A Shapley Value Representation

of Potential Games*

Takashi Ui

Institute of Policy and Planning Sciences

University of Tsukuba

Tuskuba, Ibaraki 305-8573, Japan

First Draft: December, 1995
This Version: February, 1999

Proposed Running Head: representation of potential games

Mailing Address:

Takashi Ui

Institute of Policy and Planning Sciences

University of Tsukuba

1-1-1 Tennodai

Tsukuba, Ibaraki 305-8573

Japan

E-mail: oui@shako.sk.tsukuba.ac.jp

Tel: +81-298-53-5555

Fax: +81-298-53-5555

Abstract:

In potential games, as considered by Monderer and Shapley (1996a), each player's gain from a deviation is equal to the gain in a potential function. We prove that a game has a potential function if and only if its payoff functions coincide with the Shapley value of a particular class of cooperative games indexed by the set of strategy profiles. Also a potential function of a noncooperative potential game coincides with the potentials (cf. Hart and Mas-Colell, 1989) of cooperative games indexed by the strategy set.

JEL classification numbers: C71, C72

List of Symbols

Section 2

$\Gamma = (N, A, u)$: a strategic form game.

Section 3

V: a potential function.

Section 4

$\mathcal{N}$: the collection of subsets of N.

v: a TU game.

$v_{|T}$: a restricted TU game.

e_R: a unanimity game.

$\mathcal{G}_N$: the set of all TU games.

ϕ: the Shapley value.

P: a potential function of TU games.

DP_i: the marginal contribution of player i to P.

Section 5

$\mathcal{G}_{N,A}$: the set of all TU games with action choices.

$\{\Phi_S\}_{S \in \mathcal{N}}$: an interaction potential.

Section 6

ψ: the Myerson value.

1 Introduction

Potential games, as considered by Monderer and Shapley (1996a), are games with potential functions. Potential functions, or potentials, are defined as functions of strategy profiles such that the difference induced by a single deviation is equal to that of the deviator's payoff function. It is known that the set of Nash equilibria of a potential game coincides with that of a perfect coordination game in which each player's payoff function is identical to the potential function.

Some interesting results with respect to learning processes have been reported based upon the potential functions. Blume (1993) has considered stochastic evolutionary dynamics in a class of games in which players are located on a lattice and they play the same symmetric two player potential game with nearby players. He showed that strategy profiles that maximize the potential functions have the highest probabilities in the stationary distributions (see also Ui, 1997). Monderer and Shapley (1996b) showed that fictitious play processes converge to Nash equilibria in potential games.

These findings raise the natural question of when a game has a potential function and how it can be calculated. We provide two representations of potential games. The first provides a unified view of the relationship between potential functions of games and the Shapley value, which Monderer and Shapley (1996a), Qin (1996), and Hart and Mas-Colell (1989) suggested.

Monderer and Shapley (1996a) extensively studied basic properties of potential games. One important finding is the relationship to congestion games considered by Rosenthal (1973). They provided a representation theorem, which states that finite potential games are isomorphic to congestion games. Another important finding is the relationship to the Shapley value. They defined a simple participation game and showed that it has a potential function if and only if the payoffs are described

by the Shapley value, using a result due to Hart and Mas-Colell (1989).

Qin (1996) found a similar relationship to the Shapley value. He showed that a coalition formation game considered by Myerson (1991) has a potential function. As shown by Myerson (1977), the payoffs of the coalition formation game are also given by the Shapley value of a particular cooperative game.

Hart and Mas-Colell (1989) defined and studied potential functions of cooperative rather than noncooperative games. They considered the marginal contribution of each player to the potential function and showed that it is equal to the Shapley value, giving another interpretation of the Shapley value.

Our result explains how potential games are related to the Shapley value, and how the potential functions of Monderer and Shapley (1996a) are related to those of Hart and Mas-Colell (1989). Our result states that a game has a potential function if and only if the payoff functions coincide with the Shapley value of a particular class of cooperative games indexed by the set of strategy profiles. In addition, a potential function of the game coincides with a potential function of the class of cooperative games.

The organization of this paper is as follows. In Section 2, we provide examples of potential games. In doing so, we define a class of games called bilateral symmetric interaction games, which exhibits an important idea for our main result. In Section 3, we formally define potential games and discuss their basic properties. In Section 4, we review the Shapley value and the potential functions of cooperative games. In Section 5, we prove the main result. In Section 6, we discuss examples, demonstrating how the main result is applied. We conclude the paper in Section 7.

2   Bilateral Symmetric Interaction Games

Throughout the paper, $\Gamma = (N, A, u)$ denotes a strategic form game; $N = \{1, \ldots, n\}$ is a finite set of players; A_i is a set of strategies for player $i \in N$ and strategy space $A = (A_i)_{i \in N}$; $u_i : A \to \Re$ is a payoff function of player $i \in N$ and $u = (u_i)_{i \in N}$. We write $a = (a_1, a_2, \ldots, a_n)$, $a_{-i} = (a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$, $a \backslash a_i' = (a_i', a_{-i}) = (a_1, a_2, \ldots, a_{i-1}, a_i', a_{i+1}, \ldots, a_n)$. For $S \subseteq N$, we write $A_S = (A_i)_{i \in S}$ and $a_S = (a_i)_{i \in S} \in A_S$. We write $A_{-i} = A_{N \backslash \{i\}}$.

Consider $\Gamma = (N, A, u)$ such that, for each $i, j \in N$ with $i \neq j$, there exist functions $w_{ij} : A_i \times A_j \to \Re$ and $h_i : A_i \to \Re$, and that

$$u_i(a) = \sum_{j \in N \backslash \{i\}} w_{ij}(a_i, a_j) - h_i(a_i). \tag{1}$$

We assume that $w_{ij}(a_i, a_j) = w_{ji}(a_j, a_i)$ for any $(a_i, a_j) \in A_i \times A_j$. We call Γ a *bilateral symmetric interaction game* (BSI game) because payoff functions can be decomposed into symmetric interaction terms which are bilaterally determined together with the term depending only on the player's own strategy.

A 2×2 symmetric game is a BSI game. Let $N = \{1, 2\}$, $A_1 = A_2 = \{0, 1\}$. Consider payoff functions given by the following table.

	0	1
0	x, x	z, w
1	w, z	y, y

We can choose w_{12}, w_{21}, h_1 and h_2 such that the payoff functions are given by (1): $w_{12}(0, 0) = w_{21}(0, 0) = x$, $w_{12}(0, 1) = w_{21}(1, 0) = z$, $w_{12}(1, 0) = w_{21}(0, 1) = z$, $w_{12}(1, 1) = w_{21}(1, 1) = y + z - w$, $h_1(0) = h_2(0) = 0$, $h_1(1) = h_2(1) = z - w$. Note that $w_{12}(a_1, a_2) = w_{21}(a_2, a_1)$.

A Cournot oligopoly game with a linear inverse demand function is also a BSI game. Let N be a set of firms and $A_i \subseteq \Re^+$ be a set of possible outputs. If the

inverse demand function is $F(x) = \alpha - \beta x$ and the cost function of firm i is $c_i(a_i)$ for all $i \in N$, a profit function of firm i is

$$
\begin{aligned}
\pi_i(a) &= F(\sum_{j \in N} a_j)a_i - c_i(a_i) \\
&= (\alpha - \beta \sum_{j \in N} a_j)a_i - c_i(a_i) \\
&= -\beta \sum_{j \neq i} a_j a_i - [c_i(a_i) + \beta a_i^2 - \alpha a_i],
\end{aligned}
$$

where a_i denotes an output of firm i. Here, $w_{ij}(a_i, a_j) = -\beta a_i a_j$ and $h_i(a_i) = c_i(a_i) + \beta a_i^2 - \alpha a_i$. Note that $w_{ij}(a_i, a_j) = w_{ji}(a_j, a_i)$.

In a general BSI game, the net gain from a single deviation by firm i from a_i to b_i is

$$
u_i(a \backslash b_i) - u_i(a) = \sum_{j \in N \backslash \{i\}} [w_{ij}(b_i, a_j) - w_{ij}(a_i, a_j)] - [h_i(b_i) - h_i(a_i)].
$$

Consider the function $V : A \to \Re$ defined by

$$
V(a) = \sum_{i < j} w_{ij}(a_i, a_j) - \sum_i h_i(a_i).
$$

Then, since $w_{ij}(a_i, a_j) = w_{ji}(a_j, a_i)$ for $i, j \in N$, we find that

$$
V(a \backslash b_i) - V(a) = \sum_{j \in N \backslash \{i\}} [w_{ij}(b_i, a_j) - w_{ij}(a_i, a_j)] - [h_i(b_i) - h_i(a_i)]
$$

and thus that

$$
V(a \backslash b_i) - V(a) = u_i(a \backslash b_i) - u_i(a)
$$

for each player $i \in N$. The function V satisfying the last equation is called a *potential function*, and a game that has a potential function is called a *potential game*. A BSI game is thus a potential game. It is due to Slade (1994) that the above oligopoly game has a potential function.

It should be noted that the symmetry assumption $w_{ij}(a_i, a_j) = w_{ji}(a_j, a_i)$ is crucial for the existence of a potential function. Our main result will show that

every potential game has a symmetric structure similar to those in BSI games. This allows payoff functions of potential games to be described in terms of the Shapley value.

3 Potential Games

Formally, potential games are defined as follows:

Definition 1 (Monderer and Shapley, 1996a) $\Gamma = (N, A, u)$ *is called a potential game if there exists a potential function* $V : A \rightarrow \Re$ *such that*

$$u_i(a \backslash b_i) - u_i(a) = V(a \backslash b_i) - V(a) \tag{2}$$

for any $i \in N$, $b_i \in A_i$, $a \in A$.

The following provides the simplest necessary and sufficient condition for potential games.

Theorem 1 (Slade, 1994; Facchini et al., 1997) $\Gamma = (N, A, u)$ *is a potential game if and only if there exist functions* $V : A \rightarrow \Re$ *and* $Q_i : A_{-i} \rightarrow \Re$ *such that*

$$u_i(a) = V(a) + Q_i(a_{-i})$$

for any $i \in N$, $a \in A$. V *is a potential function.*

Proof. Sufficiency is apparent. Suppose that Γ is a potential game. (2) implies that for each player $i \in N$ the function $Q_i : A_{-i} \rightarrow \Re$ with $Q_i(a_{-i}) = u_i(a) - V(a)$ for any $a_{-i} \in A_{-i}$ and arbitrary $a_i \in A_i$ is well-defined. Thus, necessity is also true. ∎

It is clear that if V is a potential function then $V + c$ is also a potential function where c is an arbitrary constant. A stronger claim is true.

Lemma 1 (Monderer and Shapley, 1996a) *Let $\Gamma = (N, A, u)$ be a potential game and V and V' be potential functions of Γ. Then there exists a constant c such that $V(a) = V'(a) + c$ for any $a \in A$.*

Potential functions are useful tools in analyzing potential games due to the following result.

Lemma 2 (Monderer and Shapley, 1996a) *Let $\Gamma = (N, A, u)$ be a potential game with a potential function V. Let $\bar{\Gamma}$ be the game $(N, A, (V)_{i \in N})$ in which every player's payoff function is V. Then the set of Nash equilibria of Γ coincides with that of $\bar{\Gamma}$.*

Thus, in a potential game, a single potential function can be used to find all the Nash equilibria, which makes analysis of the game very simple. Note that if A_i is a compact set for all $i \in N$ and the game has a continuous potential function, then the game has at least one pure strategy Nash equilibrium since $a^* \in \arg\max_{a \in A} V(a)$ is a Nash equilibrium.

Monderer and Shapley (1996a) discussed how to verify whether a game has a potential function (in Lemma 4.4 and Theorem 4.5). Suppose that, for all $i \in N$, A_i is an interval of $\Re$ and that payoff functions are twice continuously differentiable. Then, (2) is equivalent to

$$\frac{\partial u_i}{\partial a_i} = \frac{\partial V}{\partial a_i}.$$

This is true if and only if

$$\frac{\partial^2 u_i}{\partial a_i \partial a_j} = \frac{\partial^2 u_j}{\partial a_i \partial a_j} = \frac{\partial^2 V}{\partial a_i \partial a_j}. \tag{3}$$

Conversely, if (3) holds, the function V can be calculated by integrating (3). Monderer and Shapley (1996a) discussed similar conditions for non-differentiable payoff functions by replacing 'differentials' with 'differences'.

4 The Shapley Value

Let $\mathcal{N}$ be the collection of subsets of N. A cooperative game with transferable utility, or a TU game, is defined as a function $v : \mathcal{N} \to \Re$ such that $v(\emptyset) = 0$. The set of all TU games is denoted by $\mathcal{G}_N$.

For each $R \in \mathcal{N}$, define a unanimity game $e_R \in \mathcal{G}_N$ so that

$$e_R(S) = \begin{cases} 1 & \text{if } S \supseteq R \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3 (Shapley, 1953) *Any TU game v is described by a unique linear combination of a collection of unanimity games $\{e_R\}_{R \in \mathcal{N}}$, i.e.*

$$v(S) = \sum_{R \in \mathcal{N}} v^R e_R(S)$$

where

$$v^R = \sum_{T \subseteq R} (-1)^{|R \backslash T|} v(T).$$

For $v \in \mathcal{G}_N$ and $T \in \mathcal{N}$, let the restricted TU game $v_{|T} \in \mathcal{G}_N$ be such that

$$v_{|T}(S) = \begin{cases} v(S \cap T) & \text{if } S \cap T \neq \emptyset \\ 0 & \text{otherwise,} \end{cases}$$

or equivalently,

$$v_{|T}(S) = \sum_{R \in \mathcal{N}, R \subseteq T} v^R e_R(S).$$

The Shapley value (Shapley, 1953) is defined as a map $\phi : \mathcal{G}_N \to \Re^N$ such that

$$\phi_i(v) = \sum_{S \in \mathcal{N}, i \in S} \frac{(|S| - 1)!(|N| - |S|)!}{|N|!} (v(S) - v(S \backslash \{i\}))$$

for all $i \in N$. It is known that ϕ is a linear map which satisfies

$$\phi_i(e_R) = \begin{cases} 1/|R| & \text{if } i \in R \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$\phi_i(v) = \sum_{R \in \mathcal{N}} v^R \phi_i(e_R) = \sum_{R \in \mathcal{N}, i \in R} v^R/|R|$$

where $v^R/|R|$ is called Harsanyi's dividend to the members of R (Harsanyi, 1959).

Hart and Mas-Colell (1989) define a potential function for TU games. For a function $P : \mathcal{G}_N \rightarrow \Re$ and for all $i \in N$, the function $DP_i : \mathcal{G}_N \rightarrow \Re$ is defined as the marginal contribution of player i to P, i.e.

$$DP_i(v) = P(v) - P(v_{|N \setminus \{i\}}).$$

Then P is a potential function of TU games if it satisfies

$$\sum_{i \in N} DP_i(v) = v(N).$$

Hart and Mas-Colell (1989) showed that P is uniquely given by

$$P(v) = \sum_{R \in \mathcal{N}} v^R/|R|, \tag{4}$$

and that, for all $i \in N$, the marginal contribution of player i coincides with his payoff according to the Shapley value:

$$DP_i(v) = \sum_{R \in \mathcal{N}, i \in R} v^R/|R| = \phi_i(v). \tag{5}$$

5 Representation Theorems

Let a set of players N and a strategy space A be given. We consider a collection of TU games $\{v_a\}_{a \in A}$ such that $v_a(S) = v_b(S)$ if $a_S = b_S$. We call $\{v_a\}_{a \in A}$ a TU game with action choices. Notice that the value of a coalition is determined by its members and the strategies of its members, but not by the strategies of the players outside the coalition. Let

$$\mathcal{G}_{N,A} = \{\{v_a\}_{a \in A} \,|\, v_a(S) = v_b(S) \text{ if } a_S = b_S\}$$

denote the set of all TU games with action choices.

The main result is the following representation theorem, of which proof is given later.

Theorem 2 $\Gamma = (N, A, u)$ *is a potential game if and only if there exists* $\{v_a\}_{a \in A} \in \mathcal{G}_{N,A}$ *such that*

$$u_i(a) = \phi_i(v_a) \tag{6}$$

for all $i \in N$ *and* $a \in A$. *A potential function is given by*

$$V(a) = P(v_a). \tag{7}$$

As potential functions of TU games are described in terms of Harsanyi's dividends, so are potential functions of potential games due to this result. Concerning Harsanyi's dividends of $\{v_a\}_{a \in A}$, we have the following lemma.

Lemma 4 $\{v_a\}_{a \in A} \in \mathcal{G}_{N,A}$ *if and only if* $a_S = b_S$ *implies* $v_a^S = v_b^S$ *for any* $S \in \mathcal{N}$.

Proof. Suppose that $a_S = b_S$ implies $v_a^S = v_b^S$. For any $S \in \mathcal{N}$,

$$v_a(S) = \sum_{R \in \mathcal{N}} v_a^R e_R(S) = \sum_{R \subseteq S} v_a^R.$$

Thus $a_S = b_S$ also implies $v_a(S) = v_b(S)$. Therefore, $\{v_a\}_{a \in A} \in \mathcal{G}_{N,A}$.

Conversely, suppose that $\{v_a\}_{a \in A} \in \mathcal{G}_{N,A}$. Then

$$v_a^S = \sum_{T \subseteq S} (-1)^{|S \setminus T|} v_a(T)$$

and, for $T \subseteq S$, $v_a(T) = v_b(T)$ if $a_S = b_S$. Thus $a_S = b_S$ implies $v_a^S = v_b^S$. ∎

A collection $\{\Phi_S\}_{S \in \mathcal{N}}$ such that $\Phi_S : A_S \to \Re$ is called an interaction potential, which is a term from a similar concept in random field theory (see Guyon,

1995). Due to the above lemma, the following equation provides a one-to-one correspondence between $\{v_a\}_{a \in A}$ and $\{\Phi_S\}_{S \in \mathcal{N}}$:

$$\Phi_S(a_S) = v_a^S / |S|. \tag{8}$$

Due to this, Theorem 2 is equivalent to the following theorem, which, in some situations, gives a practical test of when a game is a potential game and also shows how the potential function can be calculated.

Theorem 3 $\Gamma = (N, A, u)$ *is a potential game if and only if there exists an inter-action potential* $\{\Phi_S \mid \Phi_S : A_S \to \Re,\ S \in \mathcal{N}\}$ *such that*

$$u_i(a) = \sum_{S \in \mathcal{N}, i \in S} \Phi_S(a_S) \tag{9}$$

for all $i \in N$ *and* $a \in A$. *A potential function is given by*

$$V(a) = \sum_{S \in \mathcal{N}} \Phi_S(a_S). \tag{10}$$

Proof of Theorem 3. Assume that there exists an interaction potential $\{\Phi_S\}_{S \in \mathcal{N}}$ such that $u_i(a) = \sum_{S \in \mathcal{N}, i \in S} \Phi_S(a_S)$. Define $V(a) = \sum_{S \in \mathcal{N}} \Phi_S(a_S)$. Then

$$
\begin{aligned}
V(a \backslash b_i) - V(a) &= \sum_{S \in \mathcal{N}} \Phi_S((a \backslash b_i)_S) - \sum_{S \in \mathcal{N}} \Phi_S(a_S) \\
&= \sum_{S \in \mathcal{N}, i \in S} \Phi_S((a \backslash b_i)_S) - \sum_{S \in \mathcal{N}, i \in S} \Phi_S(a_S) \\
&= u_i(a \backslash b_i) - u_i(a).
\end{aligned}
$$

Thus Γ is a potential game with a potential function V.

Conversely, suppose that Γ is a potential game with a potential function V. Let $Q_i(a_{-i}) = u_i(a) - V(a)$. For $S \in \mathcal{N}$, define

$$
\Phi_S(a_S) = \begin{cases}
V(a) + \sum_{i \in N} Q_i(a_{-i}) & \text{if } S = N \\
-Q_i(a_{-i}) & \text{if } S = N \backslash \{i\} \text{ for some } i \in N \\
0 & \text{if } |S| \leq |N| - 2.
\end{cases}
$$

13

Then, for each $i \in N$, $S \in \mathcal{N}$ and $a_S \in A_S$,

$$
\begin{aligned}
\sum_{S \in \mathcal{N}, i \in S} \Phi_S(a_S) &= \sum_{j \in N \setminus \{i\}} \Phi_{N \setminus \{j\}}(a_{N \setminus \{j\}}) + \Phi_N(a) \\
&= -\sum_{j \in N \setminus \{i\}} Q_j(a_{-j}) + V(a) + \sum_{j \in N} Q_j(a_{-j}) \\
&= V(a) + Q_i(a_{-i}) = u_i(a).
\end{aligned}
$$

Also, for each $S \in \mathcal{N}$ and $a_S \in A_S$,

$$
\begin{aligned}
\sum_{S \in \mathcal{N}} \Phi_S(a_S) &= \sum_{j \in N} \Phi_{N \setminus \{j\}}(a_{N \setminus \{j\}}) + \Phi_N(a) \\
&= -\sum_{j \in N} Q_j(a_{-j}) + V(a) + \sum_{j \in N} Q_j(a_{-j}) \\
&= V(a).
\end{aligned}
$$

This completes the proof. ∎

Proof of Theorem 2. We prove the equivalence of Theorem 2 and Theorem 3.

(8) provides a one-to-one correspondence between $\{v_a\}_{a \in A}$ and $\{\Phi_S\}_{S \in \mathcal{N}}$. (4) and (5) imply that

$$
\phi_i(v_a) = \sum_{S \in \mathcal{N}, i \in S} \Phi_S(a_S) \tag{11}
$$

and

$$
P(v_a) = \sum_{S \in \mathcal{N}} \Phi_S(a_S). \tag{12}
$$

(11) and (12) provides one-to-one correspondences between (6) and (9) and between (7) and (10) respectively. Due to these correspondences, Theorem 2 and Theorem 3 are equivalent. ∎

It should be noted that $\{v_a\}_{a \in A}$ and $\{\Phi_S\}_{S \in \mathcal{N}}$ are not uniquely determined for a given potential game. Of course, this is because the Shapley value is a map from a $(2^{|N|} - 1)$-dimensional space to an $|N|$-dimensional space and thus is not one to one.

The condition in Theorem 3 can provide an alternative definition of potential games. The following can apply to games when N is not a finite set:

Definition 2 *Let $\mathcal{N}$ be the collection of finite subsets of N. $\Gamma = (N, A, u)$ is called a* potential game *if, for all $S \in \mathcal{N}$, there exists a function $\Phi_S : A_S \to \Re$ such that*

$$u_i(a) = \sum_{S \in \mathcal{N}, i \in S} \Phi_S(a_S).$$

In Definition 1, V must be well-defined. However, in general, $V(a) = \sum_{S \in \mathcal{N}} \Phi_S(a_S)$ is not well-defined if N is not a finite set. Definition 2 does not need V. For example, the games considered by Blume (1993) have a countable number of players. They are potential games in the sense of Definition 2, but not in the sense of Definition 1.

6 Examples

The Contracting Model of Hart and Moore (1990)

Theorem 2 directly shows that a model introduced by Hart and Moore (1990) can be regarded as a potential game (Hideshi Itoh suggested this example).

Consider a set of players N and a set of physical assets B. Assets are owned by some coalition. The assignments of ownership rights are specified by a control structure, a function $\beta : \mathcal{N} \to 2^B$. β assigns to each coalition $S \in \mathcal{N}$ ownership rights of a set of assets $\beta(S) \subseteq B$. It is assumed that $S \subseteq S'$ implies $\beta(S) \subseteq \beta(S')$.

At date 0, each player $i \in N$ makes an investment $a_i \in A_i = [0, \bar{a}_i]$ in human capital at cost $c_i(a_i)$. At date 1, a cooperative game is played and the Shapley value determines the players' payoffs. When a control structure is β and an investment profile is a, the value of a coalition S is given by $v_{\beta,a}(S)$ where $v_{\beta,a} \in \mathcal{G}_N$. The payoff function is given by

$$u_i(a) = \phi_i(v_{\beta,a}) - c_i(a_i)$$

for all $i \in N$.

The object of the Hart-Moore model is to investigate how alternative assignments of ownership rights affect individual investments in human capital. It is assumed that the investment decisions are chosen in the noncooperative game $\Gamma = (N, A, u)$.

Hart and Moore (1990) assumed that

$$\frac{\partial v_{\beta,a}(S)}{\partial a_i} = 0 \text{ if } i \notin S.$$

This implies that $\{v_{\beta,a}\}_{a \in A} \in \mathcal{G}_{N,A}$. Let $\{v'_{\beta,a}\}_{a \in A} \in \mathcal{G}_{N,A}$ be such that

$$v'_{\beta,a}(S) = v_{\beta,a}(S) - \sum_{i \in S} c_i(a_i).$$

Then it is straightforward to see that

$$u_i(a) = \phi_i(v_{\beta,a}) - c_i(a_i) = \phi_i(v'_{\beta,a}).$$

Due to Theorem 2, Γ is a potential game with a potential function

$$V(a) = P(v'_{\beta,a}).$$

The Coalition Formation Game of Myerson (1991)

The second example is a coalition formation game considered by Myerson (1991). Qin (1996) showed that this game has a potential function.

Let g be a collection of unordered pairs in N. We call g a *cooperative structure*, in which $\{i, j\} \in g$ implies that i and j cooperate bilaterally. Noting that (N, g) can be regarded as an undirected graph, let N/g be a partition of N such that each element of N/g consists of the players in a connected component of (N, g):

$$N/g = \{S \subseteq N \mid i, j \in S \text{ iff there exists a path between } i \text{ and } j \text{ in } (N, g)\}.$$

N/g is a coalition structure under the cooperative structure g.

Let g_S denote the restriction of g to S:

$$g_S = \{\{i, j\} \mid \{i, j\} \in g \text{ and } i, j \in S\}.$$

Similar to N/g, let S/g_S be a partition of S such that each element of S/g_S consists of the players in a connected component of (S, g_S).

Let v be a TU game. The Myerson value $\{\psi_i(v, g)\}_{i \in N}$ is the unique allocation rule such that:

1. It is feasible in the sense that, for any $S \in N/g$,

$$\sum_{i \in S} \psi_i(v, g) = v(S).$$

2. It is fair in the sense that, for any $i, j \in N$ with $i \neq j$,

$$\psi_i(v, g) - \psi_i(v, g \setminus \{\{i, j\}\}) = \psi_j(v, g) - \psi_j(v, g \setminus \{\{i, j\}\}).$$

Myerson (1977) showed that

$$\psi_i(v, g) = \phi_i(v_g)$$

where $v_g \in \mathcal{G}_N$ is such that

$$v_g(S) = \sum_{T \in S/g_S} v(T)$$

for any $S \in \mathcal{N}$.

Myerson (1991) considered a coalition formation game $\Gamma = (N, A, u)$ such that:

1. For all $i \in N$, define $A_i = \{S \subseteq N \mid i \in S\}$. Each $a_i \in A_i$ denotes a set of players with whom player i wishes to cooperate bilaterally.

2. Two players actually cooperate if they both wish to cooperate with each other, i.e. $i \in a_j$ and $j \in a_i$. For $a \in A$, the cooperative structure is given by

$$g(a) = \{\{i, j\} \in N \times N \mid i \in a_j \text{ and } j \in a_i\}.$$

A restriction of $g(a)$ to S is

$$g_S(a) = \{\{i, j\} \in S \times S \mid i \in a_j \text{ and } j \in a_i\}.$$

3. Let a TU game v be given. The payoff function is given by the Myerson value:

$$u_i(a) = \psi_i(v, g(a))$$

for all $i \in N$.

We apply Theorem 2 to this game. Note first that if $a_S = b_S$ then $S/g_S(a) = S/g_S(b)$. Thus

$$v_{g(a)}(S) = \sum_{T \in S/g_S(a)} v(T) = \sum_{T \in S/g_S(b)} v(T) = v_{g(b)}(S).$$

This implies that $\{v_{g(a)}\}_{a \in A} \in \mathcal{G}_{N,A}$ and $u_i(a) = \psi_i(v, g(a)) = \phi_i(v_{g(a)})$. Thus the game is a potential game due to Theorem 2. Its potential function is given by $V(a) = P(v_{g(a)})$.

The Neural Network Model of Hopfield (1982)

Consider $\Gamma = (N, A, u)$ such that $A_i = \{-1, 1\}$ and

$$u_i(a) = \sum_{j \in N \setminus \{i\}} w_{ij} a_i a_j - h_i a_i$$

where w_{ij} and h_i are constants for $i, j \in N$ with $i \neq j$. If $w_{ij} = w_{ji}$ then Γ is a BSI game and thus is a potential game. The potential function is

$$V(a) = \sum_{i < j} w_{ij} a_i a_j - \sum_i h_i a_i.$$

The condition $w_{ij} = w_{ji}$ that Γ has a potential function is closely related to the finding of Hopfield (1982) in neural network theory. A neural network is a

mathematical model of the brain. In general, neural networks have no well-defined energy functions. Hopfield (1982), however, found a class of neural networks that have energy functions. It can be shown that a neural network is, mathematically, isomorphic with Γ, and it has an energy function if and only if Γ has a potential function. We briefly discuss this issue by defining a neural network with the same notations as those of Γ.

The brain consists of about 10^{11} *neurons* (nerve cells). Neurons are connected by nerve fibers, through which signals are transmitted from one neuron to another. Its mathematical model, a neural network, also consists of mathematically defined neurons. Let N be a set of neurons, and $A_i = \{-1, 1\}$ be a set of possible states of neuron i. Neuron i is either "firing" denoted by $a_i = 1$, or "not firing" denoted by $a_i = -1$.

A neuron sends a signal to another. The signal is determined not only by a state of the neuron but also by how two neurons are connected. Let w_{ij} be a connection parameter from neuron i to neuron j. It is assumed that when a state of neuron j is a_j, neuron j sends a signal $w_{ij}a_j$ to neuron i. Note that if $w_{ij} = 0$ then neuron j sends nothing to neuron i.

A state of the neural network is given by $a = (a_i)_{i \in N} \in \{-1, 1\}^N$. It changes over time according to a Markov chain in which only one neuron is selected at each time and allowed to change its state depending upon the signals the neuron receives. The selected neuron fires if and only if the sum of the signals exceeds a threshold. Formally, if $\sum_{j \neq i} w_{ij}a_j > h_i$ then neuron i fires; if $\sum_{j \neq i} w_{ij}a_j \leq h_i$ then it does not fire. In other words, neuron i chooses $a_i = 1$ if $u_i(1, a_{-i}) > u_i(-1, a_{-i})$ and it chooses $a_i = -1$ if $u_i(1, a_{-i}) \leq u_i(-1, a_{-i})$.

Hopfield (1982) assumed symmetric connections, i.e. $w_{ij} = w_{ji}$, which is the same as the condition that Γ has a potential function. Then he defined the energy

function

$$H(a) = -1/2 \sum_{i \neq j} w_{ij} a_i a_j + \sum_i h_i a_i,$$

which is equal to $-V$, and showed that H monotonically decreases in the process.

He suggested that a neural network with an energy function is isomorphic with a spin glass model, a model in statistical physics. Because of his finding, analytical tools for spin glass models have been applied to neural networks. Anderson and Rosenfeld (1988, page 458) write of "one of the most important new techniques to have been proposed in neural networks."

7   Concluding Remarks

A potential game, literally, mixes game theory with a concept that is common in physics. The result of this paper has implications concerning both.

First, our result implies that potential games are those allowing a natural extension of the Shapley value to noncooperative games. Though potential games are a very limited class, our result could help with finding other examples of potential games and economic applications.

Second, our result implies that a game has a potential function if it has such symmetry as that of BSI games. The nature of the symmetry is the same as that which prevails in physics, of which a special case is discussed in the example of the Hopfield model. This means that some techniques in physics might be applicable in studying potential games, and might bring new insights into game theory.

References

[1] Anderson, J. A., and Rosenfeld, E. (Eds.) (1988). *Neurocomputing: Foundations of Research.* Cambridge, MA: MIT Press.

[2] Blume, L. (1993). "The Statistical Mechanics of Strategic Interaction," *Games Econ. Behav.* **5**, 387–424.

[3] Facchini, G., van Megen, F., Borm, P., and Tijs, S. (1997). "Congestion Models and Weighted Baysian Potential Games," *Theory and Decision* **42**, 193–206.

[4] Guyon, X. (1995). *Random Fields on a Network: Modeling, Statistics, and Applications.* New York, NY: Springer-Verlag.

[5] Harsanyi, J. C. (1959). "A Bargaining Model for the Cooperative n-Person Game," in *Contributions to the Theory of Games* IV (A. Tucker and R. Luce, Eds.), pp. 325–356, Princeton, NJ: Princeton Univ. Press.

[6] Hart, S., and Mas-Colell, A. (1989). "Potential, Value, and Consistency," *Econometrica* **57**, 589–614.

[7] Hart, O., and Moore, J. (1990). "Property Rights and the Nature of the Firm," *J. Polit. Economy* **98**, 1119–1158.

[8] Hopfield, J. J. (1982). "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. Nat. Acad. Sci. USA* **79**, 2554–2558.

[9] Monderer, D., and Shapley, L. S. (1996a). "Potential Games," *Games Econ. Behav.* **14**, 124–143.

[10] Monderer, D., and Shapley, L. S. (1996b). "Fictitious Play Property for Games with Identical Interests," *J. Econ. Theory* **68**, 258–265.

[11] Myerson, R. B. (1977). "Graphs and Cooperation in Games," *Math. Operations Res.* **2**, 225–229.

[12] Myerson, R. B. (1991). *Game Theory: Analysis of Conflict.* Cambridge, MA: Harvard Univ. Press.

[13] Qin, C-Z. (1996). "Endogenous Formation of Cooperation Structures," *J. Econ. Theory* **69**, 218–266.

[14] Rosenthal, R. W. (1973). "A Class of Games Possessing Pure-Strategy Nash Equilibria," *Int. J. Game Theory* **2**, 65–67.

[15] Shapley, L. S. (1953). "A Value for n-Person Games," in *Contributions to the Theory of Games* II (H. Kuhn and A. Tucker, Eds.), pp. 307–317, Princeton, NJ: Princeton Univ. Press.

[16] Slade, M. E. (1994). "What Does an Oligopoly Maximize?," *J. Ind. Econ.* **42**, 45–61.

[17] Ui, T. (1997). "Quantal Responses, Potentials, and Stochastic Stability of Equilibria," Ph.D. thesis. Stanford University.