What Kinds of Human Negotiation Skill Can Be Acquired by Changing Negotiation Order of Bargaining Agents?

Keiki Takadama, Atsushi Otaki, Keiji Sato, Hiroyasu Matsushima, Masayuki Otani, Yoshihiro Ichikawa, Kiyohiko Hattori, and Hiroyoki Sato

The University of Electro-Communications
1-5-1 Chofugaoka, Chofu, Tokyo,182-8558 Japan
{keiki@,atsushi.otaki@cas.,keiji@cas.,matsushima@cas.,
masa-o@cas.,yio@cas.,hattori@,sato@}hc.uec.ac.jp

Abstract. This paper focuses on developing human negotiation skills through interactions between a human player and a computer agent, and explores its strategic method towards a human skill improvement in enterprise. For this purpose, we investigate the negotiation skill development through bargaining game played by the player and an agent. Since the acquired negotiation strategy of the players is affected by the negotiation *order* of the *different types* of agents, this paper aims at investigating what kind of the negotiation strategies can be learned by negotiating with different kinds of agents in order. Through an intensive human subject experiment, the following implications have been revealed: (1) human players, negotiating with the human-like behavior agent firstly and the strong/weak attitude agent secondly, can neither obtain the large payoff nor win many games, while (2) human players, negotiating with the strong/weak attitude agent firstly and the human-like behavior agent secondly, can obtain the large payoff and win many games.

Keywords: human skill development, agents, interaction, subject experiment, bargaining game.

1 Introduction

Recently, the human skill development has been much attention on and has been regarded as very important issue for enterprise growth. To achieve it, many approaches for training staffs, such as e-learning, are explored as the progress of the information technology (IT). According to Oshima, the human skill development in enterprise that employs wider IT is categorized in the following four stages [5]: (1) in stage 1, enterprises attempt to improve training efficiency by using IT (*e.g.*, e-learning), which is the most basic approach that enterprises take; (2) in stage 2, enterprises optimize training course by excluding useless training course and including useful training course; (3) in stage 3. enterprises attempt to connect training

G. Salvendy, M.J. Smith (Eds.): Human Interface, Part II, HCII 2011, LNCS 6772, pp. 335–344, 2011.

course with actual works, which enables the enterprises to utilize IT as the communication and knowledge management tools in actual works; and (4) in stage 4, enterprises develop human skills for management by using IT to support employee career development. What should be noted here is that most enterprises have reached at the stage 2, but they have not yet reached at the stages 3 and 4. This is because a practical human skill depends on human experience or intuition, which makes it difficult to train such employees strategically, in comparison with the training of the new employees for general knowledge such as compliance through the e-learning. Like the reached stage of enterprises, most studies on human skill development using IT is related to e-learning [1][7] as the approach in the stages 1 and 2, while a systematic approach to the stages 3 and 4 without depending the human intuition and experience have not yet been fully studied.

To overcome this problem, our previous research [6] focused on *negotiation skill* as one of important skill and explored its skill development through human and agent interaction. We take this approach because the skills trained by an interaction with agents have the potential of being a highly important ability for the employee in administrative or managerial positions (in particular, enterprise managers having little management experience are useful for gaining such experience, which contributes to fostering the leader in organizations). From this perspective, our previous research addressed the negotiation skill development through *bargaining game* as the first step toward our goal. Specifically, we conducted the subject experiments by negotiating human players with the following two different kinds of the agents: (a) strong/weak attitude agents making aggressive/defensive proposals in advantageous/dis- advantageous situations; (b) the human-like behavior agents making mutually agreeable proposals as the number of games increases. However, human subjects in this experiment negotiated with the *same* agents. Since the acquired negotiation strategy has the possibility of being affected by the negotiation *order* of *different* type of agents, this paper aims at investigating what kind of the negotiation strategies can be learned by negotiating with different kinds of agents in order.

This paper is organized as follows. Section 2 explains the bargaining game as an example for the negotiation skill development. An agent implementation of the bargaining game is described in Section 3. The subject experiment and its results are described in Sections 4. Finally, our conclusions are made in Section 5.

2 Bargaining Game

2.1 Bargaining Game

Bargaining game [8], in which two players aim at dividing money through negotiation between them, have been studied in the context of bargaining theory [3] as one of the major subjects in game theory [4]. The bargaining game is proposed for investigating when and what kind of offers of an individual player can be accepted by the other players. From this feature, the game requires skill of determining when and what kind

of offers are needed to win the game. Concretely, the bargaining game can investigate (1) whether the players learn situational advantages and disadvantages in negotiation and (2) whether they also learn appropriate strategy based on these situations.

This game is generally a one-shot game as one of typical examples of negotiation. However, it is not a realistic as social activity due to the fact that two or more negotiations are usually conducted. From this fact, we employ the sequential bargaining game as shown in Figure 1. In this example, the player 1 starts by offering 30% of reward R (R=10 in the example), then the player 2 counter offers 20% by refusing the player 1's offer. Through such negotiations, the player 2 finally offers 40%, which the player 1 accepts, meaning that the player 1 acquires 40% and the player 2 acquires 60%. If this game is not completed within the maximum number of negotiation size (MAX_STEP) which is predetermined in advance and shared among players, both players cannot acquire any rewards, meaning that the negotiation fails.

Figure 2 shows the final negotiation step in which the player 2 offers 10% to the player 1. When the player 1 refuses it, both players acquire none of the reward as shown in the upper figure, but when the player 1 accepts it, the players 1 and 2 respectively acquire 10% and 90% rewards as shown in the lower figure. If the players are rational, then the player 1 acquires even small offer from the player 2 because 10% reward is larger than 0% reward, meaning that anything is better than nothing. This indicates that the player making the last offer has an advantage in comparison with the other player because the last offerer can acquire a large reward by making a small offer to the other agent if the other player is rational. In this example, the player 2 stands in the advantageous situation, while the player 1 stands in the disadvantageous situation. Note that the game result can be calculated beforehand through game theory analysis if both players are rational. However, it goes without saying that humans sometimes take irrational behaviors, which cannot predict the result, *i.e.*, the subject experiments are needed to understand the results.

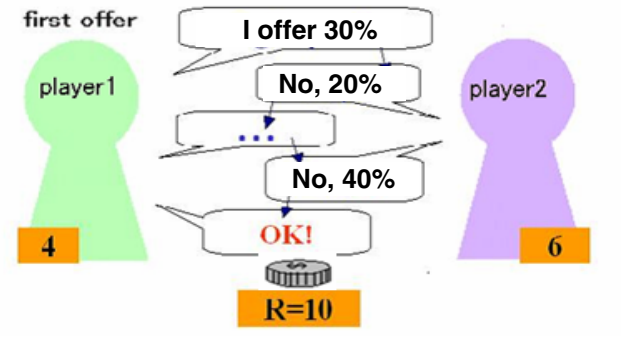


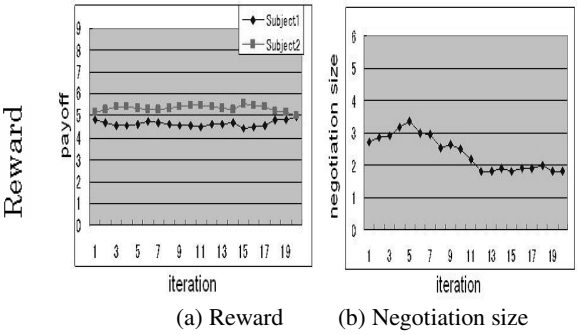
Fig. 1. Example of sequential bargaining game

Fig. 2. Final negotiation

2.2 Previous Subject Experiment

Our previous research [2] conducted the subject experiments of the sequential bargaining game and obtained the results shown in Figure 3, where the vertical and horizontal axes respectively indicate the acquired reward and iterations (*i.e.*, the number of games) in the left graph, while the vertical and horizontal axes respectively indicate the negotiation size (*i.e.*, negotiation counts until the offer is accepted from the other player) and iterations in the right graph. The left graph suggests that rewards of both human players are close to 5:5 (50%:50%), while the right graph suggests that the negotiation size increases in early iterations and decreases after several iterations because of the following reasons: (1) the negotiation process size increases in the early several iterations because both players do not know their strategies each other, which promotes them to explore possibilities of obtaining a larger reward by competing with each other which requires further negotiations (*i.e.*, a larger negotiation process size is required to explore a larger reward); and (2) the negotiation process size decreases in the late several iterations because both players find a mutually agreeable payoff by knowing their strategies each other, which decreases the motivation of players to negotiate again (*i.e.*, a few negotiation size is enough to determine their rewards) [10]. We call this tendency as the decreasing trend.

Hereafter, we use the terms “payoff” and “agent” instead of the terms “reward” and “player” for their more general meanings in the bargaining game.



(a) Reward (b) Negotiation size

Fig. 3. Result of previous subject experiments [2]

3 Modeling Agents

This section explains the agent proposed in our previous research [2], which can replicate the tendency of the payoff and negotiation process size shown in Figure 3. This agent is based on Q-learning [11] in the reinforcement learning context [9].

3.1 Knowledge

The agent has knowledge of the negotiation strategy represented by Q-tables composed of many Q-values (*i.e.*, Q(s,a)), each of which indicates the expected payoff (*i.e.*, reward) when the agent executes its action "a" in situation "s". Q(3,2) in the circle in Figure 4, for example, means that the expected payoff of offering 2 (20%) to the other agent when 3 (30%) is offered by the opponent agent is 7.7.

action

situation	Acc	1	2	3	4	5	6	7	8	9
1	1.0	8.9	7.9	6.5	5.9	4.6	3.9	2.7	1.6	0.9
2	2.0	8.8	7.8	6.6	5.6	4.9	3.7	2.8	1.8	0.8
3	3.0	8.6	7.7	6.6	5.7	4.8	3.8	2.8	1.7	0.9
4	4.0	8.7	7.6	6.8	5.7	4.9	3.6	2.9	1.7	0.8
5	5.0	8.9	7.8	6.7	5.8	4.8	3.8	2.6	1.9	0.9
6	6.0	8.7	7.7	6.9	5.8	4.7	3.9	2.6	1.8	0.8
7	7.0	-	-	-	-	-	-	-	-	-
8	8.0	-	-	-	-	-	-	-	-	-
9	9.0	-	-	-	-	-	-	-	-	-

Fig. 4. Q-table

The Q-value is used to determine the action of the agent (*i.e.*, the action with the high Q-value is selected with a high probability) and its value is updated by Eq. (1), where α, r, γ, s', and a' indicate the learning rate, payoff, discount factor, the next state, and the next action.

$$Q(s,a) \leftarrow Q(s,a) + \alpha(r + \gamma \max_{a' \in A(s')} Q(s',a') - Q(s,a)) \tag{1}$$

Each agent has the same number of Q-tables based on the number of the maximum negotiation size as shown in Figure 5, *i.e.*, the next state s' is used in the next Q-tables (not in the same Q-table). This is unique modeling of our agents compared to the conventional Q-learning agent. Specifically, the number of Q-tables each agent has is MAX_STEP/2 (*e.g.*, three when MAX_STEP is six). Note that t represents the negotiation times (*i.e.*, the number of negotiations).

3.2 Action Selection

The following action selections are employed in the agent:

- **ε-greedy selection:** This method selects an action randomly in probability ε($0 \leqq \varepsilon \leqq 1$), while it selects the action with the largest Q-value in probability 1-ε.

- **Boltzmann distribution selection:** This method selects an action stochastically based on the Q-value. The probability of selecting action "a" is calculated by Eq. (2), where T is the temperature parameter adjusting action selection randomness. When T is low, the agent's behavior becomes rational, *i.e.*, the agent selects his action with the maximum Q-value.

$$p(a \mid s) = e^{Q(s,a)/T} \Big/ \sum_{a \in A_i} e^{Q(s,a_i)/T} \tag{2}$$

To replicate the tendency of the human-like behavior (i.e., the decreasing trend described in Section 2.2), our previous research introduced the random decreasing parameter, *changeRate* ($0 < changeRate < 1$), to gradually decrease randomness as shown in Eq. (3).

$$T = T \times (1 - changeRate) \text{ in each iteration} \tag{3}$$

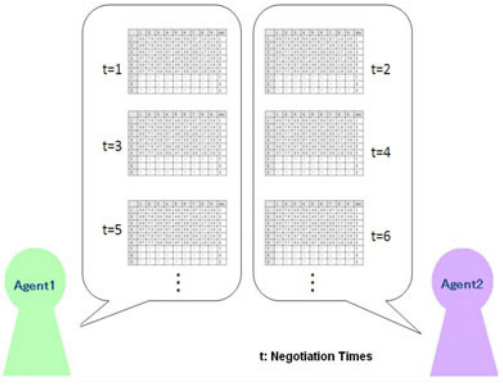


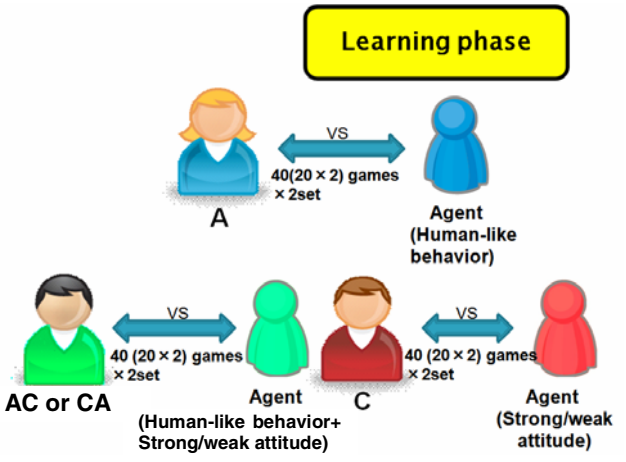
Fig. 5. Q-tables in agents

4 Subject Experiments

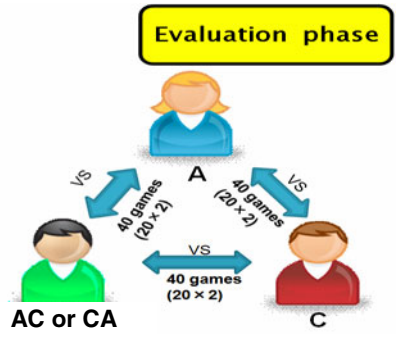
4.1 Outline

Subject experiments involved two phases, *i.e.*, the *leaning* phase in Figure 6 (a) and the *evaluation* phase in Figure 6 (b). The learning phase aims at making the subjects learn their agent strategies, while the evaluation phase aims at evaluating an effectiveness of the acquired strategies. In the learning phase, the six human subjects who are unfamiliar with the bargaining game enter different six rooms, and each subject plays the bargaining game with one kind of the agents. In detail, the subject A negotiates with a human-like agent (described in the next section), the subject C negotiates with a strong/weak attitude agent (described in the next section), and the subject AC (or CA) negotiates with the human-like agent firstly (secondly) and the strong/weak attitude agents secondly (firstly) as shown in Figure 6 (a). In the evaluation phase, on the other hand, the player plays the game with another subject

who has already negotiated with a different kind of the agent via the computer network. Specifically, the subject A negotiates with the subjects C and AC (or CA), the subject C negotiates with the subjects A and AC (or CA), and the subject AC (or CA) negotiates with the subjects A and C as shown in Figure 6 (b).



(a) Learning phase



(b) Evaluation phase

Fig. 6. Outline of subject experiment

In the experiment, human subjects are only informed that negotiation opponents change after 40 games (*i.e.*, 20 games starting with the first offerer and 20 games starting with the second offerer). For the game setting, the total number of games for one subject is set as 160 games, meaning that each subject negotiates with the same agent in two sets of 40 games (*i.e.*, 80 games in total) and negotiates with the other two subjects negotiating with different type of agents in 40 games (*i.e.*, 80 games in total negotiating with the other two subjects). The point of this experiment is that opponents are changed from computer agents to subjects without the change being noticed because they are in separate rooms and negotiate with the opponent via the

computer network. After all experiments, all subjects answer the questionnaire to investigate the negotiation skill trained by different kinds of agents.

Subjects are categorized into groups A, C, and AC (or CA) based on agent type. We conducted this subject experiment twice with the six different subjects, *i.e.*, the first and second experiments were conducted by two of A, C, and AC subjects, and by two of A, C, and CA subjects, respectively.

4.2 Two Bargaining Agents

To investigate an effectiveness of the training through the negotiation with the agent, the following agents are employed as the same as our previous research [2]:

(1) **Strong/weak attitude agent:** This agent makes aggressive/defensive proposals in an advantageous/disadvantageous situation using ε-greedy action selection with ε=0.1. For example, this agent obtains a large (around 70%) payoff in an advantageous situation and acquires a small (around 30%) payoff in a disadvantageous situation as shown in the left of Figure 7.
(2) **Human-like behavior agent:** This agent makes mutually agreeable proposals as iteration increases using randomness decreasing Boltzmann distribution selection with T=1000 and *changeRate*=0.00001. For example, this agent acquires around the 50% payoff in early several iterations and acquires around the 70%/30% in advantageous/disadvantageous situation in late several iterations as shown in the right of Figure 7, which shows the decreasing trend of the negotiation process size like human behavior.

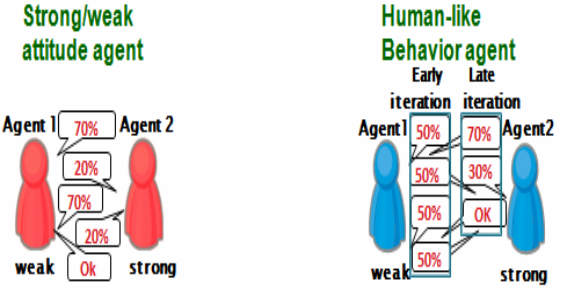


Fig. 7. Characteristics of two kind of bargaining agents

4.3 Results

Figure 8 shows the subject experiment results of the subjects A, C, AC, and CA in the evaluation phase. In this figure, the vertical axis indicates the average ranking of the win and payoff, while the horizontal axis indicates the subjects A, C, AC, and CA, each of values are averaged from four subjects A and C and two subjects of AC and CA. The gray, white and black bars in the figure indicate the win ranking, payoff ranking, and average of win and payoff ranking averaged from (a) and (b), respectively. In detail, Figure 8 (a) shows the win and payoff ranking in the *total* games, Figure 8 (b) shows the win and payoff ranking in *each set* game, and Figure 8

(c) shows the average ranking of win and payoff, all of which are averaged from (a) and (b). Note that the win and payoff rankings in the *total* games are determined by the order of "the total number of wins" and "the total acquired payoff", while those in *each set* game are determined by the order of "the number of wins in one set game" and "the acquired payoff in one set game".

Figure 8 shows the same tendency of the subjects A and C, found in our previous research [6], *i.e.*, (1) the subjects C acquired the largest payoff from the total game viewpoint; and (2) the subjects A win the game more than the other subjects from each set game viewpoint. These results suggest that (1) the subjects C learn the strategy to *acquire the large payoff* through negotiation with the strong/weak attitude agent; and (2) the subjects A learn the strategy to *win the game* through negotiation with the human-like behavior agent, because the number of the draw, *i.e.*, the negotiation failure (which exceeds the maximum number of negotiation size) in the subjects A is larger than that in the subjects C, meaning that the subjects A win or draw the games while the subjects C mostly win the game but sometimes lose it with few draws.

In comparison with the subjects A and C, (1) the subjects AC, negotiating with the human-like behavior agent firstly and the strong/weak attitude agent secondly, can neither acquire the large payoff nor win many games, while (2) the subject CA, negotiating with the strong/weak attitude agent firstly and the human-like behavior agent secondly, can win more games than the subjects C and acquire the larger payoff than the subjects A, which contributes to deriving the highest ranking of the average of win and payoff from the total and each game viewpoints as shown in Figure 8 (c). This indicates that the negotiation *order* of the *different types* of agents gives a big influence on the learning of the negotiation strategies.

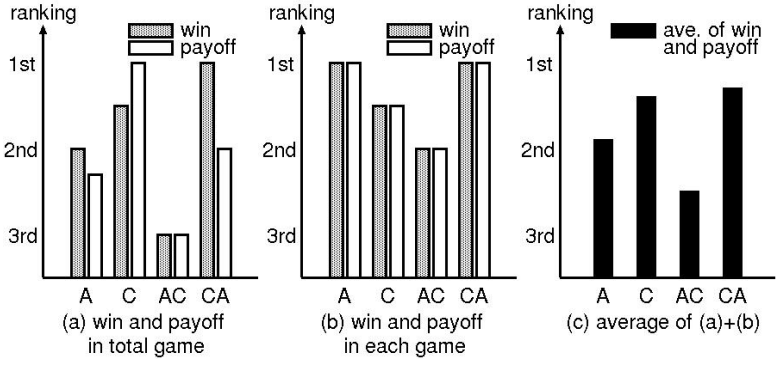


Fig. 8. Result on subject experiments

5 Conclusions

This paper focused on developing human negotiation skills through interactions between a human player and a computer agent, and explored its strategic method towards a human skill improvement in enterprise. For this purpose, we investigated

the negotiation skill development through bargaining game played by the player and an agent. In detail, we focused on an influence of the negotiation *order* of the *different types* of agents, and investigated what kind of the negotiation strategies can be learned by negotiating with different kinds of agents in order. Through an intensive human subject experiment, the following implications have been revealed: (1) human players, negotiating with the human-like behavior agent firstly and the strong/weak attitude agent secondly, can neither obtain the large payoff nor win many games, while (2) human players, negotiating with the strong/weak attitude agent firstly and the human-like behavior agent secondly, can obtain the large payoff and win many games.

Since this work is the first stage toward the strategic human skill development, the following issues must be resolved in the near future: (1) conducting more subject experiments to improve a reliability of the found implications; and (2) employing other agents that have different features to generalize the found implications.

References

1. CIPD. People Management and Technology: Progress and Potential, pp. 14-28 (2005)
2. Kawai, T., Koyama, Y., Takadama, K.: Modeling Sequential Bargaining Game Agents Towards Human-like Behaviors: Comparing Experimental and Simulation Results. In: The First World Congress of the International Federation for Systems Research (IFSR 2005), pp. 164–166 (2005)
3. Muthoo, A.: Bargaining Theory with Applications. Cambridge University Press, Cambridge (1999)
4. Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press, Cambridge (1994)
5. Oshima, A.: The Progress and Issues of IT based Human Resource Development in Japanese corporations. In: The 14th Symposium on Social Information Systems Symposium, pp. 99–104 (2008) (in Japanese)
6. Otaki, A., Hattori, K., Takadama, K.: Towards Strategic Human Skill Development Through Human and Agent Interaction: Improving Negotiation Skill by Interacting with Bargaining Agent. Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII) 14(7), 831–839 (2010)
7. Parry, E., Tyson, S.: Technology in HRM: The Means to Become Strategic Business Partner? In: Storey, J. (ed.) Human Resource Management, Thomson Learning, pp. 235–249 (2007)
8. Rubinstein, A.: Perfect Equilibrium in a Bargaining Model. Econometrica 50(1), 97–109 (1982)
9. Sutton, R.S., Bart, A.G.: Reinforcement Learning -An Introduction. MIT Press, Cambridge (1998)
10. Takadama, K., Kawai, T., Koyama, T.: Micro- and Macro-Level Validation in Agent-Based Simulation: Reproduction of Human-Like Behaviors and Thinking in a Sequential Bargaining Game. Journal of Artificial Societies and Social Simulation (JASSS) 11(2) (2008)
11. Watkins, C.J.C.H., Dayan, P.: Technical note: Q-learning. Machine Learning 8, 55–68 (1992)