ARTICLE IN PRESS
JID: PATREC [m5G;September 6, 2016;10:35]
Pattern Recognition Letters 000 (2016) 1–13
Contents lists available at ScienceDirect
Pattern Recognition Letters
journal homepage: www.elsevier.com/locate/patrec

Graph based skill acquisition and transfer Learning for continuous reinforcement learning domains☆

Farzaneh Shoeleh*, Masoud Asadpour

Faculty of Electrical and Computer Engineering, University of Tehran, Tehran, Iran

A R T I C L E I N F O

Article history:
Available online xxx

Keywords:
Reinforcement learning
Skill acquisition
Transfer learning
Graph learning

A B S T R A C T

Since reinforcement learning algorithms suffer from the curse of dimensionality in continuous domains, generalization is the most challenging issue in this area. Both skill acquisition and transfer learning are successful techniques to overcome such problem that result in big improvements in agent learning performance. In this paper, we propose a novel graph based skill acquisition method, named GSL, and a skill based transfer learning framework, named STL. GSL discovers skills as high-level knowledge using community detection from connectivity graph, a model to capture not only the agent's experience but also the environment's dynamics. STL incorporates skills previously learned from source task to speed up learning on a new target task. The experimental results indicate the effectiveness of the proposed methods in dealing with continuous reinforcement learning problems.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Reinforcement Learning (RL) is one of the main machine learning domains which allows autonomous agents to learn and improve their performance through the obtained experiences while interacting with an unknown environment. The applicability of RL methods on real world problems, especially problems with continuous state-space, is restricted by the required learning time and the *curse of dimensionality*, i.e. the computational complexity of learning grows exponentially with size of state space [28]. In the last decade, a vast number of RL studies have been carried out to overcome this problem. According to the literature, it is believed that state abstraction methods and hierarchical architectures can improve the required learning time and lessen the hampering effect of the curse of dimensionality.

Much recent research in reinforcement learning (RL) has focused on hierarchical RL methods. One of the most successful techniques in hierarchical reinforcement learning (HRL) to speed-up the learning is Option framework using temporal abstraction [29], i.e. instead of choosing an action in each step, the agent chooses a time extended activity that is performed in more than one time step. Option framework adds high level-skills (called options) to the RL framework principle methods for learning. So, an important research goal is the development of methods to answer the most challenging question in HRL domain, "How do we discover abstract skills automatically?". A lot of efforts have been made to find a proper answer to this question [7,13,23,30]. To the best of our knowledge, although several methods exist for creating new options in discrete domain, none are immediately extensible to, or have been successfully applied in continuous domains. There are few works developed for continuous domain which an agent can discover useful new skills autonomously, and thereby construct its own high-level skill hierarchies. To discover abstract skills automatically from continuous state space, the better agent models its experiences and environments dynamics, the higher performance is achieved. Graph based representation facilitates modeling such problem properties, but whats more, graph mining techniques can help agent to discover abstract skills. In this paper, we present a graph based skill learning named GSL that allows agent to use graph representation in order to model its environment and find skills autonomously. Graph representations have been used in discrete reinforcement learning domains as basis for both autonomous skill discovery and representation learning simultaneously. These abilities are also highly relevant for learning in domains which are structured. However, since graphs are inherently discrete structures, the extension of these approaches to continuous domains is not straight-forward and graphs should be seen as discrete model of a continuous domain. Based on this intuition, we define Connectivity Graph which can model both the agent's experience and environment's dynamics. By exploring the environment the agent is able to construct connectivity graph and then by detecting the community of this graph it can find high-level skills automatically and finally by defining the main components of

☆ This paper has been recommended for acceptance by Cheng-Lin Liu.

* Corresponding author.
E-mail address: f.shoeleh@ut.ac.ir, farzane.shoaleh@gmail.com, f.shoeleh@itrc.ac.ir (F. Shoeleh).

http://dx.doi.org/10.1016/j.patrec.2016.08.009
0167-8655/© 2016 Elsevier B.V. All rights reserved.

Please cite this article as: F. Shoeleh, M. Asadpour, Graph based skill acquisition and transfer Learning for continuous reinforcement learning domains, Pattern Recognition Letters (2016), http://dx.doi.org/10.1016/j.patrec.2016.08.009

skills, such as Termination condition, Initiation Set, Reward function, function approximator, agent can utilize hierarchal RL to learn the problem in continuous RL domain more accurately and efficiently. The novelty of proposed method is introducing the concept of connectivity graph and its usage in discovering skills. We use function approximation in each skill to learn the policy of agent in the region of corresponding skill. In other words, in contrast to previous works on skill learning in continuous domain where graph is used as a representation for planning and learning, We use our proposed connectivity graph just for detecting skills and use linear function approximation for planning and learning. The promising results demonstrate that our Graph based Skill Learning (GSL) approach not only can find appropriate skills but also results in notable improvements in the learning performance of the agent.

In addition to skill learning, the other key idea to address the curse of dimensionality in continuous RL domains is leveraging transfer learning techniques. Although, the insight behind HRL is generalization within a task, transfer learning results in generalization not only within the task but also across tasks. In fact, transfer learning is the answer to “How to apply the learned knowledge from one task, called the source task, to the related but different task, called the target task?”. Recently, transfer learning has attracted many researchers in the fields of artificial intelligence and machine learning [8,19,24,25,35]. In almost all machine learning and data mining algorithms, the main assumption is that both train and test samples are driven from not only the same feature space, but also the same distribution. However, this assumption does not hold in many problems [24]. Hence, researchers have focused on using the acquired knowledge from the previous tasks, which are related to but different from the current task, in order to improve the efficiency of learning methods in terms of performance and learning time. Generally, the transferred knowledge within tasks can be categorized into two groups: low-level and high-level knowledge. Note that the source and target tasks must be more related, while using low level knowledge compared to using higher level knowledge [16,31]. Thereby, techniques based on extraction and transferring high-level knowledge are more advantageous.

In this paper, the objective is to achieve better learning curve in target task by automatically extracting and utilizing skills as high-level knowledge from an auxiliary source task. Here, we propose a transfer learning framework named Skill based Transfer Learning (STL) in continuous RL domain. In this framework skills are evaluated based on their impact on leading the agent towards the goal state. So, two mechanisms are proposed to calculate the fitness of each previously learned skill. The first mechanism relies on how much a skill is successfully executed in target task and the other relies on how much a learned skill can cover the subspace of a raw skill in target task. In former, if a previously learned skill can act successfully in target task, this skill would be candidate to be transferred and in later, if a previously learned skill can cover the subspace of a raw skill maximally, the function approximator of the learned skill would be transferred into the raw skill. Consequently, the novelty of this work is guarding against negative learning, where the transferred knowledge may mislead the agent. In other words, our framework does not reuse all learned skills in target task, it only transfers those skills which are certainty useful in new target task by measuring their fitness. The results from experiments demonstrate that the proposed method is able to transfer the obtained abstract skills effectively and improve the performance of agent in target task.

The rest of this paper is organized as follows. Section 2 presents an overview of the related work. In Section 3, the proposed approaches for automatic skill acquisition and transfer learning are described. Experiments and results are reported in Section 4, and Section 5 contains the conclusion and direction for future works.

2. Related work

In both [31] and [16], different frameworks have been investigated to categorize the transfer learning approaches in RL domain. According to these researches, transfer learning algorithms can be categorized along three main dimensions: the setting, the transferred knowledge and the objective. Among these categories, the definition of transferred knowledge and the specific transfer process are the main aspects to characterize a transfer learning algorithm.

As the type of transferred knowledge can be primarily characterized by its specificity [31], the possible knowledge transfer approaches can be classified into two main categories: low-level knowledge transfer and high-level knowledge transfer. Low-level information can be considered as (s, a, r, s') tuples, an action-value function Q, a policy π, or a full model of the task, whereas high-level information can be considered as a subset of all actions used in some situations or partial policies, skills or options, rules, important features for learning, proto-value functions [20], shaping rewards, or subtask definition.

Unlike low-level knowledge, which could all be directly leveraged to initialize a learner in the target task, high-level information may not directly be applicable in the transfer learning algorithms to fully define an initial policy for the agent in the target task. However, such information would guide the agent during the learning of the target task. Besides, transfer learning algorithms using high-level knowledge assist the agent in learning a new task more effectively than lower-level information. Hence, as claimed in [31,32], it makes intuitive sense that high-level knowledge may transfer better across tasks since they can be obtained more independently compared to low-level information.

Since the proposed method is a skill based transfer learning method which is applicable in continuous domain, this section presents some of the related work in transfer learning in RL context and contrasts it with the proposed approach. For a broader review of Transfer Learning in RL domains the reader may refer to the comprehensive surveys in [16,31,32].

Lazaric and Restelli [17,18] demonstrate that source task instances can be usefully transferred between tasks. After learning the source task, the agent gathers some experiences in the target task to be compared to instances from the source task. Judging the distance and alignment, the most similar source instances are transferred. Then, in the target task, the agent uses a batch learning method for training with both source and target samples to achieve higher reward and a jumpstart. The idea of transferring similar regions among tasks was firstly proposed in [18]. In [18], the similar regions are determined using the similarity between samples in source and target, indeed using low-level knowledge. In contrast, our proposed method tries to transfer similar regions identified with high-level knowledge, namely skills which is defined using a community detection algorithm on the connectivity graph. Asadi and Huber [2,3] present an agent that transfers options between different tasks. The agent tries to find subgoals in the source task by identifying states that are “locally from a significantly stronger *attractor* for state space trajectories”. Like these works, almost all the option-transfer methods consider discrete Markov Decision Processes (MPDs), whereas here we aim at proposing a transfer learning method which is applicable to continuous RL domains.

Although all option-based transfer learning algorithms share the same structure, one of the critical steps is to identify these options. Automatic skill or option discovery can be divided into two main categories: graph based and frequency based approaches. In the first category, the transition graph is built then the sub-goals are found using graph theory analyses. The graph based approaches are different in processing and analysis of such

Download English Version:

https://daneshyari.com/en/article/4970291

Download Persian Version:

https://daneshyari.com/article/4970291

Daneshyari.com