique that significantly improves machine learning algorithms by training on a sequence of tasks rather than a single task in isolation. However, there is currently no systematic method for deciding how to construct such a sequence of tasks. In this paper, I propose that while humans are well-suited for the task of curriculum development, significant research is still necessary to better understand how to create effective curricula for machine learning algorithms.

Transfer Learning

Traditional machine learning algorithms often require a large amount of data to solve a given task, even when similar tasks have already been solved. The insight of *transfer learning* (TL) is to make use of data from one or more previous tasks in order to learn the current task with less data, which may be expensive to gather. Generalization is possible not only within tasks, but also *across tasks*. This insight is not new; transfer has been studied in the psychological literature (Thorndike & Woodworth 1901; Skinner 1953) for many years. Transfer learning has been gaining in popularity in recent years as researchers have successfully applied TL techniques to classification tasks (Thrun 1996; Rosenstein *et al.* 2005; Shi, Fan, & Ren 2008), within cognitive architectures (Choi *et al.* 2007), when learning Markov Logic Networks (Mihalkova, Huynh, & Mooney 2007; Davis & Domingos 2008), and in reinforcement learning (RL) domains (Torrey *et al.* 2005; Taylor, Stone, & Liu 2007).

Despite the wide number of settings in which TL can be applied, the high-level transfer goals remain the same. Namely, the goal of transfer is to improve learning, relative to learning without transfer, in a *target task* after learning one or more *source tasks*. In classification tasks, transfer can achieve higher accuracy with fewer labeled target task data, while in RL settings, this could mean accruing higher reward in a target task with fewer environmental interactions (see Figure 1 for examples of possible benefits from transfer in an RL setting).

Task Selection

Sequentially ordering tasks by increasing difficulty, similar to the psychological definition of *shaping* (Skinner 1953), is common practice in education. For instance, if a student's ultimate goal is to learn calculus, she would begin with arithmetic, then learn algebra, etc., until finally taking a calculus class. With similar motivation, our own research (Taylor, Stone, & Liu 2007) has shown that learning on a series of tasks and sequentially transferring knowledge between them can be much faster than directly learning the final task. However, if the goal is to minimize the total training time (or to minimize the total data required to learn), the order in which the tasks are presented to the learner may be critical. While some intuition may be gained from human curriculum development (Ritter *et al.* 2007), optimally ordering tasks for machine learning algorithms is currently an open problem.

An additional danger related to improper task selection is that of *negative transfer*. In some situations, learning one or more source tasks before a target task may decrease the agent's ability to learn due to an incorrect bias. Such an effect has been documented by researchers in multiple settings (Rosenstein *et al.* 2005; Taylor, Stone, & Liu 2007; Shi, Fan, & Ren 2008). Currently, no general technique exists to prevent negative transfer.

Under submission to AAAI 2009 Spring Symposium: Agents that Learn from Human Teachers. Please do not distribute.

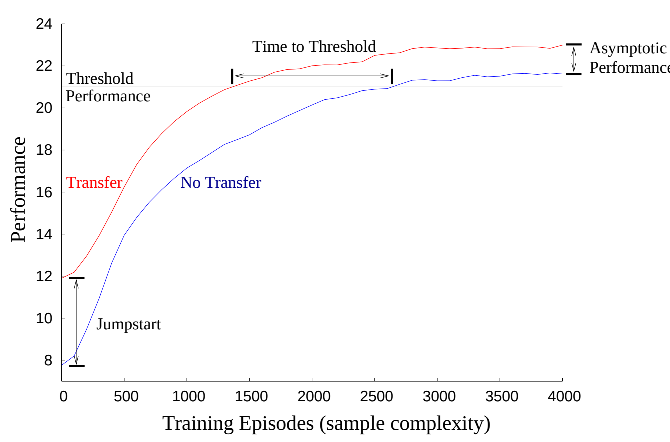

Figure 1: Many different metrics to evaluate TL algorithms are possible. This graph, taken from transfer results in a reinforcement learning domain (Taylor & Stone 2007), shows some of the possible benefits of transfer to reinforcement learning, including jumpstart, improved asymptotic performance, increased total reward (the area under a learning curve), the reward area ratio (the ratio of the area under the transfer to the area under the non-transfer learning curve), and fewer training data required to reach a prespecified threshold performance.

In addition to selecting an appropriate sequence of tasks and avoiding negative transfer, a third complication relates to deciding how much time to spend training on source tasks. Put another way, the agent should learn to reach some performance level in a task before moving to the next task in the sequence in order to minimize the total learning time. Continuing the calculus example, an appropriate curriculum would specify how much time a student should spend learning arithmetic before moving to the next class. Too little time and the next class may prove overwhelming; too much time and the student may be forced to spend longer than necessary on the total curriculum before finishing calculus. Furthermore, while some training in a source task may be beneficial, training for a very long time may actually delay learning in the target task (Taylor, Whiteson, & Stone 2007; Torrey *et al.* 2008), similar to overtraining in traditional machine learning algorithms.

The Human Element

Although some work has focused on autonomous transfer (c.f., Taylor, Kuhlmann, and Stone (2008)), the majority of transfer work implicitly relies on humans for task selection. My proposal is twofold: make such task selection an explicit part of the transfer process, and begin a new branch of research specifically studying how to best design machine learning curricula.

I envision at least three roles for human designers, each of which requires a different amount of background knowledge:

Common Sense: Humans have a large common sense knowledge, including facts such as "water is wet," which are not typically available to machine learners (Lenat 1995). For instance, when learning a soccer tasks, human knowledge about the facts that a ball can move faster than a person and that an opponent needs to be near a ball to take it away may be utilized to better design a task sequence.

Domain Experts: People who have experience with the domain in question will more likely be able to design better task sequences than those without such knowledge. For instance, someone familiar with soccer may know that learning when to pass or when to shoot is an important skill, and that more opponents on a field make the the game more difficult. Such knowledge may allow the human to design a sequence of tasks for an agent which is better than either an algorithm with no background or a human with no knowledge of soccer. Likewise, medical researchers may have more accurate intuitions than a lay person on how to best order gene classification tasks.

Algorithmic Experts: Researchers familiar with the specifics of the learning algorithm may provide further insight into designing task sequences. Such experts may be able to design curricula on their own, with only a very basic understanding of the domain, or may be more useful when their knowledge is combined with that of a domain expert.

In current TL research, the person designing the curriculum is typically the TL-algorithm designer, playing the part of all three roles. My hope is that human studies will further refine what type of backgrounds are most beneficial for successful shaping. Such a program would likely involve researchers from education, psychology, and machine learning. Human trials would require people with different levels of knowledge to design curricula and then test how well the curricula enable learning improvements. Furthermore, it may be possible to extract general principles regarding efficient curricula design by analyzing both the humans' design processes and the agents' learning performances.

If shaping techniques can be improved so that relatively little knowledge is required in practice, such techniques may be useful even at the consumer level. For instance, it is likely that a human could successfully select a sequence of tasks for a learning-enabled Roomba[1] so that it could learn to better vacuum his apartment due to the human's common sense knowledge and a background in the art of vacuuming.

One existing approach to such curriculum design is presented in the game of *NERO* (Stanley, Bryant, & Miikkulainen 2005), where agents train with a neuroevolutionary technique in a series of tasks selected by a human player (see Figure 2). People with only a limited knowledge of the domain are able to successfully train agents to accomplish a variety of tasks. However, the authors perform no analysis to find successful training strategies or to analyze how increased human background knowledge may improve agents' performance. Games such as NERO may provide an appropriate platform to conduct tests, or even more simple domains may need to be developed to better isolate the key components for successful task sequencing.

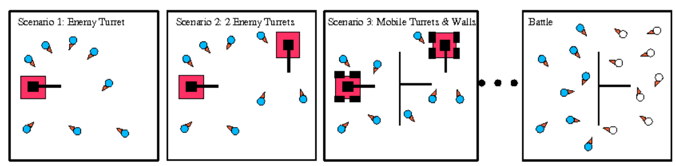

Figure 2: These diagrams of NERO (Stanley, Bryant, & Miikkulainen 2005) show a sequence of tasks selected by a human designer to successfully train a group of agents to engage a group of enemy agents.

Conclusion

Current work has demonstrated that transfer is a viable way to learn a sequence of tasks faster than learning directly on a complex final task. However, the process of selecting a sequence of tasks that successfully improves learning is poorly understood. This paper suggests that explicitly studying how humans can form successful machine learning curricula, and what background knowledge is necessary for them to do so, will further our understanding of what is needed for successful transfer, as well as how to better integrate humans into the machine learning process.

[1] www.irobot.com

References

Choi, D.; Konik, T.; Nejati, N.; Park, C.; and Langley, P. 2007. Structural transfer of cognitive skills. In *Proceedings of the Eighth International Conference on Cognitive Modeling*.

Davis, J., and Domingos, P. 2008. Deep transfer via second-order markov logic. In *AAAI Workshop: Transfer Learning for Complex Tasks*.

Lenat, D. B. 1995. CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM* 38(11):33–38.

Mihalkova, L.; Huynh, T.; and Mooney, R. 2007. Mapping and revising markov logic networks for transfer learning. In *Proceedings of the 22nd AAAI Coneference on Artificial Intelligence*.

Ritter, F. E.; Nerb, J.; Lehtinen, E.; and O'Shea, T. M. 2007. *In Order to Learn: How the Sequence of Topics Influences Learning (Oxford Series on Cognitive Models and Architectures)*. Oxford University Press.

Rosenstein, M. T.; Marx, Z.; Kaelbling, L. P.; and Dietterich, T. G. 2005. To transfer or not to transfer. In *In NIPS's Workshop, Inductive Transfer: 10 Years Later*.

Shi, X.; Fan, W.; and Ren, J. 2008. Actively transfer domain knowledge. In *ECML-PKDD*, 342–357.

Skinner, B. F. 1953. *Science and Human Behavior*. Colliler-Macmillian.

Stanley, K. O.; Bryant, B. D.; and Miikkulainen, R. 2005. Real-time neuroevolution in the NERO video game. *IEEE Transactions on Evolutionary Computation* 9(6):653–668.

Taylor, M. E., and Stone, P. 2007. Cross-domain transfer for reinforcement learning. In *Proceedings of the Twenty-Fourth International Conference on Machine Learning*.

Taylor, M. E.; Kuhlmann, G.; and Stone, P. 2008. Autonomous transfer for reinforcement learning. In *The Seventh International Joint Conference on Autonomous Agents and Multiagent Systems*.

Taylor, M. E.; Stone, P.; and Liu, Y. 2007. Transfer learning via inter-task mappings for temporal difference learning. *Journal of Machine Learning Research* 8(1):2125–2167.

Taylor, M. E.; Whiteson, S.; and Stone, P. 2007. Transfer via inter-task mappings in policy search reinforcement learning. In *The Sixth International Joint Conference on Autonomous Agents and Multiagent Systems*.

Thorndike, E., and Woodworth, R. 1901. The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review* 8:247–261.

Thrun, S. 1996. Is learning the n-th thing any easier than learning the first? In *Advances in Neural Information Processing Systems*, 640–646.

Torrey, L.; Walker, T.; Shavlik, J. W.; and Maclin, R. 2005. Using advice to transfer knowledge acquired in one reinforcement learning task to another. In *Proceedings of the Sixteenth European Conference on Machine Learning*, 412–424.

Torrey, L.; Shavlik, J.; ad Pavan Kuppili, S. N.; and Walker, T. 2008. Transfer in reinforcement learning via markov logic networks. In *AAAI Workshop: Transfer Learning for Complex Tasks*.