Advanced Robotics, Vol. 13, No. 3, pp. 283–284 (1999)

Behavior learning of a face robot based on the characteristics of human instruction

FUMIYA IIDA [1] and FUMIO HARA [2]

[1] *Department of Computer Science, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[2] *Department of Mechanical Engineering Science, University of Tokyo, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan*

E-mail: iida@ifi.unizh.ch
Keywords: Man–machine interaction; reinforcement learning; learning simulation; human instruction.

For a face robot (Fig. 1), which has many degrees of freedom in action in order to express facial expressions, rotate its face and eyes, and speak to human partners, it is necessary to learn its behavior that can be preferable for an individual human partner. We have found it difficult for the face robot to learn human-friendly behaviors through usual human instruction. The difficulty is mainly centered spotted at the learning algorithm of the face robot not taking account of characteristics of human instruction [1].

In usual human instructions, a human partner gives positive rewards when the learning proceeds toward obtaining correct behaviors, even if the robot does not execute completely correct behaviors. However, in the case when we employ the conventional reinforcement learning algorithm in the face robot instruction experiment with human partners, the face robot cannot learn its behavior by human instructions. In this study, therefore, we propose new reinforcement learning algorithms, i.e. OR-reinforcement and Comparative-reinforcement learning algorithms, based on this characteristics of human instruction.

In the conventional reinforcement learning algorithm, i.e. AND-reinforcement, rewards given by a human partner reinforce the executed action (Fig. 2), in which human rewards do not reinforce 'neighbor' behavior. In the OR-reinforcement, given rewards not only reinforce the executed action, but also the neighbor actions that include the same elements as those of the executed action as shown in Fig. 2. In the Comparative-reinforcement learning, on the other hand, given rewards also reinforce the element actions, but the actions changed in comparison with the action executed one step before.

We conduct computer simulation experiments to examine the effectiveness of the proposed learning algorithms. In the simulation, human rewards are assumed to be given as a positive/negative ones, whose values are set to $+1$ and -1 respectively. We use two models of human instruction that determine which reward is to be given in every step of a human instruction, i.e. model-A and model-B. In model-A, the human gives a positive

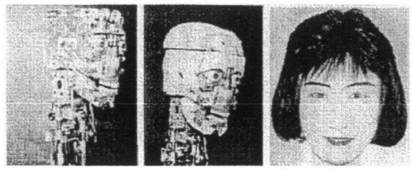

Figure 1. Appearance of the face robot.

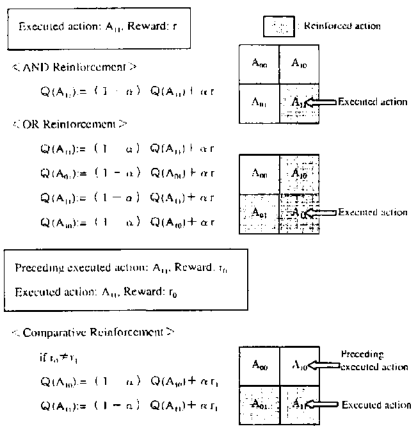

Figure 2. Proposed reinforcement method for each learning algorithm (case of 2 d.o.f.).

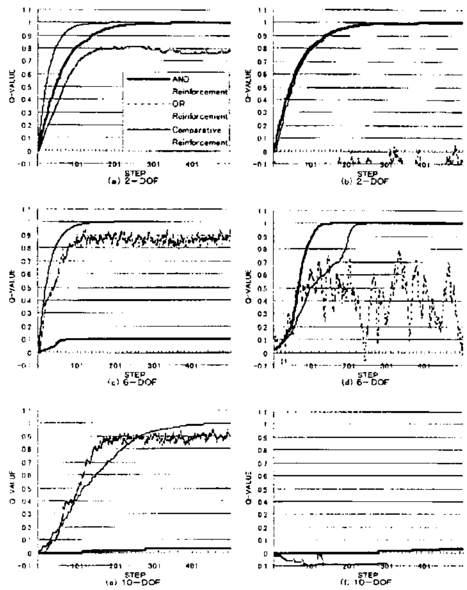

Figure 3. Result of simulation.

reward if robot learning proceeds, i.e. the number of correct element actions increases. In model-B, the human gives a positive reward only if the robot action is a completely correct one, i.e. all of the element actions are the same as those of the correct action. A robot is assumed to have 2, 6 and 10 d.o.f. The results are shown in Fig. 3. For model-A, OR- and Comparative-reinforcement learning algorithms can learn effectively as the number of degrees-of-freedom increases, i.e. the tendency that robot takes a correct action in OR- and Comparative-reinforcement converges earlier than those in AND-reinforcement. On the other hand, OR-reinforcement hinders the correct action learning in the case of model-B, because the OR-reinforcement learning algorithm may reinforce the correct behavior negatively.

In conclusion, human instruction may have a possibility to include both model-A and model-B characteristics, and it can be expected that the Comparative-reinforcement learning algorithm is more effective for a learning by human instructions, which will also be experimentally clarified in a future study.

REFERENCE

1. F. Iida, M. Tabata and F. Hara, Generating personality character in a face robot through interaction with human, in: *Proc. 7th IEEE Int. Workshop on Robot and Human Communication*, pp. 481–486, 1998.