称号及び氏名	博士（情報学）　福島　卓弥
学位授与の日付	2022 年 3 月 31 日
論　　文　　名	Machine Learning Methods for Agent-Based Real-Time Systems （エージェントベース実時間システムに対する機械学習手法）
論文審査委員	主　査　　中島　智晴
	副　査　　太田　正哉
	副　査　　岩村　幸治
	副　査　　楠木　祥文

Abstract

In recent years, the advancement of Artificial Intelligence (AI) thanks to the improvement of computers' performance has drawn the interests of people. Especially, in the domain of board games such as Go and Shogi, computers outperformed professional human players, and it is argued that computers have better decision–making abilities than human beings. In general, AI can make better decisions than humans when limited to specific tasks. However, the performance of AI's decision becomes weak in complicated environments that different from those used for a training phase. "Narrow AI" has been used to refer such weakness of AI.

On the other hand, AI that has a broad capability to self-adapt to changes in their goals or circumstances has called as Artificial General Intelligence（AGI）. Many studies have been conducted on AGI. General Game Playing（GGP）is one of the domains to contribute to the realization on AGI. In particular, GGP has been widely researched since 2005, thanks to the start of the GGP international competition. Such game-based AI competitions are utilized for the benchmarks of the AI performance because AI can be compared with another one in the same environments. In the competitions, a single AI plays some games that are difficult from each other. Then, the scores calculated by the results of the games are competed. AGI is required foe GGP competitions because the AI specialized for a particular game cannot improve the score foe other types of games.

There is also a domain called General Video Games AI（GVGAI）, where the environments

are limited to video games. Since video games are simulation environments, it is easy to introduce machine learning approaches such as reinforcement learning. Furthermore, the game environments are provided as researchers to study more actively.

On the other hand, although there are many papers on GGP, few technologies of AI are applied to the real-world for practical use. This is because those AI researches have been successful in "toy-problem" environments where no missing information, no noise, sufficient computation time, and high performance computers are always available, but not in real-world environments where there are many constraints. This is the main reason why it is difficult for AI to achieve good results in real-world environments. As a result, the ideas and works of AI studies have not been applied to practical cases such as real-world environments, but have been limited to discussed only in "toy-problem" environments.

Based on the above background, this study tackles the practical application of AI to real-world environments. The ultimate goal of this study is to realize AGI that can be applied to real-world environments. This goal can be decomposed into two sub-objectives: (1) Obtaining generality of AI and (2) Applying to real-world environments.

The research question of (1) is as follows: Can AI handle various situations? Many of the current works are for "narrow AI", which is capable to make decisions more effectively than humans when limited to specific environments. On the other hand, the sub-objective (1) is for AI to acquire the ability that can make a decision appropriately in diverse environments. In the environment employed for GVGAI, it is difficult to make a consistent evaluation because each game has a different way of calculating scores. In addition, in GVGAI, it is necessary to learn the outside of tactical decision-making such as domain-specific operations (i.e. how to move characters/agents). In order to focus only on the improvement of decision-making ability, this paper adopts an environment that requires resourcefulness and does not require recognition and learning of operation methods.

The research question in (2) is as follows: Can AI be used in a realistic environment? It is important to produce results in a more realistic environment in order to apply AI researchers to practical cases. Among the differences of environments in AI tasks, some of the main differences that hinder the realization of practical applications are listed as the following items:

1. Observable Information: if sensors give agents access to the complete state of the

environment at each point in time, the task environment is called as "fully observable". On the other hand, an environment is "partially observable" due to noisy and inaccurate sensors or because parts of the state are simply missing from the sensor data. Many "toy-problem" environments employ "fully observable" environments.

2. Agents: There are a lot of types of agent environments. One is "single agent" such as solving a puzzle. Playing chess can be assigned in "two agents" and also "competitive" environments. In the taxi-driving environment, it is a "cooperative" multi-agent environment. Therefore, team sports such as soccer are a "cooperative" and "competitive" multi-agent environment that is one of the most difficult in the agent-environments to learn AI policies.

3. State Change: If the next state of the environment is completely determined by the current state and the action executed by the agent, the environment is called "deterministic"; otherwise, it is "stochastic". Moreover, if the environment can change while an agent is deliberating, the environment is called "dynamic" for that agent; otherwise, it is "static". In real-world problems, the states change "stochastically" and "dynamically".

4. State Expression: The "discrete"/"continuous" distinction applies to the state of the environment, to the way time is handled, and to the percepts and actions of the agent. The number of states in "discrete" environments is finite even if it is very large. On the other hand, there is a virtually infinite number in the "continuous" states. The time in "continuous" environments is represented by discretizing into a very short time in simulation.

In order to tackle the sub-objective (1), it is desirable to be "competitive". This is because, in an adversarial environment, it is always unknown what actions the opponents take. Therefore, highly general decision-making is required. Good performance of Ai in the "competitive" environments implies that its decision-making ability has high generality. On the other hand, to achieve the sub-objective (2), an environment must be "partially observable", "cooperative" multi-agent, "stochastic", "dynamic", and "continuous" as defined in the above items.

To tackle the sub-objectives (1) and (2), this paper employs team sports, especially soccer, as an environment and improves the decision-making of the agents. Soccer satisfies all aforementioned environments. Moreover, in soccer, there are tactics, and the winner of a match highly depends on the characteristics of the tactics. Therefore, it is important to consider the opponents' tactics in soccer and to switch tactics while a game in order to win stably. In addition, learning the behavior of soccer agents can greatly contribute to the realization of AGI as described above. This is because switching the tactics of one's own team according to the tactics of an unknown opponent is similar to changing the decision-making policy in a different environment of AGI. This is the significance of learning the soccer agents' behaviors for the realization of AGI. Therefore, RoboCup Soccer Simulation 2D League is employed in this study.

The main contributions of this dissertation are listed as follows.

In Chapter 2, a method was proposed to obtain an expert's action sequences by using supervised learning. A tactic is represented by the combination of the tree search and the evaluation of action nodes. For this purpose, a multi-layered perception was employed to model the expert player's decision making in a form of an evaluation function. The multi-layered perception was trained by using positive and negative episodes of computational experiments that the proposed method reproduced the behavior of a given expert team, and the model with some settings was more successful than the evaluation function designed by human beings.

In Chapter 3, this paper proposed an idea of intentional substitution for training a machine learning model. The assumption is that the training dataset is complete with no missing values in the training phase while there are missing values during the prediction phase. The proposed intentional substitution replaces the feature value with some value, even if the feature value is known and the value is actually available, Mathematical analysis on the performance of the proposed is presented in terms of the expected in terms of the expected error for the test dataset. It was shown that a trivial substitution using a zero value and an average value may lead to poor prediction performance. In addition, it was shown that there exist appropriate values for the substitution that can be obtained in an ideal situation. This paper discussed the prediction performance of the proposed method using two simple benchmark functions. The experimental results successfully showed the validity of the mathematical analysis and the estimation method of the optimal substitution value in the IVS learning for two- and three-dimensional problems. As the results of computational experiments, the validity of the robust model has shown against the loss for unknown data by estimating the optimal substitution values.

In Chapter 4, the distance between kick distributions for dissimilarity analysis of action trajectories was calculated. This paper employed Earth Movers' Distance, L^2 distance, and Jensen-Shannon divergence as metrics for the distance between kick distributions. In order to show the validity of kick distributions for the similarity analysis, the proposed similarity analysis methods were compared with the human subjective evaluations. The human subjective evaluations for the similarity of the action trajectories were calculated by using a paired comparison method. It was shown that the similarity analysis methods have a positive correlation with the human subjectivity. Thus, it was found that the proposed method has the validity for the similarity analysis. Another contribution of the paper was that the calculation time could be reduced by using the continuous kick probability distributions. Moreover, a method was proposed for analyzing team tactics in RoboCup Soccer using fuzzy inference. For the purpose of improving the performance of the tactical analysis, kick probability distribution and kick direction distribution were employed. The effectiveness of the proposed method was presented by comparing with the previous method on the term-classification performance.

Chapter 5 summarized the obtained results of this paper and discussed future works.

List of Publications

No.	Title	Authors	Journal/Conference name	Chapters in PhD Thesis
1	Evaluation-function modeling with multi-layered perception for RoboCup soccer 2D simulation	T. Fukushima, T. Nakashima, H. Akiyama	Journal of Artificial Life and Robotics, Vol. 25, pp. 440-445, 2020.	Chapter 2
2	Optimal Value Estimation of Intentional-Value-Substitution for Learning Regression Models	T. Fukushima, T. Nakashima, T. Hasegawa, V. Torra	Journal of Advanced Computational Intelligence and Intelligent informatics, Vol. 25, No. 2, pp. 153-161, 2021.	Chapter 3
3	A Study on Intentional-Value-Substitution Training for Regression with Incomplete Information	T. Fukushima, T. Nakashima, T. Hasegawa	Proceedings of the ICML Workshop on the Art of Learning with Missing Values (Artemiss), Online, 9 pages, July 2020.	Chapter 3
4	Similarity Analysis of Action Trajectories based on Kick Distributions	T. Fukushima, T. Nakashima, H. Akiyama	Proceedings of the RoboCup Symposium 2019, 13pages, Sydney, Australia, July 2019.	Chapter 4
5	Team Classification with Tactical Analysis Using Fuzzy Inference in RoboCup Soccer	T. Fukushima, T. Nakashima, V. Torra	Proceedings of the Joint 11th International Conference in Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS&ISIS),Online, December 2020	Chapter 4