An Objective Function based on Fuzzy Preferences in Dynamic Decision Making

Yuji Yoshida[1], Masami Yasuda[2], Jun-ichi Nakagami[2],
Masami Kurano[3], and Satoru Kumamoto[1]

[1] Faculty of Economics and Business Administration, the University of Kitakyushu
4-2-1 Kitagata, Kokuraminami, Kitakyushu 802-8577, Japan
[2] Faculty of Science, Chiba University 1-33 Yayoi-cho,Inage,Chiba 263-8522, Japan
[3] Faculty of Education, Chiba University 1-33 Yayoi-cho,Inage,Chiba 263-8522, Japan

Abstract. This paper presents a mathematical model for dynamic decision making with an objective function induced from fuzzy preferences. The fuzzy preference is related to decision making in artificial intelligence, and this paper models human behavior based on his fuzzy preferences. A reasonable criterion based on fuzzy preferences is formulated for the dynamic decision making, and an optimality equation for this model is derived by dynamic programming.

1 Introduction

Utility functions are applied as decision maker's personal criteria in economics and management science when expected values of objectives are monotonically corresponding to economic utilities([2, 7]). On the other hand, in system control, engineering, etc., the distances are used when we need to minimize the errors between the target and the controlled objects([13]). When these objective functions are applied, they are given under invariant value criteria. This paper deals with decision maker's personal fuzzy preferences in dynamic behavior instead of these objective functions([3]). The fuzzy preference is related to decision making in artificial intelligence([1, 8, 11, 12]). This paper presents a dynamic decision making model with fuzzy preferences, and we discuss a reasonable criterion based on fuzzy preferences in dynamic decision making. By using dynamic programming, we also discuss an optimality equation in a situation where the decision maker is accustomed to his environment. Mathematical models simulating human behavior based on his decision making are applicable to various fields: robotics, customers' behavior analysis in marketing, linguistic learning modeling, multi-agent systems and so on([5]). The dynamic decision making model presented in this paper is designed as follows:

Step 1 (Diagnosis): Obtain input data from perceived state.
Step 2 (Cognition): Approximate the perceived state by cognizable states based on the experience set which consists from initial states and past cognized states, and decide current cognized state.

Step 3 (Decide preference): Decide preference for the state and memorize the new state into the experience set.

Step 4 (Look-ahead): Calculate transition probability at current state with actions.

Step 5 (Estimation): Estimate the ranking utility with preferences and decide an objective function.

Step 6 (Determination): Choose an action to optimize the objective function and move to the next state. Go to Step 1.

2   Preference and ranking

In this section, we introduce basic properties of fuzzy relations and fuzzy preferences, and we discuss a ranking method based on them. Further, we consider an extension of fuzzy preferences and the ranking method for a dynamic decision making model. Let C be a sigma-compact convex subset of some Banach space. The attributes of the states/objects can be represented as the d-dimensional coordinates when the Banach space is taken by d-dimensional Euclidean space $\mathbb{R}^d$. *States* are given by elements of C. We consider two kinds of states: One is *perceived states* and the other is *cognized states*. A perceived state represents a state outside the system and a cognized state represents a state inside the system. In this section, we deal with cognized states. Let $\mathcal{S}$ be a subset of C, which $\mathcal{S}$ has finite elements. A map $\mu : \mathcal{S} \times \mathcal{S} \mapsto [0,1]$ is called a fuzzy relation on $\mathcal{S}$. Fuzzy preferences are defined by fuzzy relations on $\mathcal{S}$ ([3,6]): A fuzzy relation μ on $\mathcal{S}$ is called a *fuzzy preference relation* if it satisfies the following conditions (a) - (c):

(a) $\mu(a, a) = 1$ for all $a \in \mathcal{S}$.
(b) $\mu(a, c) \geq \min\{\mu(a, b), \mu(b, c)\}$ for all $a, b, c \in \mathcal{S}$.
(c) $\mu(a, b) + \mu(b, a) \geq 1$ for all $a, b \in \mathcal{S}$.

Here, $\mu(a, b)$ means the degree that the decision maker likes a than b. We introduce a ranking method of states from a viewpoint of fuzzy preference, which is called a *score ranking function* ([3]). For a fuzzy preference relation μ on $\mathcal{S}$, the following map r on $\mathcal{S}$ is called a score ranking function of states induced by the fuzzy preference relation μ:

$$r(a) = \sum_{b \in \mathcal{S}: b \neq a} \{\mu(a, b) - \mu(b, a)\} \tag{1}$$

for $a \in \mathcal{S}$. If $\mathcal{S}$ is a linearly ordered set by a linear order $\succsim$, then a fuzzy relation μ on $\mathcal{S}$ defined by

$$\mu(a, b) := \begin{cases} 1 & \text{if } a \succsim b \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

is a fuzzy preference relation, and for $a, b \in \mathcal{S}$, it holds that $a \succsim b \iff r(a) \geq r(b)$. Now we consider a subset $\mathcal{C} := \{c^i | i = 1, 2, \cdots, n\}$ of C such that $\mathcal{C}$ has n

elements and a linear order $\succsim$, where $n \geq 2$. Let μ be a fuzzy relation on $\mathcal{C}$ and let r be the score ranking function induced by μ:

$$r(c^i) = \sum_{b \in \mathcal{C}} \{\mu(c^i, b) - \mu(b, c^i)\} = \sum_{j=1}^{n} r^{ij}, \tag{3}$$

where $r^{ij} := \mu(c^i, c^j) - \mu(c^j, c^i)$ $(i, j = 1, 2, \cdots, n)$. Here, the score ranking function r takes values in $[-n+1, n-1]$. By using ranking methods r, we can consistently extend the ranking on $\mathcal{C}$ to one on $\mathcal{C}'$ which has finite elements and satisfies $\mathcal{C} \subset \mathcal{C}' \subset C$. In the next section, we introduce a dynamic model where the number of cognizable states increases with time. Then, we need a scaling of the score ranking function r to normalize its value region which expands with time and the number of elements in $\mathcal{C}$. Since C has n elements, we introduce a *scaling translation* $f_n : [-n+1, n-1] \mapsto [0, 1]$ by $f_n(x) := \frac{x}{2(n-1)} + \frac{1}{2}$ for $x \in [-n+1, n-1]$.

Next we consider about a fuzzy relation and a score ranking function on an extended state $\mathcal{C}'$. Let $\mathcal{A}$ denote the family of fuzzy sets a which are represented as

$$a = \sum_{i=1}^{n} w^i c^i \tag{4}$$

with some weight vector $(w^1, w^2, \cdots, w^n)$ satisfying $\sum_{i=1}^{n} w^i = 1$. In this paper, the system makes decision making by itself. We define a fuzzy relation μ' on an extended set $\mathcal{C} \cup \{a\}$ as follows: $\mu' = \mu$ on $\mathcal{C} \times \mathcal{C}$, $\mu'(a, a) = 1$,

$$\mu'(a, b) := \sum_{i=1}^{n} w^i \mu(c^i, b) \quad \text{and} \quad \mu'(b, a) := \sum_{i=1}^{n} w^i \mu(b, c^i) \tag{5}$$

for $b \in \mathcal{C}$. The corresponding score ranking function r' for the state a is

$$r'(a) = \sum_{b \in \mathcal{C}'} \{\mu'(a, b) - \mu'(b, a)\} = \sum_{b \in \mathcal{C}} \{\mu'(a, b) - \mu'(b, a)\}. \tag{6}$$

Then, for the extended score ranking function r' induced from μ', it holds that

$$r'(a) = \sum_{i=1}^{n} w^i r(c^i) = \sum_{i=1}^{n} \sum_{j=1}^{n} w^i r^{ij}. \tag{7}$$

If the weights take values in $[0, 1]$, then $\mathcal{A}$ is the set of convex linear combinations of $\mathcal{C}$. In this paper, we consider a case where we accept that the weights w^i $(1 = 1, 2, \cdots, n)$ take values not only in the interval $[0, 1]$ but also outside the interval. This extension enables to cognize new objects outside the past knowledge. Thus, it will be possible to learn new objects in much wider scope together with time. We need to deal with fuzzy relation taking values in real numbers outside $[0, 1]$. However, the scaling of the fuzzy relations will be done totally for

the score ranking function at each time when we consider a criterion based on fuzzy preference in Section 4 from estimation results of score ranking in Section 3. We also note that this extension is applicable even when the order $\succsim$ is a partial order on $\mathcal{C}$.

3   Dynamic decision making model and score ranking

In this section, we introduce a dynamic decision making model with fuzzy preferences and a time space $\{0, 1, 2, \cdots, T\}$. Next, we estimate the score ranking function to establish a scaling function. The estimation is needed to define an objective function in the next section. Let $\mathcal{S}_0$ be a subset of C such that $\mathcal{S}_0 := \{c^i | i = 1, 2, \cdots, n\}$ has n elements and a partial order $\succsim$. $\mathcal{S}_0$ is called an *initial state space* and it is given as a *training set* in a learning model. Let μ_0 be a fuzzy preference relation on $\mathcal{S}_0$ such that for for $a, b \in \mathcal{S}_0$

$$\mu_0(a, b) := \begin{cases} 1 \text{ if } a \succsim b \\ 0 \text{ if } a \not\succsim b \\ \gamma \text{ if } a \text{ and } b \text{ are incomparable} \end{cases} \tag{8}$$

with some $\gamma \in [0, 1]$ given by the decision maker. When we deal with actual data, if a fuzzy relation given by the decision maker does not satisfy the transitive condition (b) in the definition of fuzzy preferences, one of the reasonable methods is to apply its transitive closure([3]). Let $t(= 0, 1, 2, \cdots, T)$ be a current time. An action space A_t at time $t(< T)$ is given by a compact set of some Banach space. We deal with two kinds of states. One is *perceived states* outside the system and the other is *cognized states* inside the system since generally there exists some difference between them. The cognized state is computed from the perceived state by approximation computation on the basis of the initial states and the past states. At time t, a current *cognized state* is denoted by s_t. An initial state s_0 is given by an element in $\mathcal{S}_0$. Define a family of states until time t by $\mathcal{S}_t := \mathcal{S}_0 \cup \{s_1, s_2, \cdots, s_t\} = \{c^1, c^2, \cdots, c^n, s_1, s_2, \cdots, s_t\}$ for $t = 1, 2, \cdots, T$. For $t = 0, 1, 2, \cdots, T$, $u_t(\in A_t)$ means an *action* at time t, and $h_t = (s_0, u_0, s_1, u_1, \cdots, s_{t-1}, u_{t-1}, s_t)$ means a *history* with states $s_0, s_1, \cdots, s_t$ and actions $u_0, u_1, \cdots, u_{t-1}$. Then, a *strategy* is a map $\pi_t : \{h_t\} \mapsto A_t$ which is represented as $\pi_t(h_t) = u_t$ for some $u_t \in A_t$. A sequence $\pi = \{\pi_t\}_{t=1}^{T-1}$ of strategies is called a *policy*. Let $\bar{\rho}$ be a nonnegative number. We deal with the case where a current cognized state s_t is represented by a linear combination of the initial states $c^1, c^2, \cdots, c^n$ and the past states $s_1, s_2, \cdots, s_{t-1}$:

$$s_t = \sum_{i=1}^{n} \bar{w}_t^i c^i + \sum_{j=1}^{t-1} \bar{w}_t^{n+j} s_j, \tag{9}$$

for some weight vector $(\bar{w}_t^1, \bar{w}_t^2, \cdots, \bar{w}_t^{n+t-1}) \in \mathbb{R}^{n+t-1}$ satisfying $-\bar{\rho} \leq \bar{w}_t^i \leq 1 + \bar{\rho}$ $(i = 1, 2, \cdots, n + t - 1)$ and $\sum_{i=1}^{n+t-1} \bar{w}_t^i = 1$, where we put

$$\bar{w}_0^i := \begin{cases} 1 \text{ if } s_0 = c^i \\ 0 \text{ if } s_0 \neq c^i \end{cases} \tag{10}$$

for $i = 1, 2, \cdots, n$. The equation (9) means that the current cognized state s_t is understandable from the past states $\mathcal{S}_{t-1} = \{c^1, c^2, \cdots, c^n, s_1, s_2, \cdots, s_{t-1}\}$, which we call an *experience set*. Then, $\bar{\rho}$ is called a *capacity factor* regarding the range of cognizable states. The cognizable range of states becomes bigger as the positive constant $\bar{\rho}$ is taken greater in this model. The range is measured by $\bar{\rho}$ observing the interval $-\bar{\rho} \leq \bar{w}_t^i \leq 1 + \bar{\rho}$. If $\bar{\rho} = 0$ for all $t = 1, 2, \cdots, T$, the system is conservative and the cognizable range of states at any time t is the same as the initial cognizable scope, which is the convex full of $\mathcal{S}_0 = \{c^1, c^2, \cdots, c^n\}$.

Let a *perceived state* at time t by $\tilde{o}_t (\in C)$, which depends on the action u_{t-1} taken at the previous time $t-1$ since the action u_{t-1} affects to the surroundings and the state $\tilde{o}_t$ to be perceived at time t. To determine a cognized state s_t in the form (9) from observed data regarding the perceived state $\tilde{o}_t$, we use fuzzy neural networks: First, we give an input data from the perceived state $\tilde{o}_t$ by $\{(x^1, \alpha^1), (x^2, \alpha^2), \cdots, (x^L, \alpha^L)\} \subset C \times [0, 1]$ such that $\tilde{o}_t(x^l) = \alpha^l$ for $l = 1, 2, \cdots, L$, and next we determine the weight vector $(\bar{w}_t^1, \bar{w}_t^2, \cdots, \bar{w}_t^{n+t-1})$ in (9) so as to minimize the following error between the data and a cognizable value in (9):

$$\sum_{l=1}^{L} \left(\alpha^l - \left(\sum_{i=1}^{n} \bar{w}_t^i c^i + \sum_{j=1}^{t-1} \bar{w}_t^{n+j} s_j \right) (x^l) \right)^2. \tag{11}$$

From the structure of the optimization problem, a fuzzy regression method using neural networks is applicable to (11)([4]).

Let $t(= 1, 2, \cdots, T)$ be a current time. By (9) we define a fuzzy relation μ_t on $\mathcal{S}_t$ by induction on t as follows: $\mu_t := \mu_{t-1}$ on $\mathcal{S}_{t-1} \times \mathcal{S}_{t-1}$, $\mu_t(s_t, s_t) := 1$,

$$\mu_t(s_t, a) := \sum_{i=1}^{n} \bar{w}_t^i \mu_t(c^i, a) + \sum_{j=1}^{t-1} \bar{w}_t^{n+j} \mu_t(s_j, a), \tag{12}$$

$$\mu_t(a, s_t) := \sum_{i=1}^{n} \bar{w}_t^i \mu_t(a, c^i) + \sum_{j=1}^{t-1} \bar{w}_t^{n+j} \mu_t(a, s_j) \tag{13}$$

for $a \in \mathcal{S}_{t-1}$.

To simplify the problem, we introduce a translation of weights. For $i = 1, 2, \cdots, n$, we define a sequence of weights $\{w_{t+1}^i\}_{t=0}^{T}$ inductively by $w_0^i := \bar{w}_0^i$ and

$$w_t^i := \bar{w}_t^i + \sum_{j=0}^{t-1} \bar{w}_t^{n+1+j} w_j^i \tag{14}$$

$(t = 1, 2, \cdots, T)$. Then, we can easily check $\sum_{i=1}^{n} w_t^i = 1$. The computation rule for the extended fuzzy relations at time t is given as followings: For a current time $t(= 0, 1, 2, \cdots, T)$ and an initial state or a past state $a(\in \mathcal{S}_{t-1})$, it holds that

$$\mu_t(s_t, a) = \sum_{i=1}^{n} w_t^i \mu_t(c^i, a) \quad \text{and} \quad \mu_t(a, s_t) = \sum_{i=1}^{n} w_t^i \mu_t(a, c^i). \tag{15}$$

In this paper, we use the sequence of weights $\{w_t^i\}_{t=0}^T$ in (15) rather than the sequence of weights $\{\bar{w}_t^i\}_{t=0}^{n+T-1}$ in (9). The following equation gives a computation rule regarding capacities. Define a sequence of capacities $\{\rho_t\}_{t=1}^T$ by

$$\rho_{t+1} = \rho_t + \bar{\rho}(1 + t + t\rho_t) \qquad (16)$$

for $t = 1, 2, \cdots, T$. Then, it holds that $-\rho_t \leq w_t^i \leq 1 + \rho_t$ for $i = 1, 2, \cdots, n$ and $t = 1, 2, \cdots, T$. The capacity term ρ_t is an increasing function of t, and the increase is corresponding to the facts that the range of cognizable states s_t expands with time t. Here by weights regarding the score $r_t(s_t)$ we obtain the following representation:

$$r_t(s_t) = \sum_{i=1}^n \sum_{j=1}^n w_t^i r^{ij} + \sum_{m=1}^{t-1} \sum_{i=1}^n \sum_{j=1}^n w_t^i w_m^j r^{ij} \qquad (17)$$

for $t = 1, 2, \cdots, T$, where r^{ij} is given by $r^{ij} := \mu_0(c^i, c^j) - \mu_0(c^j, c^i)$, $i, j = 1, 2, \cdots, n$. We put an upper bound of $|r_t(s_t)|$ by $K(n, t) := (n-1)(2\rho_t + 1) + (2\rho_t + 1) \sum_{m=1}^{t-1} (2\rho_m + 1)$.

4 The optimality in dynamic decision making with fuzzy preferences

We consider a decision process on the results of the previous sections. By using fuzzy random variables, we can model the vagueness factors regarding future states where the decision maker can seize and current state where he will be actually confronted with. In this section, by dynamic programming, we discuss an optimality equation in the model where the decision maker is accustomed to his environment.

Now we introduce a scaling function for the score r_t and we define an expected value from the results regarding score ranking in the previous section. For $t = 1, 2, \cdots, T$, we define a scaling function

$$\varphi_t(x) := \frac{x}{2K(n, t)} + \frac{1}{2} \qquad (18)$$

for x satisfying $|x| \leq K(n, t)$. Then we have $\varphi_t(x) \in [0, 1]$. Using the scaling function φ_t, we can take a balance among the score rankng functions r_t ($t = 0, 1, \cdots, T$). Let $t(= 0, 1, 2, \cdots, T)$ be a current time. To derive an optimality equation, we introduce total values $V_t^\pi(h_t)$ at time t by

$$V_t^\pi(h_t) := \sum_{m=t}^T \varphi_m(r_m(\tilde{X}_m^\pi)). \qquad (19)$$

Define the optimal total values $V_t(h_t)$ at time t by

$$V_t(h_t) := \sup_\pi V_t^\pi(h_t). \qquad (20)$$

Then, we obtain the following optimality equation.

Theorem 1. *It holds that*

$$V_t(h_t) = \sup_{\pi}\{\varphi_t(r_t(s_t)) + V_{t+1}(h_t, u_t, \tilde{X}_{t+1}^{\pi})\} \tag{21}$$

for $t = 0, 1, 2, \cdots, T-1$, and $V_T(h_T) = \varphi_T(r_T(s_T))$ at terminal time T.

5 Conclusion

In this paper, we have discussed the followings:

- A method to extend the cognizable range at each step.
- A criterion based on fuzzy preferences in dynamic decision making.
- An optimality equation for this model derived by dynamic programming.

Mathematical models simulating human behavior with his decision making are applicable to problems in various fields.

- For example, robotics, customers' behavior analysis in marketing, linguistic learning modeling and so on.

References

1. Coubon, J.-Cl., Dubois, D., Roy, B.: Autour de l'aide à la décision et de l'intelligence artificielle. Rapport LAFORIA/IBP 94/01 (1994).
2. Fishburn, P.C.: Utility Theory for Decision Making. John Wiley and Sons, New York (1970).
3. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multi-Criteria Decision Support. Kluwer Academic Publishers, Dordrecht Boston London (1994).
4. Ishibuchi, H., Tanaka, H.: Fuzzy regression analysis using neural networks. Fuzzy Sets and Systems **50** (1992) 257-265.
5. Kaminka, G.A., Lima, P.U., Rojas, R. (eds.): RoboCup 2002: Robot Soccer World Cup VI. Lecture Notes in Artificial Intelligence 2752, Springer, Heidelberg (2003).
6. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, London (1995).
7. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944).
8. Newell, A., Simon, H.A.: Human Problems Solving. Prentice-Hall, Englewood, Cliffs (1972).
9. Pomerol, J.-C.: Artificial intelligence and human decision making. European Journal of Operational Research **99** (1997) 3-25.
10. Puri, M.L., Ralescu, D.: The concept of normality for fuzzy random variables. Ann. Prob. **13** (1985) 1373-1379.
11. Simon, H.A.: The New Science of Management Decision. Prentice-Hall, Englewood, Cliffs (1963).
12. Simon, H.A.: The Sciences of the Artificial. MIT Press, Cambridge (1969).
13. Slotine, J.J.E., Li, W.: Applied Nonlinear Control. Prentice-Hall, Englewood, Cliffs (1991).
14. Yoshida, Y.: Continuous-time fuzzy decision processes with discounted rewards. Fuzzy Sets and Systems **139** (2003) 33-348.