Chapter 6

Awareness-Based Recommendation:
Toward the Human Adaptive and Friendly Interactive Learning System

Tomohiro Yamaguchi
Nara National College of Technology, Japan

Takuma Nishimura
Nara National College of Technology, Japan & NTT WEST, Japan

Keiki Takadama
The University of Electro-Communications, Japan

ABSTRACT

This chapter describes the interactive learning system to assist positive change in the preference of a human toward the true preference. First, an introduction to interactive reinforcement learning with human in robot learning is given; then, the need to estimate the human’s preference and to consider its changes by interactive learning system is described. Second, requirements for interactive system as being human adaptive and friendly are discussed. Then, the passive interaction design of the system to assist the awareness for a human is proposed. The system behaves passively to reflect the human intelligence by visualizing the traces of his/her behaviors. Experimental results show 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 show that the system improves the efficiency for deciding the most preferred plan for both heavy users and light users.

DOI: 10.4018/978-1-4666-4225-6.ch006
INTRODUCTION

Interactive Reinforcement Learning with Human

In field of robot learning (Kaplan 2002), interactive reinforcement learning method, reward function denoting goal, is given interactively and 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 a fixed and built-in function in the main feature of previous reinforcement learning methods. So the user can sophisticate reinforcement learner’s behavior sequences incrementally.

Shaping (Konidaris 2006; Ng 1999) 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 shaping methods (Marthi 2007; Ng 1999) have three assumptions on reward functions as following:

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

However, these assumptions will not be true on interactive reinforcement learning with an end-user. Main reason is that 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 assumes 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.

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 interactive reinforcement learning, an optimal solution may change according to the interactive reward function. Furthermore, in interactive reinforcement learning with human, various optimal solutions will occur since the optimality criteria depend on human’s preference.

Then the objective of this research is to recommend preferable solutions of each user. The main problem is “how to guide to estimate the user’s preference?”. Our solution consists of two ideas. One is to prepare various solutions by every-visit-optimality (Satoh 2006), another is the coarse to fine recommendation strategy (Yamaguchi 2008).

Requirements for Interactive System as Being Human Adaptive and Friendly

Recent years there are many researches on recommender systems to present information items that are likely of interest to the user. A recommender system is one of the major intelligent and interactive systems with a human user. However, it sometimes seems to be officious for the user. Main problem is that it is active but less intelligent (Yamaguchi 2012).

Table 1. Characteristics on interactive reinforcement learning

<table>
<thead>
<tr>
<th>Type of reinforcement learning</th>
<th>an optimal solution</th>
<th>reward function</th>
<th>optimality criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>standard</td>
<td>fixed</td>
<td>fixed</td>
<td>fixed</td>
</tr>
<tr>
<td>interactive</td>
<td>may change</td>
<td>interactive</td>
<td>fixed</td>
</tr>
<tr>
<td>interactive with human</td>
<td>various optimal</td>
<td>may change</td>
<td>human’s preference</td>
</tr>
</tbody>
</table>