Discussion Paper No. 685

ON COMMUNICATION, BOUNDED COMPLEXITY AND COOPERATION

by

Eitan Zemel*

November 1985
Revised June and November 1986

---

*Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60201.

Support, in part, was provided by NSF Grant No. ECS8121741.

ON COMMUNICATION. BOUNDED COMPLEXITY AND COOPERATION

Abstract

Neyman has shown that bounded rationality can lead to cooperation in the Finitely Repeated Prisoner's Dilemma Game, if the game is conducted by finite automata of fixed size. We view Neyman's work from the perspective of communication. In particular we present an extremely simple, universal and effective communication protocol that can be used for virtually any finitely repeated game, and for any set of individual rational actions, yielding this set as an equilibrium, provided a wide enough communication channel is available. When the communication channel is narrow, say binary, we give tight upper and lower bounds on the number of states which still allow for cooperation. The communication scheme can be used in situations other than finitely repeated games, in general allowing players to effectively diminish the computational power of their opponents.

The effect of bounded computational power on the behavior of players in a competitive situation has been recently analyzed in a variety of settings [A2, AR, B, KS, MW, N, R]. In particular, Neyman [N] has shown that bounded complexity can lead to cooperation in the Finitely Repeated Prisoner's Dilemma Game (FRPDG) where the play is conducted by finite automata of fixed size. In this note we view Neyman's results from the perspective of communication. Specifically, we demonstrate how the ability to communicate, in the context of bounded rationality, enables players to reach equilibria which are not attainable otherwise. To demonstrate the broad principle, we start with the following example.

Consider a card game such as "Black Jack." It is well-known that players who "count" in such games can achieve higher payoffs than could be sustained in a long term equilibrium. One could prevent counting by frequent reshuffling of the deck or by expelling apparent offenders, but there is also another way. Specifically, one can require players to perform, from time to time, some simple memory or arithmetic tasks such as repeating long sequences of digits, adding several numbers, etc. Such tasks, while easily done on their own, can be designed such that they are excessively difficult if one is also concentrating on counting cards. Thus, a successful performance of the task can serve as a proof that player is not counting. On the whole, the existence of such a proof can benefit all parties involved.

We follow in this note closely the basic framework and notation of Neyman [N]. The only new feature we add here is equipping the players with a formal channel of communication which allows them to send messages to each other, independently of the actual moves of the game. As will be revealed shortly, this ability offers several advantages. First, it allows players to achieve complete cooperation, avoiding the waste which is inherent in the scheme of [N]. Also, the approach can yeild extremely natural and simple strategies, employing an effective communication scheme which is independent of the game being played and which can be used universally in situations in which it is desirable to waste a certain fraction of one's opponents' computational power. As an added bonus, the analysis is quite simple and one can get exact (as opposed to asymptotic) results.

We begin by introducing the results of [N]. In particular, consider the Finitely Repeated Prisoners' Dilemma Game (FRPDG), whose basic stage game is given by the matrix G below.

	D_2	F_2
D_1	1,1	4,0
F_1	0,4	3,3

Let G^N denote the N stage repeated version of G. For a general repeated game, denote the set of actions available to a player at each stage by A^i, to his opponents by A^{-i}, and to the entire set of players by A. A finite automaton for player i is a four-tuple $FA^i = \langle S^i, q^i, f^i, g^i \rangle$ where S^i is a finite set, $q^i \in S^i$, $f^i: S^i \to A^i$ and $g^i: S^i \times A^{-i} \to S^i$. Intuitively, S^i is the set of possible states of the automaton, q^i is the initial state, $f^i(q)$ is the action taken by player i when in state q, and g^i describes the transition from state to state; if at state q the other players choose the action tuple a^{-i}, the automaton's next state is $g^i(q,a^{-i})$. The size of the finite automaton is the number of states.

For the prisoners' dilemma game, let the possible actions A^i in each stage be F and D ("Friendly" and "Deviate"). For any N = 1,2,... and positive integers s_1, s_2, define the two person game $G^N(s_1,s_2)$ as follows: the pure strategies of player i (i = 1,2) are all the finite automata of size s_i. We use as payoff for the repeated game the sum of payoffs in each of the stages. However, the results hold virtually unchanged for any reasonable definition of payoffs, for example discounted sum. It is well known that the only equilibrium strategy in G^N is to play D continuously. However, for $G^N(s_1,s_2)$, Neyman has shown:

A.1 If $2 \leq s_1, s_2 \leq N - 1$, then there are equilibrium strategies in $G^N(s_1,s_2)$ which result in the play (F,F) at each stage.

A.2 If either s_1 or s_2 is at least N, there are no equilibrium strategies in $G^N(s_1,s_2)$ which result in the play (F,F) at each stage.

Another simple observation of a similar nature is in the spirit of Proposition 2.3 of Megiddo and Wigderson [MW]:

A.3 If both s_1 and s_2 are at least N, then <u>no fixed</u> trajectory of moves, except for the constant play of (D,D), can be achieved as a result of an equilibrium of $G^N(s_1,s_2)$.

The main contribution of [N] is an elegant and surprising result for the case of machines with more than N states. Its effect is, asymptotically, to mitigate A.2 and A.3 considerably:

A.4 For any integer k there is N_0, such that if $N \geq N_0$, and $N^{1/k} < s_1, s_2 < N^k$, there is a mixed strategy equilibrium in which the payoff to each player is at least $3 - 1/k$.

A.1 is due to the fact that the only possible profitable deviation from the friendly strategy is at the last stage of the game where no reprisal is possible; however, a machine with less than N states cannot recognize that stage and thus cannot implement such a deviation. The basic idea behind A.4 is the following. Each player (machine) plays according to a complex pattern of F's and D's which requires it to spend a large fraction of its computational resources (states) just for determining at each stage what is the next move. Any deviation from the required pattern immediately leads the opponent to stop cooperation and play D throughout. The pattern of F's and D's is chosen so that the number of states which remain free is seen to be too small to allow for deviating precisely at the last stage (counting to N).

This leads naturally to the basic idea of this note, namely, equipping the machines with a separate channel of communication which is independent of the actual "play" of the game. Specifically, at each stage of the game, in addition to choosing an actual action a_i, each machine also sends a message chosen from a given domain, M. Both the action, a_i, and the message, m_i, become available to all players and both could be used as a basis for their response in subsequent stages.

We start the analysis by considering the case of a wide communication channel (large message space). Subsequently, we analyze the case of a small (binary) message space.

2. Large Message Space

We address in this section the case of a large message space. We start by analyzing the situation for two players, playing (FRPDG).

A.5 Let $s_2 \geq s_1 \geq 2$. Then, for $N > \max \{ 3, 2.5 + (s_2 - 2/|M| \}$ there exists an equilibrium of $G^N(s_1,s_2)$ which results in the constant F play for both players.

<u>Proof</u>: Let $M_1 \subseteq M$, and $M_2 \subseteq M$ be two specified subsets of the message space. Each machine i chooses randomly a message $m_{-i} \in M_{-i}$ and sends it to its opponent at stage 1, together with a friendly F move. The receiving machine (machine -i) is expected to repeat message m_{-i}, together with the play F, at each subsequent stage of the game. Any other combination of message-action pair, at any stage of the game, causes machine i to stop cooperation and play D through the end of the game. To show that these strategies are in equilibrium, we have to demonstrate two facts:

(a) Each player can implement his part of the strategy using a machine with s_i states.

(b) No machine with s_i states can achieve a higher expected payoff by using a different strategy.

First, we note that player i can implement the strategy using a machine with $|M_i| + 2$ states: one beginning state, one "deviating" state (to which the machine reverts if it detects a deviation by its opponent), and a set of $|M_i|$ states, one for each possible message sent to him by his opponent. Next examine the number of states needed to implement the strategy of deviating at the last stage without incurring <u>any</u> risk of being detected by the opponent. By the Myhill-Nerode Theorem (see, e.g., [AHU]), this number is equal to the different number of equivalence classes of histories faced by each player, where two histories are called equivalent only if any possible continuation of these two histories requires the same response for both. Since a player deviating at the last stage is effectively counting to N, no two histories of different lengths $1 \leq t_1 < t_2 \leq N-1$ could be equivalent; also, for each stage $2 \leq t \leq N - 1$, there are $|M_i|$ possible nonequivalent histories, one for each possible message received by player i. This comes to a total of $2 + (N - 2)|M_i|$. A player with a smaller number of states, which is nevertheless counting to N, must run a certain risk of sending the wrong message at some stage during the game, causing his opponent to stop cooperating. Such a risk is worth taking only if it results in expected payoff of more than 3N. A short reflection will reveal that such a risk should be taken only at the N - 1st stage, and only if the probability of being detected does not exceed 1/2. This allows one to count with up to $(1/2)|M_i|$ fewer states, i.e., a total of at least $2 + (N - 5/2)|M_i|$. We thus need to show that M_i can be chosen such that $|M_i| + 2 \leq s_i < 2 + (N - 5/2)|M_i|$. This can be done by letting $|M_i|$ smallest integer such that $s_i < (N - 5/2)|M_i|$. []

To demonstrate the scope of A.5, consider the message space consisting of potential telephone numbers, i.e. 7 decimal digits. Any pair of machines with $3 \leq s_1 \leq s_2$ states for any length of game $N > 2.5 + (s_2 - 2)10^{-7}$, can achieve complete cooperation using the very simple protocol of A.5.

It is instructive to examine the role of randomization in the equilibrium achieved in A.5. As can be easily verified, it is limited to the "communication" part of the strategy (namely, a random choice of a message). It is not possible to remove randomization completely, even for a large message space M. In fact, one can establish the following analog of A.3:

A.6. If both s_1 and s_2 are at least N, the only pure equilibrium of $G^N(s_1,s_2)$ is the constant D play for both players, even if messages are allowed.

Proof: Any predetermined set of moves and messages, of size not exceeding N can be achieved with N states by a machine by simply labeling the states 1,...,N, and defining the function f so that the right action-messsage is sent at each stage. In particular, it is always possible for each machine to deviate at the last stage, and the standard argument of the finitely repeated game applies. □

Clearly, A.5 can be generalized to a general two person finitely repeated game. It can be interpreted as asserting that, in the presence of a wide communication channel, cooperation can be maintained in the sense that the constant play of any pair of actions whose payoffs are above the individually rational levels can be made into an equilibrium. For the case of several players, we basically have the same result, provided some joint randomization, [Al], can be achieved prior to the beginning of the game. Below we demonstrate this result for games with at least four players. Specifically, let V be a given r-person, $r \geq 4$, game and let $F = (F^1,\ldots,F^r)$ be a set of actions. For each player i, let $P_i = (P_i^1,\ldots,P_i^r)$ be an action set which achieves player i individually rational level. We refer to the set P_i as "punishment" for i. We assume that for each player, either F^i is a best response to F^{-i} (such players have no incentives to deviate from F), or that his payoff from F exceeds his individually rational level (such players can be "punished"). For each player in the latter set, we assume that the gain from deviation (relative to the payoff from F) is at most equal to the loss due to punishment. (This assumption is made for convenience and so that the results parallel those of A.5. In its absence, some of the constants of A.7. may have to be slightly modified.)

A.7: Let $s_r \geq s_{r-1} \geq \dots \geq s_1 \geq 2r + 1$, $r \geq 4$. Then for $N \geq \max \{ 2.5 + (s_1 - 2)/(s_1 - 2r), 2.5 + (s_r - 2)/|M| \}$ there exists an equilibrium of $V^n(s_1,\dots,s_r)$ which results in the constant F play.

Proof: In the action space, each player is expected to play constantly according to F, with any deviation resulting in all players reverting to the punishment strategy P_i.

In the rest of the proof we examine the communication protocol. Let the indices $1,\dots,r$ be regarded as integers modulo r, so that one could think of the players as positioned on a circle. For each player i, we nominate a committee, C(i), composed of all players excluding i himself with player i - 1 acting as chairman. Let $M_i \subseteq M$ be a specified subset of the communication domain. Before the game starts, each committee C(i) agrees on a random message, $m_i \in M_i$, unknown to i. At stage one, each chairman i - 1 announces the message m_i chosen by the committee he chairs. In subsequent stages, each player is expected to repeat the message announced by his chairman. If i fails to repeat the correct message at any stage, then all players revert to P_i. If a chairman i - 1 sends the wrong message, $m_i' \neq m_i$ in stage 1, then all players in C(i) send an agreed upon message m_i^*, which identifies i - 1 as an offender. In the subsequent stage, all players punish i - 1 by playing P_{i-1}. If the members of C(i) do not agree on the message m_i^*, one uses majority rule. In case of more than one violator, one uses any consistent tie-breaking rule, e.g., always punish the lowest indexed offender. The number of states required for implementing this protocol is as follows. Each nondeviating player needs one starting stage, a set of $|M_i|$ playing states, a set of r punishing states, and a set of r − 1 states for identifying a deviating chairman. This comes to a total of $|M_i| + 2r$ states. On the other hand, a player who contemplates counting to N needs at least $2 + (N - 5/2)|M_i|$ states. Thus, the theorem is proved if subsets $M_i \subseteq M$, $i = 1,\ldots,r$, can be chosen so that, for each player:

$$|M_i| + 2r \leqslant s_i < 2 + (N - 5/2)|M_i|$$

This can be achieved by letting $|M_i|$ be the smallest integer such that the right inequality holds. []

To demonstrate the scope of A.7, consider the case of 10 players. Then, for $|M| = 10^7$, $s_1 \geqslant 25$ one can get cooperation for $N \geqslant 5 + (s_r - 2)\,10^{-7}$.

3. Binary Message Space

We now consider the case of a small message space, say $M = \{0,1\}$. We refer back to FRPDG with $s_2 \geqslant s_1$. Naturally, one can "blow up" the size of M by using messages which extend throughout several stages of the game. Using strings of length ℓ, one gets a message space of size 2^ℓ, so that one expects a version of A.5 to work as long as $s_2 \leqslant 2^{O(N)}$. This is basically true:

A.8. Let $6 \leqslant s_1 \leqslant s_2 \leqslant 2^{(s_1-1)/2}$. Then, if $N \geq \min\{ s_1 , 4 \log s_2 + 7 \}$ there exists an equilibrium in $G^N(s_1,s_2)$ which

A.8. Let $6 \leqslant s_1 \leqslant s_2 \leqslant 2^{(s_1-1)/2}$. Then, if $N \geq \min\{ s_1 , 4 \log s_2 + 7 \}$ there exists an equilibrium in $G^N(s_1,s_2)$ which results in the constant F action at each stage.

Proof:

We analyze below the communication protocol.

(a) First, assume that $s_1 < N \leqslant s_2$, i.e., player 1 cannot count to N. Thus, we need not worry about him deviating at the last stage. To make sure player 2 does not deviate, player 1 sends k binary bits, $m_1,\ldots,m_k$ for the first k stages of the game, repeating these bits in a cycle after that (i.e., in stage kr + j the bit m_j is sent again. Player 2 is required to send 0 for the first k periods, and then repeat the cycle of messages, $m_1,\ldots,m_k$ to the end of the game. The number of states required by player 1 for this policy is 2k + 1 (k states for the first k stages, additional k states for the cycle, and one state in case the opponent deviates). On the other hand, player 2 needs at most $2 + 2^k$ states if he is not counting and at least $(N - 2k + 2)2^k - 2$ states if he wishes to count to N without any risk of being detected. ($1 + 2 + \ldots + 2^{k-1}$ states for the first k stages, in which the message is being transmitted, $2^{k-1} + 2^{k-2} + \ldots + 2 + 1$ for the last k stages, and 2^k states for each of the middle N - 2k stages. Allowing for a risk of at most one-half of being detected before stage N reduces this number by at most $2^k/2$.) Thus, if an integer $k \geqslant 2N$ can be found such that

$$2 + 2^k \leqslant s_2 < (N - 2k + 1.5)2^k - 2$$

$$1 + 2k \leqslant s_1$$

then the cooperative strategy is in equilibrium. A choice which works is $k = \lfloor \log(s_2 - 2) \rfloor$.

If $s_1 \geq N$, then both machines need to monitor each other. Let K and k be constants to be determined later such that $K \geq k$. Let x and y be random binary vectors of sizes k and K and chosen by player 2 and 1, respectively. In stages 1,...,K, player 1 sends the message y and player 2 sends the message $\bar{0}$,x, where $\bar{0}$ is a sequence of K - k zeroes. Then, player 1 is expected to repeat x and player 2 is expected to repeat y cyclically throughout the game. The number of states needed for doing this is $(k + 1)\cdot 2^K$ and $(K - k) + (K + 1)\cdot 2^k$ by players 2 and 1, respectively. (Player 1, for instance, can achieve this with the following states: $K - k$ for the first $K - k$ stages in which he is receiving no information; $1 + \ldots + 2^{k-1}$ for the following k stages, in which he is gradually being informed about the message x; $K\cdot 2^k$ for the following $N - K$ stages, in which he needs to monitor y and repeat x, and 1 state in case player 2 deviates).

On the other hand, if a player expects to increase his expected payoff by "counting" to N, then the required number of states is at least $(N - 2K + 1.5)\cdot 2^K - 2$ for player 2 and at least $(N - 2K + 1.5)2^k + 2(K - k - 1)$ for player 1. Thus, the theorem is proved if we can find k and K such that

$$(K + 1)2^k + (K - k) \leq s_1 < (N - 2K + 1.5)\cdot 2^k + 2(K - k - 1)$$

$$(k + 1)\cdot 2^K \leq s_2 < (N - 2K + 1.5)\cdot 2^K - 2$$

This can be achieved by letting K be the largest integer such that $(K + 1)2^K \leq s_2$ and k be the largest integer such that $(K + 2)2^k + (K - k) \leq s_1$. []

The protocols of A.5, A.7 and A.8 are extremely simple, basically requiring each player to repeat indefinitely a sequence of bits sent to him at the beginning stages of the game. Naturally, one can devise much more complex protocols, which should be harder to compute. A natural question to ask is whether a complex protocol could be used to "waste" more states than allowed by A.8. This may be the case, but the room for improvement is rather small. Below we establish that $0(2^{N/2})$ for $M = \{0,1\}$ is the maximal number of states which still allow cooperation.

A.9. Let $K_n = 2 \cdot 2^{N/2} - 2$ for n even and $3 \cdot 2^{(N-1)/2} - 2$ for n odd. Then if $s_2 \geqslant K_n$, there exist no pattern of communication over $M = \{0,1\}$ which yields the trajectory (F,F) as an equilibrium. In fact, the only fixed trajectory which can be in equilibrium is the constant (D,D) play.

Proof: Denote the history of messages sent by player i up to stage t by h_t^i. We can think of h_t^i as an integer in the range 0 to $2^{t-1} - 1$ (i.e., the integer whose binary representation is h_t^i). Let the communication protocol be summarized as follows. In stage t, if the messages sequence, h_t^{-i} has been observed by machine i, it is required to respond with message $g^i(h_t^{-i}) \in \{0,1\}$. Clearly g^i can be implemented directly by f^i if machine i possesses one state for each possible message history. That requires a total of $S = \sum_{j=1}^{N} 2^{j-1} = 2^N - 1$ states. To get away with fewer states, we have to combine several equivalent histories into each state. This can be done as follows. Obviously, for the last stage a player needs to keep only one state, since he plans to deviate anyhow and the message he sends is of no consequence. For the $N - 1^{st}$ stage a player needs only two states, one "for each" of the two possible messages 0 and 1. For the N - 2 stage 4 states are needed, representing the message to send in that stage and which of the two possible N - 1 stage states to enter upon receiving a message of 0 and 1 in state N - 2. Continuing in this fashion we get that the total number of states required is

$$1 + 2 + \ldots 2^{(N/2)-1} + 2^{(N/2)-1} + \ldots 2 + 1 = 2(2^{N/2} - 1)$$

for n even, and $3 \cdot 2^{(N-1)/2} - 2$ for the case of n odd. □

We note that for the case of $s_1 < N$, one can acheive cooperation as long as $s_2 \leqslant O(2^{N/2})$ which is the same order of magnitude as the upper bound of A.9. Thus we have a rather sharp definition of what can or cannot be achieved, using a binary communication protocol. It is still an open question whether for the case of $s_1 \geqslant N$, one can achieve cooperation for the case of $s_1 > O(2^{N/4})$. I conjecture that this can be done iff $\log s_1 + \log s_2 \leqslant O(N/2)$.

References

[A1] R. J. Aumann, "Subjectivity and Correlation in Randomized Strategies," J. of Math. Econ., 1974, pp. 67-96.

[A2] R. J. Aumann, Survey of repeated games, in "Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern," Bibliographisches Institut Mannheim/Wien/Zurich, 1981.

[AHU] A. V. Aho, J. E. Hopcroft and J. D. Ullman, "The Design and Analysis of Computer Algorithms," Addison Wesley, Reading, Massachusetts, 1974.

[AR] D. Abreu and A. Rubinstein, "The Structure of Nash Equilibria in Repeated Games with Finite Automata," manuscript, 1986.

[AS] R. J. Aumann and S. Sorin, Bounded rationality forces cooperation, manuscript, 1985.

[B] E. Ben Porath, Repeated games with bounded complexity, manuscript, 1986.

[KS] E. Kalai and W. Stanford, "Finite Rationality and Interpersonal Complexity in Repeated Games," Discussion Paper No. 679, Center for Mathematical Studies in Economics and Management Science, Northwestern University, 1986.

[MW] N. Megiddo and A. Wigderson, On play by means of computing machines, manuscript, 1985.

[N] A. Neyman, Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma, manuscript, 1985.

[R] A. Rubinstein, Finite automata play--the repeated prisoners' dilemma, Research Report No. 149, Department of Economics, The Hebrew University, Jerusalem, Israel, 1985.