

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1271

Note: Cowles Foundation Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. Requests for single copies of a Paper will be filled by the Cowles Foundation within the limits of the supply. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

FAULTY COMMUNICATION

Stephen Morris

August 2000

Faulty Communication*

Stephen Morris
Cowles Foundation,
Yale University,
P.O.Box 208281,
New Haven CT 06520, U. S. A.
`stephen.morris@yale.edu`

July 2000
revised August 2000

Abstract

The electronic mail game of Rubinstein (1989) showed that a lack of common knowledge generated by faulty communication can make coordinated action impossible. This paper shows how this conclusion is robust to having a more realistic timing structure of messages, more than two players who meet publicly but not as a plenary group, and strategic decisions about whether to communicate.

JEL: C72, D8

1 Introduction

When a group of individuals must coordinate on a risky course of action, it is valuable to collect the group together and publicly agree on how to proceed. By meeting together, it is possible for the underlying facts motivating their joint action and the details of their plan of action to become *common knowledge* among the group. Because everyone was present in the room together, each person is confident not only of the underlying facts and plan of action,

*This research has been supported by N.S.F. grant #9709601.

but is also confident that those facts and that plan are common knowledge among the group.

Risky coordinated action *requires* some degree of common knowledge. Without a public meeting, it is hard to achieve the requisite degree of common knowledge. Unless there is a very tightly prescribed protocol specifying how information (the underlying facts, the details of the plan of action) is transmitted to the group members, common knowledge breaks down very easily. A famous example of Rubinstein (1989) suggests why this might be the case.

A pair of players must decide whether to make a costly investment. They must make their choices simultaneously, but they have the opportunity to communicate about their plans before hand. If a player's private state is bad, he will lose two million dollars if he chooses to invest. If his private state is good, he will make a net profit of a million dollars if he invests, but only if the other player invests: there are strategic links between their projects and if only one of the players invests, that player will lose two million dollars, regardless of his private state. If either entrepreneur decides not to invest, his profits are zero.

Player 2's state is good for sure. The ex ante probability that player 1's private state is good is $\frac{1}{2}$. If player 1's state is good, he passes that information on - by electronic mail - to player 2. (If his state is bad, he doesn't send any message). If player 1's message arrives safely, player 2 then sends a confirmation to player 1, informing him that she has received his message. Player 1 in turn sends a confirmation back to player 2, and so on. At each stage, there is a probability ε that the message will get lost. How many messages are required before the players are prepared to invest?

Clearly, player 1 does not invest if his state is bad and player 2 is not prepared to invest if she has not received a message from player 1. But how many messages are enough? The surprising answer is that they should never be prepared to invest. Suppose the players followed the rule that as long as a player knows that $k \geq 1$ messages have been successfully sent and received, he should invest. But then consider a player who has successfully received the k th message and has sent off the $(k + 1)$ th message but has not received a confirmation. He will assign probability at least a half to the possibility that his reply was lost (the other possibility is that his message arrived, but the confirmation was lost). But if his reply was lost, then the other player knows only that $k - 1$ messages have been successfully sent and received, and will therefore not invest. But given the payoffs (a loss of two million if the

other player invests, a gain of one million if he does), it does not pay player 1 in terms of expected revenue to invest in that circumstance. A version of this argument by contradiction can be used to show that no investment ever takes place, however many messages end up being sent (this result is more formally reviewed in section 2).

The result suggests why coordination might be tricky with faulty communication. There seems to be a general logic at work: as players communicate using a faulty technology, every communication sent generates new uncertainty about whether that communication arrived. In order to show that a strategy profile is not an equilibrium, it is enough to find one type of one player who is supposed to invest under the putative equilibrium but who is sufficiently uncertain about what information has reached the other player that he is not prepared to risk it.

But however general the logic, the example is obviously highly stylized. This paper examines the robustness of the argument to three elements.

While it is natural to assume that there is a positive probability that any message gets lost, it was also assumed that there is no upper bound on the number of messages that could conceivably be sent. A more satisfying assumption might be that there is a finite amount of time available for communication before a decision must be made, and each message takes a random length of time to arrive. In section 3, such a timing game with real time messages is analyzed. Conditions on the message arrival technology that imply no coordinated action are identified, and a natural example satisfying them is provided.

The e-mail game involves two players who never meet face to face: if the two players ever met face to face, they could immediately generate the requisite common knowledge to co-ordinate their behavior. A perhaps more relevant scenario is when a large number of players must co-ordinate their behavior, but while subgroups meet together publicly, they never get to meet together as a body. In the meetings game of section 4, N players meet together publicly and often in groups of m players. In order to co-ordinate their actions, each player must be confident that at least n of the N will invest. If $m \geq n$, one meeting is enough to coordinate their behavior. If $m < n$ and there is significant uncertainty about what meetings will take place in the future, then under weak conditions it is possible to rule out coordinated actions, however many meetings take place.

The players of the e-mail game had no choice but to keep sending out confirmation after confirmation until a message got lost: the communication

process was exogenously given. Would the same difficulty in coordination arise if players could decide how many confirmations to send? If players could commit ex ante to a rule for sending messages, they would have an incentive to agree on a protocol where player 1 sends out one message only about his private state and player 2 sends at most one confirmation. If the probability of message loss was sufficiently small, such a scenario would allow both players to invest with positive probability in equilibrium and thus be better off. Although players have a slight conflict of interest over the exact numbers of confirmations,¹ they have a common interest in containing the number of anticipated confirmations. The underlying problem created by anticipated confirmations is that if I send you a message, anticipating a confirmation, and I do not receive a confirmation, then I assign probability (about) $\frac{1}{2}$ to my message never having reached you; however, if I had not anticipated a confirmation, I would assign probability ε (the probability of a message getting lost) to my message never having reached you. Thus anticipated confirmations reduce confidence in the current communication.

However, if players cannot commit ex ante to the number of confirmations, the situation becomes very different. While *anticipated* confirmations are damaging, *unanticipated* confirmations are Pareto-improving. If I did not expect to receive a confirmation from you, a failure to receive one will not discourage me from investing. But if I do receive an unexpected confirmation, I am all the more confident that you will invest and therefore even more likely to invest myself. This provides you with an incentive to send the unanticipated confirmation. Of course, there are no unanticipated confirmations in equilibrium and the effect of the incentive to provide unanticipated confirmations is to put us back in a world where many (anticipated) confirmations get (strategically) sent and no coordinated investment takes place. In the strategic communication game of section 5, a formal analysis of this intuition is presented.²

¹This issue is discussed in more detail in section 5.

²Binmore and Samuelson (1999) present a similar model of strategic communication in the e-mail game, with more optimistic conclusions about the possibility of coordinated action. Their model and conclusions are discussed in section 5.

2 The Electronic Mail Game

If a player's private state is good (G), he has an incentive to invest if his opponent invests. His payoffs are given by the following matrix (the row represents a player's action and the column represents his opponent's action):

G	Invest	Not Invest
Invest	1	$-c$
Not Invest	0	0

where $c > 0$. If a player's private state is bad (B), he has a dominant strategy to not invest, with payoffs given by the following matrix.

B	Invest	Not Invest
Invest	$-c$	$-c$
Not Invest	0	0

Player 2's state is good for sure, but the ex ante probability that player 1's state is good is $\frac{1}{2}$. If the state is good, player 1 sends a message to player 2, informing her of this fact. The message gets lost with exogenous probability $0 < \varepsilon < \frac{1}{2}$. If player 2 receives the message, she sends a confirmation to player 1, stating that she received the message. This message also gets lost with probability ε . If player 1 receives the confirmation, he sends a re-confirmation, stating that he received the confirmation. This message also gets lost with probability ε . And so on.³

After all the messages have been sent (eventually a message will get lost) each player chooses an action (invest or not invest) simultaneously. Observe that from the communication stage, there are an infinite number of possible

³This differs from the game analyzed by Rubinstein (1989). First, this version is "private values," so each player is certain of his own payoffs. Second, this version gives player 1 a dominant strategy for one private state (bad), while in the original game players simple faced the problem of coordinated on an efficient Nash equilibrium that varied across states. Both changes simplify the presentation of later results, but do not effect the qualitative conclusions.

states of the system:

Player 1's State	Total Messages Sent	Messages Received by Player 1	Messages Received by Player 2	Probability
bad	0	0	0	$\frac{1}{2}$
good	1	0	0	$\frac{1}{2}\varepsilon$
good	2	0	1	$\frac{1}{2}(1-\varepsilon)\varepsilon$
good	3	1	1	$\frac{1}{2}(1-\varepsilon)^2\varepsilon$
good	4	1	2	$\frac{1}{2}(1-\varepsilon)^3\varepsilon$
good	5	2	2	$\frac{1}{2}(1-\varepsilon)^4\varepsilon$
.
good	$2n$	$n-1$	n	$\frac{1}{2}(1-\varepsilon)^{2n-1}\varepsilon$
good	$2n+1$	n	n	$\frac{1}{2}(1-\varepsilon)^{2n}\varepsilon$
.

Proposition 1 *If $c > 1 - \varepsilon$, this electronic mail game has a unique equilibrium: both players never invest.*

Note that for small ε , this reduces to the requirement that $c \geq 1$, i.e., no investment is the risk dominant equilibrium of the game.

PROOF. If player 1's state is bad, he has a dominant strategy to not invest. If player 2 receives no message from player 1, she assigns probability

$$\frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2}\varepsilon} = \frac{1}{1 + \varepsilon}$$

to the possibility that player 1's state is bad. Thus her expected payoff to investing is at most

$$\frac{1}{1 + \varepsilon}(-c) + \frac{\varepsilon}{1 + \varepsilon}(1) = \frac{\varepsilon - c}{1 + \varepsilon} < 0,$$

so she cannot be investing in any equilibrium. If player 1 receives no confirmation from player 2, he assigns probability

$$\frac{\frac{1}{2}\varepsilon}{\frac{1}{2}\varepsilon + \frac{1}{2}(1-\varepsilon)\varepsilon} = \frac{1}{2-\varepsilon}$$

to the possibility that his message never arrived. Thus his expected payoff to investing is at most

$$\frac{1}{2 - \varepsilon} (-c) + \frac{1 - \varepsilon}{2 - \varepsilon} (1) = \frac{1 - \varepsilon - c}{2 - \varepsilon} < 0,$$

so he cannot be investing in equilibrium. Now let k be the smallest number greater than or equal to 1 such that a player who knows that k messages have been sent and received invests with positive probability. A player who receives the k th message but has not received a reply to the $(k + 1)$ th message assigns probability

$$\frac{\frac{1}{2} (1 - \varepsilon)^{k-1} \varepsilon}{\frac{1}{2} (1 - \varepsilon)^{k-1} \varepsilon + \frac{1}{2} (1 - \varepsilon)^k \varepsilon} = \frac{1}{2 - \varepsilon}$$

to the possibility that his message never arrived. If it did not arrive, then his opponent knows only that $k - 1$ messages have been sent and received, and therefore will not be investing. Thus his expected payoff to investing is at most

$$\frac{1}{2 - \varepsilon} (-c) + \frac{1 - \varepsilon}{2 - \varepsilon} (1) = \frac{1 - \varepsilon - c}{2 - \varepsilon} < 0,$$

so he cannot be investing in equilibrium. This argument demonstrates by contradiction that there is no investment in equilibrium. ■

The result is tight in the following sense. If $\varepsilon \leq c \leq 1 - \varepsilon$, this game has an equilibrium where player 1 always invests when his state is good and player 2 invests whenever she receives at least one message. If $c \leq \varepsilon$, this game has an equilibrium where player 1 always invests if his state is good and player 2 always invests.

3 Timing

The timing game is identical to the electronic mail game of the previous section, with one change. Instead of a message getting lost with exogenous probability ε , assume instead that it may take a long time to arrive. Let $F(\tau)$ be the probability that a message will take less than or equal to τ minutes to arrive (and write $f(\cdot)$ for the density corresponding to $F(\cdot)$). At time 0, player 1 learns whether his state is good; if it is good, he sends a message to player 2; if the message arrives, she sends a confirmation; and

so on. At time T , however many messages have been sent and received, the players must make their investment decisions.

Suppose you receive your last message at time t . What probability do you assign to your message having been the last? If you have just received a message at date t , the probability that your message never arrives is

$$N(t) \equiv 1 - F[T - t];$$

the probability that your message arrives but you never receive a reply is

$$A(t) \equiv \int_{\tau=0}^{T-t} f(\tau) (1 - F(T - t - \tau)) d\tau.$$

The likelihood ratio of these two probabilities is

$$L(t) \equiv \frac{A(t)}{N(t)} = \frac{\int_{\tau=0}^{T-t} f(\tau) (1 - F(T - t - \tau)) d\tau}{1 - F[T - t]}.$$

Thus the probability that your message was the last sent, conditional on you not receiving a confirmation is

$$\frac{N(t)}{N(t) + A(t)} = \frac{1}{1 + L(t)}.$$

Proposition 2 *If $c > 1 - F(T)$ and $c > L(t)$ for all $t \in [0, T]$, the timing game has a unique equilibrium: both players never invest.*

PROOF. Clearly, player 1 will not invest if his state is bad. If player 2 receives no messages, she assigns probability $\frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2}(1 - F(T))}$ to player 1's state being bad and probability $\frac{\frac{1}{2}(1 - F(T))}{\frac{1}{2} + \frac{1}{2}(1 - F(T))}$ to player 1's state being good but player 1's first message never arriving. Thus her expected payoff is at most

$$\frac{\frac{1}{2}(1 - F(T))}{\frac{1}{2} + \frac{1}{2}(1 - F(T))} (1) - \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2}(1 - F(T))} (c) = \frac{1 - F(T) - c}{2 - F(T)} < 0.$$

Now let t^* be the earliest date for which a player who has sent a message at date t^* (but never received a confirmation) is prepared to invest with positive probability. He will assign probability $\frac{L(t^*)}{1 + L(t^*)}$ to his message having arrived (but not the confirmation); and he assigns probability $\frac{1}{1 + L(t^*)}$ to his

message having never arrived. If his message never arrived, his opponent's last message (if any) must have been sent before date t , so his opponent cannot be investing. Thus his expected payoff at most

$$\frac{L(t^*)}{1+L(t^*)} (1) - \frac{1}{1+L(t^*)} (c) = \frac{L(t^*) - c}{1+L(t^*)}.$$

This is negative since $c > L(t^*)$. ■

This result is tight in the following sense. If $1 - F(T) \leq c \leq L(0)$, the timing game has an equilibrium where player 1 always invests when his state is good and player 2 invests exactly if she receives at least one message. If $c \leq 1 - F(T)$, the timing game has an equilibrium where player 1 always invests when his state is good and player 2 always invests.

To illustrate Proposition 2, suppose the density is exponential with $f(t) = \lambda e^{-\lambda t}$. Then

$$\begin{aligned} F(t) &= 1 - e^{-\lambda t} \\ N(t) &= e^{-\lambda(T-t)} \\ A(t) &= \int_{\tau=0}^{T-t} \lambda e^{-\lambda \tau} e^{-\lambda(T-t-\tau)} d\tau = \lambda e^{-\lambda(T-t)} (T-t) \\ L(t) &= \lambda(T-t) \end{aligned}$$

Thus if $c > \lambda T$ and $c > e^{-\lambda T}$, then there is no equilibrium where investment takes place. Observe that λT is the expected number of messages that can be sent and delivered, so this is a rather restrictive condition. However, consider a density that is a sum of exponentials: $f(t) = \varepsilon \underline{\lambda} e^{-\underline{\lambda} t} + (1 - \varepsilon) \bar{\lambda} e^{-\bar{\lambda} t}$. As $\underline{\lambda} \rightarrow 0$ and $\bar{\lambda} \rightarrow \infty$, this generates essentially the e-mail information system (with probability $1 - \varepsilon$, the message arrives very quickly; with probability ε , it takes a very long time to arrive). In this case we have,

$$\begin{aligned} F(t) &= 1 - \varepsilon e^{-\underline{\lambda} t} - (1 - \varepsilon) e^{-\bar{\lambda} t} \\ N(t) &= \varepsilon e^{-\underline{\lambda}(T-t)} + (1 - \varepsilon) e^{-\bar{\lambda}(T-t)} \\ A(t) &= \varepsilon^2 \underline{\lambda} e^{-\underline{\lambda}(T-t)} + (1 - \varepsilon)^2 \bar{\lambda} e^{-\bar{\lambda}(T-t)} + \varepsilon(1 - \varepsilon) \left(\frac{\bar{\lambda} + \underline{\lambda}}{\bar{\lambda} - \underline{\lambda}} \right) \left(e^{-\underline{\lambda}(T-t)} - e^{-\bar{\lambda}(T-t)} \right) \\ L(t) &= \frac{\varepsilon^2 \underline{\lambda} e^{-\underline{\lambda}(T-t)} + (1 - \varepsilon)^2 \bar{\lambda} e^{-\bar{\lambda}(T-t)} + \varepsilon(1 - \varepsilon) \left(\frac{\bar{\lambda} + \underline{\lambda}}{\bar{\lambda} - \underline{\lambda}} \right) \left(e^{-\underline{\lambda}(T-t)} - e^{-\bar{\lambda}(T-t)} \right)}{\varepsilon e^{-\underline{\lambda}(T-t)} + (1 - \varepsilon) e^{-\bar{\lambda}(T-t)}} \end{aligned}$$

Since $L(t)$ is decreasing, if $L(0) < c$ there is no investment ever. Figure 1 plots the values of $L(0)$ as $\underline{\lambda}$ varies, setting $T = 1$, $\varepsilon = \frac{1}{10}$ and $\bar{\lambda} = 100$.

insert figure 1 around here

4 Many Players Meetings

There are N players. Each will have to decide simultaneously whether to invest or not. A player's payoff from investing is 1 if a public state is good and at least n players (including himself) invest, where $1 < n < N$. If the state is bad, or if less than n players end up investing, then the payoff to investing is $-c$. The payoff to not investing is always 0.

The ex ante probability that the state is good is $\frac{1}{2}$. Every so often, the players meet in groups of m players to discuss what they should do. The probability that there will be a total of k meetings is $\varepsilon(1 - \varepsilon)^k$, where $\varepsilon > 0$. Thus conditional on k meetings having occurred, the conditional probability that at least one more meeting will occur is always $1 - \varepsilon$. The collection of players gathered together at the j th meeting is random: each subset is equally likely. Thus the probability than any one player attends any fixed meeting is always $\frac{m}{N}$. If a first meeting occurs (this happens with ex ante probability $1 - \varepsilon$), the players at that first meeting are informed whether the state is good or bad. At each meeting, each player gets to learn all the information about the state and previous meetings that was available at previous meetings.⁴

Eventually, all the meetings are concluded. Each player must then decide whether to invest or not, based on the history of meetings attended.

Proposition 3 *If $m < n$ and*

$$c > \max \left\{ 1, \frac{(1 - \varepsilon) \frac{N-m}{N}}{1 - (1 - \varepsilon) \frac{N-m}{N}} \right\},$$

then the meetings game has a unique equilibrium: no player ever invests.

⁴This assumption makes most sense is $m > \frac{N}{2}$, so that every meeting contains at least one person who was present at every previous meeting. However, if $m \leq \frac{N}{2}$, we could instead imagine that a record of previous meetings is kept and made available to current attendees.

Note that for very small ε , the constraint on c reduces to the requirement that $c > \max \left\{ 1, \frac{N-m}{m} \right\}$.

PROOF. No player will invest if he knows that the state is bad. Consider a player who has never attended a meeting. The fact that he has never attended a meeting conveys no information about the state. Thus he assigns probability $\frac{1}{2}$ to the state being good. Thus his payoff to investing is at most

$$\frac{1}{2}(-c) + \left(\frac{1}{2}\right)(1) = \frac{1}{2}(1-c) < 0.$$

Now consider a player who knows that the state is good and who has just attended a meeting. With probability ε , no further meetings will take place. With probability $(1-\varepsilon)^k \left(\frac{N-m}{N}\right)^k \varepsilon$, exactly k more meetings will take place, but he will not be in attendance. Thus the probability that there is at least one more meeting, but no meetings where he is in attendance, is

$$\frac{(1-\varepsilon) \left(\frac{N-m}{N}\right) \varepsilon}{1 - (1-\varepsilon) \left(\frac{N-m}{N}\right)}.$$

So the probability that any player assigns to the last meeting he attended being the last meeting ever will always be

$$\frac{\varepsilon}{\varepsilon + \frac{(1-\varepsilon) \left(\frac{N-m}{N}\right) \varepsilon}{1 - (1-\varepsilon) \left(\frac{N-m}{N}\right)}} = 1 - (1-\varepsilon) \left(\frac{N-m}{N}\right)$$

Now let k^* be the smallest $k \geq 1$ such that a player who knows that the state is good and knows about exactly k meetings invests with positive probability. If his last meeting was the last meeting that anyone attended, he knows that at most m players will choose to invest (since only m players attended the last meeting). Thus his payoff to investing is at most

$$\left(1 - (1-\varepsilon) \left(\frac{N-m}{N}\right)\right) (-c) + (1-\varepsilon) \left(\frac{N-m}{N}\right) (1) < 0. \blacksquare$$

This result is tight in the following sense. If $c \leq 1$, there is an equilibrium where all players invest as long as they do not know that the state is bad. If $c \geq 1$ and *either* $m \geq n$ *or* $c \leq \frac{(1-\varepsilon) \frac{N-m}{N}}{1 - (1-\varepsilon) \frac{N-m}{N}}$ and ε is sufficiently small, then there is an equilibrium where all players invest if they know that the state is good.

5 Strategic Communication

The strategic communication game is identical to the electronic mail game of section 2, with one change. When players in the electronic mail game were forced to send messages, players in the strategic communication game can choose whether to send the message or not. This gives rise to the somewhat elaborate extensive form game illustrated in figure 2.

insert figure 2 around here

Also fix $c > 1$ and $\varepsilon < \frac{1}{1+c}$. In this game, equilibrium strategies uniquely determine beliefs at off-the-equilibrium information sets. Information sets where players send a message (S) or don't send a message (D) are all singletons. When a player is (out-of-equilibrium) called upon to choose between investing (I) and not investing (N), he knows that his opponent has sent one more message than expected, and his out-of-equilibrium belief over his opponent's investment choice can be calculated from the opponent's strategy. We will look at perfect Bayesian equilibria in this game.

Behavioral strategies can be described by probabilities at each of the binary choices facing the players. The easiest way to describe them is to label the extensive form with these probabilities, as in figure 3.

insert figure 3 around here

Thus z is the probability that player 1 invests if his state is bad and x_0 is the probability that player 2 invests if she never receives a message. For all $k \geq 1$, the triple (π_k, x_k, y_k) describes the behavioral strategy of the player who has the option of the sending the k th message in the game: π_k is the probability that he chooses to send the message; x_k is the probability that he invests if he sends the k th message in the game but does not receive a reply; and y_k is the probability that he invests if he chose not to send the k th message. Thus $\left(z, (\pi_k, x_k, y_k)_{k=1,3,5,\dots}\right)$ describes the behavioral strategy of player 1 and $\left(x_0, (\pi_k, x_k, y_k)_{k=2,4,6,\dots}\right)$ describes the behavioral strategy of player 2. Since $z = 0$ and $x_0 = 0$ in any equilibrium (because player 1 has a dominant strategy to not invest if investment conditions are bad, and player 2 assigns

probability at least $\frac{1}{2}$ to player 1's state being bad), we focus attention on $(\pi_k, x_k, y_k)_{k=1,2,3,\dots}$ in the following analysis.

Write $f^*(\pi)$ for the set of mixed strategy best responses if investment conditions are good and you expect your opponent to invest with probability π . Thus

$$f_0(\pi) = \begin{cases} \{0\}, & \text{if } \pi < \frac{c}{1+c} \\ [0, 1], & \text{if } \pi = \frac{c}{1+c} \\ \{1\}, & \text{if } \pi > \frac{c}{1+c} \end{cases}$$

Figure 4 plots this correspondence.

insert figure 4 around here

We can use f_0 to state the equilibrium conditions on players' investment choices, taking as given their message sending strategies: for each $k = 1, 2, \dots$

$$\begin{aligned} y_k &\in f_0(x_{k-1}) \\ x_k &\in f_0\left(\frac{\varepsilon x_{k-1} + (1-\varepsilon)(1-\pi_{k+1})y_{k+1} + (1-\varepsilon)\pi_{k+1}\varepsilon x_{k+1}}{\varepsilon + (1-\varepsilon)(1-\pi_{k+1}) + (1-\varepsilon)\pi_{k+1}\varepsilon}\right) \end{aligned} \quad (1)$$

Equilibrium conditions for message sending are more complicated to state in general but are simple enough to check in practise.

One class of equilibria are “message threshold” equilibria. For some $\bar{k} \geq 1$, only $\bar{k} \geq 1$ messages get sent and each player is prepared to invest as long as he knows that at least \bar{k} messages have been sent.⁵ This equilibrium is illustrated in figure 5, for the case where $\bar{k} = 2$ (the diagram includes the beliefs at each information set implied by equilibrium strategies).

insert figure 5 around here

The person who sends the \bar{k} th message is prepared to invest, even if he does not receive a confirmation, because under the equilibrium strategies he was not anticipating a confirmation. Because of this, he assigns probability $1 - \varepsilon$ to his message arriving and thus has expected payoff $1 - \varepsilon - \varepsilon c > 0$.

⁵Thus $\pi_k = 1$ for all $k \leq \bar{k}$ and $\pi_k = 0$ for all $k > \bar{k}$; $x_k = 0$ for all $k \leq \bar{k} - 1$ and $x_k = 1$ for all $k > \bar{k} - 1$; $y_k = 0$ for all $k \leq \bar{k}$ and $y_k = 1$ for all $k > \bar{k}$.

This equilibrium relies on the fact the player who receives the \bar{k} th message chooses not to send a confirmation. Under the equilibrium strategies, he is indifferent between whether to send a confirmation or not. But notice that if he fails to send a confirmation, he knows that the other player believes that he will invest with probability $1 - \varepsilon$. On the other hand, if he sent an unanticipated confirmation then (in this equilibrium) he knows that the other player will believe that he will invest with probability 1. This suggests that he should break his indifference in favor of sending a confirmation.

To model this intuition formally, consider a perturbed version of the game where a player does strictly prefer that his opponent is more confident that he will invest. In particular, suppose that payoffs to player i in the good state are:

G	Invest	Not Invest	(2)
Invest	$1 - \theta_i$	$-c - \theta_i$	
Not Invest	0	0	

where θ_i is drawn according to a c.d.f. Ψ_δ with support $[-c, 1]$ and is observed by player i only immediately before he makes his investment decision. Now if player i assigns probability π to his opponent investing, he will invest only if

$$-\theta_i + \pi(1) + (1 - \pi)(-c) \geq 0.$$

The probability of this occurring is

$$f_\delta(\pi) = \Psi_\delta((1 + c)\pi - c)$$

Thus each f_δ is a strictly increasing function with $f_\delta(0) = 0$ and $f_\delta(1) = 1$. We assume that as $\delta \rightarrow 0$, Ψ_δ concentrates mass around 0: $\Psi_\delta(-\delta) < \delta$ and $\Psi_\delta(\delta) > 1 - \delta$. This implies that as $\delta \rightarrow 0$, f_δ tends to the step correspondence, f_0 , illustrated in figure 4. Figure 6 plots f_δ for small δ :

insert figure 6 around here

In this perturbed game, the equilibrium conditions for investment become:⁶ for each $k = 1, 2, \dots$

$$y_k = f_\delta(x_{k-1}) \tag{3}$$

⁶We ignore the fact that player 2 will invest with positive probability even when she receives no message (i.e., $x_0 > 0$); this simplifying assumption does not change any qualitative conclusions.

$$x_k = f_\delta \left(\frac{\varepsilon x_{k-1} + (1 - \varepsilon)(1 - \pi_{k+1}) y_{k+1} + (1 - \varepsilon) \pi_{k+1} \varepsilon x_{k+1}}{\varepsilon + (1 - \varepsilon)(1 - \pi_{k+1}) + (1 - \varepsilon) \pi_{k+1} \varepsilon} \right)$$

An equilibrium is *immune to strategic uncertainty* if there exists a sequence of equilibria in the perturbed game that converge to the candidate equilibrium as $\delta \rightarrow 0$.

The message threshold equilibria described above are not immune to strategic uncertainty. In any equilibrium of the δ -game close to the \bar{k} -equilibrium, we will have $x_{\bar{k}} \approx f_\delta(1 - \varepsilon)$ while $y_{\bar{k}+1} \approx f_\delta(1)$. Thus the player receiving \bar{k} th message will have a strict incentive to send an unanticipated confirmation in the perturbed game.

Another class of equilibria are “punishment equilibria.” For some $\bar{k} \geq 1$, only $\bar{k} \geq 1$ messages get sent and each player is prepared to invest as long as he knows that *exactly* \bar{k} messages have been sent.⁷ This equilibrium is illustrated in figure 7, for the case where $\bar{k} = 2$.

insert figure 7 around here

As in the message threshold equilibria, the person who sends the \bar{k} th message is prepared to invest, even if he does not receive a confirmation, because under the equilibrium strategies he was not anticipating a confirmation. But the person who receives the \bar{k} th message is no longer indifferent between sending an unanticipated confirmation or not. Under the punishment equilibrium strategy profile, he will be punished if he sends a $(\bar{k} + 1)$ th message: no player chooses to invest in equilibrium if he knows that a $(\bar{k} + 1)$ th message has been sent. Such punishment equilibria perhaps do capture why many confirmations are not sent in practise: over some threshold, players will become discouraged and not anticipate investment. However, one can also make a traditional renegotiation proofness argument for why the punishments would not be time-consistent: a player sending an unanticipated $(\bar{k} + 1)$ th message could point out to the other player that it is consistent with equilibrium and Pareto-optimal to always invest from that point on, and if this appeal were credible, the original equilibrium would break down. In what follows, we will focus on *ex post efficient* equilibria, where, contingent on strategies in the message sending phase, equilibrium strategies in

⁷Thus $\pi_k = 1$ for all $k \leq \bar{k}$ and $\pi_k = 0$ for all $k > \bar{k}$; $x_{\bar{k}} = 1$ and $x_k = 0$ for all $k \neq \bar{k}$; $y_{\bar{k}+1} = 1$ and $y_k = 0$ for all $k \neq \bar{k} + 1$.

the investment game maximize the amount of investment (there is such a “maximal investment” equilibrium because of the strategic complementarities in the game). The punishment equilibria are not ex post efficient. What equilibria are left?

There is the “infinite messages - no investment” equilibrium, where players always send messages and never invest. There is also a “no messages - no investment” equilibrium, where player 1 does not send his first message if his state is good, but both players do always send confirmations if (out of equilibrium) they receive messages. These are essentially the only equilibria that are ex post efficient and immune to strategic uncertainty:

Proposition 4 *If $c > 1$, all ex post efficient perfect Bayesian equilibria immune to strategic uncertainty in the strategic communication game have both players not investing after every history; every player sends a confirmation with probability close to 1, whenever they receive a message.*⁸

Formal definitions of the refinements and a proof of the proposition are presented in the appendix.

The proposition says that once a message has been received, each player must have an incentive to send a confirmation with high probability in any equilibrium. This implies no investment in equilibrium. One equilibrium has an infinite numbers of costly messages sent, and no investment, while another equilibrium has the first player never sending a message about his private state (since it will not do any good anyway). Since there presumably is a small cost of sending messages, we might imagine that this equilibrium would be played in a richer model with small costs.

The refinements (“ex post efficiency” and “immunity to strategic uncertainty”) are ad hoc. One can imagine other formal routes to reach similar conclusions. For example, standard forward induction reasoning would seem to argue against the punishment equilibria: there is no point in sending an unanticipated confirmation if you expect it to lead to no investment; if a player sends you an unanticipated confirmation, he surely does not expect it to lower the probability of you investing. The ex post efficiency criterion could also be related more formally to the renegotiation literature in game theory. However, the modest purpose here is to suggest that making message sending a strategic choice need not remove the coordination problem

⁸More precisely, $\pi_k > \frac{1-\varepsilon(1+c)}{(1-\varepsilon)^2}$ for all $k = 2, 3, \dots$; and $x_k = 0$ and $y_k = 0$ for all k . For small ε , this implies π_k close to 1.

from the e-mail game and reasonable views of rational play in the game are consistent with that view.

The above analysis formalized an intuition based on the Pareto-improving nature of unanticipated confirmations. One can also imagine quite different arguments in favor of the view that strategic message sending choices will not reduce message sending. If we look at either message threshold equilibria or punishment equilibria, the most preferred equilibria for player 2 (from an ex ante perspective) are one message equilibria, where player 1 invests whenever his state is good and player 2 invests whenever he receives one message from player 1. The most preferred equilibria for player 1 (from an ex ante perspective) are one message equilibria, where player 2 invests whenever he receives a message from player 1 and player 1 invests whenever he receives a confirmation from player 2. In the former case, player 1 bears the ex ante risk of misco-ordination (with probability $\frac{1}{2}\varepsilon$, player 1's message gets lost, player 1 invests and player 2 does not). In the latter case, player 2 bears the risk (with probability $\frac{1}{2}(1 - \varepsilon)\varepsilon$, player 2's confirmation gets lost, player 1 invests and player 2 does not). Thus each player would like (ex ante) to be the last one to receive a confirmation. One can easily imagine stories where this lead to an ex post proliferation of messages. For example, if a player could credibly make statements like "I know you were not planning to send a confirmation; however, if I do not receive a confirmation, I will not invest," he would have an incentive to do so.

Binmore and Samuelson (1999) also examined the strategic communication game. They considered a game where each player decides how many times to confirm a message received and whether to invest or not as a function of the confirmation strategies of both players. The extensive form of this paper generates a slightly richer strategic form than the Binmore-Samuelson game, since players' behavior after their own deviations is part of the description of their strategies. However, this difference is not strategically relevant.⁹

Binmore and Samuelson showed the existence of equilibria analogous to the message threshold equilibria of this paper, and they note how all such equilibria rely on players choosing weak best responses. They therefore first perturb the game and then apply evolutionary refinements in order to make tighter predictions. They consider a game with "costs of paying attention," where players must incur a cost ex ante in order to be able to receive a

⁹I am grateful to Larry Samuelson for helping me understand the relation between the games.

given number of messages. This game, too, many equilibria, but the “tacit” equilibrium where players never invest fails to satisfy an evolutionary stability criterion.¹⁰

The strategic communication version of the e-mail game has a rich set of equilibria. The results presented in this section and the work of Binmore and Samuelson suggest that one could obtain many different conclusions by suitable perturbations of the game and choices of equilibrium refinement. Fine details of the strategic communication environment may be important in determining the likelihood of co-ordinated behavior using faulty communication channels.

6 Conclusion

With faulty communication, fully rational players in a one-off strategic interaction may have great difficulty co-ordinating on an efficient outcome, even if they are able to communicate a lot. This problem arises not merely in the stylised example of the e-mail game, but in more realistic environments with many players, real time communication structures and strategic decisions about whether to communicate.

Of course, people carry out coordinated action without the benefit of public plenary meetings all the time. But they are most likely to be successful in doing so if they can commit ex ante to rules of communication that generate the approximate common knowledge required for coordination. Societies presumably evolve conventions for coordinating behavior that prevent inefficiently many risky communications. Similarly, armies have rules about who sends confirmations of what communications. The analysis in this paper suggests that such conventions and rules may not be entirely self-enforcing in each interaction, but may rely on outside constraints.

¹⁰As Binmore and Samuelson note, it is key to their results that the costs of receiving messages are incurred ex ante. The incentive to send unanticipated confirmation will never arise as long as costs are incurred ex ante. It would be interesting to see if one could combine a small ex post cost of sending messages (i.e., the cost is only incurred if you send the message in equilibrium) with the δ -payoff perturbation described above, and generate evolutionary pressure in favor of the unanticipated confirmations.

References

- [1] Binmore, K. and L. Samuelson (1999). "Coordinated Action in the Electronic Mail Game," forthcoming in *Games and Economic Behavior*.
- [2] Rubinstein, A. (1989). "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *American Economic Review* 79, 385-391.

APPENDIX

We first formally define the refinements described in the text.

Definition 5 *Perfect Bayesian Equilibrium* $(\pi_k, x_k, y_k)_{k=1,2,\dots}$ is immune to strategic uncertainty if there exists $(x_k^\delta, y_k^\delta)_{k=1,2,\dots} \rightarrow (x_k, y_k)_{k=1,2,\dots}$ as $\delta \rightarrow 0$ such that given $(\pi_k)_{k=1,2,\dots}$, $(x_k^\delta, y_k^\delta)_{k=1,2,\dots}$ satisfy the equilibrium conditions (3) for each δ .

Definition 6 *Perfect Bayesian Equilibrium* $(\pi_k, x_k, y_k)_{k=1,2,\dots}$ is ex post efficient if, given $(\pi_k)_{k=1,2,\dots}$, $(x_k, y_k)_{k=1,2,\dots}$ is the largest solution to the equilibrium conditions (1).

PROOF OF PROPOSITION 4. A player who sends the k th message but receives no confirmation believes that with probability

$$\frac{\varepsilon}{\varepsilon + (1 - \varepsilon)(1 - \pi_{k+1}) + (1 - \varepsilon)\pi_{k+1}\varepsilon},$$

his message never arrived. Suppose that he expects that his opponent will not invest if he did not receive that message and will invest if he does receive that message. Then his expected payoff is

$$\frac{(1 - \varepsilon)(1 - \pi_{k+1}) + (1 - \varepsilon)\pi_{k+1}\varepsilon}{\varepsilon + (1 - \varepsilon)(1 - \pi_{k+1}) + (1 - \varepsilon)\pi_{k+1}\varepsilon} (1) - \frac{\varepsilon}{\varepsilon + (1 - \varepsilon)(1 - \pi_{k+1}) + (1 - \varepsilon)\pi_{k+1}\varepsilon} \quad (c)$$

This expression is non-negative if

$$(1 - \varepsilon)(1 - \pi_{k+1}) + (1 - \varepsilon)\pi_{k+1}\varepsilon \geq \varepsilon c$$

or

$$\pi_{k+1} \leq \frac{1 - \varepsilon(1 + c)}{(1 - \varepsilon)^2}.$$

Thus if $\pi_k > \frac{1 - \varepsilon(1 + c)}{(1 - \varepsilon)^2}$ for $k = 2, \dots, \bar{k}$, we have by the usual inductive argument that no player invests unless he knows that at least \bar{k} messages have been sent. Conversely, if $\pi_{\bar{k}+1} \leq \frac{1 - \varepsilon(1 + c)}{(1 - \varepsilon)^2}$, there is an equilibrium where both players invest they know that at least \bar{k} messages have been sent. So we have:

Observation: $(\pi_k, x_k, y_k)_{k=1,2,\dots}$ is an ex post efficient equilibrium if and only if there exists $\bar{k} \in \{1, 2, 3, \dots\} \cup \{\infty\}$ such that $\pi_k > \frac{1 - \varepsilon(1 + c)}{(1 - \varepsilon)^2}$ for $2 \leq k \leq \bar{k}$; if $\bar{k} < \infty$, $\pi_{\bar{k}+1} \leq \frac{1 - \varepsilon(1 + c)}{(1 - \varepsilon)^2}$; $x_k = 0$ for all $k \leq \bar{k} - 1$ and $x_k = 1$ for all $k > \bar{k} - 1$; $y_k = 0$ for all $k \leq \bar{k}$ and $y_k = 1$ for all $k > \bar{k}$.

Now we check immunity to strategic uncertainty. Fix any ex post equilibrium $(\pi_k, x_k, y_k)_{k=1,2,\dots}$ of the above form with finite \bar{k} . Consider the rate of change of x_k^δ and y_k^δ with respect to δ , evaluated at $\delta = 0$. Since $f_\delta(0) = 0$ and $f_\delta(1) = 1$, this rate of change is equal to zero for all y_k^δ and all x_k^δ with $k \neq \bar{k}$ (since the limiting equations are of the form y_k or x_k equal to $f_0(0)$ or $f_1(1)$). But

$$x_{\bar{k}} = f_0(1 - \varepsilon),$$

so $\left. \frac{dx_{\bar{k}}^\delta}{d\delta} \right|_{\delta=0} = \frac{d\Psi_\delta(1 - \varepsilon(1 + c))}{d\delta} < 0$. Thus for small δ , $x_{\bar{k}}^\delta$ will be strictly less than both $x_{\bar{k}}^\delta$ and $y_{\bar{k}}^\delta$ for all $k > \bar{k}$. But this implies that a strict incentive to send the $(\bar{k} + 1)$ th message, a contradiction.

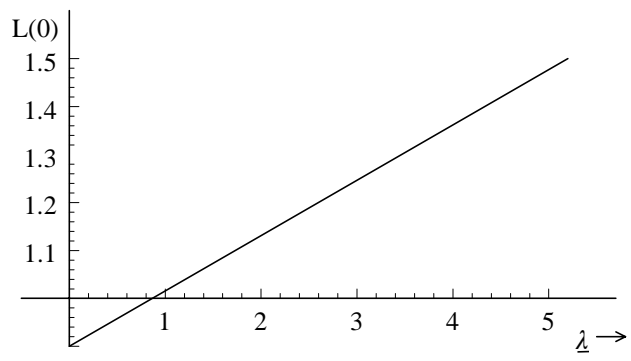


Figure 1

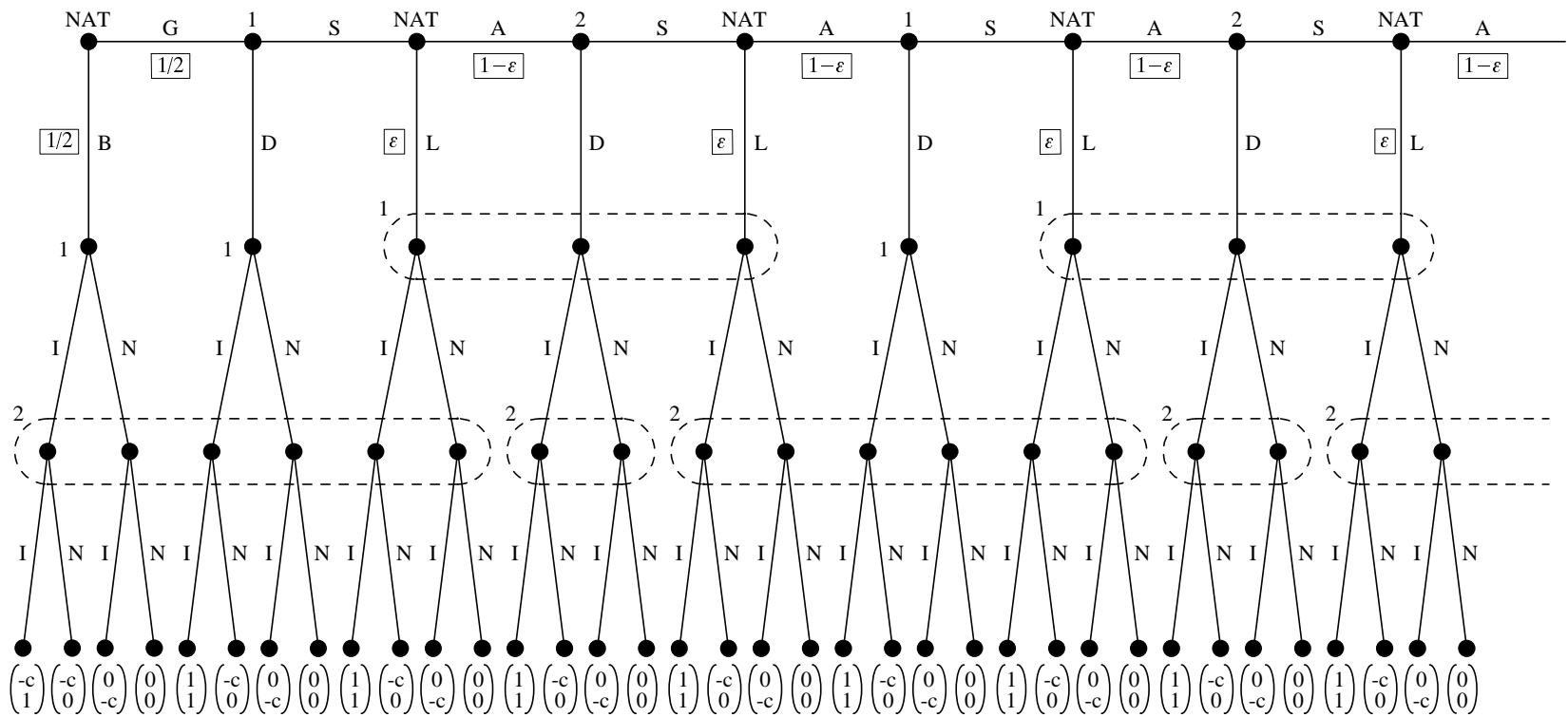


Figure 2. The Extensive Form

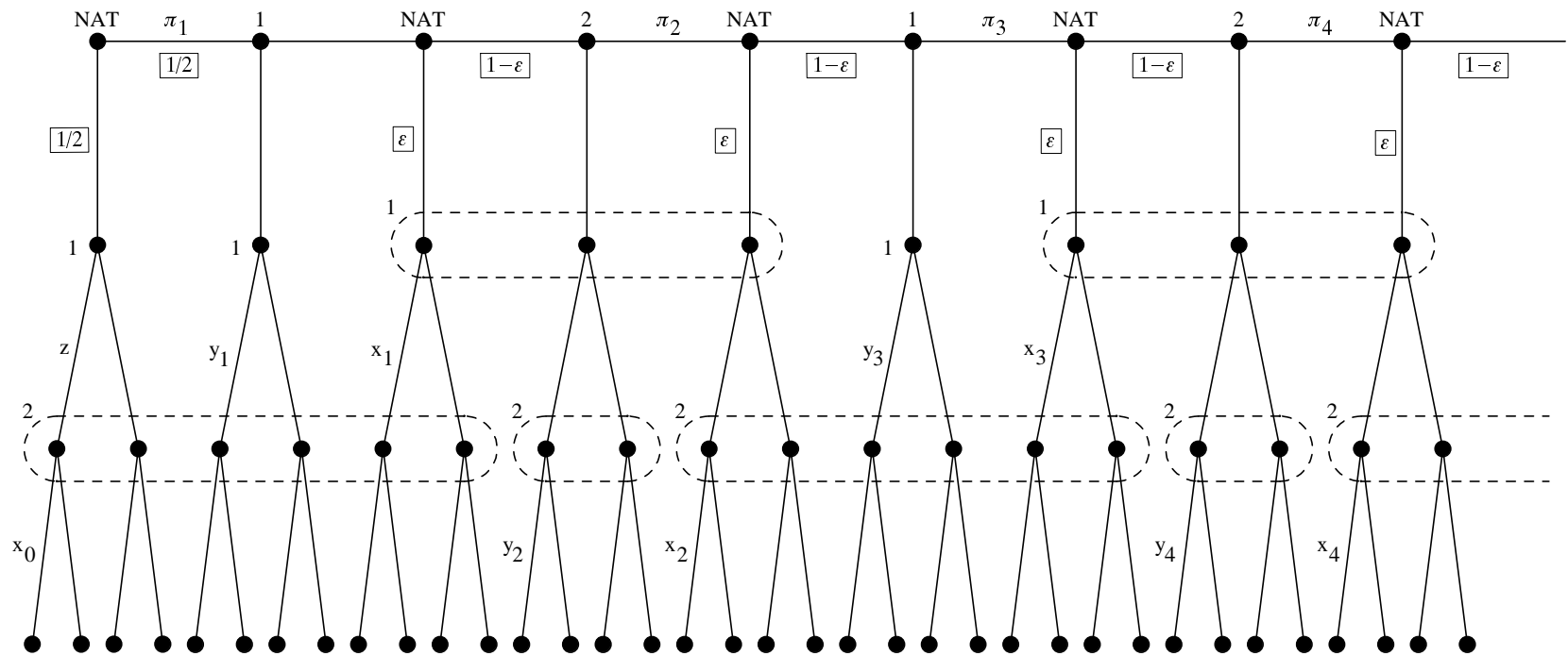


Figure 3. Strategies

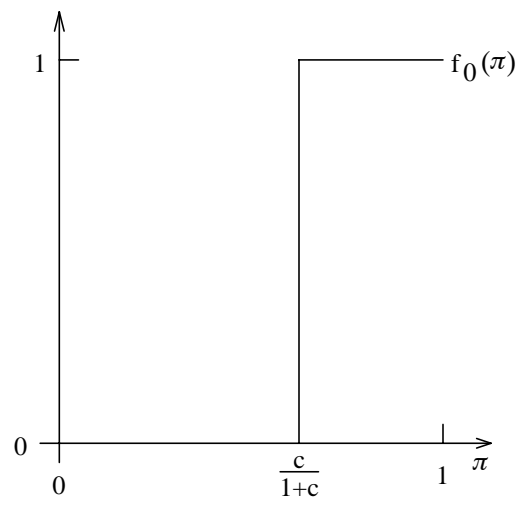


Figure 4

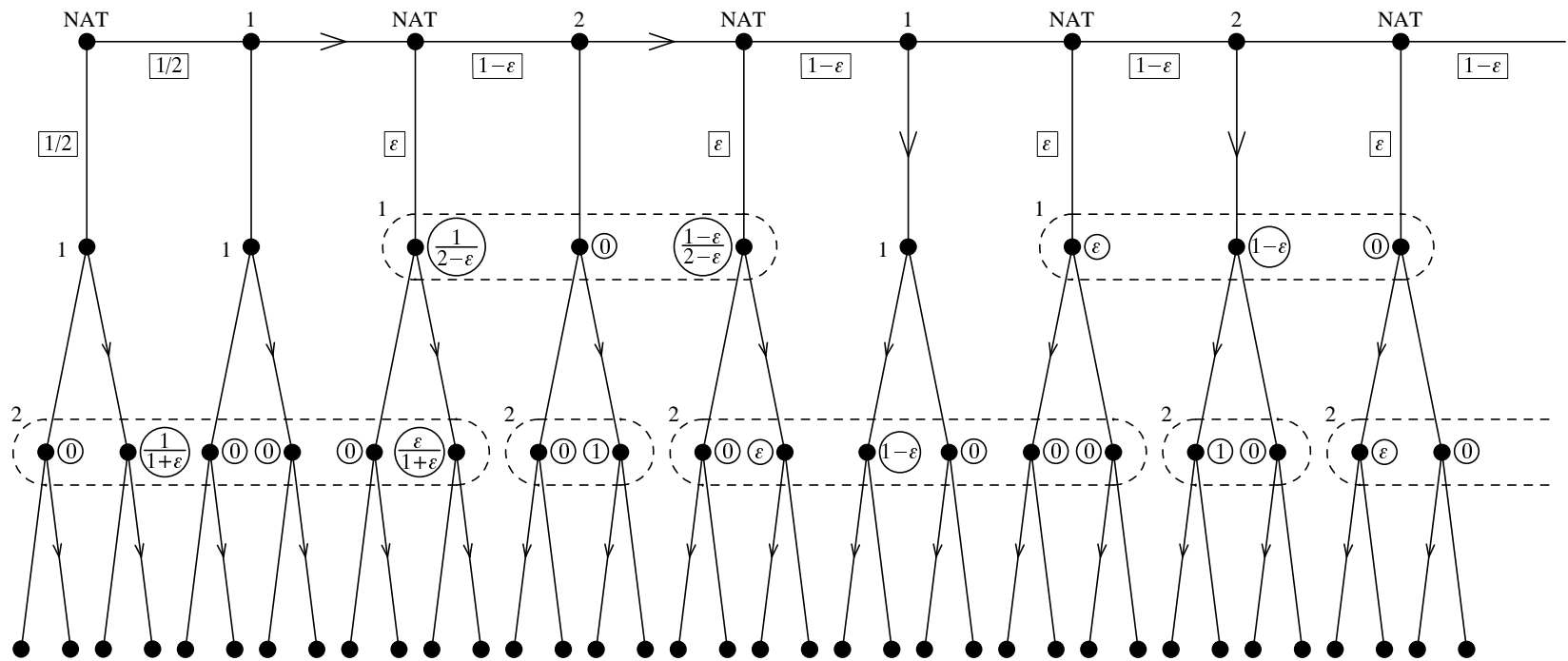


Figure 5. "Message Threshold" Equilibrium with $\bar{k} = 2$

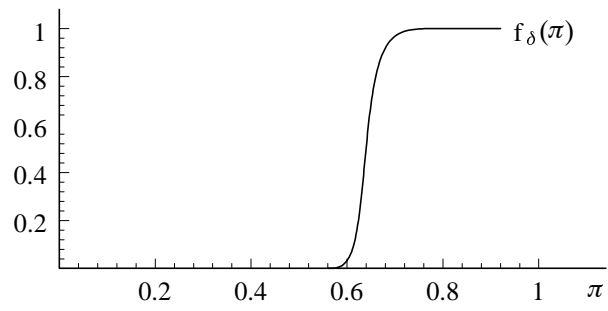


Figure 6

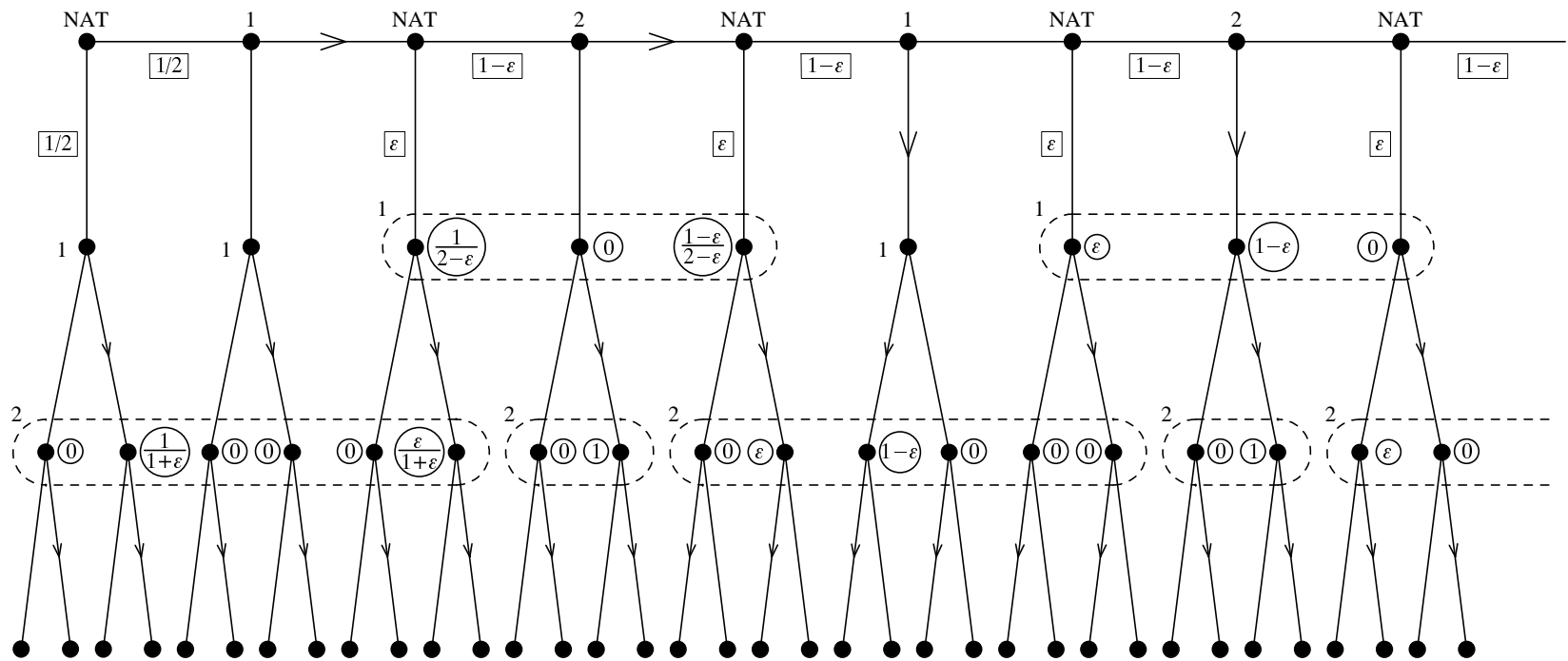


Figure 7. "Punishment" Equilibrium with $\bar{k} = 2$