44 Cooperative Games with Coalition Structures

with J. Dreze

1. Introduction

A coalition structure in an n-person game is a partition of the set of players. Coalition structures have been used in defining the various solution notions that constitute the bargaining set family, i.e. the various bargaining sets [AUMANN and MASCHLER, 1964; DAVIS and MASCHLER, 1967], the kernel [DAVIS and MASCHLER, 1965] and the nucleolus [SCHMEIDLER, 1969]; in effect, these notions are defined separately for each coalition structure. By contrast, the value [SHAPLEY, 1953], core [GILLIES, 1959] and VON NEUMANN-MORGENSTERN solutions [1944] are not a priori defined with reference to a coalition structure[1]). Moreover, much of the theory that has been developed for the bargaining set family refers to the coalition structure containing the all-player set only.

This contrast between the bargaining set family and the other solution notions is, however, merely a historical accident; it is easy to define the value, core and VON NEUMANN-MORGENSTERN solutions with respect to a given coalition structure. In this paper, we will establish theorems that connect a given solution notion defined for a coalition structure $\mathscr{B}$ with the same solution notion – applied to appropriately defined games on each of the coalitions in the coalition structure. In the case of the kernel, such a theorem has already been proved by MASCHLER and PELEG [1967].

Perhaps the most remarkable aspect of our results is that there is a single function – the function v_x^* defined in (2.4) – that plays the central role in the theorems dealing with 5 out of the 6 solution notions in question (all except the value), though each of these 5 notions is entirely different. Moreover, this function

[1]) The games in "partition function form" of THRALL and LUCAS [1963] are not analogous to games with coalition structures as used in the bargaining set family.

This chapter originally appeared in *International Journal of Game Theory* 4 (1975): 217–237. Reprinted with permission.

enters into the theorems in a completely natural way, which is essentially the same in all 5 cases. This is an extraordinary – and unusual – instance of a game theoretic phenomenon that does not depend on a particular solution notion, but holds "across" a wide class of such notions.

Section 2 collects some basic definitions. In Section 3, we define the value for a game with an arbitrary, given coalition structure $\mathscr{B}$ and relate it to the values defined separately on each element of $\mathscr{B}$. In sections 4 to 8, we present a similar analysis for the nucleolus, the core, the VON NEUMANN-MORGENSTERN solutions, the bargaining set and the kernel. The order in which these solution concepts are reviewed is motivated by convenience of exposition. In section 9, we present a condition under which a payoff vector in the core entails equal treatment for players who are substitutes but belong to different elements of the partition $\mathscr{B}$. In section 10, we show that the core of a game with a coalition structure, when not empty, is equal to the core of the superadditive cover of the game. Section 11 is devoted to two examples, of economic (and academic) interest, in which some of the results of the previous sections are applied. Section 12 is devoted to general discussion. The rationale for studying games with a coalition structure is reviewed there in some detail.

It should be made clear that we have not attempted to be absolutely comprehensive; there are important solution concepts not covered by our analysis (see for example SELTEN [1972]).

The numbering system in this paper is keyed to the numbering of the sections. Thus the theorem in Section 4 is called Theorem 4, and the corollary in Section 5 is called Corollary 5; and there is no Theorem 1 or Theorem 2.

The authors are grateful to MOSHE JUSTMAN, MICHAEL MASCHLER, BEZALEL PELEG, and DIETER SONDERMANN for critical and constructive comments on the material of this paper, including several substantial contributions.

2. Definitions

A game in characteristic function form, or simply a *game*, is a pair (N, v), where N is a finite set (the set of *players*), and v is a real-valued function on the family of subsets of N with $v(\emptyset) = 0$. The function v itself will also be called a game, or *a game on* N. The set of all games on N is denoted G^N; G^N is a EUCLIDean space of dimension $2^{|N|} - 1$, where $|N|$ is the cardinality of N.

A *payoff vector* for N is a real-valued function x on N; it may be thought of as a vector whose coordinates are indexed by the players. If $S \subset N$, write $x(S) = \sum_{i \in S} x(i)$. The set of all payoff vectors for N is denoted E^N. It is sometimes useful to constrain the set of payoff vectors under consideration to a subset X of E^N; we therefore define a *constrained game*[2]) to be a triple (N, v, X), where (N, v) is

[2]) This is by no means a new idea. The core and N-M solutions were first defined in terms of an arbitrary X by GILLIES [1959]; the nucleolus was first defined in this way by SCHMEIDLER [1969].

a game and $X \subset E^N$. When there is no constraint, then $X = E^N$; thus (N,v) may be identified with (N,v,E^N). We will use the term "game" for a constrained game as well; no confusion should result.

A *coalition structure* $\mathscr{B}$ *on* N is a partition of N, the generic element of which will be denoted B_k. A *game with coalition structure* $\mathscr{B}$ is a triple $(N,v,\mathscr{B})$. The analysis of $(N,v,\mathscr{B})$ differs from that of (N,v) in two respects:

(a) Payoff vectors associated with $(N,v,\mathscr{B})$ usually satisfy the conditions $x(B_k) = v(B_k)$ for all k (no side-payments between coalitions); in particular, these conditions are imposed by all the solution concepts considered below.

(b) In addition, the partition $\mathscr{B}$ enters directly into the definition of certain of the solution concepts (namely, the value, the bargaining set and the kernel).

The conditions stated in (a) may easily be replaced by constraints on the set of payoff vectors. Given a game (N,v), define:

$$X_{\mathscr{B}} = \{x \in E^N \colon x(B_k) = v(B_k) \text{ for all } k \text{ and } x_i \geqq v(\{i\}) \text{ for all } i\}. \quad (2.1)$$

As will be seen below, the games $(N,v,\mathscr{B})$ and $(N,v,X_{\mathscr{B}})$ are equivalent from the point of view of some, but not all, solution concepts.

We also find it convenient to define

$$X_k = \{x \in E^{B_k} \colon x(B_k) = v(B_k) \text{ and } x_i \geqq 0 \text{ for all } i \text{ in } B_k\}. \quad (2.2)$$

A 0-*normalized* game is a game for which $v(\{i\}) = 0$ for all i. If (N,v) is a 0-normalized game, then[3]) $X_{\mathscr{B}} = \times_k X_k$. In general, however, there is a distinction between the definition of $X_{\mathscr{B}}$, which includes the conditions $x_i \geqq v(\{i\})$, and the definition of X_k, which includes the conditions $x_i \geqq 0$. (See the remark in section 8.)

In section 3, we use the following definitions.

A *permutation* π of N is a one-one function from N onto itself. For $S \subset N$, write $\pi S = \{\pi i \colon i \in S\}$. If v is a game on N, define a game $\pi_* v$ on N by

$$(\pi_* v)(S) = v(\pi S).$$

Call a coalition structure $\mathscr{B} = (B_1, \ldots, B_p)$ *invariant under* π if $\pi B_j = B_j$ for all j.

Player i is *null* if $v(S \cup \{i\}) = v(S)$ for all $S \subset N$.

In sections 4 through 8, we use the following definitions. Given a vector x in E^N, the *excess* $e(x,S)$ of the coalition S is defined by

$$e(x,S) = v(S) - x(S). \quad (2.3)$$

Three solution concepts are defined in terms of excesses, namely the core, the kernel, and the nucleolus.

Given a game $(N,v,\mathscr{B})$, a payoff vector x, and a coalition B_k in $\mathscr{B}$, define a game (B_k, v_x^*) by

---

[3]) Thus, $X_{\mathscr{B}} \neq \emptyset$ implies $X_k \neq \emptyset$ for all k, a property that does not hold in general.

$$v_x^*(S) = \begin{cases} \max\limits_{T \subset N \setminus B_k} (v(S \cup T) - x(T)), & \text{for } S \subset B_k,\ S \neq \emptyset,\ S \neq B_k \\ v(S), & \text{for } S = \emptyset \text{ or } S = B_k. \end{cases} \tag{2.4}$$

Obviously, $v_x^*(S) \geqq v(S)$ for every x. Note that v_x^* need *not* be 0-normalized, even when v is.

Let $\mathscr{B} = (B_1,\ldots,B_p)$ be a partition of N. The game (N,v) is called *decomposable with partition* $\mathscr{B}$ if for all S,

$$v(S) = \sum_{k=1}^{p} v(S \cap B_k).$$

Finally, let Z be a subset of E^N and B a subset of N. For every y in the projection of Z on $E^{N\setminus B}$, we define the *section of* Z *at* y as $\{w \in E^B\colon (w,y) \in Z\}$. (See Figure 1.)

3. The Shapley Value

Fix N and $\mathscr{B}$. A $\mathscr{B}$*-value* is a function $\varphi_{\mathscr{B}}$ from G^N to E^N – i.e. a function that associates with each game a payoff vector – obeying the following conditions:

Relative efficiency: For all k, $(\varphi_{\mathscr{B}} v)(B_k) = v(B_k)$. (3.1)

Symmetry: For all permutations π of N under which $\mathscr{B}$ is invariant, $(\varphi_{\mathscr{B}}(\pi_* v))(S) = (\varphi_{\mathscr{B}} v)(\pi S)$. (3.2)

Additivity: $\varphi_{\mathscr{B}}(v + w) = \varphi_{\mathscr{B}} v + \varphi_{\mathscr{B}} w$. (3.3)

Null-Player condition: If i is a null-player, then $(\varphi_{\mathscr{B}} v)(i) = 0$. (3.4)

When $\mathscr{B} = \{N\}$, it is known that there is a unique function $\varphi_{\mathscr{B}}$ satisfying (3.1) through (3.4), namely the usual Shapley value of the game [Shapley, 1953]; it will be denoted by φ. This notation will be maintained even for games whose player set differs from N; thus if v is a game with player set M, φv is defined to be $\varphi_{\mathscr{B}} v$, where $\mathscr{B} = \{M\}$.

For each $S \subset N$, denote by $v|S$ the game on S defined for all $T \subset S$ by $(v|S)(T) = v(T)$.

Theorem 3:

Fix N and $\mathscr{B} = (B_1,\ldots,B_p)$. Then there is a unique $\mathscr{B}$-value, and it is given for all $k = 1,\ldots,p$, and all $i \in B_k$, by

$$(\varphi_{\mathscr{B}} v)(i) = (\varphi(v|B_k))(i). \tag{3.5}$$

Remark:

(3.5) asserts that the restriction to B_k of the value $\varphi_{\mathscr{B}}$ for the game (N,v) is the value φ for the game $(B_k, v|B_k)$. In other words, the value of a game with coalition structure $\mathscr{B}$ has the "restriction property": The restriction of the value is the value of the restriction of the game. An important implication of this property is that $\varphi_{\mathscr{B}}$ can be computed by computing separately $\varphi(v|B_k)$ for each k.

Proof:

The operator defined by (3.5) satisfies (3.1) through (3.4), so there is at least one $\mathscr{B}$-value. We must prove that there is only one. For each non-empty $T \subset N$, define the *T-unanimity-game* v_T *by*

$$v_T(S) = \begin{cases} 1 & \text{if } \quad S \supset T \\ 0 & \text{otherwise}. \end{cases}$$

We first show that the games v_T are linearly independent. Indeed, suppose there is a linear relation among them; let T_0 be a set of minimal cardinality such that v_{T_0} appears with non-vanishing coefficient in this linear relation. We then have $v_{T_0} = \sum \alpha_T v_T$, where all the T appearing on the right side are different from T_0 and have cardinality at least that of T_0. Therefore T_0 does not contain any of these T, and hence $1 = v_{T_0}(T_0) = \sum \alpha_T v_T(T_0) = 0$.

This shows that the v_T are linearly independent; since there are $2^{|N|} - 1$ different v_T, and $2^{|N|} - 1$ is the cardinality of G^N, it follows that they form a basis for G^N; therefore every game on N is a linear combination of the games v_T. By the additivity axiom, it then follows that if the $\mathscr{B}$-value is unique on all games of the form αv_T, where α is a constant, then it is unique.

Consider therefore a game of the form αv_T. By (3.4), $(\varphi_{\mathscr{B}}(\alpha v_T))(i) = 0$ whenever $i \notin T$. From (3.2) it follows that if i and j are in T and in the same member B_k of $\mathscr{B}$, then

$$(\varphi_{\mathscr{B}}(\alpha v_T))(i) = (\varphi_{\mathscr{B}}(\alpha v_T))(j).$$

Hence from (3.2) it follows that if $i \in B_k$, then

$$(\varphi_{\mathscr{B}}(\alpha v_T))(i) = \begin{cases} \alpha/|T| & \text{if } \quad T \subset B_k \\ 0 & \text{otherwise}. \end{cases}$$

This determines $\varphi_{\mathscr{B}}(\alpha v_T)$, and so completes the proof.

4. The Nucleolus

Let (N,v,X) be a constrained game. For each $x \in X$, let $\theta(x)$ be a vector in $E^{2^{|N|}}$, the elements of which are the excesses $e(x,S)$ for $S \subset N$, arranged in order of non-increasing magnitude; i.e. $\theta_s(x) \geqq \theta_t(x)$ whenever $t > s$. Write $\theta(y) \geq\cdot\, \theta(x)$ (or $\theta(y) >\cdot\, \theta(x)$) if and only if $\theta(x)$ is not greater (or is smaller) than $\theta(y)$ in the lexicographic order on $E^{2^{|N|}}$. The *nucleolus,* w.r.t. the set X, is then defined by

$$Nu(N,v,X) = \{x \in X \colon \theta(y) \geq\cdot\, \theta(x) \text{ for all } y \in X\}.$$

For a coalition structure $\mathscr{B}$, we define $\mathrm{Nu}(N,v,\mathscr{B}) = \mathrm{Nu}(N,v,X_{\mathscr{B}})$. In particular, when $\mathscr{B} = \{N\}$, we write $\mathrm{Nu}(N,v) = \mathrm{Nu}(N,v,\{N\})$.

When $X \neq \emptyset$, the nucleolus consists of a single element [SCHMEIDLER, 1969; KOHLBERG, 1971]; this element, as well as the set of which it is the only member, will also be called the nucleolus. Thus, like the value and unlike other solution concepts, the nucleolus assigns to each game precisely one payoff vector.

We saw in Section 3 that the restriction to B_k of the value for $(N,v,\mathscr{B})$ is the value for $(B_k, v|B_k)$. Does a similar property hold for the nucleolus? The answer is no, as the following example shows.

Example 4:

Consider the weighted majority game with $|N| = 4$, $w_1 = w_2 = w_3 = 1$, $w_4 = 2$ and

$$v(S) = \begin{cases} 1, & \sum_{i \in S} w_i \geqq 3 \\ 0, & \text{otherwise}. \end{cases}$$

Let $\mathscr{B} = \{(1),(2,3,4)\}$. Then $\mathrm{Nu}(N,v,\mathscr{B}) = (0,\frac{1}{4},\frac{1}{4},\frac{1}{2})$, whereas $\mathrm{Nu}(B_2,v|B_2) = (0,0,1)$.

The reason for this negative answer is easily understood: excesses of coalitions S not included in B_k (e.g. $S = \{1,2,3\}$ in example 4) may play a crucial role in determining the payoff vector $x|B_k \in X_k$, when x is the nucleolus[4]). The characteristic function v_x^* was defined in (2.4) in a way which captures the influence on $x|B_k$ of coalitions not included in B_k, when x is the nucleolus. (Reminder: knowledge of x is required to compute v_x^*).

Theorem 4:

Let (N,v) *be a 0-normalized game, and let* $x = \mathrm{Nu}(N,v,\mathscr{B})$. *Then* $\mathrm{Nu}(N,v,\mathscr{B})|B_k = \mathrm{Nu}(B_k,v_x^*,X_k)$.

Proof[5]):

For $S \subset B_k$ and $y \in B_k$, let

$$e^*(y,S) = v_x^*(S) - y(S)$$

and let $\theta^*(y)$ be the vector of these $2^{|B_k|}$ excesses arranged in non-increasing order. Let $x^* = x|B_k$, and let y^* in X be different from x^*. We show that $\theta^*(y^*) >\cdot\, \theta^*(x^*)$, from which it follows that $x^* = \mathrm{Nu}(B_k,v_x^*,X_k)$.

Define $y \in X_{\mathscr{B}}$ by

$$y_i = \begin{cases} y_i^* & \text{for } i \in B_k \\ x_i & \text{for } i \in N \backslash B_k. \end{cases}$$

Set

$$\alpha = \max\{e(x,R)\colon R \subset N,\ e(x,R) \neq e(y,R)\}. \tag{4.1}$$

---

[4]) In this example, $v(1,2,3) = 1$, $(v|B_2)(2,3) = 0$
$v(1,4) \quad = 1$ $(v|B_2)(4) = 0$.

The nucleolus of $(N,v,\mathscr{B})$ is determined by the excesses of the coalitions $(1,2,3)$ and $(1,4)$; the nucleolus of $(B_2,v|B_2)$ is determined by the excesses of the coalitions (2), (3), $(2,4)$ and $(3,4)$.

[5]) This proof is due to M. JUSTMAN. We are thankful for his permission to use it here.

Let there be q coordinates in $\theta(x)$ larger than α, and r coordinates equal to α. For $\varepsilon > 0$ sufficiently small, the first q coordinates of $\theta((1-\varepsilon)x + \varepsilon y)$ equal the corresponding coordinates of $\theta(x)$. Suppose that for all R with $e(x,R) = \alpha$ we have

$$e(y,R) \leqq e(x,R). \tag{4.2}$$

By the definition of α there is then at least one such R with $e(y,R) < e(x,R)$. Hence for $\varepsilon > 0$ sufficiently small, the $(q+1)$-st through $(q+r)$-th coordinates of $\theta((1-\varepsilon)x+\varepsilon y)$ are all $\leqq \alpha$, and at least one of them is $< \alpha$. Hence $\theta(x) >\cdot\, \theta((1-\varepsilon)x + \varepsilon y)$, contradicting $x = \mathrm{Nu}\,(N,v,\mathscr{B})$. Hence (4.2) is false, i.e. there is at least one coalition – call it U – with

$$e(x,U) = \alpha \quad \text{and} \quad e(y,U) > e(x,U). \tag{4.3}$$

From (4.3) and $x\,|\,N\backslash B_k = y\,|\,N\backslash B_k$ it follows that

$$(x^* - y^*)(U \cap B_k) = (x-y)(U) = e(y,U) - e(x,U) > 0. \tag{4.4}$$

Now let $(S_1,\ldots,S_l)$ be an ordering of the subsets of B_k so that

$$\theta^*(x^*) = (e^*(x^*,S_1),\ldots,e^*(x^*,S_l))$$

and if $U \cap B_k = S_p$, then

$$i < p \Rightarrow e^*(x^*,S_i) > e^*(x^*,S_p). \tag{4.5}$$

It is easy to see that if a vector z' is obtained from a vector z by arranging the coordinates of z in non-increasing order, then $z' \geq\cdot\, z$. Hence

$$\begin{aligned}\theta^*(y^*) - \theta^*(x^*) &\geq\cdot\, (e^*(y^*,S_1),\ldots,e^*(y^*,S_l)) - \theta^*(x^*)\\ &= ((x^*-y^*)(S_1),\ldots,(x^*-y^*)(S_p),\ldots,(x^*-y^*)(S_l)).\end{aligned} \tag{4.6}$$

Now if $i < p$ and $\emptyset \neq S_i \neq B_k$, then from (4.5) and $S_p = U \cap B_k$ it follows that for some $T_i \subset N\backslash B_k$, we have

$$\begin{aligned}e(x,S_i \cup T_i) &= (v(S_i \cup T_i) - x(T_i)) - x(S_i)\\ &> \max\{v(S_p \cup T) - x(T) : T \subset N\backslash B_k\} - x(S_p)\\ &\geqq v(U) - x(U\backslash B_k) - x(U \cap B_k) = e(x,U).\end{aligned}$$

Hence by (4.3) and (4.1), $e(x,S_i \cup T_i) = e(y,S_i \cup T_i)$. From this and $x\,|\,N\backslash B_k = y\,|\,N\backslash B_k$ it follows that $x(S_i) = y(S_i)$, i.e.

$$x^*(S_i) = y^*(S_i) \tag{4.7}$$

in this case. If $S_i = \emptyset$ or $S_i = B_k$, then (4.7) holds trivially. Hence (4.7) holds for all $i < p$. But by (4.4) and $S_p = U \cap B_k$, we have

$$x^*(S_p) - y^*(S_p) > 0.$$

Hence by (4.6),

$$\theta^*(y^*) - \theta^*(x^*) >\cdot 0,$$

so that $\theta^*(y^*) >\cdot \theta^*(x^*)$, as was to be proved.

Corollary 4:

Let (N,v) be a 0-normalized game, decomposable with partition $\mathscr{B} = (B_1,\dots,B_p)$. Then $\mathrm{Nu}(N,v,\mathscr{B}) = \bigtimes_{k=1}^{p} \mathrm{Nu}(B_k,v\,|\,B_k,X_k)$.

Proof:

By theorem 4, $\mathrm{Nu}(N,v,\mathscr{B}) = \bigtimes_{k=1}^{p} \mathrm{Nu}(B_k,v_x^*,X_k)$. Because (N,v) is decomposable, we have

$$v_x^*(S) = \max_{T\subset N\setminus B_k} \{v(S) + v(T) - x(T)\} = v(S) + \max_{T\subset N\setminus B_k} \{v(T) - x(T)\}.$$

It follows that $\mathrm{Nu}(B_k,v_x^*,X_k) = \mathrm{Nu}(B_k,v\,|\,B_k,X_k)$.

Remark:

A similar result holds for the SHAPLEY value; but in that case, it holds for all games, not only decomposable games (see Theorem 3).

5. The Core

The *core* of the game (N,v,X) is defined by

$$\mathrm{Co}(N,v,X) = \{x\in X\colon e(x,S) \leqq 0 \text{ for all } S\subset N\}.$$

For a coalition structure $\mathscr{B}$, we define $\mathrm{Co}(N,v,\mathscr{B}) = \mathrm{Co}(N,v,X_{\mathscr{B}})$. In particular, when $\mathscr{B} = \{N\}$, we write $\mathrm{Co}(N,v) = \mathrm{Co}(N,v,\{N\})$.

The core does not have the uniqueness property of the nucleolus. Accordingly, it could not have the "restriction property" of the value. But one could raise questions such as the following:

(i) Does $x\in \mathrm{Co}(N,v,\mathscr{B})$ imply $x\,|\,B_k\in \mathrm{Co}(B_k,v\,|\,B_k)$?

(ii) Does $y\in \mathrm{Co}(B_k,v\,|\,B_k)$ imply $y = x\,|\,B_k$ for some $x\in \mathrm{Co}(N,v,\mathscr{B})$?

The answer to question (i) is positive, but the answer to question (ii) is negative. Indeed, in example 4, $\mathrm{Co}(N,v,\mathscr{B}) = \emptyset$, whereas $\mathrm{Co}(B_2,v\,|\,B_2) = (0,0,1)$.

The definition of the characteristic function v_x^* in (2.4) is again relevant, in relating $\mathrm{Co}(N,v,\mathscr{B})$ to the cores of appropriately defined games on the B_k's.

Theorem 5:

Let (N,v) be a 0-normalized game, and let $x\in \mathrm{Co}(N,v,\mathscr{B})$. Then the section of $\mathrm{Co}(N,v,\mathscr{B})$ at $x\,|\,N\setminus B_k$ is $\mathrm{Co}(B_k,v_x^*,X_k)$.

Remarks:

1. The conclusion is perhaps most easily understood with the help of Figure 1. Thus, for all $x\in \mathrm{Co}(N,v,\mathscr{B})$, the section of the core at $x\,|\,N\setminus B_k$ defines the core

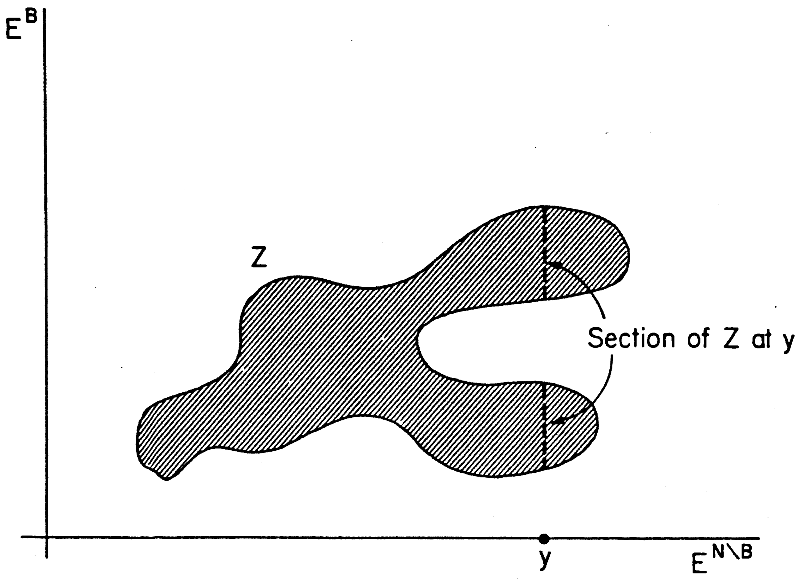

Fig. 1

of a game on B_k; however, the relevant characteristic function is not $v|B_k$, but v_x^*, which depends upon x.

2. When X consists of a single element, then the section of X at $x|N\backslash B_k$ is $x|B_k$. Our result for the nucleolus is thus of the same form as our result for the core, i.e., the conclusion in Theorem 4 could be stated in the following form: if $x \in \mathrm{Nu}(N,v,\mathscr{B})$, then the section of $\mathrm{Nu}(N,v,\mathscr{B})$ at $x|N\backslash B_k$ is $\mathrm{Nu}(B_k,v_x^*,X_k)$.
3. The conclusion does *not* include the property: $x|B_k \in \mathrm{Co}(B_k,v_x^*,X_k)$ implies $x \in \mathrm{Co}(N,v,\mathscr{B})$. Indeed $\mathrm{Co}(B_k,v_x^*,X_k)$ may well be non-empty, when $x|N\backslash B_k$ does not belong to the projection of $\mathrm{Co}(N,v,\mathscr{B})$ on $E^{N\backslash B_k}$.

Proof:

An alternative statement of the conclusion, in terms of which the proof will be presented, is:

$$x \in \mathrm{Co}(N,v,\mathscr{B}) \text{ implies:} \tag{5.1}$$
$$\text{(a) } x|B_k \in \mathrm{Co}(B_k,v_x^*,X_k);$$
$$\text{(b) } z \in \mathrm{Co}(N,v,\mathscr{B}) \text{ for all } z \in E^N \text{ for which}$$
$$z|N\backslash B_k = x|N\backslash B_k \text{ and } z|B_k \in \mathrm{Co}(B_k,v_x^*,X_k).$$

(a) From $x \in \mathrm{Co}(N,v,\mathscr{B})$, it follows that $v(S \cup T) - x(S) - x(T) \leqq 0$ for all $S \subset B_k$, $T \subset N\backslash B_k$. Hence $v_x^*(S) - x(S) \leqq 0$ for all $S \subset B_k$, and so $x|B_k \in \mathrm{Co}(B_k,v_x^*,X_k)$.

(b) From $z|B_k \in \mathrm{Co}(B_k,v_x^*,X_k)$, it follows that $v_x^*(S) - z(S) \leqq 0$ for all $S \subset B_k$. Hence

$$v(S \cup T) - z(S) - x(T) \leqq 0 \quad \text{for all} \quad S \subset B_k \quad \text{with} \quad S \neq \emptyset, \quad \text{and for all} \quad T \subset N\backslash B_k. \tag{5.2}$$

For the case $S = \emptyset$, we have

$$v(T) - x(T) \leqq 0 \quad \text{for all} \quad T \subset N \backslash B_k, \tag{5.3}$$

since $x \in \mathrm{Co}(N,v,\mathscr{B})$. From (5.2), (5.3), and $z | N\backslash B_k = x | N\backslash B_k$, it then follows that $v(S \cup T) - z(S \cup T) \leqq 0$ for all $S \subset B_k$ and $T \subset N\backslash B_k$, and so $z \in \mathrm{Co}(N,v,\mathscr{B})$.

Corollary 5:

Let (N,v) be a 0-normalized game, decomposable with partition $\mathscr{B} = (B_1,\dots,B_p)$. Then $\mathrm{Co}(N,v,\mathscr{B}) = \bigtimes_{k=1}^{p} \mathrm{Co}(B_k, X_k)$.

This result was proved by MASCHLER, PELEG, and SHAPLEY [1972], as Lemma 2.9. It also follows from Theorem 5 and decomposability upon noting that, when $x \in \mathrm{Co}(N,v,\mathscr{B})$,

$$v_x^*(S) = v(S) + \max_{T \subset N\backslash B_k} \{v(T) - x(T)\} = v(S).$$

6. VON NEUMANN-MORGENSTERN Solutions

The VON NEUMANN-MORGENSTERN solutions rely on the notion of "domination". As before, let X be a compact convex subset of E^N, and let (N,v) be a game. Let x and y be in X. Then x *dominates* y *with respect to* the coalition T if and only if $x_i > y_i$ for all $i \in T$, and $v(T) \geqq x(T)$. One then writes $x \succ_T y$. Next, x *dominates* y if there exists a $T \subset N$ for which $x \succ_T y$; one then writes $x \succ y$. A VON NEUMANN-MORGENSTERN *solution of* (N,v,X) is a set $Q \subset X$ such that

(i) there do not exist $x, x' \in Q$ with $x \succ x'$ (internal consistency);

(ii) for every $x' \in X\backslash Q$, there exists an $x \in Q$ such that $x \succ x'$ (external domination).

If $\mathscr{B}$ is a coalition structure, then a VON NEUMANN-MORGENSTERN *solution of* $(N,v,\mathscr{B})$ is a VON NEUMANN-MORGENSTERN solution of $(N,v,X_{\mathscr{B}})$.

Theorem 6.

Let (N,v) be a 0-normalized game, let Q be a VON NEUMANN-MORGENSTERN *solution of $(N,v,\mathscr{B})$, and let $x \in Q$. Then the section of Q at $x|N\backslash B_k$ is a* VON NEUMANN-MORGENSTERN *solution of (B_k, v_x^*, X_k).*

Proof:

Internal consistency and external domination in (B_k, v_x^*, X_k) follow immediately from the properties of Q and the definition of sections upon remembering that $x(B_k) = v(B_k)$ for all $x \in X$.

Thus the relationship of VON NEUMANN-MORGENSTERN solutions for $(N,v,\mathscr{B})$ to VON NEUMANN-MORGENSTERN solutions for (B_k, v_x^*, X_k) is identical to that for the core; the only difference is that a game may have many VON NEUMANN-MORGENSTERN solutions, whereas it has one core.

7. The Bargaining Set

We turn next to the bargaining set, which is defined in terms of the notion of "objection". Let (N,v,X) be a game, $\mathscr{B}$ a coalition structure. If i and j are elements of B_k, and $x \in X$, an *objection* of i against j at x consists of a payoff vector x' in E^N and a coalition $S \subset N$ such that $i \in S$, $j \notin S$, $x'_i > x_i$, $x'_l \geqq x_l$ for all $l \in S$ and $x'(S) \leqq v(S)$. A *counterobjection* of j to such an objection consists of a payoff vector x'' in E^N and a coalition $T \subset N$ such that $i \notin T$, $j \in T$, $x''_l \geqq x_l$ for all $l \in T \backslash S$, $x''_l \geqq x'_l$ for all $l \in T \cap S$ and $x''(T) \leqq v(T)$. The *bargaining set* $M(N,v,X,\mathscr{B})$ is the set of payoff vectors $x \in X$ such that for all k and all $i,j \in B_k$, j has a counterobjection to every objection of i against j at x. Define[6])

$$M(N,v,\mathscr{B}) = M(N,v,X_{\mathscr{B}},\mathscr{B}) \quad \text{and} \quad M(N,v,X) = M(N,v,X,\{N\}).$$

The following theorem is due to B. PELEG; we gratefully acknowledge his permission for publishing it here.

Theorem 7:

Let (N,v) be a 0-normalized game, and let $x \in M(N,v,\mathscr{B})$. Then the section of $M(N,v,\mathscr{B})$ at $x|N\backslash B_k$ is included in $M(B_k,v_x^*,X_k)$.

Proof:

For $x \in M(N,v,\mathscr{B})$ let (y,S) (with $y \in E^{B_k}$ and $S \subset B_k$) be an objection of i against j at $x|B_k$ in the game (B_k,v_x^*,X_k). Remember that $v_x^*(S) = v(S \cup T) - x(T)$ for an appropriate $T \subset N\backslash B_k$, so that $j \notin T$. Then (y,S) may be extended into an objection $(z, S \cup T)$ of i against j at x in the game $(N,v,\mathscr{B})$, where $z \in E^N$ is given by $z|B_k = y$, $z|N\backslash B_k = x|N\backslash B_k$. Because $x \in M(N,v,\mathscr{B})$, there exists a counterobjection (z',T') of j to i's objection $(z, S \cup T)$ in the game $(N,v,\mathscr{B})$. It is readily verified that $(z'|B_k, T' \cap B_k)$ defines a counterobjection of j to i's objection (y,S) in the game (B_k,v_x^*,X_k).

Indeed, $i \notin T' \cap B_k$, $j \in T' \cap B_k$, $z'_l \geqq x_l$ for all $l \in T'\backslash(S \cup T) \supset (T' \cap B_k)\backslash S$, $z'_l \geqq y_l$ for all $l \in (T' \cap B_k) \cap S$ and $z'(T' \cap B_k) \leqq v_x^*(T' \cap B_k)$. The last inequality follows from

$$\begin{aligned} v_x^*(T' \cap B_k) &\geqq v(T') - x(T'\backslash B_k) \geqq z'(T') - x(T'\backslash B_k) \\ &\geqq z'(T') - z'(T'\backslash B_k) = z'(T' \cap B_k). \end{aligned}$$

Remark:

The conclusion in Theorem 7 corresponds to the conclusion in Theorem 5, except for the fact that the equality in theorem 5 becomes an *inclusion* here. The reverse inclusion is false, as can be verified by means of Example 4. Indeed, in that example, the bargaining set for $(N,v,\mathscr{B})$ is defined by the condition $1/4 \geqq x_2 = x_3 \geqq 1/5$; whereas the bargaining set for (B_2,v_x^*,X_2) is defined by the condition $1/4 \geqq x_2 = x_3 \geqq 0$.

[6]) $M(N,v,\mathscr{B})$ is denoted $M_1^{(i)}(N,v,\mathscr{B})$ by AUMANN and MASCHLER [1964], and by DAVIS and MASCHLER [1967].

8. The Kernel

Finally, for the sake of completeness, we quote an earlier result of MASCHLER and PELEG [1967] regarding the kernel. It is of the same form as our result for the core.

Let $i,j \in B_k$, $x \in E^N$. Define $\delta_{ij}(x)$ to be the maximal excess, with respect to x, of a coalition containing i but not j; i.e.,

$$\delta_{ij}(x) = \max \{e(x,S): S \subset N, i \in S, j \notin S\}.$$

Define

$$K(N,v,X,\mathscr{B}) = \{x \in X: (\forall K)(\forall i,j \in B_k)(\delta_{ij}(x) \geqq \delta_{ji}(x) \quad \text{or} \quad x_i = v(\{i\}))\},$$

$$K(N,v,\mathscr{B}) = K(N,v,X_{\mathscr{B}},\mathscr{B}), \quad \text{and} \quad K(N,v,X) = K(N,v,X,\{N\}).$$

$K(N,v,\mathscr{B})$ is called the *kernel* of $(N,v,\mathscr{B})$.

Theorem 8 (MASCHLER and PELEG [1967, Theorem 2.9]):

Let (N,v) be a 0-normalized game, and let $x \in K(N,v,\mathscr{B})$. Then the section of $K(N,v,\mathscr{B})$ at $x|N\backslash B_k$ is $K(B_k,v_x^*,X_k)$.

Remark:

The definition of v_x^* is due to MASCHLER and PELEG [1967, p. 599], who call $K(B_k,v_x^*,X_k)$ the *pseudo-kernel* of (B_k,v_x^*). The kernel of (B_k,v_x^*) would require, in addition, that $x_i \geqq v_x^*(\{i\})$. One could define similarly the "pseudo-core" and "pseudo-nucleolus" of (B_k,v_x^*), to be, respectively, the core and the nucleolus of (B_k,v_x^*,X_k).

Corollary 8:

Let (N,v) be a 0-normalized game, decomposable with partition $\mathscr{B} = (B_1,\dots B_p)$. Then $K(N,v,\mathscr{B}) = \mathop{\times}\limits_{k=1}^{p} K(B_k,v|B_k,X_k)$.

Proof:

By decomposability, $v_x^*(S) = v(S) + \max\limits_{T \subset N\backslash B_k} \{v(T) - x(T)\}$. Hence

$$\max_{\substack{S \subset B_k \\ i \in S \\ j \notin S}} \{v_x^*(S) - x(S)\} - \max_{\substack{S' \subset B_k \\ i \notin S' \\ j \in S'}} \{v_x^*(S') - x(S')\}$$

$$= \max_{\substack{S \subset B_k \\ i \in S \\ j \notin S}} \{v(S) - x(S)\} - \max_{\substack{S' \subset B_k \\ i \notin S' \\ j \in S'}} \{v(S') - x(S')\}$$

and so $K(B_k,v_x^*,X_k) = K(B_k,v|B_k,X_k)$. The corollary then follows from Theorem 8.

9. Equal Treatment

Two players i and j are called *substitutes* if $v(S \cup i) = v(S \cup j)$ for all $S \subset N$ such that $i \notin S, j \notin S$. Whenever a solution concept imposes that substitutes receive the same payoff, that solution concept is said to have the *equal treatment property*. The SHAPLEY value, kernel and nucleolus of (N, v) have that property. Theorems 4 and 8 imply that the kernel and nucleolus of $(N, v, \mathscr{B})$ impose equal treatment for substitutes belonging to the same element B_k of the partition $\mathscr{B}$. Example 4 shows that the kernel and the nucleolus do *not* impose equal treatment for substitutes who belong to different elements of $\mathscr{B}$. On the other hand, it is well known that the core does not impose equal treatment for substitutes belonging to the same element[7]) of $\mathscr{B}$.

The following theorem shows that the core of $(N, v, \mathscr{B})$ imposes equal treatment for substitutes who belong to *different* elements of the partition $\mathscr{B}$. Because $\mathrm{Nu} \in \mathrm{Co}$ and $K \cap \mathrm{Co} \neq \emptyset$ whenever $\mathrm{Co} \neq \emptyset$, the condition has obvious implications for the nucleolus and the kernel as well.

Theorem 9:

Let $x \in \mathrm{Co}(N, v, \mathscr{B})$. *If* i *and* j *are substitutes in* (N, v), $i \in B_k$, *and* $j \notin B_k$, *then* $x_i = x_j$.

Proof:

We have

$$0 \geqq v(\{i\} \cup N \backslash B_k \backslash \{j\}) - x(\{i\} \cup N \backslash B_k \backslash \{j\}) =$$
$$= v(N \backslash B_k) - \{x(N \backslash B_k) - x_j + x_i\} = x_j - x_i > 0 .$$

Similarly $x_i - x_j \leqq 0$, and hence $x_i - x_j = 0$.

10. The Superadditive Cover

A game v on N is called *superadditive* if $S \cap T = \emptyset$ implies $v(S \cup T) \geqq v(S) + v(T)$. The *superadditive cover* of a game v is the game $\hat{v}$ defined by

$$\hat{v}(S) = \max \left\{ \sum_{i=1}^{p} v(S_i) : (S_1, \ldots, S_p) \text{ is a partition of } S \right\} .$$

Note that the superadditive cover is itself superadditive. In fact, if one defines a relation $\geqq$ between games on N by $v \geqq w$ if: $v(S) \geqq w(S)$ for all S, then 0 is the minimal superadditive game that is $\geqq v$.

Theorem 10:

If $Co(N, v, \mathscr{B}) \neq \emptyset$, *then* $\mathrm{Co}(N, v, \mathscr{B}) = \mathrm{Co}(N, \hat{v})$.

Proof:

1. First we prove $\mathrm{Co}(N, v, \mathscr{B}) = \mathrm{Co}(N, \hat{v}, \mathscr{B})$. Indeed, let $x \in \mathrm{Co}(N, \hat{v}, \mathscr{B})$; then $x(S) \geqq \hat{v}(S) \geqq v(S)$ and $x \in X_{\mathscr{B}}$, so $x \in \mathrm{Co}(N, v, \mathscr{B})$. Conversely let $x \in \mathrm{Co}(N, v, \mathscr{B})$; so that $x \in X_{\mathscr{B}}$. For any S, let $(S_1, \ldots, S_l)$ be a partition of S such that $\hat{v}(S) = \sum_{i=1}^{l} v(S_i)$. Because $x(S_i) \geqq v(S_i)$, $i = 1, \ldots, l$, it follows that $x(S) = \sum_{i=1}^{l} x(S_i) \geqq \hat{v}(S)$, and so $x \geqq \mathrm{Co}(N, \hat{v}, \mathscr{B})$.

---

[7]) Example: $N = \{1, 2\}$, $v(N) = 1$, $v(\{1\}) = v(\{2\}) = 0$, $\mathscr{B} = \{N\}$.

2. Next, we show that if $\mathrm{Co}(N,v,\mathscr{B}) \neq \emptyset$, then $\sum_k v(B_k) \geqq v(N)$ and $\sum_k \hat{v}(B_k) = \hat{v}(N)$. Indeed, $\hat{v}(N) \geqq \sum_k \hat{v}(B_k)$ by superadditivity. If $\hat{v}(N) > \sum_k \hat{v}(B_k)$, then there exists a partition $(S_1,\ldots,S_l)$ such that for all x in $\mathrm{Co}(N,v,\mathscr{B})$, $\sum_{i=1}^{l} v(S_i) > \sum_k \hat{v}(B_k) \geqq \sum_k v(B_k) = \sum_k x(B_k)$. Hence, for some i, $v(S_i) > x(S_i)$, contradicting $x \in \mathrm{Co}(N,v,\mathscr{B})$.

3. Finally, we show that if $\mathrm{Co}(N,\hat{v},\mathscr{B}) \neq \emptyset$, then $\mathrm{Co}(N,\hat{v},\mathscr{B}) = \mathrm{Co}(N,\hat{v})$. Indeed, let $x \in \mathrm{Co}(N,\hat{v},\mathscr{B}) \neq \emptyset$. Then $x(N) = \hat{v}(N)$ by 2., and $x(S) \geqq \hat{v}(S)$ for all S, so that $x \in \mathrm{Co}(N,\hat{v})$. Conversely, let $x \in \mathrm{Co}(N,\hat{v})$, and suppose that $x \notin X_{\mathscr{B}}$. Then, there exists a B_k such that $x(B_k) < \hat{v}(B_k)$, contradicting $x \in \mathrm{Co}(N,\hat{v})$.

11. Examples

In this section we consider two games that illustrate some of the methods and results of this paper.

1. Consider a system of p universities with a total of n professors, including m game theorists. Model this by a game with n players and a coalition structure $\mathscr{B}$, in which B_k corresponds to a university. Assume that $m < p$; that no university employs more than one game theorist; that $\mathscr{B}$ is *efficient*, i.e. $\sum_k v(B_k) = v(N)$; and that the game theorists are substitutes for each other. Number the professors and the universities so that the game theorists are numbered $1,\ldots,m$, and $i \in B_i$ for $i = 1,\ldots,m$.

There is no reason to expect that

$$v(B_i) - v(B_i \backslash \{i\}) = v(B_j) - v(B_j \backslash \{j\}),\ i,j = 1,\ldots,m\,;$$

but from efficiency it follows that

$$v(B_i) - v(B_i \backslash \{i\}) \geqq v(B_k \cup \{i\}) - v(B_k),\ i = 1,\ldots,n, k = n+1,\ldots,p\,.$$

Theorem 9 implies that for all x in $\mathrm{Co}(N,v,\mathscr{B})$,

$$\min_{i=1,\ldots,n} \{v(B_i) - v(B_i \backslash \{i\})\} \geqq x_1 = x_2 = \cdots = x_n \geqq \max_{k=n+1,\ldots,p} \{v(B_k \cup \{i\}) - v(B_k)\}.$$

In economic terms, one would say that, in the core, salaries of game theorists are equal in all universities, with an upper limit given by the productivity of the marginal game theorist employed as such and a lower limit given by the "alternative productivity" of a game theorist in a university that does not currently employ one.

2. In this example the players are again professors, but the coalitions in the coalition structure $\mathscr{B}$ are countries. We assume that v is superadditive, so that

$$v(N) - v(N \backslash B_l) \geqq v(B_l)$$

for all l; and that $\mathscr{B}$ is *in*efficient, so that strict inequality holds for at least one l, say $l = k$. For simplicity of exposition, give k a name, say "Israel" (see the discussion in Subsection 5 of Section 12). Intuitively, the strict inequality says that the total value added by Israeli professors to their country is less than the total value that they could add to the rest of the world. Or, in terms of salaries, the total salary that Israeli professors command abroad is larger than the total they command at home. For all x in $X_{\mathscr{B}}$, we have

$$v_x^*(B_k) \geqq v(B_k \cup (N \backslash B_k)) - x(N \backslash B_k) = v(N) - v(N \backslash B_k) > v(B_k) = x(B_k).$$

Let $x \backslash B_k$ be the nucleolus of (B_k, v_x^*, X_k). Assume further that, for all $S \subset B_k$, $v_x^*(S) = \sum_{i \in S} v_x^*(\{i\})$: the salaries that Israeli professors could earn abroad are unaffected by the presence there of other Israeli professors. It is readily verified that the nucleolus of (B_k, v_x^*, X_k) is such that $x_i = v_x^*(\{i\}) - c$ for all $i \in B_k$, where

$$c = \left(\sum_{i \in B_k} v_x^*(\{i\}) - v(B_k)\right) / |B_k| ;$$

this property holds whenever $v_x^*(\{i\}) \geqq c$ for all $i \in B_k$. That is: each Israeli professor would, in the nucleolus, receive a salary equal to his "opportunity cost" (the salary which he could earn abroad), minus a flat deduction which is the same for all Israeli professors.

12. Discussion

1. For a given characteristic function v, the major novel element introduced by the coalition structure $\mathscr{B}$ lies in the conditions $x(B_k) = v(B_k)$, which constrain the solution to allocate exactly among the members of each coalition the total payoff of that coalition. As a consequence of this, the bargaining over the payoff inside coalition B_k will involve a mixture of considerations which are endogenous to B_k and of considerations which are exogenous to B_k and reflect the "outside opportunities" of the members of B_k.

In so far as the value is concerned, the solution is determined entirely by considerations which are endogenous to each coalition B_k (Theorem 3). In so far as the other solution concepts reviewed here are concerned, considerations exogenous to B_k are relevant, but they are fully described (conditionally on the outside imputation $x|N \backslash B_k$) by the characteristic function v_x^* of the auxiliary game (B_k, v_x^*, X_k).

2. Whereas the *implications* of a coalition structure are quite clear, the *idea* of a coalition structure needs some clarification. On the one hand, the players are constrained to "form" the coalitions $B_1, \ldots, B_p$ that make up the structure $\mathscr{B}$. On the other hand, considerations of other coalitions, including those that "cut across" the B_k, is by no means excluded. Such coalitions are used to dominate as in the definition of core and von Neumann-Morgenstern solution, and to object as in the bargaining set and its relatives; the excesses of these coalitions enter into the definition of nucleolus and kernel. This raises the question: what, precisely, does the "constraint" to the structure $\mathscr{B}$ mean?

The scenario usually associated with the coalition structure idea is as follows: the players consider forming the coalitions $B_1, \ldots, B_p$; one may think of them as going to business lunches in p different groups, each B_k forming a group. At these lunches they negotiate the division of the payoff, on the assumption that the coalitions $B_1,\dots,B_p$ will be formed. In such negotiations, it is perfectly reasonable for each coalition B_k to base the division of the payoff on the opportunities that its members have outside of B_k. The negotiations at the lunch may of course break down, and at no time is it asserted that they will or even should succeed. What is being asserted is only that *if* the structure $\mathscr{B}$ forms, then the B_k should divide the payoff in whatever way the particular solution concept under consideration dictates.

What this scenario does not explain is why the groups B_k would form, or why the process of formation of the groups B_k should be separated from the bargaining for the payoff.

In attempting to answer that question, we will consider first some of the traditional explanations for the formation of coalition structures, and show that they do not survive close examination (Subsection 3). We will then advance three different arguments that show why coalition structures might arise. The most transparent one of these (Subsection 4) is valid only for games that are *not* superadditive. The other two arguments (Subsections 5 and 6) are more subtle; but they apply to all games, superadditive or not.

3. The arguments traditionally advanced for the formation of coalition structures include a) difficulties of communication; b) legal barriers such as anti-trust laws; c) personal, family, patriotic, geographical or professional relationships. Unfortunately, they do not survive close examination. The arguments advanced under a) and b) apply to potential coalitions S – those used in dominating, objecting, threatening, and so on – to exactly the same extent as they apply to the coalitions B_k constituting the structure. If a coalition is difficult or impossible to form because of communication difficulties, then these difficulties should also be taken into account when the players are comparing their opportunities. If antitrust laws forbid the formation of a coalition, and this coalition is thereby excluded from participating in a coalition structure, it should by the same token be excluded from considerations affecting the payoff.

There remain the groupings formed for a variety of "personal" reasons – item c) in the above list. It might be argued that one uses such considerations to choose the people one deals with, but there would be no hesitation about breaking up such a grouping if it would lead to material gain. When examined carefully, this argument is seen to imply a kind of lexicographic utility, with personal relationships coming out on the short end, i.e. counting for next to nothing. Most people value their personal relationships more highly than this indicates. But even if one accepts such a lexicographic utility, it would seem better to allow the entire lexicographic utility to determine the payoff, and not arbitrarily to use one component for the coalition structure and another component for bargaining.

4. In certain non-superadditive games, one can give a fairly straightforward explanation for the formation of coalition structures. But before we do that,

let us try to understand the phenomenon of non-superadditivity – to see how it is that a game arising in applications might actually fail to be superadditive. After all, superadditivity is intuitively rather compelling; why shouldn't disjoint coalitions, when acting together, get at least as much as they can when acting separately?

The answer is that the very act of "acting together" may be difficult, costly, or illegal, or the players may, for various "personal" reasons, not wish to do so. In other words, we return to points a), b), and c) of the previous subsection – but as explanations for non-superadditivity of the game, rather than for the formation of coalition structures. In fact, they are simply restrictions on the formations of coalitions. As such, it is perfectly natural to embody them in the definition of the characteristic function v; and in general, a non-superadditive v will result. For instance, if coalitions of n players or more are forbidden by the law, then "$v(S) = -\infty$ whenever $|S| \geqq n$" describes the situation correctly.

Another point is that there is more involved than just "acting together"; there is also the matter of side payments. If $v(S \cup T) \geqq v(S) + v(T)$, then S can transfer some of its own payoff to T. Even when communication is unrestricted, such transfers may be illegal, restricted, or subject to transaction costs. Moreover, "acting together" and sharing the proceeds may change the nature of the game. For example, if two independent farmers were to merge their activities and share the proceeds, both of them might work with less care and energy; the resulting output might be less than under independent operations, in spite of a possibly more efficient division of labour.

Suppose then, that v is a game on N, and let $\hat{v}$ be its superadditive cover (see Section 10). Call a coalition structure $\mathcal{B}$ *efficient* if $\sum_k v(B_k) = \hat{v}(N)$. Obviously $\hat{v}(N) \geqq v(N)$. If $\hat{v}(N) = v(N)$ – as is the case in all superadditive games – then one could consider it reasonable for the coalition structure $\{N\}$ to form. If, however, $\hat{v}(N) > v(N)$, then the coalition structure $\{N\}$ is inefficient, and so a major incentive for its formation is absent. In that case it is possible that an efficient coalition structure $\mathcal{B}$ will form; but since a major cause of non-superadditivity is lack of communication for one reason or another, it might well happen, too, that $\mathcal{B}$ will be inefficient. In any case, when $\hat{v}(N) > v(N)$, one may expect the formation of a $\mathcal{B}$ other than $\{N\}$. Thus one sees that points a), b), and c) in the previous subsection, though not directly valid as an explanation for coalition structures, nevertheless are involved in an indirect manner; they explain non-superadditivity, and this in turn leads to coalition structures[8]).

If the reader wishes, he may view the analysis in this subsection as part of a broader analysis, which would consider simultaneously the process of coalition

[8]) In some cases, one might wish to model the situation by means of a game in "partition function form" [THRALL and LUCAS, 1963], where the characteristic function is a vector-valued function on the family of partitions of N. For instance, if the law requires the existence of at least m coalitions, then only coalition structures with m elements or more should be considered. The theory of games in partition function form raises substantial technical difficulties, and is not considered in this paper.

formation and the bargaining for the payoff. Let a given coalition structure $\mathscr{B}$, and a given payoff x consistent with it, provide a "solution" to the game (N,v); then certainly the payoff x must provide a "solution" to the game $(N,v,\mathscr{B})$. Our analysis has been concerned with this last topic, and should thus be understood as a contribution to partial equilibrium analysis.

5. If the game (N,v) is superadditive, the arguments for the formation of the structure $\{N\}$ sound rather compelling. Must we then abandon altogether the concept of coalition structure, in the context of superadditive games?

We think not. *Coalition structures represent groupings formed for reasons that are important and weighty, but whose impact is*

a) *difficult to measure, and/or*

b) *difficult to communicate believably, and/or*

c) *consciously excluded by the players from bargaining considerations.*

As an example that illustrates all three points, consider the world academic community, which we may think of as having partitioned itself into countries. From the point of view of material payoff, this partition is inefficient. Thus the United States could probably absorb the entire academic community of a small country like Israel, and pay each of its members a salary considerably higher than the one he is now getting. The professors do not move because for personal reasons they prefer to live in their own country. In competing for payoff within their own country, though, they may very well cite the opportunities they have abroad (see Example 2 in Section 11).

Now the question arises, why do not the professors incorporate their preferences for living in their own country into their utility functions explicitly, and bargain accordingly? One answer is that it is difficult to measure such preferences on the same terms as salary, another is that because of their subjectivity it is difficult to communicate such preferences to one's colleagues, or at least to communicate them in a convincing and believable manner.

But the most important consideration, perhaps, is of the third kind. We can at least imagine a situation in which one's love for one's country can be assessed in monetary terms; it is simply a question of deciding what increase in salary abroad would make a person want to move. We could even imagine situations in which these assessments could be believably communicated to one's colleagues. Consider now the consequences of using such assessments in the bargaining: The result would be that people who value their home country highly would be penalized. Indeed, the worth $v(S)$ of coalitions including such people would tend to be lower than the worth of coalitions containing their colleagues who are relatively indifferent, and the excesses would of course behave in the same way. This is quite understandable, as the indifferent professor is in a better position to exert pressure by means of threats to leave than the professor who values highly the place in which he lives and is unwilling to consider a move. Nevertheless it is quite conceivable that all concerned would agree that this kind of consideration should not enter into salary determinations. Given such an understanding, it is clear that the situation would best be analyzed by a game with a coalition structure determined by the countries, in which the worths $v(S)$ of the coalitions are expressed in more or less "objective" monetary terms.

A coalition structure may thus reflect considerations that are excluded from the formal description of the game by necessity (impossibility to measure or communicate) or by choice. This situation may arise in the non-superadditive as well as in the superadditive case – but in the former case it is not the only possible explanation of the coalition structure. Finally, the difficulty to measure or communicate may exist at the level of the scientist and not only at the level of the players. This illustrates a basic principle of modelling; following SAVAGE [1954], one might call it the "small worlds" principle. A model cannot always be expected to take into account in a systematic and consistent manner all the complexities of a complex situation. It is often necessary – or if not necessary at least convenient – to treat basically similar phenomena in a methodologically different fashion. Thus in SAVAGE's theory of subjective probabilities, a very clear distinction is made between the concepts of "act", "consequence", and "state of the world", and this distinction is basic to his theory. But in the real world these distinctions blur, and it is sometimes difficult to distinguish between these concepts. Our situation here is similar; we cannot say exactly which considerations go into determining a coalition structure, and which go into bargaining; we are not even sure that these two elements can be clearly separated from each other; but we feel that there are situations in which the two elements are present, and are better treated separately.

6. In the previous subsection we adduced an *exogenous* argument for the formation of a coalition structure $\mathscr{B}$, i.e. we took the coalition structure to be based on factors *not* taken into account in the characteristic function v. On the other hand, in Subsection 4 we used an *endogenous* argument, i.e. we explained the formation of $\mathscr{B}$ in terms of v itself. In this subsection, we will adduce another endogenous argument; but the current argument, unlike that in Subsection 4, is valid for superadditive as well as non-superadditive games[9]).

Briefly, the point is that the coalition structure $\mathscr{B}$ might arise because for the players in some of the B_k, it might be more worthwhile to bargain in the framework of $\mathscr{B}$ than in the framework of the all-player coalition N.

For example, consider the game v on $\{1,2,3\}$ defined by

$$v(S) = \begin{cases} 0 & \text{if } |S| = 1 \\ 8 & \text{if } |S| = 2 \\ 9 & \text{if } |S| = 3. \end{cases}$$

[9]) The argument in this subsection is due to MICHAEL MASCHLER, and we are grateful to him for permission to publish it here.

Assume that there are no significant asymmetries between the players (i.e. asymmetries not included in the description of the situation by v). Suppose that, for some reason, players 1 and 2 find themselves at "lunch" (see Subsection 2) without player 3. There is little doubt that they would quickly seize the opportunity to form the coalition (1,2) and collect a payoff of 4 each. Thus the outcome would be (4,4,0), with $\mathscr{B} = \{(1,2),(3)\}$. This would happen in spite of its inefficiency. The reason is that if players 1 and 2 were to invite player 3 to the lunch, the outcome would presumably be (3,3,3). Neither would they want to risk inviting him and offering him, say, $\frac{1}{2}$ (and dividing the remaining $8\frac{1}{2}$ among themselves); because each of the two players 1 and 2 would realize that once player 3 is invited to participate in the negotiations, the situation turns "wide open" – anything can happen.

All this if players 1 and 2 happen "to find themselves at lunch". But even if this does not happen by chance, it is now fairly clear that the players would seek to form pairs for the purpose of negotiation, and not negotiate in the all-player framework.

Our example is particularly convincing because of its symmetry. Even in unsymmetric cases, though, it is clear that the framework of negotiations plays an important role in the outcome, so that individual players and groups of players will seek frameworks that are advantageous to them; and this may well lead to inefficient coalition structures. The phenomenon of seeking an advantageous framework for negotiating is also well-known in the real world at many levels – from decision making within an organization like a corporation or a university, to international negotiations.

The remarks at the end of Subsection 4 about the partial equilibrium nature of the analysis apply here as well. Incidentally, considerations of coalition formation are implicit in the VON NEUMANN-MORGENSTERN solution, even when applied only to the coalition structure $\{N\}$. But we know that the VON NEUMANN-MORGENSTERN approach to game theory is only one of a number of possible approaches, and it is worthwhile to study coalition structures explicitly and in the context of notions other than just the VON NEUMANN-MORGENSTERN solution.

References

AUMANN, R. J., and M. MASCHLER: The Bargaining Set for Cooperative Games, in: Advances in Game Theory, M. DRESHER, L. S. SHAPLEY, and A. W. TUCKER, eds., Annals of Mathematics Studies, No. 52, pp. 443–476, Princeton 1964 [Chapter 42].

DAVIS, M., and M. MASCHLER: Existence of Stable Payoff Configurations for Cooperative Games, in: Essays in Mathematical Economics in Honor of Oskar Morgenstern, M. Shubik, ed., pp. 39–52, Princeton 1967.

–: The Kernel of a Cooperative Game, Naval Research Logistics Quarterly **12**, 223–259, 1965.

GILLIES, D. B.: Solutions to General Non-zero-sum Games, in: Contributions to the Theory of Games IV, R. D. LUCE, and A. W. TUCKER, eds., Annals of Mathematics Studies, No. 40, pp. 47–85, Princeton 1959.

KOHLBERG, E.: On the Nucleolus of a Characteristic Function Game, SIAM Journal of Applied Mathematics, **20**, 62–66, 1971.

MASCHLER, M., and B. PELEG: The Structure of the Kernel of a Cooperative Game, SIAM Journal of Applied Mathematics, **15**, 569–604, 1967.

MASCHLER, M., B. PELEG, and L. SHAPLEY: The Kernel and Bargaining Set for Convex Games, International Journal of Game Theory **1**, 73–93, 1972.

SAVAGE, L. J.: Foundations of Statistics; Wiley, New York 1954.

SCHMEIDLER, D.: The Nucleolus of a Characteristic Function Game, SIAM Journal of Applied Mathematics **17**, 1163–1170, 1969.

SELTEN, R.: Equal Share Analysis of Characteristic Function Experiments, in: Contributions to Experimental Economics III, H. SAUERMANN and J. C. B. MOHR, eds., pp. 130–165, Tübingen 1972.

SHAPLEY, L. S.: A Value for *n*-person Games, in: Contributions to the Theory of Games II, H. W. KUHN, and A. W. TUCKER, eds., Annals of Mathematics Studies no. 52, pp. 307–317, Princeton 1953.

THRALL, R., and W. F. LUCAS: Games in Partition Function Form, Naval Research Logistics Quarterly **10**, 281–298, 1963.

VON NEUMANN, J., and O. MORGENSTERN: Theory of Games and Economic Behaviour, Princeton 1944.

Note added in proof: Theorem 6 (p. 128) is wrong. See Chang, C.: A Note on the von Neumann–Morgenstern Solution, International Journal of Game Theory **17**, 311–314, 1988.