Computing the Harsanyi function $H(S)$ for n -person cooperative transferable utility normal form games

MELVIN M. SAKURAI

Research Information Services, Honolulu, Hawaii

and

JERRY M. BRENNAN

University of Hawaii, Honolulu, Hawaii

An APL program is introduced for computing the Harsanyi function $H(S)$ of transferable utility games in normal form. The Harsanyi function is a descriptive model that numerically indexes the value of coalitions when they threaten sanctions against each other. This modeling approach distinguishes $H(S)$ from other competing descriptive models and creates a potential for reducing the level of descriptive error in game theory analyses when strategic threats are evident.

In the pages that follow, an APL computer program is described for calculating the *Harsanyi function* of cooperative n -person transferable utility games in normal form. Cooperative n -person game theory is primarily concerned with the normative and prescriptive analysis of multilateral conflicts of interest (von Neumann & Morgenstern, 1944). These conflicts are labeled *games*, and the contesting parties are called *players*. A game is said to be in *normal form* when the respective outcomes associated with each player's set of strategy alternatives are presented tabularly in a matrix.

The general theory is functionally divided into descriptive and solution domains. There are many descriptive function models, all of which attempt to capture critical strategic properties of a game by numerically indexing the expected value of every possible coalition—but from distinct and competing perspectives. Any one of these numeric representations can serve as the basis for solution-theory analyses aimed at prescribing how increments and decrements might be rationally apportioned among the opposing parties to resolve a conflict. A diverse range of solutions is possible, since there are also many competing solution concepts.

Currently, the von Neumann-Morgenstern *characteristic function*, v , is the most widely utilized descriptive model in all forms of game theory application (for a com-

putational program, see Sakurai & Brennan, 1988b). According to the usual derivation of v , conflicting parties can best protect their interests by acting on a presumption that each will inflict the most damaging strategy available on its opponent regardless of consequences for itself. The Harsanyi function (symbolized by H) rests on a contrary view. It contends that coalitions will not ignore the costs of implementing maximally damaging strategies (Harsanyi, 1959, 1963, 1977). This does not necessarily mean that coalitions refrain from threatening their opponents when negotiations fail to yield an acceptable settlement. But it does anticipate a more elaborate bargaining process under threats.

Threats are broadly construed as any conditional commitment to a strategy that will increase the cost of conflict for opponents. Note that nothing limits the damage potential of threats only to opponents. It is quite possible that threatening parties who actually implement such strategies may suffer comparable or even worse losses. Accordingly, the threat process can sometimes have conflicting ramifications. On the one hand, threats are clearly intended to make opponents reluctant about risking further unresolved conflict because of the increased consequential costs they may eventually suffer. Conversely, there is an erosion of the perceived bargaining position and commitment to conflict for threatening parties who must bear significant costs in order to act on a threat.

In threat situations of this sort, prudent coalitions will attempt to strengthen their returns by finding the best strategy compromise between maximizing an opponent's conflict costs and minimizing their own. It happens that this can be accomplished when the difference between a coalition's payoff and that of its opponent is maximized. There is some empirical evidence supporting the veracity of these theoretical arguments (Sakurai, 1986, in press).

This work was supported in part by National Science Foundation Grant SES84-20167 awarded by the Decision and Management Science Program (now Decision, Risk, and Management Science Program) to the first author. Melvin M. Sakurai's address is: Research Information Services, 745 Fort Street, Suite 600, Honolulu, HI 96813. Jerry M. Brennan's address is: Department of Sociology, University of Hawaii-Manoa, Porteus Hall 247, 2424 Maile Way, Honolulu, HI 96822.

THE HARSANYI FUNCTION

The following discussion applies to conflicts of interest or games in which utility is freely transferable. This means players have a common medium of exchange that permits them to negotiate and make transfers so that increments (utilities) are both proportional to the amounts transferred and conserved under reapportionment. For nontransferable utility games, the generalized model is more valid (Harsanyi, 1977, chap. 12).

Let the finite set of players N be denoted by $1, 2, \dots, n$. Each player $i \in N$ has a set of strategy choice alternatives $C_i = (c_1, c_2, \dots, c_n)$ associated with different outcomes that also depend on the strategy choices made by all other players. Every player i has a payoff function $U_i(c_k, c_l, \dots, c_n)$ that defines i 's utility payoff outcome for the combination of i 's k th selection, player j 's l th choice, $\dots$, and player n 's n th selection. Similarly, every coalition of players $S \subset N$ has a set of joint strategy alternatives $C_S = \prod_{i \in S} C_i$ that is the topological product of individual member strategy sets.

Now define $\theta_S = \prod_{i \in S} C_i$ as a joint threat strategy to which coalition S commits against its complement $-S = (N-S)$. Similarly, represent $-S$'s joint threat strategy against S by θ_{-S} . If these threat strategies are actually implemented, their intersection will determine the outcome or quantity of payoff that each coalition can distribute among its members; for S , this is determined by

$$U_S = \sum_{i \in S} U_i(\theta_S, \theta_{-S}).$$

Harsanyi argues that if U_S is given for every $S \subset N$, the players can use some form of marginal dividend payment rule to settle on a final apportionment of payoffs among themselves. The specific dividend rule chosen by Harsanyi is expressed as

$$x_i = \sum_{\substack{i \in S \\ S \subset N}} \frac{(s-1)! (n-s)!}{n!} [U_S - U_{-S}]$$

where x_i is each player's final payoff and s is the number of members constituting S . Some readers will recognize this as the Shapley value solution (Shapley, 1953).

If the marginal dividend rule advanced by Harsanyi is actually used to determine individual player payoffs, S can increase its collective outcome by implementing a threat strategy that optimally maximizes the difference $D_S = U_S - U_{-S}$. The same logic suggests that in order to protect its interests coalition $-S$ must simultaneously minimize D_S by selecting an optimal counter strategy. It follows that the selection of optimal threat strategies can be properly viewed as a zero-sum threats subgame between complementary coalitions. An important consequence of this conclusion is that the subgame then has a unique solution, which corresponds validly with the *minimax value* of D_S . That solution defines the Harsanyi function value H of coalition S . Symbolically:

$$H(S) = \max_{\theta_S \in C_S} \min_{\theta_{-S} \in C_{-S}} [U_S - U_{-S}]$$

where C_S and C_{-S} are the respective sets of joint strategy choice alternatives available to S and $-S$.

Although this formulation increases the Harsanyi function's sensitivity to asymmetries in threat potential, it also causes H to be *nonsuperadditive*. Consequently, there will always be certain coalitions in which members are collectively worse off than if they had remained independent players. *Superadditivity* is a general plausibility condition that assures players of doing at least as well in coalitions as they would in aggregately equivalent but nonaligned subgroups; or, more formally, where $v(S \cup T) \geq v(S) + v(T) \forall S, T \subset N, S \cap T \neq \emptyset$. Nonsuperadditive game models violate this condition for one or more coalitions. As a result, certain coalitions in the Harsanyi representation of a game will always have less ascribed stability (because they are costly to form) than might be attributed by other competing superadditive descriptive models. This suggests that H will tend to depict games as being somewhat less uniformly stable than will normally superadditive models.

To derive the Harsanyi function, each coalition, S , is pitted against its complement in a threats subgame. S 's Harsanyi function value is the average of a strictly cooperative (sum) and a strictly competitive (difference) minimax evaluation of the subgame. The cooperative analysis is functionally simple. It merely determines the maximum wealth (designated S -MAX) available to complementary coalitions if they join together in a grand union. The competitive analysis is more complex.

Represent the threats subgame between S and $-S$ by a matrix A_S with row dimensions defined by C_S , column dimensions by C_{-S} , and outcomes given tabularly. Cell outcome entries are computed by $a_{\sigma\tau} = \sum_{i \in S} u_i(\sigma, \tau) - \sum_{k \in -S} u_k(\sigma, \tau)$ where σ is a selected strategy from C_S and τ is a strategy chosen from C_{-S} . Sometimes S can select a pure joint strategy to maximize its competitive advantage in such a game; other circumstances may require a probability mix of strategy selections.

Let P be a set of probability weights (summing to one) over the pure joint-strategy choices of coalition S , such that p_σ is the probability that S selects joint strategy σ . If σ is an optimal pure joint strategy, then $p_\sigma = 1$. A similar set of probability weights Q can also be defined for $-S$. Every system of this type has a unique optimal strategy solution or minimax *saddle point* given by $\max_{p \in P} \min_{q \in Q} \sum_\sigma \sum_\tau p_\sigma a_{\sigma\tau} q_\tau$. This saddle point represents the competitive minimax difference value of S in a threats subgame with $-S$ and is designated $\text{DIFF}_{S, -S}$.

If an optimal pure joint strategy is available to S , the competitive threats subgame is said to have a *pure strategy saddle point*. This happens when there is at least one element in A_S that is simultaneously minimal in its row and maximal in its column. Thus, if A_S is an $m \times n$ matrix, a_{ih} is a pure saddle point if

$$a_{lh} = \min_{j=1, \dots, n} (a_{lj}) = \max_{i=1, \dots, m} (a_{ih}),$$

where

$$\min_{j=1, \dots, n} (a_{lj})$$

is a minimal element of the collection $(a_{l1}, \dots, a_{lj}, \dots, a_{ln})$, which corresponds with the l th row of A_S , and

$$\max_{i=1, \dots, m} (a_{ih})$$

is the maximal element of column h . Any pure strategy saddle point satisfying this criteria is designated $\text{DIFF}_{S, -S}$.

When there are no optimal pure saddle points, S must use a mixed probability strategy to maximize its competitive advantage. In that case, S 's competitive minimax difference value is given by the primal or dual of a linear program (Charnes, 1953):

$$\begin{aligned} \text{DIFF}_{S, -S} &= \max \gamma, & \text{DIFF}_{S, -S} &= \min \delta, \\ \text{such that } \gamma - \sum_{\sigma} p_{\sigma} a_{\sigma\tau} &\leq 0 & \text{such that } \delta - \sum_{\tau} a_{\sigma\tau} q_{\tau} &\geq 0 \\ \sum_{\sigma} p_{\sigma} &= 1 & \sum_{\tau} q_{\tau} &= 1 \\ p_{\sigma} &\geq 0 & q_{\tau} &\geq 0. \end{aligned}$$

The maximum value (primal) solution is subsequently designated $\text{DIFF}_{S, -S}$.

Recall that the Harsanyi function indexes coalition S 's expected worth by averaging S 's strictly cooperative and strictly competitive value in a threats subgame between S and its complement $-S$. Thus, $H(S)$ and $H(-S)$ are computed by using the formula:

$$H(S) = 1/2(S - \text{MAX} + \text{DIFF}_{S, -S})$$

$$H(-S) = 1/2(S - \text{MAX} - \text{DIFF}_{S, -S}).$$

The grand coalition's Harsanyi function value is taken as $S - \text{MAX}$.

THE APL PROGRAM

This program automatically computes the Harsanyi function for all normal form games having up to nine players. Each individual player is permitted a maximum of nine pure strategy alternatives and every player can have a different number of strategy alternatives. Negative payoff outcome values are also allowed.

The program first examines the overall normal form game matrix to determine $S - \text{MAX}$, the largest row sum. Next a reduced threats subgame (competitive difference) matrix is constructed for each complementary set of coalitions S and $-S$. Cell values are computed as described above. Each matrix is evaluated for pure strategy saddle points that can be designated as $\text{DIFF}_{S, -S}$. Multiple saddle points always have the same numeric value (Charnes & Cooper, 1961) and any one can be selected as $\text{DIFF}_{S, -S}$.

If there are no pure saddle points, the threats subgame is automatically transformed into a linear program (corrected for negative outcome values). A maximum value

(primal) solution is then computed using Smillie's (1983) simplex program and designated as $\text{DIFF}_{S, -S}$.

The respective $S - \text{MAX}$ and $\text{DIFF}_{S, -S}$ values are subsequently used to compute the Harsanyi function value of corresponding coalitions as formulated above.

Input

To execute the program: (1) indicate the number of strategy alternatives for each respective individual player; (2) type HS; and (3) specify the overall normal form game outcome matrix. More than one line may be used for input.

Program syntax requirements are as follows: (1) label individual players using upper case letters (e.g., A, B, ..., N); (2) indicate coalitions by grouping constituent members with no space between player labels (e.g., AC is the coalition between Player A and Player C); (3) number each player's set of pure strategy alternatives sequentially (i.e., 1, 2, ..., n); (4) represent the various pure strategy combinations resulting from individual player strategy selections by vectors (e.g., the vector 1 1 2 signifies that Player A and Player B select their respective first strategy alternatives while Player C selects his/her second alternative); and (5) use lowercase letters corresponding to player labels in combination with appropriate strategy vector subscripts to identify individual player outcomes associated with any particular pure strategy selection combination (e.g., a_{121} represents Player A's outcome for the strategy combination 1 2 1).

The number of strategy alternatives for each player is collectively specified by a vector of appropriate dimension (leaving blank spaces between the elements). For example, the vector 3 2 4 describes a 3-person game in which Player A has three strategy alternatives, Player B two alternatives, and Player C four alternatives.

A vector is also used to specify the normal form game outcome matrix. The vector is constructed by sequentially stringing together successive rows of the outcome matrix. This assumes that rows of the outcome matrix are defined by a complete logically ordered listing of all possible pure strategy selection combinations and columns by alphabetically ordered player labels. Cells within the matrix indicate corresponding outcomes for each player from the respective strategy selection combinations (row labels). For example, Table 1 represents the normal form outcome

Table 1
A 3-Person Two-Choice Normal Form Game Outcome Matrix

Strategy Selection Combinations			Corresponding Outcomes		
A	B	C	A	B	C
1	1	1	a_{111}	b_{111}	c_{111}
1	1	2	a_{112}	b_{112}	c_{112}
1	2	1	a_{121}	b_{121}	c_{121}
1	2	2	a_{122}	b_{122}	c_{122}
2	1	1	a_{211}	b_{211}	c_{211}
2	1	2	a_{212}	b_{212}	c_{212}
2	2	1	a_{221}	b_{221}	c_{221}
2	2	2	a_{222}	b_{222}	c_{222}

matrix of a 3-person game in which each player has two strategy alternatives. This matrix is input by specifying the vector $a_{111} b_{111} c_{111} a_{112}, \dots, b_{122}, \dots, c_{222}$. To compute the Harsanyi function of the example game simply specify: 2 2 2 HS $a_{111} b_{111}, \dots, c_{222}$.

Example/Output

A sample program output is listed in the Appendix. Notice that it adopts the program syntax described above. In the sample game, Player A has three strategy choices while Players B and C each have two choices. This is the same game used in previous articles (Sakurai & Brennan, 1988a, 1988b) to illustrate computation of the characteristic function and constrained game function (Charnes, 1953; Charnes & Sorensen, 1974). Interested readers will find that both competing models differ appreciably from the Harsanyi function in their representation of the same sample game.

Program outputs begin by listing the overall normal form outcome matrix for user confirmation. The program also prints a reduced game matrix for every player and coalition to support the design of experimental test games. Analytic results include: (1) determination of $H(N)$ and associated strategy combination(s); (2) pure strategy saddle points; and (3) linear program maximum value (primal) solutions with strategy selection probability weights (negative value correction coefficients are also printed as indicated).

Smillie's simplex program identifies individual or joint strategy alternatives as consecutively numbered (equation) variables beginning with Variable 2. Variable 1 is reserved for the maximum value solution and only variables having probability weights greater than zero are listed. For example, suppose coalition AB has four joint strategy alternatives but only the first joint strategy (i.e., Variable 2) and third joint strategy (Variable 4) are required for an optimal mixed probability strategy. In that case, the program will only list results for Variables 1 (the primal solution), 2, and 4. Whenever a reduced game matrix contains outcomes of less than zero value, the printed negative correction coefficient must be subtracted from each listed "LEVEL" value to obtain correct strategy selection probabilities.

Program outputs end with a consolidated factorial listing of the Harsanyi function (Searle, 1971). For example,

the Harsanyi function for a 3-person game will be listed: $H(A), H(B), H(AB), H(C), H(AC), H(BC), H(ABC)$.

Language and System Requirements. The program requires STSC/APL*PLUS for the IBM-PC and compatibles with 640K and one floppy disk drive.

Availability. A program listing is available from the first author. In addition, a transfer copy will be made for users who send a blank DS/DD floppy disk formatted for the IBM PC and compatibles.

REFERENCES

- CHARNES, A. (1953). Constrained games and linear programming. *Proceedings of the National Academy of Sciences*, **39**, 639-641.
- CHARNES, A., & COOPER, W. W. (1961). *Management models and industrial applications of linear programming* (Vol. 2). New York: Wiley.
- CHARNES, A., & SORESENSEN, S. (1974). Constrained n-person games. *International Journal of Game Theory*, **3**, 141-158.
- HARSANYI, J. C. (1959). A bargaining model for the cooperative n-person game. In A. W. Tucker & R. D. Luce (Eds.), *Contributions to the theory of games: Vol. IV. Annals of Mathematics Studies*, **40**, 325-355.
- HARSANYI, J. C. (1963). A simplified bargaining model for the n-person cooperative game. *International Economic Review*, **4**, 194-220.
- HARSANYI, J. C. (1977). *Rational behavior and bargaining equilibrium in games and social situations*. London: Cambridge University Press.
- SAKURAI, M. M. (1986). *Threat strategy descriptive models under varying outcome parity conditions: A competitive test of the Harsanyi function, Shapley value function with quarrelling, and characteristic function* (RM-NSF-864). Honolulu, HI: Research Information Services.
- SAKURAI, M. M. (in press). Modelling strategic threats: A competitive test of the Harsanyi function $H(S)$ and characteristic function $v(S)$. *Journal of Conflict Resolution*.
- SAKURAI, M. M., & BRENNAN, J. M. (1988a). *Computing the constrained game function CGF(S) for n-person cooperative transferable utility normal form games*. Manuscript submitted for publication.
- SAKURAI, M. M., & BRENNAN, J. M. (1988b). Computing the von Neumann-Morgenstern characteristic function $v(S)$ for cooperative n-person transferable utility normal form games: LP and saddlepoint solutions. *Behavior Research Methods, Instruments, & Computers*, **20**, 367-371.
- SEARLE, S. R. (1971). *Linear models*. New York: Wiley.
- SHAPLEY, L. S. (1953). A value for n-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games: Vol. II. Annals of Mathematics Studies*, **28**, 307-317.
- SMILLIE, K. W. (1983). *Statpack3 and APL statistical package* (3rd ed.). Edmonton, Alberta, Canada: Department of Computer Science, University of Alberta.
- VON NEUMANN, J., & MORGENSTERN, O. (1944). *The theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

APPENDIX

Sample Output Computing the Harsanyi Function of a 3-Person Normal Form Game

OPTSADLP pLPTAB

Input is a flat matrix or vector.

3 Person game with choices per person: 3 2 2

Game Matrix

ABC	A	B	C
111	-1	100	20
112	13	8	56
121	37	0	37
122	1	-11	8
211	39	9	39
212	21	12	38
221	67	15	11
222	5	3	-6
311	65	17	82
312	9	8	7
321	37	37	1
322	5	9	16

Largest row sum(S-MAX) =164 for 311

Harsanyi Reduced Difference Matrices

GAMMA BC,A DIFF

BC A 1 2 3

11	121	9	34
12	51	29	6
21	0	-41	7
22	-4	-8	20

Input to simplex is:

1	0	0	0	0	0	0	0	0	0
1	-162	-50	-75	1	0	0	0	0	0
1	-92	-70	-47	0	1	0	0	0	0
1	-41	0	-42	0	0	1	0	0	0
1	-37	-33	-61	0	0	0	1	0	0
0	1	1	1	0	0	0	0	1	1

Smillie simplex program results with 41 subtracted.

MAXIMUM VALUE 1.00000

VAR.	LEVEL
1	1.00000
4	-40.00000
5	-8.00000
6	-36.00000
8	-22.00000

GAMMA AC,B DIFF

AC B 1 2

11	-81	74
12	61	20
21	69	63
22	47	-4
31	130	1
32	8	12

Input to simplex is:

1	0	0	0	0	0	0	0	0	0
1	0	-155	1	0	0	0	0	0	0
1	-142	-101	0	1	0	0	0	0	0
1	-150	-144	0	0	1	0	0	0	0
1	-128	-77	0	0	0	1	0	0	0
1	-211	-82	0	0	0	0	1	0	0
1	-89	-93	0	0	0	0	0	1	0
0	1	1	0	0	0	0	0	0	1

Smillie simplex program results with 81 subtracted.

MAXIMUM VALUE 10.83636

VAR. LEVEL

1	10.83636
2	-80.70909
3	-80.29091
4	-62.92727
5	-59.90909
6	-27.09091
8	-53.30909

GAMMA AB,C DIFF

AB C 1 2

11	79	-35
12	0	-18
21	9	-5
22	71	14
31	0	10
32	73	-2

SADDLE = 14

GAMMA C,AB DIFF

C A	1	2	2	3	3
B	1	2	1	2	1

1	-79	0	-9	-71	0	-73
2	35	18	5	-14	-10	2

SADDLE = -14

GAMMA B,AC DIFF

B A	1	1	2	2	3	3
C	1	2	1	2	1	2

1	81	-61	-69	-47	-130	-8
2	-74	-20	-63	4	-1	-12

Input to simplex is:

1	0	0	0	0	0	0	0	0	0
1	-211	-69	-61	-83	0	-122	1	0	0
1	-56	-110	-67	-134	-129	-118	0	1	0
0	1	1	1	1	1	1	0	0	1

Smillie simplex program results with 130 subtracted.

MAXIMUM VALUE -10.83636

VAR. LEVEL

1	-10.83636
5	-129.92727
7	-129.07273

GAMMA A,BC DIFF

A B	1	1	2	2
C	1	2	1	2

1	-121	-51	0	4
2	-9	-29	41	8
3	-34	-6	-1	-20

Input to simplex is:

1	0	0	0	0	0	0	0	0	0
1	0	-70	-121	-125	1	0	0	0	0
1	-112	-92	-162	-129	0	1	0	0	0
1	-87	-115	-120	-101	0	0	1	0	0
0	1	1	1	1	0	0	0	1	1

Smillie simplex program results with 121 subtracted.

MAXIMUM VALUE -1.00000

VAR. LEVEL

1	-1.00000
4	-120.00000
6	-120.00000
7	-79.00000

81.5 76.58182 89 75 87.41818 82.5 164