U. S. AIR FORCE
PROJECT RAND

TWO-PERSON COOPERATIVE GAMES

J. F. Nash

P-172

31 AUGUST 1950

The RAND Corporation
1500 FOURTH ST. • SANTA MONICA • CALIFORNIA

In this paper we shall define the concept of a general two-person cooperative game and develop a concept of a solution for such games. There are two different ways in which the solution may be derived. One is by the use of a set of axioms describing general properties a solution should possess and from which it can be deduced that there is but one possibility in each case. The other proceeds by setting up a model of the negotiation process which the players go through in deciding upon a course of action. This is done in such a way as to obtain a non-cooperative game. [1]

By using the term cooperative we mean to imply that the players have complete freedom of communication and complete information on the structure of the game. Furthermore, there should be the possibility of making enforced agreements, binding either one or both players to a certain agreement or policy.

It is assumed that either player may secure a commitment (enforced policy contract) upon himself if he so desires. But each player is supposed not to have any commitments upon himself before entering into the negotiation involved in this game, or at least none relevant to the situation. And we assume that while one player is obtaining a commitment upon himself the other may do likewise. Finally, we assume that the situation may be regarded as an isolated incident in the life of each player, and not one where a player's behavior could set up an advantageous or disadvantageous precedent.

The possible usefulness of a commitment is fairly clear. If one player can announce to the other that he is bound to accept only the most favorable sort of proposal for an agreement and the other is uncommitted, then the committed player should have an advantage, provided the other is rational and the commitment still allows some mutually profitable arrangement.

The mathematical description of the game is as follows: each player has a convex compact metric space S_i of mixed strategies s_i; there is similar space J of joint strategies; each pair (s_1,s_2) corresponds to a certain joint strategy, σ, and this correspondence induces a mapping $S_1 \times S_2 \rightarrow J$ which is linear on each space S_i; for each joint strategy σ there are two payoffs, $p_1(\sigma)$ and $p_2(\sigma)$, which are linear continuous functions on J. Since the joint strategies are to be employed when agreement has been reached the only significant properties which one of them possesses are its utilities to the players, that is the numbers $p_1(\sigma)$ and $p_2(\sigma)$ corresponding to it. Consequently we need only to know the set of utility pairs, (u_1,u_2) which correspond to some such joint strategy. This set will be simply a set π in the plane whose coordinates are the utility functions of the players, and π will be compact and convex.

The mappings $S_1 \times S_2 \rightarrow J$ and $J \rightarrow \pi$ induce a mapping $S_1 \times S_2 \rightarrow \pi$ which is linear on each space S_i, or bilinear. It may be written $(s_1,s_2) \rightarrow (p_1(s_1,s_2),\ p_2(s_1,s_2))$.

In the actual negotiation model we restrict each player to a special class of commitments which appears to contain enough variety to enable a player to bring all the strong points of his position into the negotiation so that a greater range of possibilities would be useless to a player.

The arrangement of the negotiation in a two stage form with two simultaneous moves by the two players in each stage appears at first to be a very artificial device. It is really simply a convenience, since a one stage form could be used, but would be essentially equivalent to putting the two stage game in normal form, and hence messier to handle.

Now for the formal model:

Stage 0. Players are informed of the situation, may talk it over, if they like.

Stage 1. Each player goes to his attorney and arranges to be forced to play a certain mixed strategy t_i if the two do not eventually reach agreement; t_i is called player i's "threat".

Stage 1.5. The players return from their attorneys and display the commitments they have made.

Stage 2. They return to their attorneys and each commits himself to a "demand" d_i which is a point on his utility scale, the idea being that player i will accept no deal which has utility less than d_i to him.

The payoffs in this model are defined as follows: If there is a point (u_1, u_2) in π such that $u_1 \geq d_1$ and $u_2 \geq d_2$ then the payoff to each player is his demand, d_i. If not, then the payoffs are $p_1(t_1, t_2)$ and $p_2(t_1, t_2)$. The interpretation is that if their demands are compatible they should get what they demanded, but otherwise they must execute their threats. This method of defining payoffs makes each player want his demand to be as large as possible without loss of consistency.

Since the demands are made knowing the threats, the demand game may be considered separately. Let N be the point $(p_1(t_1, t_2),\ p_2(t_1, t_2))$

in π, and let the utilities here be (u_{1o}, u_{2o}).

The game's payoffs may be described as follows:

to player 1 $\quad d_1 g + u_{1o}(1-g)$

to player 2 $\quad d_2 g + u_{2o}(1-g)$

where g is a function $g(d_1, d_2)$ which is 1 for compatible demands and 0 for incompatible demands.

Now this game has discontinuous payoff functions and an infinite number of equilibrium points in pure strategies, in general. Consequently we cannot use these equilibrium points effectively for prediction purposes. However it is possible to obtain more insight into the situation by investigating the stability of these equilibrium points.

To do this we "smooth" the game so as to obtain a continuous payoff function, study the resulting game, and observe the limiting behavior of the equilibrium points of the smoothed games as the amount of smoothing is reduced to zero.

We consider a certain general class of smoothing methods. Various other apparently different smoothing procedures are actually equivalent.

To smooth the game we merely replace the discontinuous function g by a continuous one, so that instead of having either consistency or inconsistency for each pair of demands we have a probability of consistency, represented by g. This sort of thing would arise whenever there is uncertainty as to the precise shape of the set π. Let us assume g is a differentiable function and that $g=1$ on π, tapering off towards 0 as one moves away from π.

The payoff functions are now simply

$$p_1 = d_1 g + (1-g) u_{1o}$$

and

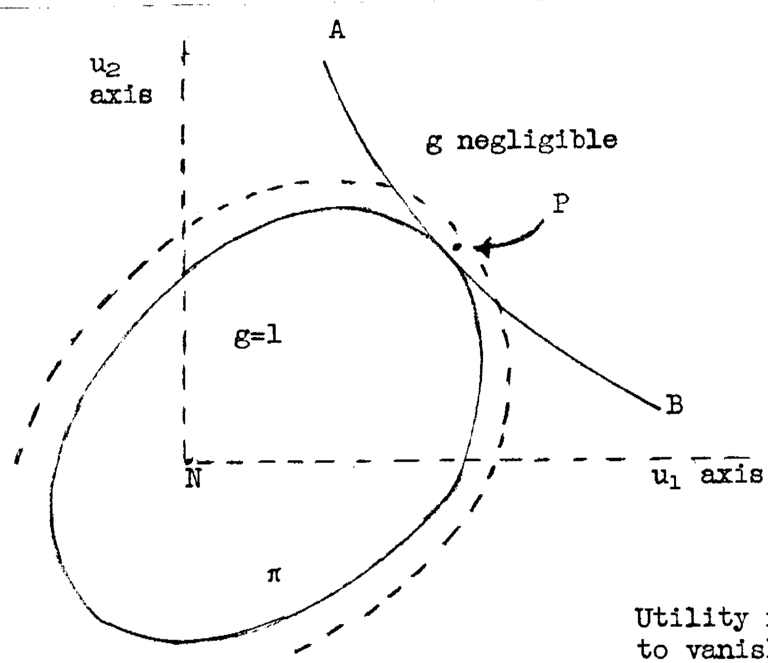
$$p_2 = d_2 g + (1-g) u_{2o} .$$

We may assume the utility functions, so chosen that $u_{1o} = u_{2o} = 0$ which gives us

$$p_1 = d_1 g, \qquad p_2 = d_2 g.$$

Now consider a point at which $d_1 d_2 g$ assumes its maximum value and where $d_1, d_2 \geq 0$. Here we will have $d_1 g$ maximized for fixed d_2 and $d_2 g$ for fixed d_1. Consequently (d_1, d_2) will be an equilibrium point. Furthermore, if g is free of irregularities, it will be the only equilibrium point. If ρ is the maximum of $u_1 u_2$ on π then we must have $d_1 d_2 g \geq \rho$ and hence $d_1 d_2 \geq \rho$ since $0 \leq g \leq 1$.

Figure 1 shows the set π, the point N, the point P where $d_1 d_2 g$ is a maximum, and the hyperbola AB which touches π at the point Q where $u_1 u_2$ is a maximum over π (remember u_{1o} and u_{2o} are now zero)



Utility functions chosen so as to vanish at N.

Figure 1.

P must lie above the hyperbola AB but near enough to π so that g is near one.

Now as less and less smoothing is used g will decrease more and more rapidly on moving away from π and hence P must be nearer and nearer to π. In the limit P must approach Q since it's always above AB. Consequently we take Q for the solution. The point Q gives the appropriate demands, and the values, for the game by its coordinates.

If (d_{10}, d_{20}) are the coordinates of Q they form an equilibrium point of the unsmoothed game which is the only necessary limit of equilibrium points of the smoothed games.

Having treated the second part of our two stage negotiation game we may use this solution to obtain payoff functions for the first stage and then proceed to solve that.

It is clear that the first stage, the presentation of threats, determines the point N and that this in turn determines the outcome of the second stage.

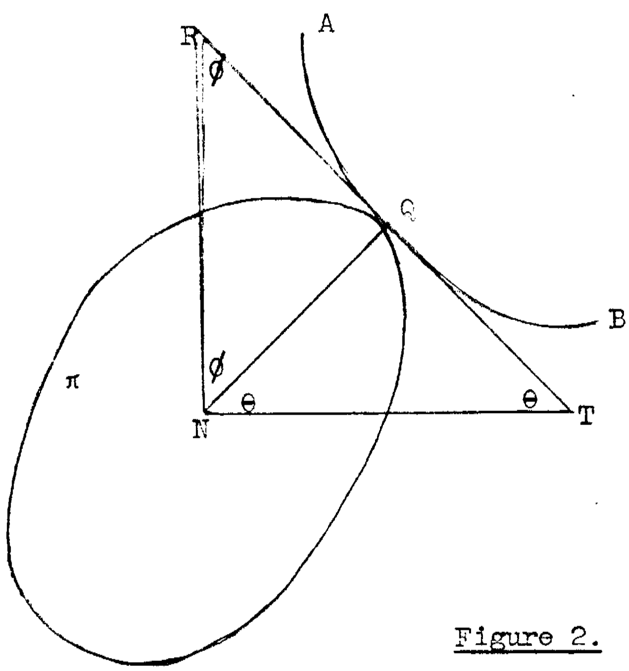


Figure 2.

The point Q determined by a point N has certain geometrical properties. Q bisects the segment of the tangent to the hyperbola AB at Q cut off by the vertical and horizontal axes through N. Also the lines QN and RT have equal, though opposite, slope.

This gives us a method for determining, given a point Q, which points, N, in π would give Q as a solution point. One takes any contact line through Q to π (in general this will be unique) and then draws a line through Q of equal but opposite slope. Any point, N, on the segment of the latter line cut off by π will lead to Q as final point.

We have been glossing over a special degenerate case. If π has a flat part on top (u_2 maximized at more than one point on π) then we must have a convention to determine what Q to take if N lies on the top of π.

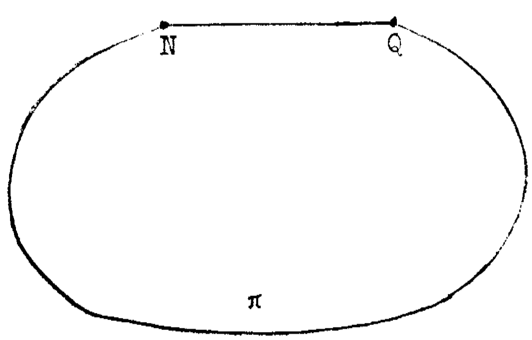


Figure 3.

We set up the convention that Q shall be taken always so that no Q' has $u_i(A') \geq (Q)$ and $u_1(Q') + u_2(Q') > u_1(Q) + u_2(Q)$ [no weak dominance] . This makes Q continuously dependent on N, assuming a similar convention on vertical flat spots. We shall use this convention only as a convenience, and it will not affect our final conclusion.

Now the method described before for finding the points N which lead to a particular Q induces a ruling of the set π by straight line segments which may intersect only on the upper-right edge of π. On each of these lines a payoff may be assigned for each player by taking his utility for the point Q corresponding to that line. Thus a payoff to each player is assigned to each point N and these payoffs vary continuously with N.

But N is just $(p_1(t_1,t_2),\ p_2(t_1,t_2))$ and hence these payoffs are determined by t_1 and t_2. We shall now show that the payoff functions are quasi-concave. This means that if $\mathcal{P}_1$ and $\mathcal{P}_2$ are the induced payoff functions then

$$\min \left[\mathcal{P}_1(t_1,t_2),\ \mathcal{P}_1(t_1',t_2)\right] \leq \mathcal{P}_1(\alpha t_1 + (1-\alpha)t_1',t_2)$$

for any t_1, t_1', t_2 and $0 \leq \alpha \leq 1$; and a similar condition on $\mathcal{P}_2$. These properties follow easily from consideration of the geometrical definition of $\mathcal{P}_1$ and $\mathcal{P}_2$.

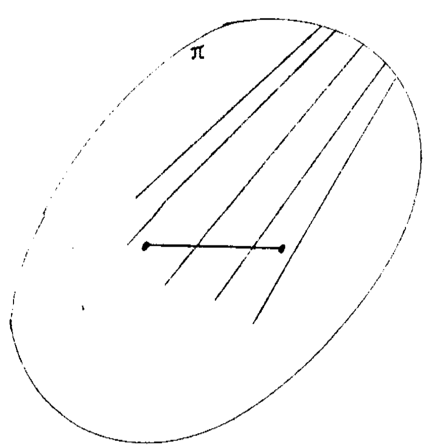


Figure 4.

The set of points N determined by threat pairs of the form $(\alpha t_1+(1-\alpha)\ t_1', t_2)$ will be a straight line segment since p_1 and p_2 are linear. Consequently, since π is ruled by non-intersecting straight lines one end of this segment will be on the lowest and furthest right of the rulings (worst for player two) and hence here $\mathcal{P}_2$ will be at its minimum over the whole segment. (See Figure 4.) Thus it is clear that $\mathcal{P}_2$ (and $\mathcal{P}_1$ similarly) must be quasi-concave.

Now when the payoff functions are continuous and quasi-concave we may use the Kakutani theorem to show there is an equilibrium point. For any t_{1o} the set of t_2's such that $\mathcal{P}_2(t_{1o},t_2)$ is maximized will be a closed convex subset of S_2 from the quasi-concavity of $\mathcal{P}_2$ in t_2. Two such threats t_{1o} and t_{2o} would determine two sets of counter strategies and their product would be a compact convex subset of $S_1 \times S_2$. Karlin's generalized Kakutani fixed point theorem shows then that there must be an equilibrium pair of threats (t_{1o},t_{2o}) such that neither player can improve by changing his threat. We now observe that over all combinations of threats $\mathcal{P}_1$ and $\mathcal{P}_2$ are directly interrelated, so that if one is known so also is the other. Each is a monotone decreasing concave function of the other. Since we have equilibrium, $\mathcal{P}_2(t_{1o},t_2) \leq \mathcal{P}_2(t_{1o},t_{2o})$ for any t_2, hence $\mathcal{P}_1(t_{1o},t_2) \geq \mathcal{P}_1(t_{1o},t_{2o})$ and therefore t_{1o} <u>assures</u> player one of the payoff $\mathcal{P}_{1o} = \mathcal{P}_1(t_{1o},t_{2o})$, and similarly for t_{2o}. Since no higher payoff could be assured, t_{1o} has the maximin property. Since the second player can assure $\mathcal{P}_{2o} = \mathcal{P}_2(t_{1o},t_{2o})$, but can do no better against t_{1o}, t_{1o} has the minimax property as well.

These threats are actually pure strategies in the negotiation game and so a question arises as to whether some other equilibrium point in mixed strategies (or pure too) might give a player more than $\mathcal{P}_{10}$. However the payoffs at such a point would have to define a point of π and $(\mathcal{P}_{10}, \mathcal{P}_{20})$ is a point on the upper-right edge of π which makes it such that neither player can be better off than at $(\mathcal{P}_{10}, \mathcal{P}_{20})$ without making the other worse off. Therefore for any equilibrium point the payoffs will be $(\mathcal{P}_{10}, \mathcal{P}_{20})$ and hence we take these numbers as the values of the game.

A slight qualification should be put in relevant to the degenerate cases. The convention we introduced concerning a flat top or side of π will have effect only when one player has a threat which guarantees him, without any subsequent bargaining, as high a utility as he could possibly get. In other words he can force N to be a point which is as desirable to him as any other point in π. Thus he has no real incentive for cooperation although this might be helpful to the other player. Here the convention says that he will make the other player as well off as he can without hurting himself any.

We believe this is really artificial and that the value of the game to the other should be determined only by the range between this "automatic charity" solution and the value he can assure himself by some threat, upon the assumption that his opponent will use a threat which assures him (the opponent) directly his maximum obtainable utility.

Axiomatic Treatment

In this approach we shall use a set of axioms which lead to the "automatic charity" solution mentioned before in the degenerate cases. Axioms which lead to a solution concept which is less defined in the degenerate cases may be set up based on the continuity properties of this sort of solution, but they are much more cumbersome.

We use the same representation of the problem as before, with the sets S_1, S_2, and π. The solution will be a point, P, in π. The axioms follow.

(1) For each game $\left\{S_1, S_2, \pi\right\}$ there is a unique solution $P(S_1, S_2, \pi) \in \pi$.

(2) If $P' \in \pi$ and $u_1(P') \geq u_1(P)$ and $u_2(P') \geq u_2(P)$ then P' is P.

(3) A linear transformation of the form $u_i' = au_i + b$, $a > 0$, applied to either utility function does not change the result.

(4) The determination of the solution does not depend on which player is called 1, which 2. (The symmetry axiom.)

(5) If $S_1' \subset S_1$ then $u_1 \left[P(S_1', S_2, \pi)\right] \leq u_1 \left[P(S_1, S_2, \pi)\right]$, and similarly for player 2.

(6) If $\pi' \supset \pi$ and $\pi \supset P(S_1, S_2, \pi')$ then $P(S_1, S_2, \pi') = P(S_1, S_2, \pi)$.

(7) There is a pair (s_1, s_2) of strategies, one from S_1 and one from S_2 such that $u_1 \left[P(S_1, S_2, \pi)\right] \leq u_1 \left[P(s_1, s_2, \pi)\right]$, and similarly for player 2.

Thus, for either player, there is some way in which the two players' ranges of strategy choices can be restricted so as to make him no worse off. (The other may be injured by this.)

These axioms involve the weak dominance principle, (Ax. 2), and for this reason they lead to the "automatic charity" solution. One shows that they give this solution by first considering the case when each space S_i consists of but one point. Here the axioms are essentially those used in the paper "The Bargaining Problem" [2] and lead to the solution developed there.

Since either player has an "optimal threat", a particular strategy such that no matter what specific strategy is chosen for the other, he will always obtain p_{10} (or p_{20}) it follows from Axiom 7 that P is (p_{10}, p_{20}).

REFERENCES

[1] Theory of Non-cooperative Games, by J. Nash; reproduced at RAND.

[2] The Bargaining Problem, J. F. Nash, Jr., Econometrica Vol. 18, No. 2, April, 1950.

je