Edinburgh School of Economics
Discussion Paper Series
Number 125

Communication in Games of Incomplete Information: The Two-player Case

R Vijay Krishna (University of Edinburgh)

Date
September 2004

Published by
School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361
http://www.ed.ac.uk/schools-departments/economics

THE UNIVERSITY *of* EDINBURGH

Communication in Games of Incomplete Information: The Two-player Case

R. Vijay Krishna[*][†]

September, 2004

Abstract

We study the effect of communication in two-person games of incomplete information. We show that any rational mediated communication mechanism satisfying a Nash domination condition can be implemented as the perfect Bayesian equilibrium of a communication extension of the original game and ends in finite time with probability 1.

1 Introduction

Since the seminal paper of Crawford and Sobel [6], it has been well understood that communication in games of incomplete information can expand the set of equilibrium outcomes that can be achieved and provide the players with Pareto improving outcomes. In this paper, we examine the question of how much the set of equilibrium outcomes can be expanded via cheap communication procedures in two-person games of incomplete information.

The question is obviously moot if the players can write enforceable contracts wherein they report their private information to each other and subsequently take an action. We will refer to such situations as consisting of *perfect commitment*. The set of such contracts exhaust all the achievable outcomes. But in many economic situations of interest, there is no

[*]This paper is based on chapter 3 of my doctoral dissertation at The Pennsylvania State University. I would like to thank my thesis advisors Profs. Kalyan Chatterjee and Tomas Sjöström for their support and discussions on, among other things, noisy mechanisms, Prof. Françoise Forges for very detailed comments and Prof. James Jordan for some extremely helpful conversations. All errors that remain are my own.

[†]School of Economics, University of Edinburgh, William Robertson Building, 50, George Square, Edinburgh EH8 9JY, UK. *Email:* `vijay.krishna@ed.ac.uk`.

commitment at all—players have unverifiable private information, there is no enforcement agency to which the players can appeal and they cannot commit to any course of action. Indeed, the Crawford-Sobel results are of such interest precisely because even though the players cannot commit to telling the truth, there is some sharing of information between a Sender (who has the private information) and a Receiver (who takes the only payoff relevant action).

In Bayesian games (i.e. games of incomplete information) there is another construct which is extremely useful. This is the notion of the mediated communication mechanism. The idea is that there is a disinterested party, known as the *mediator* to whom the players report their private information and the mediator then makes private, incentive compatible recommendations. It can be shown (see Myerson [16]) that even in instances where there is no commitment, the set of incentive compatible mediated solutions is exactly the same as the set of incentive compatible enforceable contracts in the case of perfect commitment. This is because although the players cannot commit, it is the disinterested mediator who makes the recommendations and his being disinterested guarantees that the players trust him. The players make honest reports to the mediator because they trust him to make recommendations according to the mediation plan. They then follow his recommendations because it is incentive compatible to do so. Following Myerson (pp. 261-262 in Chapter 6 of [16]), we say that "… incentive-compatible mediation plans [mediation communication mechanisms] are the appropriate generalisation of the concept of correlated equilibria in Bayesian games with communication."

Of course, this still leaves unresolved the case where there is no commitment and no disinterested mediator. In finite games of one-sided incomplete information, it is well known that extended cheap talk can increase the set of equilibrium outcomes. It is also known that the more general conversations that can be considered are conversations without an exogenous deadline. Indeed, there is an example (due to Simon [17]) of a game where finite conversations of any length do not transmit any information and they neither increase the set of equilibrium outcomes nor improve on the babbling outcome. But unbounded conversations produces outcomes that are Pareto superior for both types of the player with the private information and the uninformed player.

It is known that in finite games of incomplete information with three or more players, there is a communication mechanism which implements any rational mediated outcome. In other words, the players can talk amongst themselves and achieve any outcome achievable in the presence of a mediator or if there was perfect commitment. A distinguishing feature of this mechanism is that it satisfies three important desiderata:

C1 The equilibria from the communication game induced by the protocol be perfect Bayesian;

C2 The communication be cheap, non-binding and non-verifiable and

C3 The mechanism be unmediated.

Before we discuss our desiderata, we shall briefly discuss the meaning of the terms used above. By cheap communication, we mean that the communication itself is costless. By non-binding, we mean that by taking part in the mechanism, the players do not lose any strategic options that were available to them as part of the original Bayesian game. By non-verifiability, we mean that there can be no verification in the future of anything that was said in the past. In particular, there is no verifiability of past events. (Note that this strongly resembles our intuitive idea of what constitutes "talk.") Of course, the first requirement seems natural as we would not like our outcomes to depend on incredible threats. Hence, requirements C1 and C2 can be said to constitute a robust mechanism that represents our intuitive idea of cheap talk. But requirements C1 and C2 are not enough. These are satisfied by mediated communication mechanisms. Indeed, in mediated communication mechanisms, there is no cost to communicating with the mediator and the players still retain all of their strategic options when they are about to take an action. Thus, requirement C3 is crucial, that there be no other player.

Unfortunately, there is no protocol that satisfies the above three requirements in the two player case. This is made precise below. (The reader will find all the relevant definitions in §3.)

Theorem 0. *There exists a Bayesian game and a mediated communication mechanism which satisfies a Nash domination condition such that there is no communication mechanism satisfying requirements C1-C3 above which can approximate the mediated communication mechanism outcome.*

Remark. It should be mentioned that in the sequel, we shall take all communication mechanisms to be of order type ω. For more on the so-called *transfinite conversations*, the reader is referred to Krishna [12].

Proof. We will demonstrate an example of a generic game with the property that no communication of denumerable length will simulate the presence of a mediator. Indeed, in doing so, we will also not insist on requirement C1. It will suffice to provide a Sender-Receiver game as an example. (Recall that a Sender-Receiver game is a game where the Sender has some private information and the Receiver takes the only payoff relevant action.) By Theorem A of Aumann and Hart [1], any communication in a Sender-Receiver game can be thought of as a canonical conversation, where a canonical conversation is a conversation where in odd periods, the Sender sends a signal which could potentially give the Receiver some information about the Sender's type and in even periods, the players compromise about future courses of action via joint lotteries. But now consider the example of Forges [7] who gives an example of a Sender-Receiver game where the set of mediated communication mechanism outcomes is strictly larger than the set of outcomes from unbounded canonical conversations (i.e. conversations that have denumerable length). She also demonstrates the existence of a mediated communication mechanism outcome which cannot be approximated by unmediated cheap talk. □

As mentioned in the proof of Theorem 0, we do not get very far by relaxing requirement C1. Nevertheless, as we shall see below, we can do much better if we relax requirement C2. In our Main Theorem below, we show that any rational mediated communication mechanism outcome (see the definition in §3) that that satisfies a Nash domination condition can be implemented via cheap communication mechanism that satisfies requirements C1 and C3. We are not stipulating that there somehow be any sort of commitment—all we are asking is that in the extensive form game which represents an instance of the mechanism, the players be allowed, on occasion, to verify their coordinates in the game-tree. (This will be made clear below, via an example in §4.)

From the point of view of applications, this raises some very interesting questions. It has long been thought that the lack of commitment in an economic environment poses some very real problems by reducing the set of achievable outcomes. But we have shown that this is just because all the possible communication mechanisms have not been explored. It should be emphasised that there is no enforcement of any kind and the players' participation in the protocol (which induces the extended game) is completely voluntary. They do not have to take part in the protocol, but do so because it is in their best interests to do so. This demonstrates that the problem of lack of commitment can be greatly ameliorated with communication and more so with increasingly sophisticated communication mechanisms.

In fact, one could view our result as the culmination of what can be done without a mediator or equivalently, when there is no commitment. We show that given sufficient latitude in the choice of mechanism, there is no loss of efficiency at all. This implies that in any application where there is limited commitment, the choice of mechanism is not a benign choice. Indeed, we believe it is incumbent on the modeller to explain his choice of mechanism. For instance, consider Bester and Strausz [5] who study the case of limited commitment in principal agent problems. They characterise the incentive efficient equilibria with one round of signalling from the agent. But why should this be the appropriate mechanism to consider? Why not consider multiple rounds of cheap talk or infinitely many rounds of cheap talk, or a more sophisticated mechanism (such as the one we introduce in

§§4 and 5)? We believe that these are important questions that the modeller should not shy away from.

The remainder of this paper is structured as follows. In §2 we describe the related literature and how it relates to the present paper. In §3 we introduce the model, all the relevant definitions and our Main Theorem. In §5 we provide a proof of the Main Theorem. We conclude with a brief discussion in §6.

2 Related Literature

The usefulness of cheap-talk in Bayesian games was first pointed out by Crawford and Sobel [6]. The idea of mediated solutions is discussed in great detail in Chapter 6 of Myerson [16]. A version of the mediated solution in extensive form games is described by Myerson in [15]. In Chapter 6 of [16], Myerson shows that noisy mechanisms (a precise definition is provided in §6) can help agents achieve outcomes that are otherwise not attainable. He also alludes to an equivalence between noisy communication mechanisms in complete information games and noisy mechanisms in Sender-Receiver (and incomplete information) games. Although there are no theorems to this effect, the ideas presented therein have turned out to be immensely valuable. It is this equivalence which enables us to prove our Main Theorem in §5. We adapt a mechanism due to Ben-Porath [3] which uses urns and balls to implement correlated equilibria in a large class of bi-matrix games of perfect information to our environment. In the context of equivalence of noisy mechanisms in games of incomplete information and games of perfect information, we should mention a recent paper by Urbano and Vila [18] where it is shown that if the players have bounded computational capabilities, they can achieve any correlated equilibrium in bi-matrix games. Their result is based on the fact that certain algebraic operations (exponentiation and taking logarithms) in prime fields are extremely hard to compute.

Barany [2] studies how a communication protocol can implement rational correlated equilibria in games of complete information. He shows that if there are at least four players, then any rational correlated equilibrium can be implemented via a communication protocol. Forges [8] extends this study to Bayesian games. She shows that when there are at least four players, every *ex-ante* mediated outcome can be implemented as a correlated equilibrium of a cheap talk extension of the original game. Now using the result of Barany, it is straightforward to show that every rational mediated outcome can be implemented as a Nash equilibrium. Here, ex-ante means that all the communication occurs before the players learn their types. Gerardi [10] extends this result to the standard interim case, where players communicate after learning their types. Unfortunately, none of these results use procedures that are sequentially rational.

The problem of sequential rationality is addressed by Ben-Porath [4] who also requires only three players. He shows that in any Bayesian game with at least three players, any rational mediated outcome satisfying certain individual rationality constraints can be implemented via a communication protocol which satisfies requirements (1) and (2) above and ends in finite time. (Requirements (1) and (2) were, in fact, adapted from [4].)

We should also mention a result due to Lehrer and Sorin [13]. They point out that in general, mediated communication mechanisms need not be stochastic. They introduce a deterministic mediated mechanism which can implement any rational mediated mediated communication mechanism (which, in general, is not deterministic). Their result applies to both complete information games (wherein they implement correlated equilibria and Bayesian games. It should be noted that the mechanism we use in our example and in the proof of our Main Theorem in §§4 and 5 respectively, seems as much a descendant of the Lehrer-Sorin mechanism as of the Ben-Porath mechanism.

There is also a literature which studies outcomes from pure cheap talk alone (i.e. communication which satisfies our requirement (2)) in two-player games. The main paper in this area is Aumann and Hart [1] who characterise the set of all Nash equilibrium outcomes from pure cheap talk with unbounded conversations in two-person Bayesian games with one-sided incomplete information. The example in Forges [7] is based on this theory. Krishna [12] studies a special class of such games, namely the so-called Sender-Receiver games. He shows that if a Sender-Receiver game satisfies a certain condition, then, among other things, the set of equilibria from unbounded conversations is the same as the set of equilibria from conversations that are almost surely finite.

Pure cheap talk is used in a variety of applications and we shall only mention two instances. Prime among these is the Crawford-Sobel model. It and it's variants have been the workhorse in a large number of political economy models (see [11]). Another example of cheap talk at work is Matthews and Postlewaite [14], who show that with preplay communication in a standard k-double auction, the set of equilibrium outcomes is independent of the parameter k.

3 The Model and Main Result

A *game of incomplete information* (which we shall also refer to as a *Bayesian game*) is characterised by $\Gamma^b := (N, (C_i)_{i\in N}, (T_i)_{i\in N}, P, (u_i)_{i\in N})$, where N is the set of players and for each player $i \in N$, the set of possible actions is C_i, the set of possible types is T_i, P is a probability measure on T and i's utility is given by u_i. If we let $C := \times_i C_i$ and $T := \times_i T_i$, then $u_i : C \times T \to \mathbb{R}$ is a von Neumann-Morgenstern utility function. We say that Γ^b is *finite* iff

N is finite and C_i and T_i are finite for each $i \in N$. (We will mainly be interested in the case where $N = 2$.)

Each player i first learns his own type and then forms beliefs over the types of the other players. His belief over the types of the other players is denoted by $p_i : T_i \longrightarrow \Delta(T_{-i})$ where $\Delta(T_i)$ is the set of probability distributions over T_{-i} and is calculated via Baye's rule so that

$$p_i(t_{-i}|t_i) := \frac{P(t)}{\sum\limits_{s_{-i}\in T_{-i}} P(s_{-i}, t_i)}.$$

Thus, we can also say that beliefs are *consistent* since all the conditional beliefs p_i can be derived from a single probability measure P. Player i's strategy σ_i is a function $\sigma_i : T_i \longrightarrow \Delta(C_i)$. A profile of strategies $\sigma := (\sigma_i)_{i\in N}$ is a *Bayesian Nash equilibrium* if for each player i, given his type $t_i \in T_i$, $\sigma_i(t_i)$ maximises his expected utility given his information and the other players' strategies. In other words,

$$\sum_{t_{-i}\in T_{-i}} p(t_{-i}|t_i)u_i(\sigma_i(t_i), \sigma_{-i}(t_{-i}), t) \geqslant \sum_{t_{-i}\in T_{-i}} p(t_{-i}|t_i)u_i(\widetilde{\sigma}_i(t_i), \sigma_{-i}(t_{-i}), t)$$

for all strategies $\widetilde{\sigma}_i \in \Delta(C_i)^{T_i}$.[1]

Now suppose there is a disinterested mediator[2] with whom the players can communicate. This gives us a Bayesian game with communication wherein the players communicate after they learn their types but before they choose an action. Let each player report his type confidentially to the mediator. The mediator, upon receiving messages from all the players, makes private probabilistic recommendations of actions to take to each player. Specifically, player i sends a message $m_i \in T_i$ to the mediator. Then the mediator makes a recommendation according to $\mu : T \longrightarrow \Delta(C)$. Thus, a strategy for player i is a pair (m_i, δ_i) where m_i is the type report that player i sends to the mediator and $\delta_i : C_i \longrightarrow \Delta(C_i)$ is such that player i, upon receiving the recommendation c_i, takes the action $\delta_i(c_i)$. We will require that $\sum_{\tilde{c}\in C} \mu(\tilde{c}|t) = 1$ and $\mu(c|t) \geqslant 0$ for all $c \in C$ and $t \in T$. Indeed, any such function $\mu : T \longrightarrow \Delta(C)$ will be referred to as a *mediated communication mechanism*. Recall that such a mechanism still satisfies C1 and C2. We shall call a mediated communication mechanism *rational* if $\mu(c|t)$ is rational for each $c \in C$ and for every $t \in T$.

Let us consider the instance where each player reports his type honestly to the mediator and obeys the recommendation of the mediator. Then, player i's expected utility from the mediation mechanism μ is

$$U_i(\mu|t) := \sum_{t_{-i}\in T_{-i}} \sum_{c\in C} p_i(t_{-i}|t_i)\mu(c|t)u_i(c, t).$$

[1] If X and Y are sets, Y^X denotes the set of all mappings from X to Y.

[2] The discussion on mediated mechanisms is based on Chapter 6 of Myerson [16].

That we can restrict attention to truth-telling strategies in the reporting stage is because of the Revelation Principle (see, for instance, [16]). The mediation plan μ therefore induces a communication game Γ^b_μ wherein each player first chooses a reporting strategy m_i (which depends on his type alone) and a choice of action $\delta_i(c_i)$ upon receipt of the recommendation c_i. Thus a strategy is a pair (m_i, δ_i). The type sets are the same as in Γ^b and the utility functions in Γ^b_μ are derived in the obvious way. Now if player i were to use strategy (m_i, δ_i), his expected utility from the mechanism μ is given by

$$U_i(\mu, \delta_i, m_i|t_i) := \sum_{t_{-i}\in T_{-i}} \sum_{c\in C} p_i(t_{-i}|t_i)\mu(c|t_{-i}, m_i)u_i(c_{-i}, \delta_i(c_i), t)$$

if all the other players remain honest and obedient. We will say that a mediated communication mechanism μ is *incentive compatible* iff being honest while reporting their type to the mediator and obedient while following the recommendation of the mediator is a Bayesian Nash equilibrium of the game with communication. Thus, μ is incentive compatible if, for all $i \in N$, $t_i \in T_i$, $m_i \in T_i$ (the report of player i) and $\delta_i \in \Delta(C_i)^{C_i}$ (the choice of action for player i when the recommendation is c_i), it is the case that

$$U_i(\mu|t) \geqslant U_i(\mu, \delta_i, m_i|t_i).$$

The mediator represents, in an indirect way, commitment among the players. Although they cannot write binding contracts to reveal their information and act according to the resulting prescription (possibly because the information is not verifiable and there is no enforcement mechanism), they can nonetheless achieve the same outcomes through a mediator. Thus, the set of incentive compatible mechanisms represents all the outcomes that could potentially be realised as the equilibrium of any other mechanism or protocol.

Recall that an outcome function is a function $\psi : T \to \Delta(C)$. (Then $\psi(t)$ is a probability measure on C and $\psi(t)(c)$ is the probability of the outcome being c.) Consider now any other cheap communication extension of Γ^b represented by Γ^b_c. Here, Γ^b_c is an extensive form game where players communicate after learning their types. At some endogenously determined point in time, the players simultaneously choose their actions from the original game Γ^b. A strategy profile $\sigma = (\sigma_i)_{i\in N}$ in Γ^b_c induces an outcome function $\psi^\sigma : T \to \Delta(C)$. We will say that an outcome function ψ can be *implemented* via a communication mechanism if there exists a cheap communication extension Γ^b_c and a perfect Bayesian equilibrium σ which induces ψ. We will therefore call ψ (following Ben-Porath [4]) a *communication equilibrium outcome* if for all $i \in N$, $t_i, t_i' \in T_i$ and $\delta \in \Delta(C_i)^{C_i}$, it is the case that

$$\sum_{t_{-i}\in T_{-i}} \sum_{c\in C} p(t_{-i}|t_i)\psi(t)(c)u_i(c, t) \geqslant \sum_{t_{-i}\in T_{-i}} \sum_{c\in C} p(t_{-i}|t_i)\psi(t_i', t_{-i})u_i(\delta(c_i), c_{-i}, (t_i, t_{-i})).$$

The right hand side of the equation above represent the two kinds of deviations that any player could have, namely by pretending to be another type t_i' or by choosing an action according to $\delta_i(c_i)$ instead of choosing c_i.

As mentioned above, the set of mediated communication mechanism outcomes encompasses everything that is achievable as the equilibrium outcome of some mechanism. We are interested in implementing these mediated outcomes via communication mechanisms that are unmediated. In other words, our goal is to find a mechanism so that for any mediated communication mechanism outcome, there is a perfect Bayesian equilibrium of the game induced by the mechanism that implements the mediated mechanism. Let us introduce some more notation. For any Bayesian game Γ^b, let $\mathrm{NE}(\Gamma^b)$ be the associated set of Bayesian Nash equilibrium payoffs to the various players and their type and $\mathrm{co}(\mathrm{NE}(\Gamma^b))$ its convex hull. We need one more definition before we continue.

3.1 Definition. *Let Γ^b be any Bayesian game. A mediated communication mechanism μ is **Nash dominating for player** m if there exists a payoff vector $\alpha \in co(NE(\Gamma^b))$ with the payoff to type j of player m given by α_m^j such that $U_m(\mu|t_j) > \alpha_m^j$ for all $t_j \in T_m$. We shall call the equilibrium which generates such an outcome Nash dominated for player m.*

We are now ready to state our Main Theorem.

Main Theorem. *Let Γ^b be any Bayesian game with two players and let ψ be an outcome function that is induced by some mediated communication mechanism μ. Suppose μ is Nash dominating for player m. Then, there exists a communication mechanism satisfying C1 and C3 that is cheap and non-binding that implements ψ with probability 1.*

Remarks. Note that the only criterion we have dropped is that the communication mechanism be non-verifiable. Also note that for the Nash domination condition, we have used the convex hull of the set of Bayesian Nash equilibria of the original game. We will describe below how this can be achieved.

We will need a few more ideas in order to move towards our end. We will use the idea of a *joint lottery*. A joint lottery simulates a public randomising device and works in a straightforward manner for lotteries of the form $(\lambda, 1-\lambda)$ where $\lambda \in \mathbb{Q} \cap [0,1]$. Consider the case where there are two players and suppose they wish to pursue a particular course of action called outcome 1 with (rational) probability λ and some other outcome, 2, with the complementary probability. To achieve this, we assume that $\lambda = \frac{p}{q}$ where $p, q \in \mathbb{Z}_+$ and are coprime. Now let each player have $\mathbb{Z}_q := \mathbb{Z}/q\mathbb{Z} = \{0, \ldots, q-1\}$ as a message space. The players simultaneously send messages z_1 and z_2 respectively to each other from the message space with each message being drawn according to the uniform distribution. Then, we will stipulate that outcome 1 obtains if $z_1 + z_2 \pmod{q} \in \mathbb{Z}_p$. Note that $z_1 + z_2$

$\pmod q$ is uniformly distributed over $\mathbb{Z}_q$. It is now easy to see that outcome 1 obtains with probability $\lambda = \frac{p}{q}$ and it is also the case that neither player has an incentive to unilaterally choose the message according to some other distribution (which constitutes a change in strategy). We will also use the coding tools of envelopes with notes in them and larger envelopes which contain these envelopes. (These correspond respectively to the balls with messages in them and the urns containing the balls in Ben-Porath [3]. Our choice of this alternative terminology stems from æsthetic concerns as we will be using envelopes within envelopes (which we shall refer to as *nested envelopes*) and we find the image of urns within urns somewhat unappealing.)

Now to see how to convexify the set of Bayesian Nash equilibria. For simplicity, consider the case where the players want to generate a convex combination of two equilibria of the original game. In particular, suppose that they want to implement the first equilibrium with probability $\rho \in [0, 1]$ and the second with complementary probability. Achieving this boils down to conducting a lottery with the respective probabilities. But the procedure described above, namely the joint lotteries, only works for rational probabilities. Nevertheless, a multi-step version of the procedure can be adopted to produce any probability.

Let each player have $\{a, b\}$ as a message set. With probability $\frac{1}{2}$, the players simultaneously send the message a to each other. Then, let the players record the outcome as 0 if they send each other the same message and 1 if the messages are different. Repeat this procedure again. After n repetitions, the players will have generated a string of 0's and 1's, which can be written as $0.\,o_1 o_2 \ldots o_n$ (where $o_i \in \{0, 1\}$). We will regard this as the binary expansion of a real number. The players will stop and implement the first outcome if the number they generate is less than ρ and the second outcome if the number if greater than ρ. Also, they will stop in finite time a.s. For example, consider the case where $\rho = 0.0101010101\ldots$ in base 2. Then, there will be a first digit which does not match the expansion. In particular, suppose after four rounds of communication, the players generate 0.0100, they will stop and play the first outcome. They could not have stopped after three rounds as the first three digits of the number they generated were the same as the first three digits in the binary expansion of ρ. But the fourth round told them that the number they were generating was definitely less than ρ.

This gives us a procedure to create arbitrary convex combinations. The players can now choose to play different outcomes with different probabilities. In particular, they can also play any convex combination of the Bayesian Nash equilibria of this game.

4 An Example

We now consider an example that illustrates the workings of our mechanism. Consider the following game Sender-Receiver from §6.7 of Myerson [16]. The Sender has two types, t_1 and t_2 with the prior probability of t_1 being $q \in [0, 1]$. The Receiver has three possible actions, namely x, y and z. Consider the case where $q = \frac{1}{2}$.

		x	y	z
$[q]$	t_1	2,3	0,2	$-1, 0$
$[1 - q]$	t_2	1,0	2,2	0,3

Figure 1: Example

We leave it to the reader to verify that with one round of cheap talk (as in Crawford-Sobel), there is no transmission of information and the only equilibrium is the so-called *babbling equilibrium* (also known as the equilibrium of the *silent game* in the terminology of Aumann and Hart [1]). In this equilibrium, the Sender's expected payoffs (for the respective types) is (0,2), while the Receiver gets an expected payoff of 2.

It turns out that any finite number of stages also do not increase the set of equilibrium outcomes. (For the initiated reader, this is because the graph of equilibrium payoffs of the silent game as a multi-function over the set of probability distributions over types, which in this case is $[0, 1]$, is *bi-convex*.) But there are mediated solutions that provide Pareto superior payoffs (in expectation) for both the Sender and the Receiver.

Mediated Communication Mechanism. Consider

$$\mu(x|t_1) = \frac{2}{3}, \quad \mu(y|t_1) = \frac{1}{3}, \quad \mu(z|t_1) = 0,$$

$$\mu(x|t_2) = 0, \quad \mu(y|t_2) = \frac{2}{3}, \quad \mu(z|t_2) = \frac{1}{3}.$$

which maximises the Receiver's payoffs. This gives expected payoffs

$$U_S(\mu|t_1) = 1\frac{1}{3}, \quad U_S(\mu|t_2) = 1\frac{1}{3}, \quad U_R(\mu) = 2\frac{1}{2}.$$

To see truthful reporting by the Sender, consider when t_1 lies. If he reports t_2 his payoff is

$$\frac{2}{3} \times 0 + \frac{1}{3} \times (-1) < 1\frac{1}{3}.$$

If the Sender is of type t_2, then by reporting t_1 he gets

$$\frac{2}{3} \times 1 + \frac{1}{3} \times 2 = 1\frac{1}{3} = U_s(\mu|t_2)$$

which is the same payoff as reporting truthfully. Thus, neither type has an incentive to report untruthfully. Now consider the Receiver's posterior probabilities upon receiving the various recommendations. If the Receiver gets the recommendation x, he knows that the true state of the world is t_1 as $\mu(x|t_2) = 0$. Conditional on his new posterior, his best action is the recommended action x. Similarly, for the case when he receives the recommendation z, his posterior probability assigns probability 1 to state t_2 which makes z the optimal action for him. If he receives y as a recommendation, he assigns probability $\frac{2}{3}$ to the state being t_2. But with these posteriors, $U_R(y|y) = 2 = U_R(z|y) > U_R(x|y)$, making the recommended y an optimal choice. We will now consider a cheap, unmediated communication mechanism which will implement the above mediated plan.

The Mechanism

Step 1 We shall assume that the players have access to sealable envelopes and pieces of paper on which they can write messages. Let the Receiver write messages and place each message in an envelope according to the mediated solution. This means that there are six envelopes, two of which contain the message x, three of which contain the message y and the last one contains the message z. The Receiver then places the envelopes in larger envelopes according to the type of the Sender. In particular, he takes two large envelopes and labels them t_1 and t_2. In the envelope labelled t_1, he places the two envelopes containing x and one containing y. In the envelope labelled t_2, he places all the other envelopes containing messages. Note that this mimics the distribution which is the mediated solution.

Step 2 We now come to the random monitoring stage. Our protocol requires the players to verify the contents of all the envelopes with some probability, ρ (which can be taken to be rational). This is accomplished by a joint lottery as described in §3. If the outcome of the lottery requires the verification of the contents of all the envelopes, then this is performed. If there is a deviation from the equilibrium by the Receiver, then the Sender decides to babble (which results in a payoff of 2 for the Receiver). If there is no deviation, then the Receiver repeats Step 1. If however, the outcome of the joint lottery was that there be no verification, the players move to Step 3.

Step 3 The Receiver then hands the two labelled envelopes to the Sender. The Sender privately takes a message out of the envelope corresponding to his type and hands it to the Receiver. The Receiver opens the envelope and takes the prescribed action. All the other envelopes remain unopened and are destroyed publicly.

We have not yet specified the random monitoring probability, ρ. We shall do so now. Let α_R denote the Receiver's payoffs from the babbling equilibrium and let W denote the maximal payoff available to him in the game. Also, let $U_R(\mu)$ be his expected payoff from following the mechanism. (Note that in this case, $\alpha_R = 2$, $W = 3$ and $U_R(\mu) = 2\frac{1}{2}$. Thus, $U_R(\mu) > \alpha_R$ which means that μ is Nash dominating for the Receiver.) If the Receiver deviates, his payoff is bounded above by

$$\rho\alpha_R + (1-\rho)W.$$

As mentioned above, his expected payoff from following the mechanism is $U_R(\mu)$. He will not deviate if

$$U_R(\mu) > \rho\alpha_R + (1-\rho)W. \tag{1}$$

But $U_R(\mu) > \alpha_R$, so there exists a $\overline{\rho}$ so that (1) is satisfied. We can then take any $\rho > \overline{\rho}$ such that $\rho \in \mathbb{Q} \cap [0,1]$. It is easy to see that the mechanism ends in finite time with probability 1.[3] It is also easy to see that the equilibrium constructed is perfect Bayesian. In §5 below, we show that our mechanism for the Sender-Receiver game is a special case of a general scheme.

5 Proof of Main Theorem

We shall now provide a generalisation of the protocol described in §4. Let us assume that the players wish to implement a mediated communication mechanism μ, as described in §3.

The Mechanism

Step 1 Suppose player 1 is of type $t_i \in T_1$ and player 2 is of type $\tau_j \in T_2$. Player 2 makes envelopes for this state ($t = (t_i, \tau_j)$) according to $\mu(t)$. In particular, suppose $\mu(c|t) = \frac{1}{2} = \mu(c'|t)$ (where $c = (c_1, c_2)$). Then player 2 makes two envelopes for state t. In the first envelope, say the envelope corresponding to the action c, there will be envelopes marked *Player 1* and *Player 2*. The envelope marked *Player 1* will have the action c_1 inside on a piece of paper and similarly for the other envelope. To recap, an envelope for state t has inside it more envelopes containing envelopes with messages. These are the *nested envelopes* referred to in §3.

Now player 2 makes envelopes for all states $t \in T$. But the labels on the envelope for a particular state are as follows: state $t = (t_i, \tau_j)$ will be marked t_i, m where $m \in |T_2|$ and there is a bijection $\xi : T_2 \to |T_2|$ such that $m = \xi(\tau_j)$. Note that the bijection $\xi \in \Xi$, the space of all bijections from T_2 to $|T_2|$, is chosen by player 2 at random according to the uniform distribution. Then the players proceed to Step 2.

[3] The probability that it takes at least n iterations being ρ^n which becomes small very quickly as n becomes larger. Also, the protocol runs for an expected time of $\frac{\rho}{(1-\rho)^2}$.

Step 2 This is the random monitoring stage. The players conduct a joint lottery where with a probability ρ, the contents of all the envelopes are checked. If there is a deviation by player 2, i.e. if he has not placed the messages in the envelopes and the nested envelopes according to the mediated solution μ, then they play the Nash dominated equilibrium for player 2. If there is no deviation, they go back to Step 1. If the joint lottery does not require them to open all the envelopes, the players proceed to Step 3.

Step 3 Suppose that player 2's true type is τ_1. He then removes all the envelopes corresponding to states $t = (t_i, \tau_j)$ where $\tau_j \neq \tau_1$ and hands the remaining envelopes to player 1. In other words, he tells player 1 that he is of type $m = \xi(\tau_1)$, but since player 1 does not know ξ and all the ξ's are equally likely, his beliefs about player 2's type remain unchanged. Player 1 now picks the envelope corresponding to his type and picks one of the envelopes in it. This envelope contains the envelopes marked *Player 1* and *Player 2*. He hands player 2 his envelope and they play the recommended action.

Remark. The reason for player 2 making envelopes corresponding to all states $t \in T$ is that if he just made it for all $t = (t_i, \tau_j)$ where τ_j is his true type and the players are required to inspect the contents of the envelopes in the monitoring stage, then player 1 can, in principle, infer the type of player 2.

All that remains to be done is to determine a monitoring probability, ρ. This will be done as in §4. Recall that α_2^j is the payoff to type τ_j of player 2. Now consider the case where player 2 of type τ_j deviates. Then, assuming that the maximum payoff in the game is W, his payoffs are bounded from above by

$$\rho\alpha_2^j + (1-\rho)W.$$

But his expected payoff from following the protocol is $U_2(\mu|\tau_j)$ and we require that

$$U_2(\mu|\tau_j) > \rho\alpha_2^j + (1-\rho)W. \tag{2}$$

But $U_2(\mu|\tau_j) > \alpha_2^j$ by assumption, therefore there exists a $\overline{\rho}$ such that (2) is satisfied and we can find a $\rho \geqslant \overline{\rho}$ such that $\rho \in \mathbb{Q} \cap [0,1]$. We now consider the issue of perfection. At each zero-probability node at which the players takes an action, their beliefs are exactly their priors. More precisely, the players' beliefs about each other's types are always updated using Baye's rule, and since deviations do not provide any information, the priors remain the same. Given an envelope with a recommendation, player 2 doesn't update his priors about the type of player 1 before he takes an action. Any updating to be performed is done by Baye's rule and is possible for any allocation of messages in envelopes in the various bigger envelopes by player 2. Thus, the equilibrium constructed is perfect Bayesian.

6 Discussion

In this paper, we have discussed the notion of a mediated solution and demonstrated that any mediated solution which satisfies a Nash domination condition can be implemented via a cheap communication mechanism. In this section, we will discuss some other ideas related to that of the mediator before we conclude.

In §§4 and 5, we adapted a mechanism due to Ben-Porath [3]. Our main technique is to use a *menu* of nested envelopes to simulate the role of the mediator. The key to a mediated mechanism is the observation that even upon receipt of the recommendation from the mediator, the players are unsure of the type of their opponent. This is so even after the players update their beliefs using Baye's rule. Thus, the essential role of the mediator is to move the players' posterior beliefs about each other's types around in an incentive compatible way. This can also be achieved by a noisy communication medium (see, for instance, chapter 6 of [16]). In other words, we can also think of the mediator as someone who introduces noise in the communication. Let us consider this more abstractly.

Let M_1 and M_2 represent the set of messages send to some machine (which could also be a disinterested third party, the mediator) and let X_1 and X_2 represent the set of messages that the players receive. Now define, $M := M_1 \times M_2$ and $X := X_1 \times X_2$. Let $(X, \mathscr{X})$ be a measurable space and with a slight abuse of notation, let $\Delta(X)$ denote the space of probability measures on this measurable space. This is the abstract representation of any single stage communication mechanism where the players send messages and receive randomised messages after which they take an action. (It is straightforward to extend the definitions to multi-stage communication mechanisms, but we shall not do so here in the interests of notational simplicity.) More precisely, we can state the following.

6.1 Definition. *A **communication mechanism** is a tuple $(M, (X, \mathscr{X}), \nu)$ where $\nu : M \to \Delta(X)$.*

Thus, the signals that the players receive is a random variable drawn according to the probability measure $\nu(m)$ over $(X, \mathscr{X})$. If we assume that X is a topological space and $\mathscr{X}$ is the corresponding Borel σ-algebra, then the support of a measure P is a closed set $C =: \text{supp}(P)$ such that (i) $P(C) = 1$ and (ii) $P(C') < 1$ for all closed sets $C' \subset C$. The support exists if X is a sufficiently nice topological space, an assumption we shall make. This leads us to the following condition and definition.

Condition T. *For all $m \in M$, supp$(\nu(m))$ is a singleton.*

6.2 Definition. *A communication mechanism is **noiseless**[4] iff it satisfies Condition T. A com-*

[4]Lehrer and Sorin [13] call a noiseless mechanism *deterministic*. We use this alternative terminology to emphasise the distinction from noisy mechanisms.

munication mechanism that is not noiseless will be called ***noisy****.*

Remark. It is pertinent to point out Lehrer and Sorin's result at this stage. They show that by suitably expanding M_1 and M_2, any rational mediated solution can be achieved through a noiseless (deterministic) albeit mediated communication mechanism.

It is instructive to consider the role of the mediator here in the simple case of Sender-Receiver games, as in §4. Consider a mediated mechanism $\mu(\cdot|t)$. For each type t, the Sender is promised that the outcome chosen will be drawn according to the probability distribution $\mu(\cdot|t)$. The Sender truthfully reveals his type t to the mediator if he believes that the outcome will, in fact, be generated by the said probability distribution. The mediator then make a recommendation to the Receiver by randomising over the elements in $\text{supp}(\mu(\cdot|t))$ according to the probability distribution $\mu(\cdot|t)$ (with the recommendation being the outcome of the randomisation). Of course, the Sender has to truthfully reveal his type and the Receiver has to comply and take the recommended action, but this is what the incentive compatibility constraints guarantee. Thus, a mediated mechanism is an example of a noisy communication mechanism.

Another example of a noisy communication mechanism is given by the game in §4 above. Recall that there is no transmission of information even with unbounded cheap talk. However, Myerson shows that if the Sender has access to pigeons that get lost with some probability, then there can be some transmission of information. This is possible because when a pigeon doesn't arrive, the Receiver is not sure if it was because the Sender did not send the pigeon or because the pigeon got lost. Note that the message that the Receiver receives is the arrival of the pigeon and not the message the pigeon may bear. Myerson then shows that with more general noisy mechanisms (i.e. mediated mechanisms), even more equilibrium outcomes can be realised as communication equilibria. But a common feature of all these noisy mechanism is that they provide some degree of protection for the players. Specifically, in the game in §4, with just signalling, type t_1 of the Sender always wants to pretend to be type t_2. However, the introduction of a noisy communication channel provides him with more protection from exposure.

As mentioned before, our mechanism is a descendant of Ben-Porath's [3]. We modify his mechanism by allowing for a menu of nested envelopes from which one of the players chooses. As in [3], the envelopes with messages provide enough uncertainty to simulate the noise that a mediator brings. The main connection between our mechanism and the one used by Ben-Porath in [4] is the use of random monitoring. Otherwise, his mechanism ends in finite time and has no possibility of verification (as in requirement C2 above). We believe that with our Theorem 0 and Main Theorem and Ben-Porath's [4] result, we have a complete taxonomy of everything that can be achieved through unmediated talk. We also hope that our and the aforementioned results will persuade game theoretic modelers to pay more attention to the reasons for their particular choice of mechanism in their applications.

References

[1] Aumann, Robert J. and Sergiu Hart (2003). "Long Cheap Talk," *Econometrica, 71(6)*, 1619–1660.

[2] Barany, Imré (1992). "Fair Distribution Protocols or How the Players Replace Fortune," *Mathematics of Operations Research, (17),* 329–340.

[3] Ben-Porath, Elchanan (1998). "Communication without Mediation: Expanding the Set of Equilibrium Outcomes by "Cheap" Pre-play Procedures," *Journal of Economic Theory, 80,* 108–122.

[4] Ben-Porath, Elchanan (2003). "Cheap Talk in Games with Incomplete Information," *Journal of Economic Theory, 108,* 45–71.

[5] Bester, Helmut and Roland Strausz (2001). "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica, 69(4),* 1077–1098.

[6] Crawford, Vincent and Joel Sobel (1982). "Strategic Information Transmission," *Econometrica, 50,* 579–594.

[7] Forges, Françoise (1990). "Equilibria with Communication in a Job Market Example," *Quarterly Journal of Economics, 105*, 375–398.

[8] Forges, Françoise (1990). "Universal Mechanisms," *Econometrica, (58),* 1341–1364.

[9] Forges, Françoise (1994). "Non-Zero Sum Repeated Games and Information Transmission," in *Essays in Game Theory in Honor of Michael Maschler*, 65–95, ed. Nimrod Megiddo, Springer-Verlag, New York.

[10] Gerardi, Dino (2000). "Interim Pre-play Communication," mimeo, Northwestern University.

[11] Grossman, Gene M. and Elhanan Helpman (2001). *Special Interest Politics,* MIT Press, Cambridge, Massachusetts.

[12] Krishna, R. Vijay (2004). "Extended Conversations in Sender-Receiver Games," working paper, Pennsylvania State University.

[13] Lehrer, Ehud and Sylvain Sorin (1997). "One-Shot Public Mediated Talk," *Games and Economic Behavior, 20*, 131–149.

[14] Matthews, Steven and Andrew Postlewaite (1989). “Preplay Communication in Two-Person Sealed-bid Double Auctions,” *Journal of Economic Theory, (48),* 239–263.

[15] Myerson, Roger B. (1986). “Multistage Games with Communication,” *Econometrica, 54*, 323–358.

[16] Myerson, Roger B. (1991). *Game Theory: Analysis of Conflict,* Harvard University Press, Cambridge, Massachusetts and London, England.

97–110.

[17] Simon, Robert Samuel (2002). “Separation of Joint Plan Equilibrium Payoffs from the Min-max Functions,” *Games and Economic Behavior, 41*, 79–102.

[18] Urbano, Amparo and Jose E. Vila (2002). “Computational Complexity and Communication: Coordination in Two-Player Games,” *Econometrica, 70(5)*, 1893–1927.