OXFORD REVIEW OF ECONOMIC POLICY, VOL. 18, NO. 4

COORDINATION, COMMUNICATION, AND COMMON KNOWLEDGE: A RETROSPECTIVE ON THE ELECTRONIC-MAIL GAME

STEPHEN MORRIS
Yale University[1]

Common knowledge plays an important role in coordination problems and coordination problems are central to many areas of economic policy. In this paper, I review some common-knowledge puzzles culminating in the electronic-mail game. These puzzles may seem distant from practical concerns. However, I then argue why insights derived from this literature are useful in interpreting empirical evidence of how people coordinate under uncertainty and in understanding the role of communication in coordinating behaviour.

I. INTRODUCTION

Let us start by considering a classic common-knowledge puzzle. Alice and Bob each have an envelope, containing a certain number of dollar bills. Both know that one envelope contains one more dollar than the other, with Alice having an odd number of dollars and Bob having an even number of dollars. Each individual observes the amount of money in his or her envelope, but not the amount in the other individual's envelope. First, Alice and Bob are each asked to announce if he or she knows that he or she is richer, i.e. has the larger amount of money. Of course, neither knows this, so neither makes an announcement. Now, it is publicly announced that no envelope contains a thousand dollars or more. Initially, nothing happens. But eventually one of the two announces that he or she is richer. Why?

If Alice in fact had $999, she would immediately realize that Bob must have $998, and so she would announce that she was the richer. But now if Bob has $998 and does not hear Alice announce that she is the richer, he will be able to infer that Alice must

[1] This work is based on a talk prepared for the 2001 European Meetings of the Econometric Society in Lausanne.

have $997; thus he will announce that he is richer. Now we may argue by backward induction, and infer that eventually the richer individual will become aware that he/she is the richer individual. The key to the story is that the announcement is *public*. If Alice and Bob were each privately informed that no envelope contained $1,000 or more, then—unless Alice has exactly $999—neither individual would have learned anything useful. But the public announcement generates *common knowledge* among Alice and Bob that no one has more than $1,000. That is, Alice and Bob both know it, both know that both know it, and so on. The story illustrates the important role that common knowledge plays in interactive reasoning.

The importance of common knowledge, and the distinctive role that it plays, was also highlighted by the electronic mail (e-mail) game, described by Rubinstein (1989). Rubinstein considered a setting where—if there is common knowledge of payoffs—two players of a game can coordinate on an efficient outcome. However, suppose that there is merely a high number of levels of knowledge, but not common knowledge. For example, suppose that each player knows the true payoff matrix; each player knows that both players know the true payoff matrix; but one player is not sure if the other player knows that he knows the true payoff matrix. In this case, Rubinstein exhibits a setting where the unique equilibrium has both players forced to an inefficient outcome because of the lack of common knowledge. And this happens no matter how many finite levels of knowledge of payoffs there are, as long as there is not common knowledge. The e-mail game is reviewed in detail below.

Rubinstein suggested that this result was rather counter-intuitive: surely players with a high number of levels of knowledge would choose to act as if they had common knowledge? This suggests that the modelling challenge is to come up with a plausible model of bounded rationality that might limit this extreme sensitivity to very high levels of knowledge.

However, a different reading of the e-mail game example is that it provides a useful, if extreme, illustration of the logic by which higher-order beliefs and knowledge might influence outcomes in strategic settings. Informal arguments about the importance of higher-order beliefs in settings such as financial markets, bank runs, and exchange-rate crises are endemic, but standard models tend to play down the effect of higher-order beliefs (this is done implicitly, by assuming common knowledge of payoffs or modelling uncertainty with simple type spaces where higher-order uncertainty is not significant). The e-mail game has played an important role in thinking about how to model higher-order beliefs in applied settings. In particular, Carlsson and van Damme (1993) introduced the study of a class of 'global games' where each player observes the true payoffs of the game with a very small amount of noise. While their motivation for studying this class of games was quite different, it turns out that we can understand the striking equilibrium selection results that occur in these games by working out the higher-order beliefs that the information structure generates, and then the logic of the equilibrium selection is in fact closely related to what occurs in the e-mail game. Morris and Shin (1998) argued that this type of reasoning could be used to understand currency crises, and a number of authors since then have used such models in different contexts—see the article by Heinemann in this issue and the survey of Morris and Shin (2000).

But which reading is correct? Do the common-knowledge 'paradoxes' culminating in the e-mail game illustrate the need for developing models of bounded rationality that remove the paradoxical outcomes? Or do they represent a sensible starting point for modelling the role of higher-order beliefs in applied settings? In this article, I want to discuss two arguments in defence of the latter point of view. First, believing that players without common knowledge might behave according to rational predictions does not entail the belief that they reason to very high levels (as Alice and Bob were required to do in the envelopes example at the beginning of this paper). It is possible to describe heuristics for rational behaviour in risky coordination problems without common knowledge that are very simple and are fully consistent with rational behaviour. The second defence is more prosaic: I argue that the type of lack of common knowledge that arises in the stylized information structure of the e-mail game will arise much more generally in settings where there is imperfect but plentiful communication.

The paper is structured as follows. I give a brief review of the common-knowledge paradoxes, cul-

Figure 1

Bob

Alice		0	2	4	·	$2n$	$2n + 2$	·
	1	⊗	⊗		·			·
	3		⊗	⊗	·			·
	5			⊗	·			·
Alice	·	·	·	·	·	·	·	·
	$2n + 1$				·	⊗	⊗	·
	$2n + 3$				·		⊗	·
	·	·	·	·	·	·	·	·

minating in the e-mail game. My purpose here is to introduce them to those who do not know them, and also to describe a number of key paradoxes in a common language, in order to bring out their connections. The treatment is informal and non-technical. The article by Board in this issue gives a more detailed treatment of the concepts that lie behind these puzzles. Then I attempt to provide some countervailing intuition to the view that rational behaviour in the e-mail game is counter-intuitive. Finally, I review some recent research looking at how the e-mail game generalizes to more realistic communication structures. The ideas reviewed here have been discussed elsewhere. My hope is that by using the always provocative e-mail game as a unifying theme, it is possible to bring out some interesting connections.

II. COMMON-KNOWLEDGE PARADOXES: A VERY BRIEF REVIEW

In this section, I very briefly review some common-knowledge paradoxes culminating in the e-mail game. I have changed original versions of the paradoxes to try and make them more comparable. At the end of each example, I provide some very brief bibliographical notes. For histories and surveys of these issues, see Geanakoplos (1994) and Fagin *et al*. (1995, ch. 1).

(i) Envelopes

First, we see how we can represent the envelopes example discussed in the introduction. Each row in Figure 1 represents the number of dollars that Alice has in her envelope (an odd integer). Each column represents the number of dollars that Bob has in his envelope (an even integer). The symbol ⊗ appears in boxes corresponding to *possible states*, i.e. where the difference between the amount of money in Alice's envelope and the amount of money in Bob's envelope is 1.

How can knowledge and common knowledge be studied in this diagram? The event that no one has more than or equal to six dollars is shown below:

	0	2	4	6
1	*	*		
3		*	*	
5			*	*
7				

Now the event that everyone knows that no one has more than or equal to six dollars is:

	0	2	4	6
1	*	*		
3		*	*	
5			*	
7				

Figure 2

Bob

		0	2	4	·	$2n$	$2n+2$	·
	1	p	$p(1-p)$		·	0	0	·
	3	0	$p(1-p)^2$	$p(1-p)^3$	·	0	0	·
	5	0	0	$p(1-p)^4$	·	0	0	·
Alice	·	·	·	·	·	·	·	·
	$2n+1$	0	0	0	·	$p(1-p)^{2n}$	$p(1-p)^{2n+1}$	·
	$2n+3$	0	0	0	·	0	$p(1-p)^{2n+2}$	·
	·	·	·	·	·	·	·	·

The event that everyone knows that everyone knows that no has more than or equal to six dollars is:

	0	2	4	6
1	*	*		
3		*	*	
5				
7				

and so on.

This example has the same logic as some classic puzzles examples involving hats, dirty faces, and cheating wives. Nalebuff (1989) discusses a version of the envelopes problem where the individuals are asked if they want to trade envelopes; this introduces complications, that we do not explore wish here, concerning how gambles are evaluated *ex ante* and *ex post*.

(ii) Agreeing to Disagree

Now suppose that the probability that the smallest envelope contains n dollars is $p(1-p)^n$. These probabilities can be represented as in Figure 2.

What probability do Alice and Bob each assign to the event that Alice is richer (i.e. that the true state is on the leading diagonal)? Independent of the amount of money in her envelope, Alice assigns probability $1/(2-p)$ to Alice being richer. Bob, on the other hand, as long as he has at least $2 in his envelope, assigns probability $(1-p)/(2-p)$ to Alice being richer. Of course, if Bob has no money in his envelope, he knows that Alice is richer.

Now suppose that in fact Bob has $\$2n$ and Alice has $\$(2n+1)$. Both Alice and Bob know that they assign probabilities $1/(2-p)$ and $(1-p)/(2-p)$, respectively, to Alice being richer. Both know that both know this. In fact, we can use the diagram to confirm that they both know that they both know . . . ($n+1$ times) . . .that Alice is richer. But it is not common knowledge that they have these different posteriors. Aumann (1976) showed that if individuals shared a common prior, and there is common knowledge of their posterior beliefs of an event, their posterior beliefs must be equal. This example shows that it is possible to have an arbitrarily high number of levels of knowledge that individuals have different posteriors, without them being the same. Geanakoplos and Polemarchakis (1982) first gave an example making this point.

Now suppose that Alice and Bob successively announced what probability they assigned to Alice being richer. For a long time, they would stick to their initial beliefs $1/(2-p)$ and $(1-p)/(2-p)$. But eventually, one individual would realize that the other's failure to revise their beliefs gave crucial information to him or her, and they would converge to common-knowledge posteriors. Geanakoplos and Polemarchakis (1982) showed that this always happens on finite spaces (Washburn and Teneketzis (1984) provide the infinite state analogue).

(iii) Coordinated Attack

For the next two examples, we change the story but keep the type space and figures.

Now suppose that Alice and Bob are deciding whether to make a risky investment decision. Each would like to invest if he or she (i) has the resources to invest and (ii) *knows* that the other individual is investing. It is common knowledge that Alice has the resources to invest, but initially Alice does not know whether Bob has resources or not. If Bob is able to invest, he sends an e-mail message to Alice informing her of this fact. However, there is a possibility that the message will be lost. If the message arrives, Alice sends a message back to Bob, confirming that she received the first message. Again, the message may be lost. If then Bob receives the confirmation, he sends another message in turn. And so on.

We can use Figure 1 to represent this situation. The rows represent possible types for Alice in this scenario, while the columns represent types for Bob. If Bob is type 0, this means he does not have resources for investment (and thus he does not send a message). If Alice is type 1, it means that she has received no messages. If Bob is type 2, it means he has resources for investment and he has sent a message, but he has not received a confirmation. If Alice is type 3, it means that she has received a message informing her that Bob has resources for investment, has sent a confirmation, but has not received a confirmation of her confirmation. And so on. Notice that if an individual is type n, then n is the maximum number of messages that that individual thinks may have been sent and that individual thinks the other individual is either type $n-1$ or type $n+1$.

Can we come up with a rule specifying when each individual should invest as a function of his or her type, such that (i) no individual ever invests alone and (ii) an individual who is supposed to invest always has resources to invest? The answer is no, and the proof is elementary. Let n_A be the lowest type of Alice who invests and let n_B be the lowest type of Bob who invests. By (ii), $n_B \geq 2$. By (i), $n_A \geq n_B + 1$ and $n_B \geq n_A + 1$ if $n_A \geq 3$. For any finite n_A and n_B, we obtain a contradiction.

This example is due to Gray (1978). His story was that Alice and Bob were two generals commanding armies on different hills, and they sent messengers to each other who might get lost. The argument generalizes to show that if coordinated action requires common knowledge of some fact, and that fact is not initially common knowledge, then no amount of communication through faulty channels can ever generate common knowledge and thus coordinated action.

(iv) The Electronic-mail Game

To turn the previous example into a game, we must add probabilities to the various type profiles and add payoffs to the various action profiles. Let the probability that Bob does not have resources to invest be p and let the probability that any message gets lost be (coincidentally) p also. One can verify that this story gives rise to the probability distribution over types of Figure 2. Let the payoffs for Alice and Bob be given by the following matrix if both have resources to invest:

	Invest	Not invest	
Invest	1,1	–2,0	
Not invest	0,–2	0,0	(1)

If Bob does not have resources to invest, then he has a dominant strategy not to invest.

Now we have a game of incomplete information. The unique equilibrium of this game has Alice and Bob never investing. To see why, first observe that if Bob is type 0, he has a dominant strategy not to invest. If Alice is type 1, she assigns probability $1/(2-p)$ to Bob being type 0. But we see from the payoff matrix (1) that if Alice assigns probability more than 1/3 to Bob not investing, she will not invest. So, since $[1/(2-p)] > (1/2) > (1/3)$, Alice does not invest if she is type 1. But now if Bob is type 2, he knows that Alice will not invest if she is type 1, and he assigns probability $1/(2-p)$ to Alice being type 1. So he will not invest. This argument iterates to ensure that no type of either player ever invests.

This is the e-mail game of Rubinstein (1989).[2] The result is more than just adding numbers to the coordinated attack problem. As Rubinstein noted, the strategic paradox relies on the fact that payoffs are asymmetric when players do not coordinate their behaviour, i.e. the player who invests alone bears the cost of miscoordination. Also, it is worth emphasizing that while Gray's paradox shows that perfect coordination is impossible without common knowledge, one can straightforwardly choose rules of behaviour such that the probability of miscoordination is small when the probability of errors (p) is sufficiently small, *if* the players could be relied upon to follow those rules. For generals commanding armies, and in Gray's applications to computer science, there should not be a problem having *ex-ante* optimal rules of behaviour followed *ex post*. But in the e-mail game, players would like to make an *ex-ante* commitment to invest, but if they are unable to do so, then not only is there no equilibrium with perfect coordination, there is also no equilibrium with any positive probability of miscoordination. Morris and Shin (1997) discuss at more length the different implications of the common-knowledge paradoxes in economics, computer science, and philosophy.

III. SHOULD WE CARE ABOUT COMMON-KNOWLEDGE PARADOXES?

In this section, I review two kinds of arguments as to why the coordination problems that arise in the e-mail game should be taken seriously: first, one can provide simple intuition for rational behaviour in the e-mail game that does not rely on complex reasoning on the part of players; second, I review some work confirming that the type of problems that emerge in the e-mail game occur more generally.

(i) Heuristics for Risky Coordination

Rubinstein's abstract for the e-mail game is the following:

> The paper addresses a paradoxical game-theoretic example which is closely related to the coordinated attack problem. Two players have to play one of two possible coordination games. Only one of them receives information about the coordination game to be played. It is shown that the situation with 'almost common knowledge' is very different from when the coordination game played is common knowledge.

The paradox is that in two apparently close situations—common knowledge of payoffs versus an arbitrarily high number of levels of knowledge of payoffs—players behave very differently. If you buy the intuition that these situations are very close in the minds of the players, then you are led to the conclusion that players should not be expected to behave rationally in the e-mail game. With this interpretation, the message of the e-mail game is that one should try to come up with a model of boundedly rational behaviour that delivers predictions that are insensitive to whether there is common knowledge or a large number of levels of knowledge.[3] While I describe an alternative possible interpretation, this remains a very interesting line of research that has not been much pursued (see Dulleck (1998) for one exception).

An important paper by Monderer and Samet (1989) suggests an alternative intuition, however. Suppose we turn around the question posed by the e-mail game and ask what definition of approximate common knowledge delivers the conclusion that rational behaviour when there is approximate common knowledge is close to rational behaviour when there is common knowledge. Say that event is *common p-belief* if everyone believes it with probability at least

[2] There are some small changes to simplify my presentation. 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 simply faced the problem of coordinating on an efficient Nash equilibrium that varied across states. Finally, the probability of lost messages and the *ex-ante* probability of different payoffs have been set equal.

[3] Rubinstein (1989) writes 'What would you do if the number on your screen is 17? It is hard to imagine that when L is slightly above M and ε is small a player will not play B. The sharp contrast between our intuition and the game-theoretic analysis is what makes this example paradoxical. This example joins a long list of games such as the finitely repeated Prisoner's Dilemma, the chain store paradox, and Rosenthal's game, in which it seems that the source of the discrepancy is rooted in the fact that in our formal analysis we use mathematical induction when reasoning. Systematic explanation of our intuition that we will play B when the number on our screen is 17 (ignoring the inductive consideration contained within Proposition 1's proof) is definitely a most intriguing question.'

p, everyone believes with probability at least p that everyone believes it with probability at least p, and so on. They establish that behaviour when there is common p-belief of payoffs, with p close to 1, is close to behaviour when there is common knowledge. Monderer and Samet (1996) and Kajii and Morris (1998) show that no weaker notion of approximate common knowledge will suffice (and both papers use examples with a similar logic to the e-mail game to show tightness). But these mathematical results—the topology generated by common p-belief is the weakest one generating continuity of equilibrium behaviour—also generates an alternative intuition about rational play in settings of risky coordination.[4]

Aumann (1976) showed that an event E is common knowledge (in the iterative sense described earlier) at some state ω if and only if there is an *evident* event (an event that everyone knows to be true whenever it is true) that is true at state ω and implies event E. An event that occurs in public is automatically evident. Monderer and Samet (1989) show that an event E is common p-belief (in the iterative sense described earlier) if and only if there is a *p-evident* event (an event that everyone believes to be true with probability at least p, whenever it is true) that is true at state ω, with the property that whenever the p-evident event is true, everyone believes the event E with probability at least p. An event is p-evident for p close to 1 if it is almost public: whenever it occurs, everyone is almost sure that it occurs. Thus the intuition suggested by the literature on common p-belief is that players will be able to coordinate on some efficient but strategically risky behaviour if and only if the true payoffs are almost public. The more risky a Nash equilibrium of a complete information game is, the more public payoffs must be in order to support play of that Nash equilibrium (this claim can be made precise). A distinctive feature of the information structure of the e-mail game is that (if the probability of errors is small), *no* non-trivial event is common p-belief for p greater than 1/2. This means the *risk-dominant* Nash equilibrium must be played in the e-mail game, because of the lack of almost-public events supporting play of the risk-dominant equilibrium (Morris *et al*., 1995).[5]

So we have two competing intuitions for how rational people might play the game described in matrix (1). One intuition says that if there is uncertainty about payoffs, you will not invest, but if there is a high number of levels of knowledge about payoffs, you should behave as if there is common knowledge and invest. Another intuition says that if there is uncertainty about payoffs, you will not invest, but if there is common p-belief of payoffs, for $p \geq 2/3$, you should invest. The latter intuition is consistent with standard theory but requires behaviour to be sensitive to how public knowledge of payoffs is. The global games introduced by Carlsson and van Damme (1993) also have the feature that there are no non-trivial almost public events, and this is the driving force behind results in that literature—the theoretical and applied literature on global games, and the relationship to higher-order beliefs, are surveyed in Morris and Shin (2000).

There has been some experimental work on this subject (this is surveyed in the article by Heinemann in this issue). Camerer (2000) reports a small experiment that is very close to the original e-mail game and Cabrales *et al*. (2002) construct a five-signal game in the spirit of the e-mail game and global games. Both sets of results are consistent with earlier work on dominance-solvable games, showing that while initial play is not equal to the unique rational prediction, as subjects learn to play the game, they converge to the unique rational outcome. Cabrales *et al*. conclude (both from the experimental results and subject interviews) that the players do not engage in iterative reasoning to any significant extent. Rather, repeated play leads to convergence to the unique equilibrium. Note that since all these games are dominance solvable, the convergence to the unique equilibrium occurs robustly in many learning and evolutionary models (see, for example, Nachbar, 1990; Milgrom and Roberts, 1990).

[4] As Rubinstein (1989) notes, 'other definitions of convergence may be useful not only as technical methods but also for expressing other intuitions of closeness'.

[5] If a symmetric two-player, two-action game has two strict Nash equilibria, the risk-dominant Nash equilibrium is one where each player is choosing a best response to a 50/50 probability distribution over the actions of his opponent (Harsanyi and Selten, 1988). In the game of matrix (1), (Not invest, Not invest) is the risk-dominant equilibrium.

The above experimental work does not directly address the competing intuitions discussed here. It would be nice to establish that—even without learning—while common knowledge of payoffs may lead to the efficient equilibrium being played, lack of common knowledge of the type induced by the e-mail game and global games leads to inefficient but risk-dominant equilibria being played. This is a subtle thing to test, because there is a tendency to play the risk-dominant equilibrium in experiments even when a lack of common knowledge is not (deliberately) included into the experimental design. Indeed, Cabrales *et al*. (2002) note that there was a strong tendency to converge to the risk-dominant equilibrium even when there was common knowledge. The related experimental work of Heinemann *et al*. (2002) on global games also finds that while the results are consistent with theory concerning environments without common knowledge, similar results hold when there is common knowledge.

An intriguing recent experimental result of Chaudhuri *et al*. (2001) supports the view that the existence of almost public events (i.e. events that are p-evident with p close to 1) may be key in generating risky coordination. They examine experimental subjects' play in the minimum effort game of Van Huyck *et al*. (1990). In this game, everyone choosing low effort is a safe equilibrium—analogous to the risk-dominant equilibrium. There are also equilibria with high levels of effort, but these are strategically risky in the sense that a player will be prepared to choose high effort only if there is a p-evident event (with high p) contingent on which everyone will choose the risky action. Chaudhuri *et al*. allow players to receive advice from a previous generation of players about how to play the game. They observe that if players privately observe the advice received by all players, coordination on the high-effort equilibrium still does not occur. However, if the advice is read out to players publicly, they are able to coordinate on the high-effort equilibrium. Since the procedure being followed is common knowledge, both these scenarios in principle generate common knowledge of the advice received by all players. However, players treat the two scenarios differently. One possible explanation is that in the first scenario, players act as if there is some probability that others do not read the advice they receive; another is that players do not realize that if there is common knowledge that everyone receives some information, there is common knowledge of that information. In either case, these experimental results provide support for the heuristic of only choosing a risk action in a coordination game if there are sufficiently public events supporting that behaviour.

Let me conclude by mentioning another useful heuristic highlighted in Morris and Shin (2000). In symmetric binary-action coordination games with a continuum of players and a global-games noise structure, the lack of almost-public events pins down a prediction of how people will play (these games are discussed in more detail in Heinemann's article in this issue). Each player will always choose a best response to a uniform belief over the proportion of players choosing each of the two actions. In other words, the unique strategy surviving iterated deletion of dominated strategies will imply identical behaviour to the very simple heuristic of acting as if you are in complete ignorance of others' strategy choices. This, too, suggests that taking seriously a lack of common knowledge in economic settings does not require a commitment to the belief that players reason to high levels of knowledge. Notice that the key here is that even very accurate information about payoffs allows for strategic uncertainty (uncertainty about others' actions) to play an important role in pinning down outcomes. This feature mirrors a key element of the common-knowledge paradoxes discussed earlier, where, for example, Alice was always almost sure about the amount of money in Bob's envelope, but the small amount of uncertainty was leveraged into a large impact.

(ii) Taking Noisy Communication Seriously

Rubinstein (1989) wrote: 'The story of the interchange of messages by electronic mail is intended only to provide a precise, albeit rather special, model of how knowledge comes to be shared by the players.'

The purpose of this section is to review work that examines in more detail how the outcome of the e-mail game depends on the communication protocol. The models remain rather stylized, but we can examine in more detail when almost public events emerge from a communication process. I start with three more straightforward questions. How does the result depend on the error probabilities? What happens when a large number of players must coordinate their behaviour, and they meet publicly, but in small groups? And what happens when we make realistic assumptions about the timing of messages? Then I look at two more subtle questions that arise when we start to endogenize the communication process. What happens if we stick to the original communication process of the e-mail game, but make it a strategic decision whether to send messages or not? And what happens when we add unprovable but public messages (cheap talk) into the communication process? The bottom line is going to be that the issues concerning communication and coordination raised by the e-mail game are fairly robust, but there are many open issues.

Error probabilities

Dimitri (2000) has examined the robustness of the e-mail game to different message-error probabilities across players. Suppose that the probabilities of Alice and Bob's messages getting lost are p_A and p_B, respectively, and the probability that Bob does not have resources to invest is q. Then, one can show that there is no investment in equilibrium if and only if

$$p_B < 2[q/(1 - q)],$$
$$p_B < 2[p_A/(1 - p_A)],$$
$$\text{and } p_A < 2[p_B/(1 - p_B)].$$

Thus if p_B is small relative to q and p_A and p_B are sufficiently close to each other, investment will not occur. If they are very different, then a player who does not receive a confirmation will be sure where it got lost. This ensures that there are sufficiently almost-public events to allow coordination.

Many-player communication

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 fewer than n players end up investing, then the payoff to investing is –2. The payoff to not investing is always 0.

The *ex-ante* probability that the state is good is 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 $p(1 - p)^k$, where $p > 0$. Thus conditional on k meetings having occurred, the conditional probability that at least one more meeting will occur is always $1 - p$. The collection of players gathered together at the jth meeting is random: each subset is equally likely. Thus the probability than any one player attends any fixed meeting is always m/N. If a first meeting occurs (this happens with *ex-ante* probability $1 - p$), 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. A player does not know how many meetings (if any) occurred after the last meeting that he attended.

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

Proposition 1: if $m < n$ and $(1 - p)[(N - m)/N] < 2/3$, then the meetings game has a unique equilibrium: no player ever invests (Morris, 2001).

This result is tight in the following sense. If *either* $m \geq n$ *or* $(N - m)/N \geq 2/3$ and p is sufficiently small, then there is an equilibrium where all players invest if they know that the state is good.

Real time communication

Consider a timing game that is identical to the e-mail game of section II(iv), with one change. Instead of a message getting lost with exogenous probability p, assume instead that it may take some 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 $f(.)$ for the density corresponding to $F(.)$). 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 that a player receives his last message at time t. What probability does he assign to his message having been the last? If he has just received a message at date t, the probability that his message never arrives is

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

the probability that his message arrives but he never receives a reply is

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

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 his message was the last sent, conditional on not receiving a confirmation, is

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

Proposition 2: if $L(t) < 2$ for all $t \in [0,T]$, the timing game has a unique equilibrium: both players never invest (Morris, 2001).

This result is tight in the following sense. If $L(0) \geq 2$, 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. Morris (2001) considers some examples that illustrate this condition.

Strategic communication decisions

The information structure in the e-mail game is clearly an unfortunate one for players who would like to coordinate on the efficient outcome. If one were designing an optimal communication system *ex ante*, it would be optimal to have only one or two messages sent (see Chwe, 1995). However, suppose that you were unable to commit to a communication system, but instead were simply playing the e-mail game of section II(iv), with the additional complication that players could choose not to send a message at any point. Binmore and Samuelson (2001) have examined this question.

A first observation is that this game has many equilibria. There is an 'infinite message–no investment' equilibrium identical to the original equilibrium of the e-mail game, where players choose to keep sending their messages (until one gets lost) and no investment takes place (in this equilibrium, if you choose not to send a message, your opponent will just assume that your message got lost). There is also a 'no message–no investment' equilibrium, where no messages get sent and no investment takes place (if you do not expect anyone ever to invest, you obviously do not have an incentive to send any messages; if you do not expect your opponent ever to invest, you do not have an incentive to invest). But there are also efficient equilibria. For example, there is a 'one message' equilibrium where Bob sends a message confirming that he has the resources for investment, but Alice sends no confirmation. Bob will then invest if and only if he has resources for investment and Alice will invest if and only if she has received a message from Bob. In this equilibrium, Alice has no incentive to confirm receipt of the message from Bob, since (i) she knows Bob will invest and (ii) Bob completely ignores her message in equilibrium anyway.

So to make any predictions in this game, it will be necessary to modify the game and/or the solution concept. Binmore and Samuelson (2001) consider a number of modifications. Their strongest results come from considering evolutionary stability in a perturbed game where there are small *ex-ante* costs to listening to each message and there is a small *ex-post* cost to sending each message. Since messages are costly, the infinite message–no investment strategy profile of the original e-mail game cannot be an equilibrium. But they also show that the no message–no investment equilibrium is not evolutionarily stable. That is, for small costs, that strategy can be invaded by the 'one message' equilibrium. Any equilibrium where messages get sent and investment takes place is evolutionarily stable.

On the other hand, Morris (2001) suggests an intuition as to why the no investment equilibrium might be robust. Recall that anticipated confirmations are damaging in the e-mail game because if I send you a message, anticipating a confirmation, and I do not receive a confirmation, then I assign probability $1/(2-p)$ to my message never having reached you; however, if I had not anticipated a confirmation, I would assign probability p (the probability of a message getting lost) to my message never having reached you. Thus, for small p, 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 potentially Pareto-improving. If I did not expect to receive a confirmation from you, a failure to receive one would 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.

Why does this intuition not work to rule out the 'one message' equilibrium in the basic e-mail game with strategic message-sending decisions? As I noted above, the problem is that Alice expects Bob to invest for sure anyway (even without the confirmation). In Morris (2001), I therefore consider a perturbed version of the game where the probability of a player investing is always strictly increasing in the probability he assigns to his opponent investing (this is achieved by adding an idiosyncratic component to each player's payoff that, with small probability, will reverse his optimal action). In this case, if players followed the strategy profile of the one message equilibrium, Alice would be almost sure that Bob would invest, if Alice does not send a confirmation. But if Bob interprets an unanticipated confirmation from Alice in the natural way as a signal that Alice plans to invest, then an unanticipated confirmation will increase the likelihood of Alice investing. So perturbing the game with idiosyncratic payoff noise and refining out-of-equilibrium beliefs in the spirit of this argument, implies that only no investment equilibria can be played.

Clearly, different perturbations and refinements lead to different conclusions about the robustness of the e-mail game to strategic message sending. We do not know enough to come to any definitive conclusions on this issue.

Adding public cheap talk

The analysis of the previous section was based on the assumption that players had a strategic choice about whether to send a message or not. But it was assumed that any messages sent were true; i.e. if a player claimed to have received a message, he really had received that message. Ideally, one would like to model a situation where communication takes place both privately and publicly and some claims are provable while others are 'cheap talk'. However, such a general analysis would be complex. In this section, I give one example to illustrate the limitations of unprovable cheap talk, even if it is public, in overcoming the lack of common knowledge problem in the e-mail game. This example comes from Baliga and Morris (2000).

Suppose that we start with the information structure of the e-mail game, described in section II(iv). But now suppose that prior to making their action choices, players have the opportunity to make some public cheap talk statements. For example, Bob could publicly announce that he has the resources for investment, or Alice could publicly announce that she has received the first message from Bob. If either of these messages were credible in equilibrium, then investment would be possible in equilibrium following the public statements. But whether they are credible depends on the incentives of Bob when he does not have the resources to invest. We have assumed that, in this case, he has a dominant strategy not to invest. But that leaves open the question of what he would like Alice to do (conditional on the fact that he is not going to invest). If he would like Alice not to invest when he does not invest, then cheap talk will solve the coordination problem. Bob could simply announce whether or not he has resources to invest, and Alice and Bob would both invest if Bob's public announcement said that he did have resources to invest.

The more subtle case arises if Bob always wants Alice to invest (even if he does not himself have resources to invest). In this case, the truth-telling equilibrium breaks down: Bob would have an incentive to report that he had resources to invest even if he did not. In fact, *every* equilibrium has no investment by any type. To see why, fix an equilibrium and let n be the lowest type of either player who ever invests with positive probability (after any message). Since Bob's type 0 has a dominant strategy not to invest, we must have $n \geq 1$. Without loss of generality, let n be even, i.e. a type of Bob. Let M^* be the set of messages that Alice might send that lead Bob's type n to invest with positive probability. Alice's type $n - 1$ sends a message in that set with probability 1 (since she knows that Bob's type $n -$

2 is not investing, she chooses her message to maximize the probability that Bob's type n invests). Now recall that *ex ante*, Bob's type n assigned probability $1/(2 - p)$ to Alice being of type $n - 1$. Conditional only on observing a message in M^*, that probability must weakly go up (Alice's type $n - 1$ always sends a message in M^*, even though type $n + 1$ may not). Thus for at least one message in M^*, Bob's type n must assign probability at least $1/(2 - p)$ to Alice being type $n - 1$, and therefore not investing. But then it is a strict best response for Bob not to invest, a contradiction.

This example suggests that once there is a lack of common knowledge of payoffs, even public cheap talk may fail to resolve it. However, it turns out that the strong result that cheap talk makes no difference to the equilibrium set is rather special. See Baliga and Morris (2000) and Baliga and Sjöström (2001) for more on this issue.

REFERENCES

Aumann, R. (1976), 'Agreeing to Disagree', *Annals of Statistics*, **4**, 1236–9.

Baliga, S., and Morris, S. (2000), 'Co-ordination, Spillovers and Cheap Talk', Discussion Paper No. 1301, Northwestern University.

— Sjöström, T. (2001), 'Arms Races and Negotiations', Northwestern University, mimeo.

Binmore, K., and Samuelson, L. (2001), 'Coordinated Action in the Electronic Mail Game', *Games and Economic Behavior*, **35**, 6–30.

Cabrales, A., Nagel, R., and Armenter, R. (2002), 'Equilibrium Selection through Incomplete Information in Coordination Games: An Experimental Study', working paper, http://www.econ.upf.es/cgi-bin/onepaper?601.

Camerer, C. (2000), 'Iterated Reasoning in Dominance-solvable Games', ch. 5 of book draft, available at http://www.hss.caltech.edu/CourseSites/Psy101/psy101.html

Carlsson, H., and van Damme, E. (1993), 'Global Games and Equilibrium Selection', *Econometrica*, **61**, 989–1018.

Chaudhuri, A., Schotter, A., and Sopher, B. (2001), 'Talking Ourselves to Efficiency: Coordination in Inter-generational Minimum Games with Private, Almost Common and Common Knowledge of Advice', available at http://homepages.nyu.edu/~as7/Chaudhuri_Schotter_Sopher.pdf

Chwe, M. (1995), 'Strategic Reliability of Communication Networks', University of Chicago, mimeo.

Dimitri, N. (2000), 'Efficiency and Equilibrium in The Electronic Mail Game', *Quaderni del Dipartimento di Economia Politica*, 295, University of Siena.

Dulleck, U. (1998), 'A Note on the Email Game—Bounded Rationality and Induction', Humboldt University, Berlin, mimeo.

Fagin, R., Halpern, J., Moses, Y., and Vardi, M. (1995), *Reasoning about Knowledge*, Cambridge, MA, MIT Press.

Geanakoplos, J. (1994), 'Common Knowledge', in R. Aumann and S. Hart (eds), *Handbook of Game Theory*, vol. 2, ch. 40, Amsterdam, Elsevier Science.

— Polemarchakis, H. (1982), 'We Can't Disagree Forever', *Journal of Economic Theory*, **28**, 192–200.

Gray, J. (1978), 'Notes on Database Operating Systems', in R. Bayer, R. Graham, and G. Seegmuller (eds), *Operating Systems: An Advanced Course*, New York, Springer.

Harsanyi, J. and Selten, R. (1988), *A General Theory of Equilibrium Selection in Games*, Cambridge, MA, MIT Press.

Heinemann, F., Nagel, R., and Ockenfels, P. (2002), 'Speculative Attacks and Financial Architecture: Experimental Analysis of Coordination Games with Public and Private Information', working paper, http://www.sfm.vwl.uni-muenchen.de/heinemann/publics/hno-1.htm

Kajii, A., and Morris, S. (1998), 'Payoff Continuity in Incomplete Information Games', *Journal of Economic Theory*, **82**, 267–76.

Milgrom, P., and Roberts, J. (1990), 'Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities', *Econometrica*, **58**, 1255–78.

Monderer, D., and Samet, D. (1989), 'Approximating Common Knowledge with Common Beliefs', *Games and Economic Behavior*, **1**, 170–90.

— — (1996), 'Proximity of Incomplete Information in Games with Common Beliefs', *Mathematics of Operations Research*, **21**, 707–25.

Morris, S. (2001), 'Faulty Communication', mimeo, available at http://www.econ.yale.edu/~sm326/faulty.pdf

Morris, S., and Shin, H. S. (1997), 'Approximate Common Knowledge and Co-ordination: Recent Lessons from Game Theory', *Journal of Logic, Language and Information*, **6**, 171–90.

— —(1998), 'Unique Equilibrium in a Model of Self-fulfilling Currency Attacks', *American Economic Review*, **88**, 587–97.

— —(2000), 'Global Games: Theory and Applications', available at http://www.econ.yale.edu/\~smorris/seattle.pdf

— Rob, R., and Shin, H. S. (1995), '*p*-Dominance and Belief Potential', *Econometrica*, **63**(1), 145–57.

Nachbar, J. (1990), ' "Evolutionary" Selection Dynamics in Games: Convergence and Limit Properties', *International Journal of Game Theory*, **19**, 59–89.

Nalebuff, B. (1989), 'Puzzles: The Other Person's Envelope is Always Greener', *Journal of Economic Perspectives*, **3**, 171–81.

Rubinstein, A. (1989), 'The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge', *American Economic Review*, **79**, 385–91.

Van Huyck, J., Battalio, R., and Beil, R. (1990), 'Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure', *American Economic Review*, **80**, 234–48.

Washburn, R., and Teneketzis, D. (1984), 'Asymptotic Agreement among Communicating Decisionmakers', *Stochastics*, **13**, 103–29.