Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)

Representations for Continuous Learning

David Isele
University of Pennsylvania
Philadelphia, PA, USA
isele@seas.upenn.edu

Abstract

Systems deployed in unstructured environments must be able to adapt to novel situations. This requires the ability to perform in domains that may be vastly different from training domains. My dissertation focuses on the representations used in lifelong learning and how these representations enable predictions and knowledge sharing over time, allowing an agent to continuously learn and adapt in changing environments. Specifically, my contributions will enable lifelong learning systems to efficiently accumulate data, use prior knowledge to predict models for novel tasks, and alter existing models to account for changes in the environment.

Introduction

Lifelong machine learning (Thrun 1996) is a branch of artificial intelligence that has developed to handle systems that can adapt and learn in ever-changing environments. In Lifelong machine learning, there are possibly many source tasks, and the objective is to optimize the representation and performance for all tasks, as new information arrives. Examples of lifelong learning include methods for constructing general representations for discrete reinforcement learning problems (Ring 1998; Rafols and others 2005), bootstrapping learned relations to infer new relations in text understanding (Mitchell and others 2015), and using sparse coding to share information across learned models (Ruvolo and Eaton 2013).

I am interested in a subclass of lifelong learning, *continuous learning*, where previous knowledge is immediately transferred to new tasks, and the learning process is allowed to continue without relearning already learned knowledge.

My dissertation views the problem of continuous learning as a problem of representation. In order to efficiently represent knowledge over the lifetime of a system, we need representations that can handle the demands of a continuous learner. The representations must accumulate knowledge in an online manner, scale to a large number of diverse tasks, discover similarities, share knowledge between tasks, adapt to changes over time, and be able to predict good initializations for novel tasks. The last point of prediction, is of particular importance to *continuous learning*, since it allows the system to run continuously in changing environments.

The initial work of my thesis has been to address the issue of predicting task models. My previous work (Isele, Rostami, and Eaton 2016) introduced a coupled dictionary that works with a sparse coding reinforcement learning framework (Bou Ammar, Eaton, and Ruvolo 2014) to learn a dictionary online. Our coupled dictionary learns to model task descriptions enabling the system to make predictions for novel tasks.

Moving forward, I plan to look into more expressive models that have a greater focus on time and changing representation. To this end, I am considering two directions: a hierarchical online learner that can accommodate tasks that change over time and a non-linear time-varying model that incorporate exploration into its predictions.

By designing representations that are able to handle the demands of continuous learning I hope to lay the foundation for lifelong learning systems that can robustly function in unstructured environments.

Past Work: Predicting Policies

A lifelong learning framework that has been shown to be versatile is the sparse coding method of Ruvolo and Eaton (2013) which was developed for classification and regression and has been extended to handle reinforcement learning problems with continuous state and action spaces (Bou Ammar, Eaton, and Ruvolo 2014). In this approach, models for different tasks are represented as sparse combinations of a shared knowledge repository of basis vectors.

I have shown that this approach can be used on robotic systems (Isele and others 2016). However a problem with this approach is that each time a new task is encountered, the system is halted to gather training data to characterize the new task. This was addressed in previous work that (Isele, Rostami, and Eaton 2016) showed already learned knowledge can be used to predict a starting policy.

In order to accomplish this, we add a coupled dictionary that learns to represent both a description of the task and the learned policy. Given the description of a new task, a policy can be predicted without needing to train on the new task. This approach enables us to predict policies that often outperform the policies learned by a single task policy gradient learner. Additionally, using the predicted policy as a jump start allows our lifelong learning system to focus specifically on the novel aspects of a new task.

Future Work: Hierarchical Temporal Models

The system described in the previous section is capable of learning and transferring information across a variety of tasks, however all models are restricted to be linear in the basis defined by the knowledge repository. This greatly limits the expressiveness of the system, preventing its application to many interesting real world problems. We are looking into ways of extending current lifelong learning techniques to accommodate more powerful learners, with representations that adapt to changing environments.

In order to achieve a hierarchical model that can share knowledge across tasks and adapt to changes over time, we consider using deep nets as a foundation. Deep nets have powerful hierarchical representations which have been shown to learn features that generalize well and can be used for transfer (Razavian and others 2014).

The training method for deep nets is already sequential, suggesting that they can be extended to online applications and will adapt to changes over time. However, a problem is that sequential data often causes the network to change too quickly - experience replay is needed to break temporal correlations and drift (Mnih and others 2013). In order to be useful for a continuous learner, experience replay must be modified to allow adaption to changing environments while holding on to past experiences that are still relevant.

Other considerations that must be addressed to extend deep nets into continuous learners center around efficiency. Experiences must be represented efficiently so that the system can scale to an arbitrary length runs. Similarly, while there has been some work related to handling multiple tasks (Rusu et al. 2016), these approaches do not scale well to many tasks.

Future Work: Exploratory Predictions

The coupled dictionary approach (Isele, Rostami, and Eaton 2016) predicts new models based on learned locally optimal models. While this can produce good initializations, the predicted models are restricted to the space of known solutions. This puts the burden on the single task learner to explore the state space, and by design the single task learner's initialization makes it unlikely that it will explore other local optima.

In the context of a sequential decision-making multi-armed bandit problem this is equivalent to only exploiting. Since we want our continuous learner to constantly improve and adapt, it would be beneficial to incorporate exploration into it's prediction model. This requires a representation that can compare the value of exploring and exploiting in a space that changes over time. One possible solution is to use the Gaussian Process Upper-Confidence Bound (GP-UCB) (Srinivas et al. 2010) with extensions to time-varying functions (Bogunovic, Scarlett, and Cevher 2016).

However, using Gaussian Processes presents a scalability challenge for continuous learning. The complexity of Gaussian Processes increase as the number of samples grows. Any continuous learner requires a way to incorporate multiple sequential tasks while bounding the complexity to prevent the system from arbitrarily increasing over time.

Fundamental Questions

My thesis aims to identify some of the fundamental aspects required for continuous learning representations while also demonstrating specific implementations that satisfy these requirements. I plan to address the following fundamental questions:

- What are the needs of a continuous learning system?
- What abstract representations address these needs?
- How can these abstract representations be realized?

By identifying the specific needs of a continuous learning system, and demonstrating example systems that satisfy these needs I hope to create a framework for the development of robust continuous learning in novel domains.

References

Bogunovic, I.; Scarlett, J.; and Cevher, V. 2016. Time-varying gaussian process bandit optimization. *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS)*.

Bou Ammar, H.; Eaton, E.; and Ruvolo, P. 2014. Online multi-task learning for policy gradient methods. *ICML*.

Isele, D., et al. 2016. Lifelong learning for disturbance rejection on mobile robots. *IROS*.

Isele, D.; Rostami, M.; and Eaton, E. 2016. Using task features for zero-shot knowledge transfer in lifelong learning. *IJCAI*.

Mitchell, T. M., et al. 2015. Never-ending learning. In *AAAI*.

Mnih, V., et al. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Rafols, E. J., et al. 2005. Using predictive representations to improve generalization in reinforcement learning. In *IJCAI*, 835–840. Citeseer.

Razavian, A. S., et al. 2014. CNN Features off-the-shelf: an Astounding Baseline for Recognition. *Computer Vision and Pattern Recognition Workshops (CVPRW)* 512–519.

Ring, M. B. 1998. Child: A first step towards continual learning. In *Learning to learn*. Springer. 261–292.

Rusu, A. A.; Rabinowitz, N. C.; Desjardins, G.; Soyer, H.; Kirkpatrick, J.; Kavukcuoglu, K.; Pascanu, R.; and Hadsell, R. 2016. Progressive neural networks. *arXiv preprint arXiv:1606.04671*.

Ruvolo, P., and Eaton, E. 2013. ELLA: An efficient lifelong learning algorithm. *ICML* 28:507–515.

Srinivas, N.; Krause, A.; Kakade, S. M.; and Seeger, M. 2010. Gaussian process optimization in the bandit setting: No regret and experimental design. *ICML*.

Thrun, S. 1996. Is learning the n-th thing any easier than learning the first? *NIPS* 640–646.