Analyzing State Space Similarities for Multi-Task Deep Reinforcement Learning in Atari Games

Sooyoung Jang
Department of Computer Engineering
Hanbat National University
Daejeon, South Korea
syjang@hanbat.ac.kr

Changbeom Choi
Department of Computer Engineering
Hanbat National University
Daejeon, South Korea
cbchoi@hanbat.ac.kr

***Abstract*—Deep reinforcement learning (DRL) has shown effectiveness across various individual tasks, but often faces challenges when applied to multi-task scenarios. This study investigates the potential for sequential multi-task DRL by analyzing the similarities within the state spaces of different Atari game environments. Using dynamic time warping and dimensionality reduction techniques, we extract meaningful features from the state representations and compare their similarities. The result provides insights into the relationships between distinct Atari games. This work aims to establish a foundation for future research focused on developing more efficient and robust multi-task DRL algorithms that can leverage shared knowledge across similar tasks.**

***Index Terms*—Deep Reinforcement Learning, State Space Analysis, Atari Games.**

I. Introduction

Deep reinforcement learning (DRL) has emerged as a powerful technique for solving complex problems across various domains, achieving superhuman performance in games like Go, chess, and Atari [1]. However, transitioning this success to multi-task learning, where an agent learns to master multiple tasks sequentially, presents significant challenges. Current DRL approaches often struggle with inefficient learning and degraded performance when switching between tasks [2].

This paper explores the potential for improving sequential multi-task DRL by investigating a fundamental question: how similar are the state spaces of different tasks? We focus on the domain of Atari games, a popular benchmark for DRL research, and analyze the inherent similarities within their state spaces.

To achieve this, we leverage dynamic time warping (DTW) to align and compare temporal sequences of state representations, capturing similarities even across varying episode lengths and action timings. We then apply principal component analysis (PCA), which is a linear dimensionality reduction technique, to extract meaningful features from these comparisons, revealing underlying relationships between different Atari game environments.

By visualizing and analyzing these relationships, we aim to characterize the state space similarities between different Atari games, identifying potential clusters of games and provide insights into the opportunities for knowledge transfer in sequential multi-task DRL within this domain.

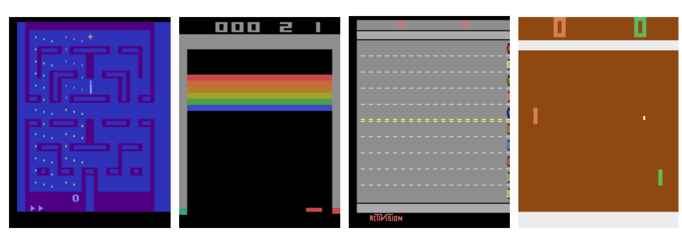
II. Methodoloy and Findings

Our methodology for analyzing the state space similarities of Atari games involves two key steps: feature extraction and similarity comparison.

A. Feature Extraction

Raw state representations in Atari games, typically image frames, are high-dimensional and complex, making direct comparison difficult. To address this, we employ fastDTW [3] in conjunction with PCA.

- **Data Collection:** We begin by collecting observations from 4 Atari games, namely, Alien, Breakout, Freeway, and Pong, through agent rollouts. The snapshots of these Atari games are presented in Fig. 1. The agent rollouts involve running an agent trained on each game for 5 episodes. Even within the same game, the lengths of the episodes vary.
- **DTW:** We apply fastDTW to align and compare temporal sequences of state representations. DTW measures the similarity between time series even if they have different lengths or are not perfectly synchronized. This is particularly relevant for game states, as their episodes can vary. FastDTW is an approximate version of DTW that enables the computation to scale with near-optimal performance.
- **PCA:** To further reduce complexity and extract meaningful features, we apply Principal Component Analysis (PCA) to the DTW distance matrices. This step transforms the high-dimensional distance information into a lower-dimensional feature space, capturing the essential similarities and differences between state sequences.



(a) Alien (b) Breakout (c) Freeway (d) Pong

Fig. 1: Snapshots of 4 Atari games utilized for the analysis.

Fig. 2 visualizes 20 state space representations, each comprising 5 episodes from 4 Atari games. The application of fastDTW and PCA has resulted in a clear clustering of the game state representations. Notably, the states of Alien, Freeway, and Pong tend to cluster closely together, while Breakout appears more dispersed.

B. Similarity Analysis

With the extracted features, we analyzed the state space similarities based on euclidean distances between the centroids of different Atari games.

Table I shows the centroid distances between different game pairs, indicating the relative closeness or separation of their state spaces. The distance between Breakout and Freeway is 18.5667, the highest among all pairs. This indicates significant differences in the state representations between these two games. On the other hand, the distance between Alien and Breakout is the smallest among all pairs and may indicate that the state representations of Alien and Breakout share more similarities with each other than with the other games.

From these results, we may induce that transferring knowledge from Alien to Breakout is relatively easier than transferring knowledge from Alien to other games. And we may design an order for sequential training to efficiently

III. Conclusion

The analysis of state space similarities in Atari games provides valuable insights for the development of multi-task DRL algorithms. By understanding the relationships between distinct game environments, we can explore strategies for transferring knowledge and improving learning efficiency when transitioning between tasks. Future work will focus on incorporating these findings into the design of more robust and adaptable multi-task DRL models.

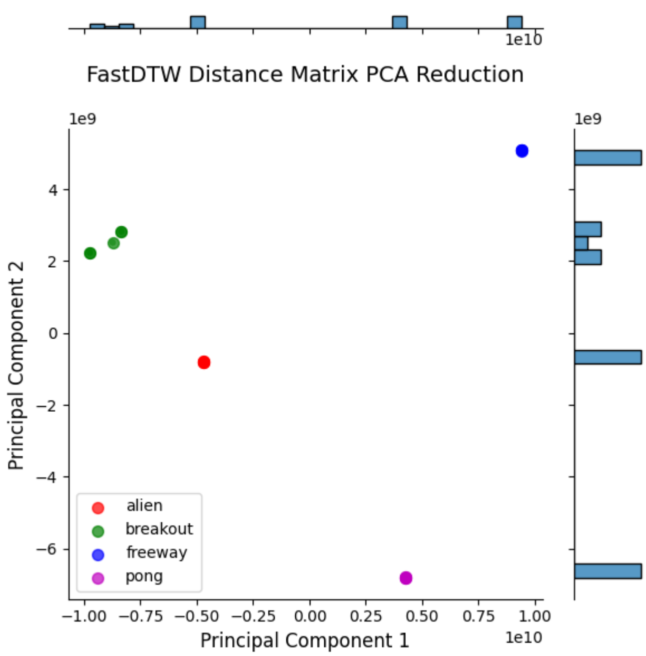


Fig. 2: Visualization of state space of 4 Atari games based on extracted features. Each point represents an episode, and their proximity reflects the similarity of their state space.

TABLE I: Centroid distances between clusters of Atari games. Smaller distances indicate greater similarity (values are in units of $\times 10^9$).

	Alien	Breakout	Freeway	Pong
Alien	0	5.44	15.27	10.77
Breakout	5.44	0	18.57	16.20
Freeway	15.27	18.57	0	12.94
Pong	10.77	16.20	12.94	0

References

[1] V. Mnih et al., “Human-level control through deep reinforcement learning,“ Nature, vol. 518, pp. 529–533, 2015.

[2] V. Varghese et al., “A Survey of Multi-Task Deep Reinforcement Learning,“ Electronics 9, no. 9: 1363, 2020.

[3] S. Salvador, and P. Chan, “FastDTW: Toward accurate dynamic time warping in linear time and space,“ Intelligent Data Analysis, vol. 11, no. 5, pp. 70–80, 2004.