

Chapter 10

Awareness Based Recommendation: Passively Interactive Learning System

Tomohiro Yamaguchi

Nara National College of Technology, Japan

Takuma Nishimura

NTT Neomeit, Japan

Keiki Takadama

The University of Electro-Communications, Japan

ABSTRACT

In Artificial Intelligence and Robotics, one of the important issues is to design Human interface. There are two issues, one is the machine-centered interaction design to adapt humans for operating the robots or systems. Another one is the human-centered interaction design to make it adaptable for humans. This research aims at latter issue. This paper presents the interactive learning system to assist positive change in the preference of a human toward the true preference, then evaluation of the awareness effect is discussed. The system behaves passively to reflect the human intelligence by visualizing the traces of his/her behaviors. Experimental results showed that subjects are divided into two groups, heavy users and light users, and that there are different effects between them under the same visualizing condition. They also showed that the authors' system improves the efficiency for deciding the most preferred plan for both heavy users and light users.

INTRODUCTION

Interactive Reinforcement Learning With Human

A long term goal of interactive learning system is to incorporate human to solve complex tasks. Reinforcement learning is the Standard behavior learning method for among robot, animal and human. In interactive reinforcement learning, there are two roles, a learner and a trainer. The input of a reinforce-

DOI: 10.4018/978-1-5225-8356-1.ch010

ment learner as a learning goal is called a *reward*, and the output of the learner as a learning result is called a *policy*. For example, as training a dog by a human trainer, Peterson (2000, 2001) showed that clicker training is an easy way to shape new behaviors. When a dog performs a new behavior to learn, the trainer clicks the clicker as a positive reward. Pryor (2006) remarks that clicker training is a method for training an animal that uses positive reinforcement in conjunction with a clicker to mark the behavior being reinforced under behavior modification principles.

In current researches of interactive reinforcement learning, there are two approaches to support a learner by giving feedback as, whether a learning goal (reward based), or a learning result (policy based). The former approach is clicker training for the robot, in that a human trainer gives a learning goal to the robot learner. In field of robot learning, Kaplan et al. (2002) showed that interactive reinforcement learning method in that reward function denoting goal is given interactively has worked to establish the communication between a human and the pet robot AIBO. The main feature of this method is the interactive reward function setup which was fixed and build-in function in the main feature of previous reinforcement learning methods. So the user can sophisticate reinforcement learner's behavior sequences incrementally.

Ng et al. (1999) and Konidaris & Barto (2006) showed that reward shaping is the theoretical framework of such interactive reinforcement learning methods. Shaping is to accelerate the learning of complex behavior sequences. It guides learning to the main goal by adding shaping reward functions as subgoals. Previous reward shaping methods have three assumptions on reward functions as following:

- Main goal is given or known for the designer;
- Marthi (2007) remarks that subgoals are assumed as shaping rewards those are generated by potential function to the main goal;
- Ng et al. (1999) showed that shaping rewards are policy invariant, it means not affecting the optimal policy of the main goal.

However, these assumptions will not be true on interactive reinforcement learning with a non-expert end-user. Main reason is discussed by Griffith et al. (2013) that human feedback signals may be inconsistent with the optimal policy. It is not easy to keep these assumptions while the end-user gives rewards for the reinforcement learning agent. It is that the reward function may not be fixed for the learner if an end-user changes his/her mind or his/her preference. However, most of previous reinforcement learning methods assume that the reward function is fixed and the optimal solution is unique, so they will be useless in interactive reinforcement learning with an end-user.

To avoid this problem, the latter approaches are that a human trainer provides a sample of learning result to the robot learner. For robot learning with human, inverse reinforcement learning proposed by Ng & Russell (2000) is the method that after the human provides demonstrations of an optimal policy, the reward function for the demonstrations is generated to learn the optimal policy. Another approach is called policy shaping proposed by Griffith et al. (2013). Instead of requiring demonstrations, it allows a human trainer to simply critique the learner's behavior ("that was right/wrong"). Thus the human's feedback is a label on the optimality of actions of each state.

To introduce our approach, we organize reinforcement learning methods. Table 1 shows the characteristics on interactive reinforcement learning. In reinforcement learning, an optimal solution is decided by the reward function and the optimality criteria. In standard reinforcement learning, an optimal solution is fixed since both the reward function and the optimality criteria are fixed. On the other hand, in

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/awareness-based-recommendation/226562

Related Content

Internet-Enabled User Interfaces for Distance Learning

We Liu, Keng Soon The, Roshan Peiris, Yongsoon Choi, Adrian David Cheok, Charissa Lim Mei-Ling, Yin-Leng Theng, Ta Huynh Duy Nguyen, Tran Cong Thien Quiand Athanasios V. Vasilakos (2009).

International Journal of Technology and Human Interaction (pp. 51-77).

www.irma-international.org/article/internet-enabled-user-interfaces-distance/2937

Twin Wiki Wonders? Wikipedia and Wikibooks as Powerful Tools for Online Collaborative Writing

Meng-Fen Grace Lin, Curtis J. Bonkand Suthiporn Sajjapanroj (2009). *Human Computer Interaction: Concepts, Methodologies, Tools, and Applications* (pp. 1262-1282).

www.irma-international.org/chapter/twin-wiki-wonders-wikipedia-wikibooks/22314

Interpretive Flexibility Along the Innovation Decision Process of the UK NHS Care Records Service (NCRS): Insights from a Local Implementation Case Study

Anastasia Papazafeiropoulouand Reshma Gandeche (2007). *International Journal of Technology and Human Interaction* (pp. 1-12).

www.irma-international.org/article/interpretive-flexibility-along-innovation-decision/2897

B is for Bias: From Rational Maximizer to Homo Heuristicus

Calin Valsan (2014). *International Journal of Applied Behavioral Economics* (pp. 35-47).

www.irma-international.org/article/b-is-for-bias/113838

Exploring the Effectiveness of the Clickers-Aided Flipped English Classroom

Zhonggen Yu (2020). *International Journal of Technology and Human Interaction* (pp. 90-102).

www.irma-international.org/article/exploring-the-effectiveness-of-the-clickers-aided-flipped-english-classroom/247039