Automatic Skill Transfer Learning Through Domain Adaptation

Farzaneh Shoeleh
Faculty of Electrical and Computer Engineering
University of Tehran,Tehran, Iran
f.shoeleh@ut.ac.ir

Masoud Asadpour
Faculty of Electrical and Computer Engineering
University of Tehran,Tehran, Iran
asadpour@ut.ac.ir

ABSTRACT

Generalization is the most challenging issue in continuous Reinforcement Learning(RL) domains to overcome the curse of dimensionality. Transfer Learning(TL) is a successful technique to result in big improvements in agent learning performance by providing generalization not only within a task, but also across different but related tasks. In this paper, we propose a novel TL approach in continuous RL problems which utilizing domain adaptation techniques.

Keywords

Learning Systems, Reinforcement Learning, Transfer Learning, Domain Adaptation

1. INTRODUCTION

The reinforcement learning (RL) paradigm is a popular way for agents to learn from experience with minimal feedback. The required learning time and the curse of dimensionality restrict applicability of RL on real world problems. Transfer learning has recently attracted many researchers in the fields of machine learning [6, 1, 4]. The insight behind transfer learning is that generalization may occur not only within tasks, but also across tasks. In the RL realm, While significant progress has been made to improve learning in a single task, the idea of transfer learning has only recently been applied to reinforcement learning tasks [9, 3]. In general, TL algorithms can be divided into heterogeneous and homogeneous TL by considering whether the feature spaces between source and target domains are the same or not. In this paper, we propose a novel heterogeneous Skill based Transfer Learning method using Domain Adaptation, named STL_{DA}, to address the curse of dimensionality in continuous RL domains by leveraging transfer learning techniques which discovers the similarity between source and target tasks to boost learning performance in target task.

2. PROPOSED METHOD

To extract and transfer high-level skills, the proposed skill based transfer learning method has three following phases.

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.*

Phase 1. Learning Source Task. Agent learns the source task using the framework introduced in [8]. In the source task, the agent firstly constructs a connectivity graph as a model to capture its experiences and the environment's dynamics. Then, it extracts and learns skills as high-level knowledge to be used in the new target task [8]. After learning the source task, the agent is provided by a collections of source samples, (D_S) in the form of $< s, a, s', r, q, skill_{ID} >$.

Phase 2. Domain Adaptation. Prior to running the second step, agent gathers a collections of target data, (D_T) in the form of $< s, a, s', r >$, while living within target task. Then, agent utilizes a widely used domain adaptation algorithm named Transfer Component Analysis (TCA) [5] to project the source and the target samples onto common latent space where the distance between the domains of samples will be reduced. TCA tries to solve the following kernel learning problem

$$\begin{aligned} \min_{W} \quad & tr(W^T KLKW) + \mu tr(W^T W) \\ s.t. \quad & W^T KHKW = I_m \end{aligned} \tag{1}$$

where $\mu > 0$ is a tradeoff parameter and K consists Gram matrices defined on source domain, target domain and cross-domain data. L and H matrices are computed as follows:

$$L_{ij} = \begin{cases} \frac{1}{n_1^2} & \text{if } x_i, x_j \in X_S \\ \frac{1}{n_2^2} & \text{if } x_i, x_j \in X_T \\ \frac{-1}{n_1 n_2} & \text{otherwise} \end{cases} \tag{2}$$

$$H = I_{n_1+n_2} - (\frac{1}{n_1 + n_2}) 1_{n_1+n_2} \tag{3}$$

where $1_{n_1+n_2} \in R^{(n_1+n_2)\times(n_1+n_2)}$ is with all 1's, and $I_{n_1+n_2} \in R^{(n_1+n_2)\times(n_1+n_2)}$ is the identity matrix.

The output of this phase is a sample similarity function which helps the agent to identify which source sample $s_S \in D_S$ is similar to a given target sample $s_T \in D_T$ in obtained latent subspace and assign the s_S skill id to the s_T's one.

Phase 3. Learning Target Task.

To successfully transfer learning, the agent utilizes the sample similarities and q-value of each sample to learn the skills in the target task offline. To do so, the final value of state-action pairs $Q(s, a)$ in source domain is used in target domain instead of direct interaction with environment. The agent would define a number of raw skills in the target task, which are equal to the number of learned skills in the source task. Then, each target sample is assigned into the corresponding raw skill using sample similarity function obtained through second phase. To avoid negative transfer, we sug-

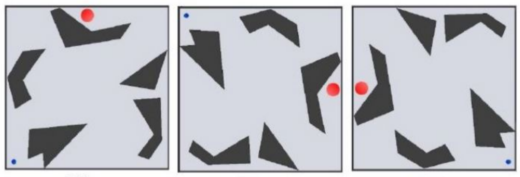

Figure 1: Test domains

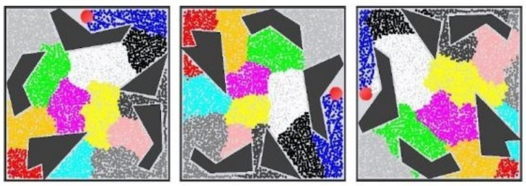

Figure 2: Extracted Skills in source tasks(left), transferred Skills in target tasks through first scenario(middle), and transferred Skills in target tasks through second scenario(right).

gest a KNN based mechanism to eliminate samples which are located in wrong skill regions. Each raw skill would be learned offline based on the *q-value* of the samples which are located within the raw skill. Note that agent does not interact with environment during learning the raw skills. In target task, these skills are firstly considered as gestating skills which are allowed to have a gestation period (e.g., ten episodes), where they cannot be selected for execution but their policies are updated using off-policy learning. After gestation period, each gestating skill would be added to the agent's action set as a learned skill.

3. RESULTS

The proposed method was assessed using Pinball, a well-known 4 dimensional continuous test domain for RL [2] (see Figure 1). Two transfer learning scenarios are considered: 1) from PB_1 to PB_2, 2) from PB_1 to PB_3. Table 1 illustrates the performance of four agents through these scenarios: 1) an agent with flat policy (SARSA), 2) a GSL agent [8], 3) a SCL agent [2], 4) an STL_{DA} agent employing skill transfer learning. Note that the first three agents do not utilize the transfer learning mechanism. Table 1 shows the comparison of aforementioned agents in terms of average number of steps to reach goal state. Besides, we apply pairwise t-test [7] at 0.05 significance level to highlight the competitiveness of our methods. The bolded results show that improvement of performance in STL$_{DA}$ is statistically significant. Table 2 shows the effectiveness of STL$_{DA}$ in terms of four metrics introduced in [9] in order to measure the benefits of a transfer learning method. It is worth mentioning that in calculation of Time to threshold metric, the threshold is defined as asymptotic performance of SCL, as used by the state-of-art skill learning method. According to this metric, using transfer learning makes GSL agent reach SCL's performance in average 358 episodes in pinball domain while SCL achieve this after 500 episodes. To make it clear, Figure 2 illustrates the skills in the source and target task.

Table 1: Comparison of different agents in terms of average number of steps to reach goal stat.

	STL_{DA}	GSL	SCL	SARSA
PB$_2$	**98** (±16)	120 (±24)	135 (±46)	2018.2 (±601)
PB$_3$	**104** (±21)	127 (±36)	161 (±65)	2187 (±731)

Table 2: The average and standard deviation of different metrics for STL$_{DA}$ in each scenarios.

	Jumpstart	Asymptotic Performance	Transfer ratio	Time to threshold
S_1	2100 (±258)	8959 (±172)	3.7% (±0.04)	358 (±27)
S_2	1710 (±345)	8891 (±105)	0.78% (±0.03)	338 (±25)

4. CONCLUSION

In this paper, we propose a novel skill based transfer learning framework using domain adaptation, named STL$_{DA}$. STL$_{DA}$ utilizes GSL [8] framework to discover skills as high-level knowledge. Then, it deploys a well known domain adaption technique, i.e. TCA [5], to find the sample similarity function which maps a given target sample into the best similar source sample and consequently assigns it the right skill *id*. Having such function, agent learns the raw extracted skills offline, without interacting with environment, using the transferring $q-value$ of source samples into target task. The experimental results indicate the effectiveness of STL$_{DA}$ agent in dealing with continuous RL problems.

REFERENCES

[1] M. Fang, Y. Guo, X. Zhang, and X. Li. Multi-source transfer learning based on label shared subspace. *Pattern Recognition Letters*, 51:101–106, jan 2015.

[2] G. Konidaris and A. A. S. Barreto. Skill discovery in continuous reinforcement learning domains using skill chaining. In *Advances in Neural Information Processing Systems*, pages 1015–1023, 2009.

[3] A. Lazaric. Transfer in Reinforcement Learning : a Framework and a Survey. *Reinforcement Learning*, 12:143–173, 2012.

[4] M. Long, J. Wang, G. Ding, D. Shen, and Q. Yang. Transfer learning with graph co-regularization. *IEEE Transactions on Knowledge and Data Engineering*, 26(7):1805–1818, 2014.

[5] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang. Domain adaptation via transfer component analysis. *Neural Networks, IEEE Transactions on*, 22(2):199–210, 2011.

[6] S. J. Pan and Q. Yang. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, oct 2010.

[7] D. J. Sheskin. Handbook of parametric and nonparametric statistical procedures, 2007.

[8] F. Shoeleh and M. Asadpour. Graph based skill acquisition and transfer learning for continuous reinforcement learning domains. *Pattern Recognition Letters*, 2016.

[9] M. E. Taylor and P. Stone. Transfer Learning for Reinforcement Learning Domains : A Survey. *Journal of Machine Learning Research*, pages 1633–1685, 2009.