APPLICATIONES MATHEMATICAE
25,4 (1999), pp. 285–294

T. RADZIK (Wrocław)

ON A NEW SOLUTION CONCEPT FOR BARGAINING PROBLEMS

Abstract. The purpose of this paper is to discuss the properties of a new solution of the 2-person bargaining problem as formulated by Nash, the so-called Average Pay-off solution. This solution of a very simple form has a natural interpretation based on the center of gravity of the feasible set, and it is "more sensitive" to changes of feasible sets than any other standard bargaining solution. It satisfies the standard axioms: Pareto-Optimality, Symmetry, Scale Invariance, Continuity and Twisting. Moreover, it satisfies a new desirable axiom, Equal Area Twisting. It is surprising that no standard solution of bargaining problems has this property. The solution considered can be generalized in a very natural and unique way to n-person bargaining problems.

1. Introduction. An *n-person bargaining situation* concerns generally n parties (individuals, players) who have the opportunity to collaborate for a mutual benefit. During a "bargaining process" they try to find a satisfactory solution which, in general, can be difficult to achieve since (most often) when one player gains, at least one of the others must lose. To rationalize that process and to simplify it, many bargaining procedures have been constructed in the literature.

Starting with Nash (1950), it has become customary to formulate an (*n-person*) *bargaining problem* as a pair (S, d), where S (the *feasible set*) is a set in the n-dimensional Euclidean space whose elements are all the possible utility n-tuples that the players can obtain by possible cooperating, and d

1991 *Mathematics Subject Classification*: 90D35, 90D12.

Key words and phrases: bargaining problem, bargaining solution, average pay-off solution, twisting.

The paper was supported by Polish KBN Grant No. 2 P03A 00710/96. Its preliminary version was presented at the conference on Game Theory in Jerusalem, July 1995.

(the so-called *disagreement point*) is a point of S that describes the utility vector for the players when they do not cooperate at all. Nash proposed the procedure of searching for a solution of the problem (i.e. a function that assigns a "fair outcome" $F(S,d)$ in S to any bargaining problem (S,d)) with some natural properties described by axioms. Ever since many other solutions have been constructed in a similar way, for different collections of axioms. Among the most known and applied, apart from the *Nash solution*, we can list the *Kalai–Smorodinsky solution* (Kalai & Smorodinsky (1975)) and the *Egalitarian solution* (see Figure 1). A full review of bargaining solutions is given in a recent paper by Thomson (1994).

The purpose of this paper is to discuss the properties of a new solution for 2-person bargaining problems, the so-called *Average Pay-off solution*. Its form is very simple and it has an interesting interpretation based on the center of gravity of the feasible sets. It seems that this solution has more desirable properties (at least in some situations) in comparison with other standard solutions. It satisfies almost all the standard axioms, and an additional new one, the *Equal Area Twisting axiom* which is a substantial modification of the *Twisting axiom* introduced by Thompson & Myerson (1980). The Equal Area Twisting axiom is very natural and it seems desirable to include it in any collection of "fair axioms". Unfortunately, it is an open problem (presented in the form of a conjecture) to find the smallest collection of axioms that ensures the uniqueness of our solution.

The solution discussed has another desirable feature. Unlike many classical solutions, it can be generalized in a unique and natural way to a solution for n-person bargaining problems.

In a recent paper, Ambarci (1995) introduced the definition of the Average Pay-off solution, but in another setting, without investigating its properties. He studied a wide class of bargaining solutions, the so-called Reference function solutions, corresponding to Reference functions (for the 2-person case with convex feasible sets), and he found an axiomatization of any such solution, provided the reference function is given. It follows that this axiomatization uniquely determines the Average Pay-off solution when the reference function is taken as the center of gravity of the feasible sets. Our approach is completely different. We study several properties of the Average Pay-off solution in detail and discuss a possible axiomatic characterization of this concept.

The organization of this paper is the following: Section 2 introduces all the needed basic definitions and axioms for the 2-person bargaining problem, and includes their discussion. Section 3 gives the construction of the new bargaining solution, and the main theorem is formulated there. The proof of the theorem is presented in Section 4. Section 5 contains some final remarks about possible generalizations to the n-person case.

2. Assumptions and axioms. In this section we rigorously formulate the definition of the bargaining problem and its solutions for the 2-person case, together with their possible properties described by axioms. We also introduce a new *Equal Area Twisting axiom* which plays a basic role in our paper.

Let us consider the 2-person case of a bargaining problem in the standardly simplified form (S, d), for which

(i) the *disagreement point* d is $(0,0) \in \mathbb{R}^2$;

(ii) the *feasible set* S contains point d and is a compact subset of $\mathbb{R}^2_+$ with nonempty interior (here $\mathbb{R}^2_+ = \{(x_1, x_2) : x_1 \ge 0,\ x_2 \ge 0\}$).

(iii) the set S is *comprehensive*; that is, if $x \in S$, $y \in \mathbb{R}^2_+$ and $x \ge y$, then $y \in S$. (By definition, $(x_1, x_2) \ge (y_1, y_2)$ iff $x_1 \ge y_1$ and $x_2 \ge y_2$.)

Let Σ denote the set of all the bargaining problems (S, d) satisfying assumptions (i)–(iii). In the sequel, in view of (i), we identify the pair (S, d) with the feasible set S.

Any function $F : \Sigma \to \mathbb{R}^2_+$ with $F(S) \in S$ is said to be a *bargaining solution*; we write

$$F(S) = (F_1(S), F_2(S)).$$

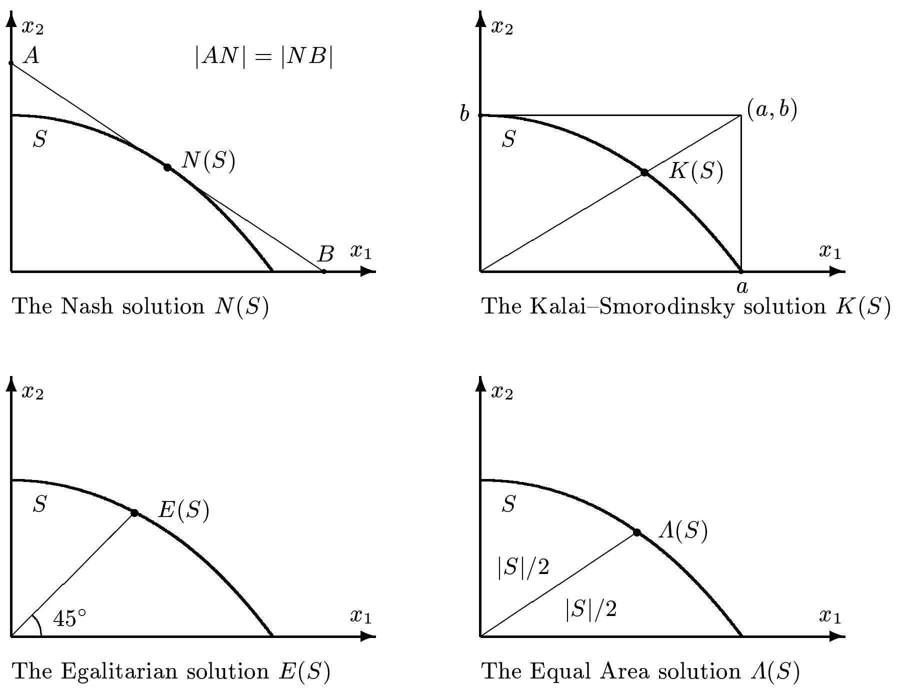


Fig. 1. Four classical bargaining solutions

Remark 2.1. The most important classical solutions, the Nash solution and the Kalai–Smorodinsky solution, are well-defined when the feasible sets S are convex. In our subsequent considerations, we only need a much weaker assumption, the comprehensivity of S.

In the literature, many various standard properties (described by axioms) have been considered for possible solutions of bargaining problems, and different collections of axioms lead to different solutions. We also consider some standard group of axioms, beginning with the following two notions of *strict domination* (in S) and the *Pareto boundary set* PO(S):

$$(y_1, y_2) > (x_1, x_2) :\Leftrightarrow (y_1 > x_1 \text{ and } y_2 > x_2),$$
$$\mathrm{PO}(S) := \{x \in S : \text{there is no } y \in S \text{ with } y \geq x \text{ and } y \neq x\}.$$

Axiom 1 (*Pareto Optimality*). There is no $x \in S$ with $x > F(S)$. If the Pareto boundary set PO(S) of the feasible set S is a connected curve, then $F(S) \in \mathrm{PO}(S)$.

Axiom 2 (*Symmetry*). If S is invariant under transposition of coordinates, then $F_1(S) = F_2(S)$.

Axiom 3 (*Scale invariance*). Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be of the form $T(x, x) = (ax, bx)$ with positive constants a and b. Then $T(F(S)) = F(T(S))$.

Axiom 4 (*Continuity*). If $S^\nu \to S$ in the Hausdorff metric, then $F(S^\nu) \to F(S)$.

Axiom 5 (*Twisting*). Let $i = 1, 2$. If $x \in S' \setminus S$ implies $[x_i \geq F_i(S)$ and $x_j \leq F_j(S)]$ and $x \in S \setminus S'$ implies $[x_i \leq F_i(S)$ and $x_j \geq F_j(S)]$, then $F_i(S') \geq F_i(S)$ (here $j = 2$ if $i = 1$, and $j = 1$ if $i = 2$).

Remark 2.2. All the five axioms are standard in the literature, and almost all the classical solutions (the Nash solution, the Kalai–Smorodinsky solution, the Equal Area solution and others) satisfy them (at least in the class of convex feasible sets S). Our paper is strongly concerned with the last Axiom 5, introduced by Thomson & Myerson (1980). It can be modified in the following way (with $|T|$ denoting the area of T).

Axiom 5′ (*Strong Twisting*). Let $i = 1, 2$ and let $|S \setminus S'| > 0$ or $|S' \setminus S| > 0$. If $x \in S' \setminus S$ implies $[x_i \geq F_i(S)$ and $x_j \leq F_j(S)]$ and $x \in S \setminus S'$ implies $[x_i \leq F_i(S)$ and $x_j \geq F_j(S)]$, then $F_i(S') \geq F_i(S)$ with $F(S') \neq F(S)$ (here $j = 2$ if $i = 1$, and $j = 1$ if $i = 2$).

Remark 2.3. Axiom 5′ is a stronger version of Axiom 5 and seems more reasonable, expressing a "greater sensitivity" of bargaining solutions to some nontrivial changes of feasible sets. Unfortunately, only the Equal Area solution (among all the classical solutions) satisfies it.

The last axiom (illustrated in Figure 2) is new and seems very desirable. It describes some kind of "deeper sensitivity" of the solution than the previous one.

Axiom 6 (*Equal Area Twisting*). Let $i = 1, 2$ and let $a_1, a_2 > 0$ and $|S \setminus S'| = |S' \setminus S| > 0$. If $x \in S' \setminus S$ implies $[x_i \geq a_i$ and $x_j \leq a_j]$ and $x \in S \setminus S'$ implies $[x_i \leq a_i$ and $x_j \geq a_j]$, then $F_i(S') \geq F_i(S)$ with $F(S') \neq F(S)$ (here $j = 2$ if $i = 1$, and $j = 1$ if $i = 2$).

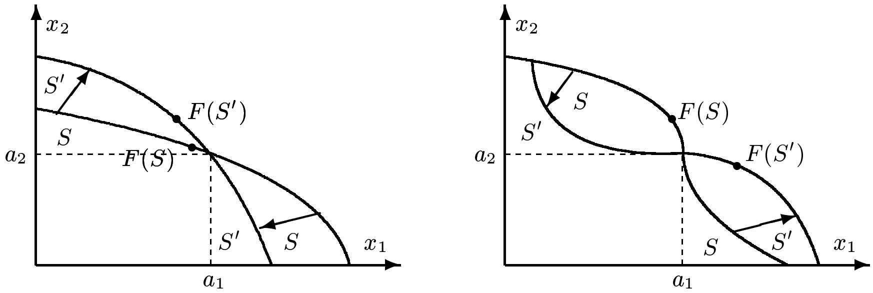


Fig. 2. Two cases of Equal Area Twisting ($|S \setminus S'| = |S' \setminus S|$)

The property expressed by Axiom 6 seems natural and desirable. Let us analyze it when $i = 2$ (the left part of Figure 2). Namely, what will happen when a feasible set S changes to another one S', preserving the assumption of that axiom? It follows that the part $S \setminus S'$ will vanish while the new part $S' \setminus S$ will arise in the feasible set. The areas of $S \setminus S'$ and $S' \setminus S$ are equal, but on the other hand, all pay-offs of Player 1 in $S \setminus S'$ are strictly greater than his pay-offs in $S' \setminus S$. Therefore, it is reasonable to expect that $F_2(S') > F_2(S)$ and $F(S') \neq F(S)$, and this is exactly what Axiom 6 says.

Remark 2.4. One could ask whether Equal Area Twisting implies Strong Twisting, or conversely. It its not difficult to show that Strong Twisting does not imply Equal Area Twisting. Namely, consider the Equal Area solution $\Lambda(S) = (\Lambda_1(S), \Lambda_2(S))$, and consider any "twisting change" of a feasible set S to S' with $|S' \setminus S| = |S \setminus S'|$ such that S coincides with S' in the part $x_1 \geq \Lambda_1(S)$. Then, obviously, $\Lambda(S') = \Lambda(S)$. Therefore, the Equal Area solution does not satisfy the Equal Area Twisting axiom though it satisfies the Strong Twisting axiom. The author does not know whether the converse implication (Equal Area Twisting $\Rightarrow$ Strong Twisting) is valid.

Remark 2.5. The Equal Area Twisting axiom expresses the very reasonable property that a solution of bargaining problems should be "strictly sensitive" to "twisting changes" of feasible sets when their areas remain constant, even if the center of twisting does not coincide with the solution. Unfortunately, no standard solution satisfies this axiom, which is rather surprising.

In view of Remarks 2.3 and 2.5, we come to the following question, basic in our paper.

Question. *Does there exist a solution $F(S)$ of the bargaining problem that satisfies Axioms* 1–4, *Axiom* $5'$ *and Axiom* 6?

The answer is YES! In the next section we give a simple construction of such a solution.

3. The Average Pay-off solution. Define

$$\lambda_1 = \frac{1}{|S|}\int_S x_1\,d\mu \quad \text{and} \quad \lambda_2 = \frac{1}{|S|}\int_S x_2\,d\mu, \tag{1}$$

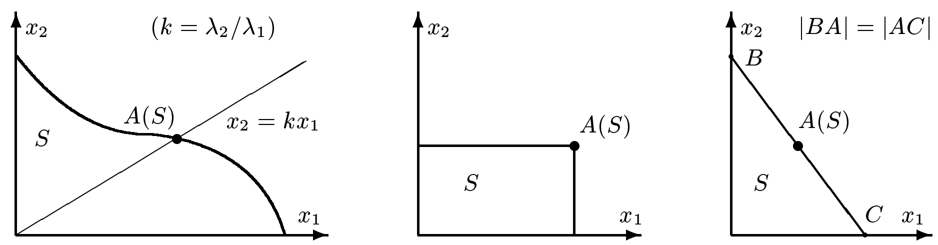
where μ is the Lebesgue measure on $\mathbb{R}^2$.

It is easily seen that λ_1 is the average pay-off of Player 1 over the feasible set S, while λ_2 is the average pay-off of Player 2 over S. Therefore, it seems reasonable to define a new bargaining solution to be "monotonically dependent" on λ_1 and λ_2, in the sense that if the feasible set changes in such a way that λ_1/λ_2 increases, then Player 1 should gain while Player 2 should lose. This principle is basic in the next definition.

Definition. For any $S \in \Sigma$, the *Average Pay-off solution* $A(S)$ is defined as the unique not strictly dominated point of S of the form

$$A(S) = (A_1(S), A_2(S)) = (\lambda_1 u, \lambda_2 u) \tag{2}$$

with maximal u.



(S arbitrary) (S a rectangle) (S a triangle)

Fig. 3. The Average Pay-off solution $A(S)$ in three cases

Remark 3.1. Formulae (1) and (2) allow us to give a simple geometric interpretation of the solution $A(S)$. Namely, $A(S)$ is the intersection point of the Pareto boundary set of the feasible set S and the line going through the disagreement point and the gravity center of S.

Now we are ready to formulate the main theorem of this paper, the proof of which will be given in the next section.

THEOREM. *The Average Pay-off solution $A(S)$ satisfies the Pareto Optimality, Symmetry, Scale Invariance, Continuity, Strong Twisting and Equal Area Twisting axioms.*

REMARK 3.2. The solution $A(S)$ is well defined on bounded comprehensive sets S. This feature allows one to apply it in more general models, where nonconvex feasible sets S can arise, and consequently, neither the Nash solution nor Kalai–Smorodinsky solution nor many other standard solutions can be used.

REMARK 3.3. It follows from the Theorem that $A(S)$ lies between the two end points of the Pareto boundary set $\mathrm{PO}(S)$ if it is a connected curve. For $\mathrm{PO}(S)$ disconnected, $A(S)$ always lies on $\mathrm{PO}(S)$ between the Dictatorial solutions. This follows from the proof of the theorem.

REMARK 3.4. In view of the Theorem, two questions arise: (a) Is $A(S)$ the only solution satisfying the axioms listed there?, and (b) What is the smallest group of axioms that uniquely determines the solution $A(S)$? The author does not know the answers, but it seems that the following conjecture is true.

CONJECTURE. *The Average Pay-off solution $A(S)$ is the unique solution that satisfies the Pareto Optimality, Symmetry, Scale Invariance, Continuity and Equal Area Twisting axioms.*

4. Proof of Theorem. *Pareto Optimality.* Let (a_1, a_2) and (b_1, b_2) be the end points (with a_1 minimal and b_2 maximal) of the Pareto boundary set $\mathrm{PO}(S)$ of a feasible set S in Σ. Without loss of generality, we can assume that

$$0 < a_1 < b_1 \quad \text{and} \quad 0 < b_2 < a_2. \tag{3}$$

Therefore, the set

$$\mathrm{RE} = S \cap \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \le b_2\}$$

is the rectangle $\{(x_1, x_2) : 0 \le x_1 \le b_1,\ 0 \le x_2 \le b_2\}$. Let λ_1 and λ_2 be defined by (1). It is not difficult to see that the Pareto Optimality of $A(S)$ will be proved if we show that

$$b_2/b_1 \le \lambda_2/\lambda_1 \le a_2/a_1. \tag{4}$$

Obviously, it suffices to show the first inequality in (4) since the second is symmetric. It is not difficult to verify that

$$b_1 \int_{\mathrm{RE}} x_2\, d\mu = b_2 \int_{\mathrm{RE}} x_1\, d\mu$$

and, by (3),

$$b_1 \int\limits_{S\setminus \mathrm{RE}} x_2\,d\mu \ge b_1 b_2 \int\limits_{S\setminus \mathrm{RE}} d\mu \ \ge b_2 \int\limits_{S\setminus \mathrm{RE}} x_1\,d\mu.$$

In view of the above, the first inequality in (4) holds since it is equivalent to

$$b_1\Big\{ \int\limits_{\mathrm{RE}} x_2\,d\mu + \int\limits_{S\setminus \mathrm{RE}} x_2\,d\mu\Big\} \ge b_2\Big\{ \int\limits_{\mathrm{RE}} x_1\,d\mu + \int\limits_{S\setminus \mathrm{RE}} x_1\,d\mu\Big\}.$$

Symmetry. Obvious.

Scale Invariance. Consider $T(x_1, x_2) = (ax_1, bx_2) = (x_1', x_2')$ with positive constants a and b, and let $S' = T(S)$. Further, let λ_1 and λ_2 be as in (1), and define λ_1' and λ_2' in the same way with S' replacing S. In view of (2), for maximal u and u' satisfying

$$(5) \qquad (\lambda_1 u, \lambda_2 u) \in S \quad \text{and} \quad (\lambda_1' u', \lambda_2' u') \in S',$$

we have

$$(6) \qquad A(S) = (\lambda_1 u, \lambda_2 u) \quad \text{and} \quad A(S') = (\lambda_1' u', \lambda_2' u').$$

Changing variables in the double integral shows that

$$\lambda_1' = \frac{1}{|S'|}\iint\limits_{S'} x_1'\,dx_1'\,dx_2' = \frac{1}{ab|S|}\iint\limits_{S} ax_1 ab\,dx_1\,dx_2 = a\lambda_1.$$

Similarly, $\lambda_2' = b\lambda_2$. Hence, by (6) and the definition of T and S', we have

$$(7) \qquad T(A(S)) = (a\lambda_1 u, b\lambda_2 u), \qquad A(T(S)) = (a\lambda_1 u', b\lambda_2 u').$$

On the other hand, since u and u' are maximal for which (5) holds, they are maximal satisfying $(a\lambda_1 u, b\lambda_2 u) \in S'$ and $(a\lambda_1 u', b\lambda_2 u') \in S'$. Therefore $u = u'$ and hence, by (7), $T(A(S)) = A(T(S))$.

Continuity. Let $I_C(x)$ denote the indicator function of a set C. Further, let λ_1 and λ_2 be as in (1), and define λ_1^ν and λ_2^ν in the same way with S^ν replacing S. All the feasible sets in Σ are compact (by the assumption in Section 2). Hence, from the convergence $S^\nu \to S$ in the Hausdorff metric, it easily follows that for any $x \in \mathbb{R}^2$, $I_{S^\nu}(x) \to I_S(x)$. Therefore, by the Lebesgue bounded convergence theorem, we get $\lambda_1^\nu \to \lambda_1$ and $\lambda_2^\nu \to \lambda_2$, whence by the definition of the Average Pay-off solution, $A(S^\nu) \to A(S)$. The simple details are omitted.

Strong Twisting. Without loss of generality, we can assume that $x \in S' \setminus S$ implies $[x_1 \ge A_1(S)$ and $x_2 \le A_2(S)]$, and $x \in S \setminus S'$ implies $[x_1 \le A_1(S)$ and $x_2 \ge A_2(S)]$. We have to show that $A_1(S') \ge A_1(S)$ with $A_1(S') \ne A_1(S)$ whenever $|S' \setminus S| > 0$ or $|S \setminus S'| > 0$. One can easily see that it suffices to show that

$$\lambda_2'/\lambda_1' \le \lambda_2/\lambda_1, \tag{8}$$

where λ_i' and λ_i are taken for S' and S, respectively, according to (1).

In view of (1), (2) and the above assumptions, it follows that

$$\begin{aligned}\lambda_2' = \frac{1}{|S'|}\int\limits_{S'} x_2\,d\mu &= \frac{1}{|S'|}\Big\{\int\limits_{S} x_2\,d\mu + \int\limits_{S'\setminus S} x_2\,d\mu - \int\limits_{S\setminus S'} x_2\,d\mu\Big\} \\ &< [\lambda_2|S| + A_2(S)|S'\setminus S| - A_2(S)|S\setminus S'|]/|S'| \\ &= \lambda_2[|S| + u|S'\setminus S| - u|S\setminus S'|]/|S'|,\end{aligned}$$

whence

$$\lambda_2' < \lambda_2[|S| + u|S'\setminus S| - u|S\setminus S'|]/|S'|.$$

In the similar way, we show that

$$\lambda_1' > \lambda_1[|S| + u|S'\setminus S| - u|S\setminus S'|]/|S'|.$$

Therefore (8) holds, and consequently, the solution $A(S)$ satisfies the Strong Twisting axiom.

Equal Area Twisting. We can repeat the proof of the previous part replacing $A_1(S)$ and $A_2(S)$ by a_1 and a_2, respectively, getting

$$\lambda_2' < (\lambda_2[|S| + a_2|S'\setminus S| - a_2|S\setminus S'|]/|S'| = \lambda_2|S|/|S'|,$$

and similarly $\lambda_1' > \lambda_1|S|/|S'|$. Thus, (8) again holds, completing the proof of the Theorem.

5. Generalization to n-dimensional case. The Average Pay-off solution has a natural and simple generalization to a solution of n-person bargaining problems (S,d), where the feasible sets S are bounded comprehensive subsets of the Euclidean space $\mathbb{R}^n$ containing the disagreement point $d = (0,\dots,0)$. Namely, it is natural to define the solution $A(S)$ as the unique undominated point

$$A(S) = (A_1(S),\dots,A_n(S))$$

of S with coordinates proportional to the components of the vector

$$(\lambda_1,\dots,\lambda_n) = \bigg(\frac{1}{|S|}\int\limits_S x_1\,d\mu,\dots,\frac{1}{|S|}\int\limits_S x_n\,d\mu\bigg),$$

where μ is the n-dimensional Lebesgue measure on $\mathbb{R}^n$.

It is an open problem how to generalize the Theorem to the n-person case.

Finally, we sum up the arguments that can encourage the study of the Average Pay-off solution $A(S)$ in more detail:

(1) $A(S)$ has a very simple form;

(2) $A(S)$ satisfies many of the standard axioms;

(3) $A(S)$ is well-defined on comprehensive feasible sets S (not necessarily convex) and therefore it can be applied in more general models;

(4) $A(S)$ is "more sensitive" to changes of feasible sets, much more than any other classical solution;

(5) $A(S)$ can be easily and in a natural way generalized to the n-person case.

Acknowledgements. The author thanks Prof. W. Thompson for several discussions on the topic, and an anonymous referee for his helpful comments.

References

N. Ambarci (1995), *Reference functions and balanced concessions in bargaining*, Canad. J. Econom. 28, 675–682.

E. Kalai and M. Smorodinsky (1975), *Other solutions to Nash's bargaining problems*, Econometrica 43, 513–518.

J. F. Nash (1950), *The bargaining problem*, ibid. 28, 155–162.

W. Thomson (1995), *Cooperative models of bargaining*, in: R. Aumann and S. Hart (eds.), *Handbook of Game Theory with Economic Applications*, Vol. II, North-Holland, 1237–1284.

W. Thomson and R. B. Myerson (1980), *Monotonicity and independence axioms*, Internat. J. Game Theory 9, 37–49.

Tadeusz Radzik
Institute of Mathematics
Technical University of Wrocław
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: tradzik@jegorek.pwr.jgora.pl

Received on 13.6.1996;
revised version on 28.11.1997