On Arbitration for the Bayesian Collective Choice Problem

By

Dao-Zhi Zeng*, Masamitsu Ohnishi**, Toshihide Ibaraki* and Ting Chen***

(Received September 27, 1994)

Abstract

This paper deals with arbitration for the Bayesian collective choice problem. A similar problem is discussed in Myerson (1979) under the assumption that the arbitrator chooses a bargaining solution, derived from the generalized Nash product of Harsanyi and Selten (1972). This paper, however, asserts that arbitration differs from pure bargaining, because an arbitrator behaves so that the fairness-utility function evaluated by himself is maximized. We argue that the functional form of the fairness-utility function is uniquely determined if the arbitrator acts according to some plausible criteria.

1. Introduction

Members of a group (players) jointly select a choice on which their payoffs depend. This is called the *collective choice problem*. Arbitration is very effective in solving this problem, because the players usually cannot reach agreement, and even if they do, the agreement may not be binding. For example, consider the problem of the Prisoner's Dilemma:

		Player 2	
		a_2	b_2
Player 1	a_1	5, 5	−1, 6
	b_1	6, −1	0, 0

Usually the choice (a_1, a_2) is considered to be the best choice, because it is better than the unique Nash equilibrium (b_1, b_2) for both players. But even though two players do reach this agreement, this choice is not self-enforcing, because every

* Department of Applied Mathematics and Physicis, Faculty of Engineering, Kyoto University, Yoshidahon-machi, Sakyo-ku, Kyoto 606-01, JAPAN.

** Department of Business Management, Faculty of Economics, Tohoku University, Kawauchi, Aoba-ku, Sendai 980-77, JAPAN.

*** Institute of Systems Engineering, Huazhong University of Science and Technology, Wuhan, 430074, P. R. CHINA.

player benefits from betraying. In this case, arbitration seems to be the only effective solution.

In the above example, the arbitrator and players have complete information about all the players' preferences and endowments. More generally, however, it may be necessary to conisider a case in which the arbitrator and players do not have complete information. Myerson [4] considers the arbitration for the following *Bayesian collective choice problem*, which will be referred to as *the bargaining problem* in this paper:

$$BP=(C, T_1, T_2, \cdots, T_n, U_1, U_2, \cdots, U_n, P). \tag{1.1}$$

In this notation, individual players are numbered 1, 2, $\cdots$, n. A finite set C is the set of choices available to the players. We suppose that they can use mixed choices, which will be described later. For each player i, T_i is the set of his possible types. For simplicity, we suppose that T_i ($i=1, 2, \cdots, n$) are finite sets. Each type $t_i \in T_i$ completely describes the characteristics of player i: his preferences, beliefs, abilities, and endowments. Let $T=T_1\times T_2\times\cdots\times T_n$; we call an element of T a type vector. Each U_i is a von Neumann-Morgenstern utility function, which measures the payoff $U_i(c, t)$ to player i if the arbitrator chooses c and if $t\in T$ is the true type vector. To the arbitrator, as well as to the players, the information about the true type vector is incomplete, and P is a probability distribution on T such that $P(t)$ is the probability, estimated by the arbitrator, that t is the true type vector. Each player i knows his own type and his estimate about other players is supposed to be consistent with that of the arbitrator, in the sense that his estimate is the conditional distribution of P given his own true type.

Harsanyi and Selten [1] consider a different model of two-person bargaining with incomplete information. They generalize the Nash product from the bargaining of complete information to the bargaining of incomplete information. The maximizer of the generalized Nash product is the unique solution of a bargaining problem that satisfies eight axioms. Following this result, Myerson defines an arbitration solution for (1.1), whose mathematical description will be given later in (2.9).

In our view, the arbitrator should have special power to conduct his own investigation and to obtain different or secret information about the true type vector of players. Therefore, his estimate may not be consistent with those of the players. Furthermore, it may happen that the arbitrator knows all the players' estimates but the players do not know the arbitrator's estimate.

We also claim that arbitration is different from bargaining. The generalized Nash product, which is used in Myerson [4], is derived from eight plausible axioms for bargaining. However, in our opinion, some axioms among those are inappropriate for arbitration. For example, one of the axioms called the " profitability " axiom for bargaining, which states that every player will receive nonnegative *expected* utility with respect to all his possible types, is suitable for the bargaining problem, but not enough for the arbitration problem. If the arbitration result turns out to be worse than no agreement to some players, the arbitrator's reputation will be spoiled after the true type vector is revealed. Therefore, a rational arbitrator will never make the result worse than no agreement for any type vector, not only better than no agreement in the sense of expected value.

Furthermore, although the generalized Nash product can define a unique solution for the model of Harsanyi and Selten [1], it is very difficult to show a similar uniqueness for model (1.1).

To clarify the difference between arbitration and bargaining problems, we introduce arbitrator's fairness-utility function U_a to form a new model. The arbitrator is supposed to choose a solution that maximizes U_a, under certain constraints that restrict the behavior of the arbitrator with incomplete information about the true type vector. The idea that the arbitrator maximizes his fairness-utility function is not strange to us. In the problem of Final-Offer Arbitration and Final-Double-Offer Arbitration ([5]), we suppose that the arbitrator chooses the offer that is closest to the point c^a he most likes. That is equivalent to maximizing $U_a(c, t)=-d(c, c^a)$ (which measures the distance between c and c^a if the type vector is t).

Formally, this paper considers the following model, which is called *the arbitration problem* in this paper:

$$AP=(C, T_1, T_2, \cdots, T_n, P_1, P_2, \cdots, P_n, U_1, U_2, \cdots, U_n, P_a, U_a). \quad (1.2)$$

In contrast with (1.1), this notation includes players' estimates P_i, the arbitrator's estimate P_a and the arbitrator's utility function $U_a(c, P_a)$. Each player i has a probability estimate about the types of other players, which can be expressed by $P_i(t_{-i}|t_i)$[1]. Players' estimates may not be consistent with the estimate P_a of the arbitrator. The domain of function $U_a(c, P_a)$ is $\mathcal{M}(C)\times\mathcal{P}(T)$, where $\mathcal{M}(C)$ is the set of all the mixed choices and $\mathcal{P}(T)$ is the set of probability distributions over T. It represents the arbitrator's evaluation of fairness for choice c if the true type vector of players is estimated according to P_a. Therefore, we call $U_a(c, P_a)$ *(arbitrator's) fairness-utility function.*

[1] For convenience, as many other researchers, we use t_{-i} to denote the vector composed of all the components of t except t_i, i. e., $t_{-i}=(t_1, \cdots, t_{i-1}, t_{i+1}, \cdots, t_n)$. Therefore, sometimes we write $t=(t_{-i}, t_i)$.

In addition to notations (1.1) and (1.2), we also assume the existence of a disagreement choice $c_0 \in C$ to represent the situation in which players fail to reach agreement. For convenience, we suppose that

$$U_i(c_0, t)=0 \text{ for all } i\in\{1, 2, \cdots, n\}. \tag{1.3}$$

The organization of the paper is as follows. In Section 2, we introduce some definitions and further illustrate our model. Section 3 discusses the fairness-utility function $U_a(c, t)$ in the case of complete information. Then we introduce a result of Kaneko and Nakamura [2], which is used to determine how U_a depends on c. Section 4 considers the fairness-utility function $U_a(c, P_a)$ in the case of incomplete information, and claims that the P_a-part of U_a has von Neumann-Morgenstern expected utility repesentation. Based on those assumptions used in Sections 3 and 4, we then show that the functional form of U_a can be uniquely determined. In Section 5, we work on an example to compare our result with that of [4], which is based on the generalized Nash product.

2. Definitions[2] and the Model

In arbitration problem (1.2), the arbitrator determines a choice, pure or mixed, for every possible type vector, and all players obtain the payoffs corresponding to the choice, after the true type vector is revealed. A mixed choice is denoted by

$$\sum_{k=1}^{m} a_k \delta_{c_k},$$

where $a_k \geq 0$, $k=1, 2, \cdots, m$ and $\sum_{k=1}^{m} a_k=1$. In this notation, δ_{c_k} means[3] to choose a pure choice $c_k \in C$ with probability 1 and therefore this mixed choice chooses c_k with probability a_k, for every $k\in\{1, 2, \cdots, n\}$. Every U_i is supposed to be a von Neumann-Morgenstern utility function, and hence we have

$$U_i\left(\sum_{k=1}^{m} a_k\delta_{c_k},\ t\right)=\sum_{k=1}^{m} a_k U_i(c_k, t) \text{ for } i\in\{1, 2, \cdots, n\}. \tag{2.1}$$

The set of all the mixed choices is denoted by $\mathcal{M}(C)$.

Since every player has private information about his own type and the arbitrator

[2] Detailed descriptions for the bargaining problem can be found in [4].

[3] As usual, we identify $c\in C$ with δ_c.

has only incomplete information about the real type vector, a solution to problem (1.2) is a *choice mechanism* π, which is a real-valued function with domain C for every $t \in T$, such that

$$\sum_{c \in C} \pi(c \mid t) = 1 \text{ for all } t, \tag{2.2}$$

$$\pi(c \mid t) \geq 0 \text{ for all } c \in C \text{ and } t \in T.$$

We interpret that for each type vector $t \in T$, choice mechanism π gives the mixed choice

$$\pi(t) = \sum_{c \in C} \pi(c \mid t) \delta_c. \tag{2.3}$$

In real arbitration, after the arbitrator gives such a choice mechanism π, player i reveals his true type t_i as t_i' and all players know the revealed type vector $t' = (t_1', \cdots, t_n')$. This results in choice $\pi(t')$ of (2.3) and leads to utility

$$U_i(\pi(t'), t) = \sum_{c \in C} \pi(c \mid t') U_i(c, t)$$

for player i, where $t = (t_1, \cdots, t_n)$.

Given a choice mechanism π, let

$$Z_i(\pi, t'_i \mid t_i) = \sum_{t_{-i} \in T_{-i}} \sum_{c \in C} P_i(t_{-i} \mid t_i) \pi(c \mid t_{-i}, t'_i) U_i(c, t), \tag{2.4}$$

$$Z_i(\pi \mid t_i) = Z_i(\pi, t_i \mid t_i),$$

where $T_{-i} = T_1 \times \cdots \times T_{i-1} \times T_{i+1} \times \cdots \times T_n$, $(t_{-i}, t'_i) = (t_1, \cdots, t_{i-1}, t'_i, t_{i+1}, \cdots, t_n)$ and $t = (t_{-i}, t_i)$. This $Z_i(\pi, t'_i \mid t_i)$ denotes the conditionally expected utility payoff to player i under choice mechanism π, given that his true type is t_i, but he tells the arbitrator that his type is t'_i when all other players tell the truth. If player i also tells the truth, his payoff is denoted by $Z_i(\pi \mid t_i)$.

Definition 2.1 A mechanism π is *incentive compatible* if it satisfies the following incentive constraints:

$$Z_i(\pi \mid t_i) \geq Z_i(\pi, t'_i \mid t_i) \text{ for all } i \in \{1, 2, \cdots, n\},\ t_i,\ t'_i \in T_i. \tag{2.5}$$

Definition 2.2 mechanism π is (*individually*) *rational* if it satisfies the following participation constraints:

$$Z_i(\pi \mid t_i) \geq 0 \text{ for all } i \in \{1, 2, \cdots, n\}, \tag{2.6}$$

(recall assumption (1.3) about c_0) and is *(individually) strongly rational* if it satisfies:

$$U_i(\pi(t), t) \geq 0 \text{ for all } t \in T \text{ and } i \in \{1, 2, \cdots, n\}. \tag{2.7}$$

Obviously, strong rationality implies rationality.

In a bargaining problem, a meaningful choice mechanism is usually required to be both incentive compatible and rational. However, in an arbitration problem, a choice mechanism is required to be both incentive compatible and strongly rational. Because, if the arbitration result turns out to be worse than no agreement to some players (i.e., $U_i(\pi(t), t) < 0$ for some i) after the true type vector is revealed, the arbitrator's reputation will be spoiled.

As discussed in Section 1, we suppose that the arbitrator evaluates the fairness of every choice $c \in \mathcal{M}(C)$ by his fairness-utility function $U_a(c, P_a)$, whose domain is $\mathcal{M}(C) \times \mathcal{P}(T)$, where $\mathcal{P}(T)$ is the set of all probability distributions over T. Given that the type vector of players is t, we denote the resulting fairness-utility function as $U_a(c, t)$. In our problem, however, the arbitrator has only incomplete information about the player's type vector, and a mechanism π may be fair for some type vectors but may not be fair for other type vectors. To evaluate the overall performance, arbitration with incomplete information can be described by:

$$\text{maximize } OBJ(AP) = U_a(\pi, P_a) \tag{2.8}$$
$$\text{subject to (2.2), (2.5) and (2.7).}$$

The optimal solution π^a of (2.8) is then announced to all players by the arbitrator. After that, the types of the players are declared by themseleves possibly strategically, and, based on them, the arbitration result is determined by the optimal mechanism and told to the players.

Myerson [4] applies the result of Harsanyi and Selten [1] to the bargaining problem BP of (1.1). To solve BP, he forms the following mathematical programming problem:

$$\text{maximize } \prod_{i=1}^{n} \prod_{t_i \in T_i} Z_i(\pi \mid t_i)^{R(t_i)} \tag{2.9}$$
$$\text{subject to (2.2), (2.5) and (2.6),}$$

where each $R(t_i)$ is the marginal distribution of $P(t)$ corresponding to player i of type t_i.

It is evident that the objective function in (2.9) is just a special form of the objective function in (2.8). However, we will introduce some axioms to model more reasonable behavior of the arbitrator, and show that the objective function of (2.9) is excluded from our model (2.8).

3. Fairness-Utility Function with Complete Information

For a given type vector $t \in T$, our arbitrator will evaluate the fairness of a mixed choice $c \in \mathcal{M}(C)$ by his fairness-utility function $U_a(c, t)$. Since fairness can be regarded as a kind of social welfare, when the type vector t is known, it should be natural to assume that the fairness-utility function $U_a(c, t)$ with respect to c will become a social welfare function as studied in Kaneko and Nakamura [2]. If we take this view point, the utility value of a mixed choice cannot be obtained as a convex combination of the utility values of pure choices, as described in Example 1.1 of [2]. In other words, we cannot expect that (2.1) will hold for U_a, i. e., $U_a(c, t)$ is not of the von Neumann and Morgenstern type with respect to the first variable c.

In Section 2, we used symbol $\sum_{k=1}^{m} \alpha_k \delta_{c_k}$ to denote a mixed choice which is composed of pure choices c_k $(k=1, 2, \cdots, m)$. Let us now extend this notation to allow c_k itself to be a mixed choice. Of course, the composed result is still a mixed strategy. According to [2], we call a vector $(U_1, U_2, \cdots, U_n)$ a *profile* if U_i is a utility function of player i over $\mathcal{M}(C)$ for $i=1, 2, \cdots, n$. Given a profile, the arbitrator can determine his fairness-utility function. This functional relation is denoted by W, i. e.,

$$U_a = W(U_1, U_2, \cdots, U_n). \tag{3.1}$$

We suppose that our arbitrator will follow the following four rationality criteria.

Axiom 3.1 (Pareto Efficiency) If $U_i(c', t) \geq U_i(c, t)$ holds for c and $c' \in \mathcal{M}(C)$ and for all $i=1, 2, \cdots, n$, in which strict inequality holds for some i, then $U_a(c', t) > U_a(c, t)$.

Axiom 3.2 (Independence of Irrelevant Alternatives with Neutral Property) Given two profiles $(U_1, U_2, \cdots, U_n)$ and $(U_1, U_2, \cdots, U_n)$, denote arbitrator's fairness-utility functions by $U_a = W(U_1, U_2, \cdots, U_n)$ and $\hat{U}_a = W(\hat{U}_1, \hat{U}_2, \cdots, \hat{U}_n)$, respectively. Let c_1, c_2, c_3, $c_4 \in \mathcal{M}(C)$ and let c_0 be the disagreement choice. Suppose, for all $i=1, 2, \cdots, n$,

$$U_i(\alpha_1\delta_{c_1}+\alpha_2\delta_{c_2}+(1-\alpha_1-\alpha_2)\delta_{c_0},\ \mathrm{t})>U_i(\beta_1\delta_{c_1}+\beta_2\delta_{c_2}+(1-\beta_1-\beta_2)\delta_{c_0},\ t)$$

if and only if

$$\hat{U}_i(\alpha_1\delta_{c_3}+\alpha_2\delta_{c_4}+(1-\alpha_1-\alpha_2)\delta_{c_0},\ \mathrm{t})>\hat{U}_i(\beta_1\delta_{c_3}+\beta_2\delta_{c_4}+(1-\beta_1-\beta_2)\delta_{c_0},\ t)$$

for all probability distributions $(\alpha_1,\ \alpha_2,\ 1-\alpha_1-\alpha_2)$ and $(\beta_1,\ \beta_2,\ 1-\beta_1-\beta_2)$. Then, $U_a(c_1,\ t)>U_a(c_2,\ t)$ if and only if $\hat{U}_a(c_3,\ t)>\hat{U}_a(c_4,\ t)$.

Axiom 3.3 (Player Symmetry) The function W does not change if we interchange any two players.

Axiom 3.4 (Continuity) Let c_1, c_2 and $c_3\in\mathcal{M}(C)$ satisfy $U_a(c_1,\ t)\geq U_a(c_2,\ t)\geq U_a(c_3,\ t)$. Then there exists some $\alpha\in[0,\ 1]$, such that $U_a(\alpha\delta_{c_1}+(1-\alpha)\delta_{c_3})=U_a(c_2,\ t)$.

The Nash social welfare function is defined as follows:

$$\begin{aligned} U_a(c,\ t) &= \sum_{i=1}^{n}\log\left(U_i(c,\ t)-U_i(c_0,\ t)\right) \qquad (3.2)\\ &= \sum_{i=1}^{n}\log U_i(c,\ t). \end{aligned}$$

The last equality is a result of assumption (1.3).

The following conclusion states that the Nash social welfare function is the unique candidate of arbitrator's fairness-utility function if we impose the above four axioms.

Lemma 3.1 (Kaneko and Nakamura [2]) The function W of (3.1) satisfies Axioms 3.1–3.4 if and only if it is the Nash social welfare function.

4. Fairness-Utility Function with Incomplete Information

In Section 3, we discussed the arbitrator's fairness-utility function under complete information. However, in the Bayesian collective choice problem, the arbitrator has only incomplete information about the player's type vector. A given mechanism may be fair for some type vectors (we call them *good type vectors*) but may be unfair for other type vectors (we called them *bad type vectors*). For a mechanism π, the arbitrator prefers the case in which good type vectors appear in large probability. That is to say, our arbitrator has a preference $\succ$ on the set $\mathcal{P}(T)$ of all probability distributions in T.

Von Neumann-Morgenstern established some axioms for the representation of the expected utility in [6], which are refined in Kreps [3]. In order to use their conclusion, the preference $>$ is supposed to conform to the following three axioms:

Axiom 4.1 $>$ must be asymmetric and negatively transitive.

This axiom says that there is no pair P^1, $P^2 \in \mathcal{P}(T)$ such that $P^1 > P^2$ and $P^2 > P^1$, and if $P^1 > P^2$, then for any third element P^3, either $P^1 > P^3$ or $P^3 > P^2$ (or both) holds.

Axiom 4.2 Suppose that P^1 and P^2 are two probability distributions such that $P^1 > P^2$, and that P^3 is some other probability distribution. Then for any $\alpha \in (0, 1)$, $\alpha P^1 + (1-\alpha) P^3 > \alpha P^2 + (1-\alpha) P^3$.

Axiom 4.3 Suppose that P^1, P^2 and P^3 are three probability distributions such that $P^1 > P^2 > P^3$. Then, there exist α, $\beta \in (0, 1)$ such that $\alpha P^1 + (1-\alpha) P^3 > P^2 > \beta P^1 + (1-\beta) P^3$.

Then for a given π, there exists a numerical representation $U_a(\pi, P_a)$ with respect to the second variable, which is in the form of von Neumann-Morgenstern expected utility representation.

Lemma 4.1 (Kreps [3]) Given a choice mechanism π, the arbitrator's expected fairness-utility function is represented by:

$$U_a(\pi, P_a) = \sum_{t \in T} P_a(t) U_a(\pi(t), t),$$

if it satisfies Axioms 4.1-4.3 with respect to its second variable. Here, $P_a(t)$ is the probability of type vector $t \in T$, estimated by the arbitrator, and $\pi(t)$ is defined by (2.3).

Theorem 4.1 Any arbitrator who acts according to Axioms 3.1-3.4 and Axioms 4.1-4.3 has the following fairness-utility function:

$$U_a(\pi, P_a) = \sum_{t \in T} P_a(t) \sum_{i=1}^{n} \log U_i(\pi(t), t). \tag{4.1}$$

Although both (4.1) and the objective function of (2.9) are related to the Nash product in bargaining theory, the two expressions are completely different. If we take the logarithm of the objective function of (2.9), the sum is taken of all players i and their possible types t_i, using marginal distribution $R(t_i)$ of P as weights. This expression is well-defined only when the probability estimates of all players are consistent. On the contrary, in (4.1), the sum is taken of all possible type vectors $t \in T$ using coefficient $P_a(t)$. In this case, it is not required that the probability estimates of all players are consistent. This is a more natural assumption in arbitration with incomplete information, since players usually do not communicate with each other in the presence of arbitration. While we admit that sometimes the arbitrator can declare his estimate after his investigation so that players may revise their estimates to become consistent with each other, we have to consider cases in which some players will not be glad if the arbitrator reveals their private information.

5. Example

In this section, we use the example of [4] to illustrate our results and compare them with those of [4].

Two players must share the cost of a public project that would benefit them both. The project costs \$100, and the two players ask an arbitrator to help them divide the cost. The arbitrator knows that the project would be worth \$90 to player 2 (i. e., player 2 has only one type, denoted by t_2, which is known by everyone), but its value to player 1 would depend on his type. If player 1 is of type t_1^0, then the project is also worth \$90 to him, but if player 1 is of type t_1^1 then the project is worth \$30 to him. Only player 1 knows for sure what his type is. It is supposd that the arbitrator and player 2 share the same probability estimate (although we allow them to have different estimates in our model), which figure that the probability of being type t_1^0 is v and that of being type t_1^1 is $1-v$. In addition, we suppose that the players have utilities U_i which are linear in money value.

This problem can be formally modeled by $C=\{c_0, c_1, c_2\}$, $T=\{t^0=(t_1^0, t_2), t^1=(t_1^1, t_2)\}$. We have $P_a(t^0)=v$, $P_a(t^1)=1-v$, and the utility functions $(U_1(c, t), U_2(c, t))$ of players 1 and 2 are given by:

type vector	c_0	c_1	c_2
$t^0=(t_1^0, t_2)$	(0, 0)	(−10, 90)	(90, −10)
$t^1=(t_1^1, t_2)$	(0, 0)	(−70, 90)	(30, −10)

The choices in C are interpreted as follows: c_0 is the disagreement choice " do not undertake the project ", which gives everyone 0 utility; c_1 is the choice " undertake the project and make player 1 pay for it "; and c_2 is the choice " undertake the project and make player 2 pay for it ". The intermediate financing option between c_1 and c_2 can be represented by the mixed choice, which corresponds to an element in $\mathcal{M}(\{c_0, c_1, c_2\})$.

For notation, let

$$\pi_j^0 = \pi(c_j \mid t^0) \text{ and } \pi_j^1 = \pi(c_j \mid t^1) \text{ for } j = 0,\ 1,\ 2,$$

and we can use a 6-dimensional vector $(\pi_0^0,\ \pi_1^0,\ \pi_2^0;\ \pi_0^1,\ \pi_1^1,\ \pi_2^1)$ to represent a choice mechanism π. The two players are supposed to have von Neumann-Morgenstern utility functions $U_1(c,\ t)$ and $U_2(c,\ t)$.

Suppose that the arbitrator uses the Nash solutions, which are the solutions to problems

$$\begin{aligned}
\text{maximize} \quad & U_1\left(\pi_0^0\delta_{c_0} + \pi_1^0\delta_{c_1} + \pi_2^0\delta_{c_2},\ t^0\right) \cdot U_2\left(\pi_0^0\delta_{c_0} + \pi_1^0\delta_{c_1} + \pi_2^0\delta_{c_2},\ t^0\right) \\
& \left(= (-10\pi_1^0 + 90\pi_2^0) \cdot (-70\pi_1^0 + 30\pi_2^0)\right) \\
\text{subject to} \quad & \pi_0^0 + \pi_1^0 + \pi_2^0 \leq 1, \\
& 0 \leq \pi_0^0 \leq 1,\ 0 \leq \pi_0^1 \leq 0.9,\ 0 \leq \pi_2^0 \leq 0.9
\end{aligned}$$

and

$$\begin{aligned}
\text{maximize} \quad & U_1\left(\pi_0^1\delta_{c_0} + \pi_1^1\delta_{c_1} + \pi_2^1\delta_{c_2},\ t^1\right) \cdot U_2\left(\pi_0^1\delta_{c_0} + \pi_1^1\delta_{c_1} + \pi_2^1\delta_{c_2},\ t^1\right) \\
& \left(= (-70\pi_1^1 + 30\pi_2^1) \cdot (90\pi_1^1 - 10\pi_2^1)\right) \\
\text{subject to} \quad & \pi_0^1 + \pi_1^1 + \pi_2^1 \leq 1, \\
& 0 \leq \pi_0^1 \leq 1,\ 0.1 \leq \pi_1^1 \leq 0.3,\ 0.7 \leq \pi_2^1 \leq 0.9,
\end{aligned}$$

respectively. The optimal mixed choices are $c^a(t^0) = (0,\ 0.5\ 0.5)$ and $c^a(t^1) = (0,\ 0.2,\ 0.8)$.

By (2.4), we have

$$\begin{aligned}
Z_1(\pi,\ t_1^1 \mid t_1^0) &= -10\pi_1^1 + 90\pi_2^1, \\
Z_1(\pi \mid t_1^1) &= -70\pi_1^1 + 30\pi_2^1, \\
Z_1(\pi,\ t_1^0 \mid t_1^1) &= -70\pi_1^0 + 30\pi_2^0, \\
Z_1(\pi \mid t_1^0) &= -10\pi_1^0 + 90\pi_2^0.
\end{aligned}$$

Player 2 has only one possible type, and

$$Z_2(\pi \mid t_2) = v(90\pi_1^0 - 10\pi_2^0) + (1 - v)(90\pi_1^1 - 10\pi_2^1).$$

Therefore, the incentive compatibility holds if and only if:

$$
\begin{aligned}
&-10\pi_0^0+90\pi_2^0\geq -10\pi_1^1+90\pi_2^1, \qquad (5.1)\\
&-70\pi_1^1+30\pi_2^1\geq -70\pi_1^0+30\pi_2^0,\\
&\pi_0^0+\pi_1^0+\pi_2^0=1,\ \pi_0^1+\pi_1^1+\pi_2^1=1,\\
&\pi_i^j\geq 0,\ i=1,\ 2,\ 3,\ j=0,\ 1.
\end{aligned}
$$

For example, it is easy to check that mechanism (0, 0.5, 0.5 ; 0, 0.2, 0.8), which consists of $c^a(t^0)$ and $c^a(t^1)$, is not incentive compatible. In addition, rationality is ensured if π satisfies

$$
\begin{aligned}
&-10\pi_1^0+90\pi_2^0\geq 0, \qquad (5.2)\\
&-70\pi_1^1+30\pi_2^1\geq 0,\\
&v(90\pi_1^0-10\pi_2^0)+(1-v)(90\pi_1^1-10\pi_2^1)\geq 0.
\end{aligned}
$$

Furthermore, adding the following constraints to the first and the second constraints in (5.2), we obtain the strong rationality :

$$
\begin{aligned}
&90\pi_1^0-10\pi_2^0\geq 0, \qquad (5.3)\\
&90\pi_1^1-10\pi_2^1\geq 0.
\end{aligned}
$$

Problem (2.9) of [4] is now described as the following nonlinear programming :

$$
\begin{aligned}
&\text{maximize } Z_1(\pi\,|\,t_1^0)^v\cdot Z_1(\pi\,|\,t_1^1)^{1-v}\cdot Z_2(\pi\,|\,t_2) \qquad (5.4)\\
&\text{subject to (5.1) and (5.2).}
\end{aligned}
$$

The optimal solution is an incentive compatible and rational mechanism

(0, 0.505, 0.495 ; 0.561, 0, 0.439)

for the case $v=0.9$. If the true type vector turns out to be t^1, the utility of player 2 is $90\times 0-10\times 0.439=-4.39<0$, which is worse than no agreement. In our opinion, this is not a good arbitration result, because the arbitrator's name will be spoiled if the type vector is really t^1.

Adding constraint (5.3) to (5.4), the resulting optimal solution, which is incentive compatible and strongly rational, is

$$\pi^*=(0,\ 0.5,\ 0.5\ ;\ 0.5,\ 0.05,\ 0.45) \qquad (5.5)$$

for all $v \geq 0.4$. When $v=0.4$, this result is not still attractive, because although the probability 0.4 of type vector t^0 is smaller than the probability 0.6 of type vector t^1, this result seems fair only for type vector t^0 (recall that the arbitrator's best choices are $c^a(t^0)=(0, 0.5, 0.5)$ and $c^a(t^1)=(0, 0.2, 0.8)$).

Finally, introduce the fairness-utility function U_a of (4.1). We have

$$U_a(\pi, t^0) = \log(-10\pi_1^0 + 90\pi_2^0)(90\pi_1^0 - 10\pi_2^0),$$
$$U_a(\pi, t^1) = \log(-70\pi_1^1 + 30\pi_2^1)(90\pi_1^1 - 10\pi_2^1).$$

And our arbitration result π^a is the optimal solution of the following problem:

$$\begin{aligned} \text{maximize} \quad & v \log(-10\pi_1^0 + 90\pi_2^0)(90\pi_1^0 - 10\pi_2^0) \\ & + (1-v)\log(-70\pi_1^1 + 30\pi_2^1)(90\pi_1^1 - 10\pi_2^1) \\ \text{subject to} \quad & (5.1),\ (5.2) \text{ and } (5.3). \end{aligned}$$

Optimal solutions of this problem are numerically computed for some v:

probability	mechanism	$U_a(\pi^a, t^0)$	$U_a(\pi^a, t^1)$
$v=0.8$	(0, 0.41279, 0.58721 ; 0.29830, 0.14432, 0.55738)	3.18297	1.69090
$v=0.5$	(0, 0.30000, 0.70000 ; 0.12500, 0.18750, 0.68750)	3.07918	1.87506
$v=0.2$	(0, 0.20020, 0.79980 ; 0.00000, 0.20020, 0.79980)	1.71654	1.99999

We find that when v decreases, the fairness in the case of type t^0 strictly decreases while the fairness in the case of type t^1 strictly increases. In contrast, the fairness of (5.5) is measured by $U_a(\pi^*, t^0)=3.20412$, and $U_a(\pi^*, t^1)=-\infty$! Therefore, as claimed in this paper, our results are more acceptable as arbitration results for the present example.

Acknowledgement

The authors are grateful to Professor M. Fukushima of Nara Institute of Science and Technology for his beneficial discussion and other help. The authors also thank Prof. S. Muto of Tohoku University for valuable comments and suggestions, especially reminding us of an important reference [2].

References

[1] Harsanyi, J. C. and Selten, R., A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information, *Management Science*, Vol. 18, No. 5, 1972, pp. 80-106.

[2] Kaneko, M. and Nakamura, K., Nash Social Welfare Function, *Econometrica*, Vol. 47, No. 2, 1979, pp. 423-435.

[3] Kreps, D. M., *Notes on the Theory of Choice*, Westview Press, Boulder, Colorado, 1988.

[4] Myerson, R. B., Incentive Compatibility and the Bargaining Problem, *Econometrica*, Vol. 47, No. 1, 1979, pp. 61-73.

[5] Zeng, D.-Z., Ohnishi, M. and Ibaraki, T., Intrinsic Gap and Final-Double-Offer Arbitration, the 6th IFAC/IFORS/IMARS Symposium on LSSTA, Preprints, Vol. 1, 47-52, Beijing, 1992.

[6] Von Neumann J. and Morgenstern, O., *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, N. J., 1944.