AgentK2: Compromising Strategy Based on Estimated Maximum Utility for Automated Negotiating Agents

Shogo Kawaguchi, Katsuhide Fujita, and Takayuki Ito

1 An Implementation of Negotiating Agents Based on Compromising Strategy

1.1 Other Agent's Analysis and Basic Strategy

In the setting of the competition, my utility space is not mutually taught the other agent. Therefore, there is a little information can be used for the strategy construction. In this thesis, it proposes the technique for expecting the best mutual agreement offer that can be get out from the other agent in the future from the statistical information of the value in which the other agent's proposal is evaluated it's own utility space. Moreover, it is assumed that the agent's behavior is decided to compromise to the best mutual agreement offer. Concretely, agent's own behavior is decided based on the following expressions (1) and expressions (2).

$$emax(t) = \mu(t) + (1 - \mu(t))d(t) \tag{1}$$

$$target(t) = 1 - (1 - emax(t))t^{\alpha} \tag{2}$$

$emax(t)$ shows the maximum utility value presumed to t which is receipt from the other agent. It calculates by average $\mu(t)$ of value in which offer of other agent who accumulates and collected is evaluated by own utility space and width $d(t)$ of action until time t. The width of the behavior of the other agent is presumed by deflection,

Shogo Kawaguchi · Takayuki Ito
Nagoya Institute of Technology
e-mail: kawaguchi@itolab.mta.nitech.ac.jp,
ito.takayuki@nitech.ac.jp

Katsuhide Fujita
Nagoya Institute of Technology / Massachusetts Institute of Technology
e-mail: fujita@itolab.mta.nitech.ac.jp

T. Ito et al. (Eds.): Complex Automated Negotiations, SCI 435, pp. 235–241.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

and whether a how much advantageous offer for me can be get out from the other agent by tempering with the average is considered. The width $d(t)$ of the action until time t is calculated by decentralization of value in which offer of other agent who accumulates and collected is evaluated by own utility space.

When assuming that other agent's offer was generated in own utility space based on uniform distribution. Decentralization can be calculated as follows.

$$\sigma^2(t) = \frac{1}{n}\sum_{i=0}^{n} x_i^2 - \mu^2 = \frac{d^2(t)}{12} \tag{3}$$

Therefore, when you calculate the width $d(t)$ of the action from decentralization

$$d(t) = \sqrt{12}\sigma(t) \tag{4}$$

The width $d(t)$ of the action of the other agent is presumed from decentralization. When the mean value of the action is located at the center of the domain of the utility value, it can be thought that the maximum value that can be get out from the other agent is harmony of $1/2$ of the average and the width of the action. However, it is possible to move only in the direction where the utility value is high when the average of the utility value in the utility space of other agent's offer is low, and when the average is high oppositely, the action can be expanded only in a low direction. Therefore, accurate presumption is done by assuming the mean value to be weight. $target(t)$ is a standard of the utility value of the proposal at time t, and α is a coefficient in which the speed of the compromise is adjusted. Ending the negotiation early doesn't have the advantage in this rule. Therefore, it is effective that it searches for other agent's trend consuming the time limit, and repeating the proposal mutually, and both search for both mutual agreement candidate ideas with high utility value. However, because it is a tournament form, own utility value is requested to be raised as much as possible. Therefore, it proposes the offer that my utility value rises, and it proposes the action that approaches the maximum value that can be get out from the other agent presumed with the time passage immediately after beginning of the negotiation.

Figure 1 is an example of $target(t)$ that changes α when $emax(t)$ is set by $\mu(t) = \frac{1}{10}t \quad d(t) = \frac{2}{5}t^2$ from 1 to 9.

1.2 Control of Approach

It is not possible to correspond when the other agent take a hard stance for a simple strategy. It concede that the other agent in the hard stance too much when approaching in the maximum.

Figure 2 is an example of $target(t)$ that changes α when $emax(t)$ is set by $\mu(t) = 0 \quad d(t) = \frac{1}{10}$ from 1 to 9. When the approach is not controlled, it concedes to the other party according to $target(t)$ without bounds as shown in figure 2.

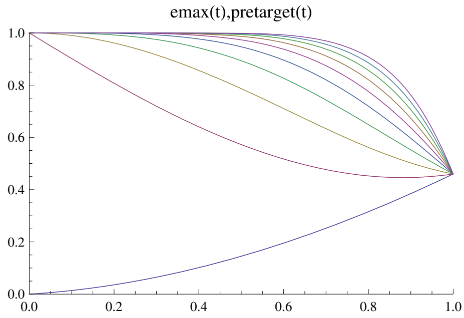

Fig. 1 Example of $target(t)$ when $emax(t)$ is set by $\mu(t) = \frac{1}{10}t \quad d(t) = \frac{2}{5}t^2$

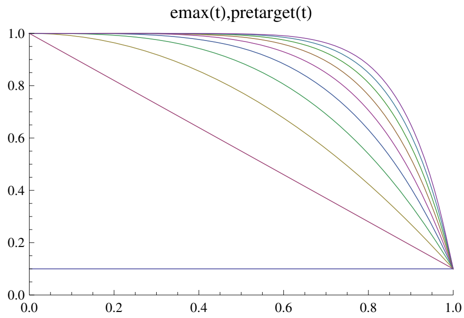

Fig. 2 Example of $target(t)$ when $emax(t)$ is set by $\mu(t) = 0 \quad d(t) = \frac{1}{10}$

Then, the the degree of concession of each is measured. When own concession degree is too larger than the width of the presumed behavior of the other agent, the approach is slowed down. The degree of a concession each other is presumed by the following expressions (5). The degree of the lowest concession at time t is assumed to be $g(t)$.

$$ratio(t) = \begin{cases} \frac{d(t)+g(t)}{1-target(t)} & \text{if } \frac{d(t)+g(t)}{1-target(t)} < 2 \\ 2 & \text{otherwise} \end{cases} \tag{5}$$

The movement of $ratio(t)$ when $emax(t)$ is set by $\mu(t) = 0 \quad d(t) = \frac{1}{10}$ is figure 3.

When I am conceding too much compared with the other party, $ratio(t)$ approaches 0 as shown in figure 3. Therefore, an excessive approach is controlled by assuming concession degree $ratio(t)$ to be weight of the target of the action. $target(t)$ is defined by using $ratio(t)$ as expression (6).

$$target(t) = ratio(t) * (1 - (1 - emax(t))t^{\alpha}) + (1 - ratio(t)) \tag{6}$$

When the other agent is amiable, it is possible to concede to the other party quickly by tempering with the degree of a concession each other and controlling own

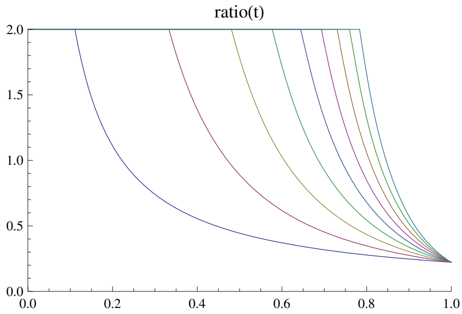

Fig. 3 Example of $ratio(t)$ when $emax(t)$ is set by $\mu(t) = 0 \quad d(t) = \frac{1}{10}$

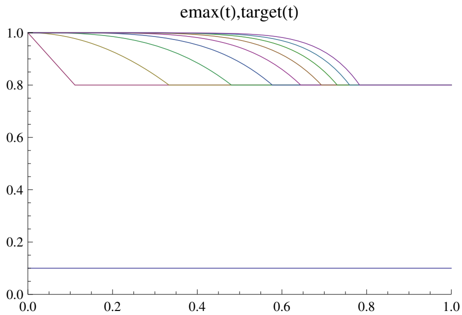

Fig. 4 $target(t)$ that introduced $ratio$

behavior. It becomes possible to model the action that doesn't concede more than constancy when opposing it oppositely.

The appearance of $target(t)$ that introduces $ratio(t)$ is figure 4. The thing that an excessive approach is controlled compared with figure 2 can be perceived in figure 4.

1.3 Coefficient Selection

Next, the function that selects coefficient α is shown. It had swinging at the first stage of the negotiation, and a steady action was designed in the last stage. Therefore, the random nature is introduced into the compromise coefficient. However, to use *target* to judge whether to accept the other agent's offer, the judgment should make sure there is nothing swing. Therefore, it is time when it makes own offer in case of occasion to judge other agent's offer and a different coefficient is used. Of each is assumed to be $\alpha(t)$ and $\beta(t)$, and the width of swinging of the action is decided as τ by the following expressions (7) and (8).

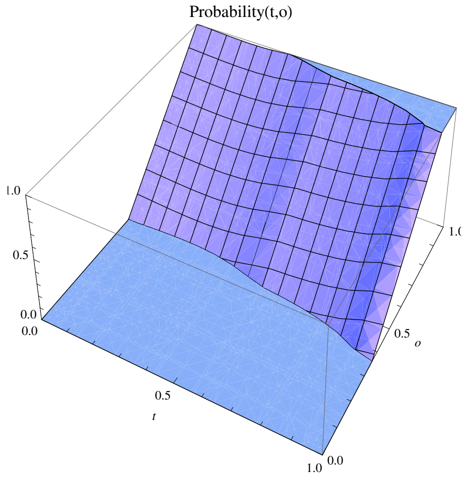

Fig. 5 Acceptance probability space

$$\alpha(t) = 1 + \tau + 10\mu(t) - 2\tau\mu(t) \tag{7}$$

$$\beta(t) = \alpha + random[0,1] * \tau - \frac{\tau}{2} \tag{8}$$

Expression (7) is a coefficient used when whether the other agent's offer is accepted is judged. Expression (8) is a coefficient used when own proposal is made. $random[0,1]$ of expression (8) generates a random value between from 0 to 1.

1.4 Decision of My Offer, and Evaluation of Other Agent's Offer

The idea which offer to select in my utility space was shown. It searches for the offer that should be proposed based on it. The search becomes difficult as the issue increases. Then, it searches for alternatives with the utility value of $target(t)$ while changing the starting position of the search at random by using iterative deepening depth-first search.

Next, evaluation whether to accept other agent's offer. Whether it accepts is judged stochastically based on the distance with target $target(t)$ and the distance from the average. The expression of the probabilistic computation of the acceptance is (9).

$$P = \frac{t^5}{5} + (\text{Offer} - emax(t)) + (\text{Offer} - target(t)) \tag{9}$$

Acceptance probability P is calculated by using time passage, value Offer in which offer by other agent is evaluated by own utility space, $target(t)$ when $\alpha(t)$ is used for coefficient of approach, and estimated maximum value $emax(t)$.

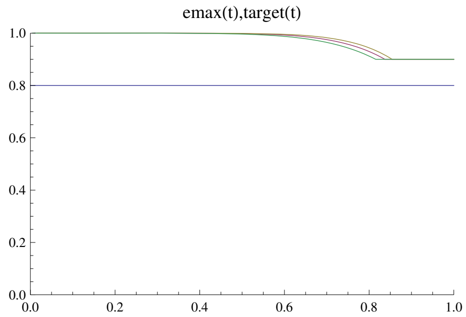

Fig. 6 Example of $target(t)$ when $emax(t)$ is set by expression 10

Figure 5 shows the acceptance probability space when the setting of $emax(t)$ is assumed to be $\mu(t) = \frac{1}{10}t \quad d(t) = \frac{2}{5}t^2$, and the axis reaches time t and utility value o of the offer from the other agent.

Even if it is a small probability, the possibility of accepting by repeating actually remains. Therefore, when acceptance probability P is 0.1 or less, it becomes 0.

1.5 Correction in the Last Stage of Negotiation

Next, the correction in the last stage of the negotiation is shown. An excessive compromise was able to be controlled by introducing presumption concession degree $ratio(t)$. However, the compromise stops when the other agent is proposing the mutual agreement candidate offer of a close utility value to its target and it doesn't change.

$$\mu(t) = \frac{4}{5}t \quad d(t) = 0 \tag{10}$$

Figure 6 is an example of $target(t)$ when $emax(t)$ is set by expression (10). The following expressions are given as a correction when the other agent is presenting the mutual agreement candidate offer of a close utility value to own target.

$$\gamma(t) = -300t + 400 \tag{11}$$

$$\delta(t) = target(t) - emax(t) \tag{12}$$

$$\varepsilon(t) = \begin{cases} \frac{1}{\delta(t)^2} & \text{if } \frac{1}{\delta(t)^2} < \gamma(t) \\ \gamma(t) & \text{otherwise} \end{cases} \tag{13}$$

$$\eta(t) = \frac{\delta(t) * \varepsilon(t)}{\gamma(t)} \tag{14}$$

$$target_2(t) = \begin{cases} target(t) - \eta(t) & \text{if } target(t) > emax(t) \\ emax(t) & \text{otherwise} \end{cases} \tag{15}$$

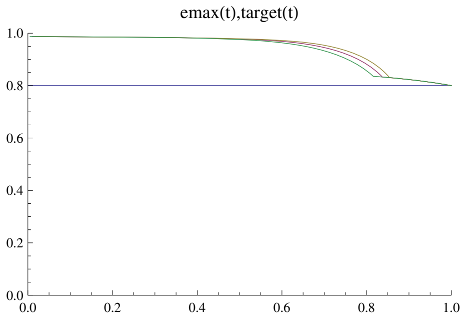

Fig. 7 Example of applying correction function to figure 6

$\gamma(t)$ is a function that adjusts the approach of the correction in the last stage of the negotiation. $\delta(t)$ shows the difference between targeted value $target(t)$ and estimated maximum value $emax(t)$. It limits it by $\gamma(t)$ though $\varepsilon(t)$ is in inverse proportion to the square of $\delta(t)$. $\eta(t)$ is a function that decides which extent to correct by using $\gamma(t)$,$\delta(t)$, and $\varepsilon(t)$ When $target(t)$ when coefficient $beta(t)$ is used is larger than $emax(t)$ correction is added to $target(t)$, and it is assumed a final target according to $\eta(t)$ function. It substitutes it by $emax(t)$ when $target(t)$ is smaller than $emax(t)$. This is assumed to be expression (15) as a final standard.

A constant approach is done when the distance with the other party is short when the correction is introduced. Therefore, the approach doesn't stop as shown in figure 6. It acts as shown in agent chart 7.

1.6 Strategy Corresponding to Discount Factor

$$DiscountTarget(t) = target(t) * DiscountRate(t) \tag{16}$$

$$DiscountTarget(t) * (1 - \mathrm{BOU}) + target(t) * \mathrm{BOU} \tag{17}$$

This is a strategy corresponding to Discount Factor. First strategy, Target is corrected by using the discount rate.(16) But, It decreases too much in this correction. Then, it adjusts it using the Best Offered opponent bid(BOU).(17) As a result, it is prevented from decreasing too much.