

Comparison of Information Structures

Olivier Gossner*

May 24, 1996

We introduce two ways of comparing information structures, say \mathcal{I} and \mathcal{J} . First we say that \mathcal{I} is richer than \mathcal{J} when for every compact game G , all correlated equilibrium distributions of G induced by \mathcal{J} are also induced by \mathcal{I} . Second, we say that \mathcal{J} is faithfully reproducible from \mathcal{I} when all the players can compute from their information in \mathcal{I} “new information” that they could have received from \mathcal{J} . We prove that \mathcal{I} is richer than \mathcal{J} if and only if \mathcal{J} is faithfully reproducible from \mathcal{I} .

*Economics Department, Universitat Pompeu Fabra, Balmes 132, 08008 Barcelona, Spain. E-mail: gossner@upf.es. I wish to thank Sylvain Sorin and Nicolas Vieille for many discussions. I am also grateful to Jean-François Mertens for very enriching comments. Part of this work was done while I was visiting the Center for Interactive Decision Theory at Jerusalem.

1 Introduction

Information structures were defined by Aumann ([Au74]) to introduce the concepts of correlated equilibrium and of correlated equilibrium distributions. Given a game G and an information structure \mathcal{I} , the game G extended by \mathcal{I} is the game in which players first receive information according to \mathcal{I} , and second play in G . The distributions on the actions of G induced by Nash equilibria of this game are called the correlated equilibrium distributions of G induced by \mathcal{I} .

For instance, it is well known that before playing the game “battle of sexes”:

	L	R
T	2,1	0,0
B	0,0	1,2

players may wish to toss a coin, then play (T, L) if heads shows, and (B, R) otherwise. These strategies are a Nash equilibrium of the game extended by the coin information structure, and the induced distribution on the actions of the game is $\frac{1}{2}(T, L) + \frac{1}{2}(B, R)$. One could wonder: “What if the players do not have a coin at disposition?”. In fact, there are many other information structures that can replace the coin, like a (fair, 6-sided) die. Consider players that observe the outcome of a die, and play (T, L) if the observed outcome is even, and (B, R) otherwise. This defines a Nash equilibrium of the game extended by the dice information structure inducing the same distribution as before. The fact that a coin can be replaced by a die does not depend on the particular (battle of sexes) game. In fact, for any game G , all correlated equilibrium distributions of G induced by the coin are also correlated equilibrium distributions of G induced by the die. We define an information structure \mathcal{I} to be richer than another \mathcal{J} whenever for any game G , all correlated equilibrium distributions of G extended by \mathcal{J} are also correlated equilibrium distributions of G extended by \mathcal{I} .

It is striking that in correlated equilibria, players condition their actions on their information about the state of nature, even though it is common knowledge that the payoffs of the game do not depend on the state of nature. This is

actually not a contradiction since, for a given player, different signals imply different beliefs on the actions of the other players, which justifies different choices of actions¹. After getting his signal, and before playing the game, a player may forget part of his information if it reveals nothing about the actions the others are about to take. For instance, a player observing the outcome of the die and who believes that the others will act only according to the parity of the outcome may keep only the information “even”, or “odd”. Renaming “even” and “odd” as “heads” and “tails”, we see that if they forget part of their information about the outcome of the die, players get “new” signals as if issued by a coin.

More generally, if players get signals issued by \mathcal{I} , they can compute “new” (or interpreted) signals that could have been sent by \mathcal{J} . An interpretation ϕ from \mathcal{I} to \mathcal{J} describes how players compute these new signals. We call an interpretation ϕ a compatible interpretation when the probability of the interpreted signals induced by the distribution of signals in \mathcal{I} and by ϕ is equal to the probability of the signals issued by \mathcal{J} . Moreover, a compatible interpretation is said to be a faithful interpretation when every player has the same conditional probability over the interpreted signal of the other players, conditional on his original signal (given by \mathcal{I}) or on his interpreted one. In other words, ϕ is faithful if no player loses information by computing his interpreted signal and forgetting his original one.

Our main result is that \mathcal{I} is richer than \mathcal{J} if and only if there exists a faithful interpretation from \mathcal{I} to \mathcal{J} .

Section 2 contains preliminaries, in Section 3 we present the different ways of comparing information structures. Section 4 contains the proof of the main result. In general, we can not assume the interpretation to be deterministic. Players may randomize to compute their new signals. In section 5 some examples are presented, as well as conditions under which the interpretations can be assumed to be deterministic.

¹See Aumann’s [Au87] article for the close relation between correlated equilibria and Bayesian rationality

2 Preliminaries

2.1 General notations

Throughout the paper, I is the finite set of players. For every collection of sets indexed by I (Z^i) $_{i \in I}$, Z represents $\prod_i Z^i$, and Z^{-i} is $\prod_{j \neq i} Z^j$. Also, for $z \in Z$, z^i is the i -th coordinate of z , and z^{-i} is $(z^j)_{j \neq i} \in Z^{-i}$. For a topological set W , $\Delta(W)$ is the set of regular probability measures over the Borel σ -algebra on W . If P is a probability measure, \mathbf{E}_P represents the expectation operator over P .

2.2 Extended games

A **compact game** $G = ((S^i)_i, g)$ is given by a compact set of strategies S^i for each player i and by a continuous payoff function g from S to \mathbb{R}^I . The set of mixed strategies for player i is $\Sigma^i = \Delta(S^i)$, and g is extended to Σ by $g(\sigma) = \mathbf{E}_\sigma g(s)$ (the product set $\prod_i \Delta(S^i)$ is identified to be a subset of $\Delta(S)$). A **Nash equilibrium** of G is a profile of mixed strategies $\sigma \in \Sigma$ such that for every player i and $\tau^i \in \Sigma^i$, $g^i(\sigma^{-i}, \tau^i) \leq g^i(\sigma)$.

In [Me94] for instance, a correlation device is given by a probability space of states of nature by and a signaling function for each player that characterizes his information about the state. Equivalently, if the signaling functions have finite range, we represent information structures as probabilities over product sets of signals. An **information structure** $\mathcal{I} = ((X^i)_i, \mu)$ is given by a family of finite sets X^i and by a probability measure μ over X . When x is drawn according to μ , i is informed about x^i . Without loss of generality, we assume that all the signals of player i occur with a positive probability (for all $x^i \in X^i$, $\mu(x^i) > 0$).

Definition 2.1 For every compact game G and information structure \mathcal{I} , $\Gamma(\mathcal{I}, G)$ represents the game G **extended by** \mathcal{I} defined as follows:

- $x \in X$ is drawn according to μ , each player i is informed about x^i
- each player i chooses $\sigma^i \in \Sigma^i$
- the payoffs are given by $g(\sigma)$

A strategy for player i is a mapping f^i from X^i to Σ^i , and the payoff function

of $\Gamma(\mathcal{I}, G)$ is given by $g_{\mathcal{I}}(f) = \mathbf{E}_{\mu}g(f(x))$

Since X is finite, $\Gamma(\mathcal{I}, G)$ is also a compact game when Σ^{iX^i} is endowed with the product topology. $D(\mathcal{I}, G)$ represents the set of correlated equilibrium distributions of G induced by \mathcal{I} . Namely, it is the set of image distributions on S of μ by Nash equilibria of $\Gamma(\mathcal{I}, G)$.

For $x^i \in X^i$, let $p(x^i) \in \Delta(X^{-i})$ denote the conditional probability of μ given x^i over X^{-i} ($p(x^i)(x^{-i}) = \frac{\mu(x^{-i}, x^i)}{\mu(x^i)}$).

Remark 2.1 *We shall use the following characterization of Nash equilibria of $\Gamma(\mathcal{I}, G)$:*

f is a Nash equilibrium of $\Gamma(\mathcal{I}, G)$ if and only if for every player i and $x^i \in X^i$:

$$f^i(x^i) \in \arg \max_{\tau^i \in \Sigma^i} \mathbf{E}_{p(x^i)} g^i(\tau^i, f^{-i}(x^{-i}))$$

Proof: This is simply a consequence of the relation:

$$g_{\mathcal{I}}^i(f) = \sum_{x^i \in X^i} \mu(x^i) \mathbf{E}_{p(x^i)} g^i(f^i(x^i), f^{-i}(x^{-i}))$$

This expresses the fact that a rational and bayesian player maximizes his expected payoff conditional to his information.

3 Comparison of information structures

In this section, we introduce two ways of comparing information structures $\mathcal{I} = ((X^i)_i, \mu)$ and $\mathcal{J} = ((Y^i)_i, \nu)$.

The first definition says that \mathcal{I} is richer than \mathcal{J} whenever \mathcal{I} induces all the correlated equilibrium distributions that are induced by \mathcal{J} .

Definition 3.1 *\mathcal{I} is richer than \mathcal{J} whenever for every compact game G , $D(\mathcal{I}, G) \supset D(\mathcal{J}, G)$.*

For the second definition, we imagine that players receive signals from \mathcal{I} , and define conditions under which they can reproduce signals that could have been issued by \mathcal{J} . An **interpretation mapping for player i** from \mathcal{I} to \mathcal{J} is an application ϕ^i from X^i to $\Delta(Y^i)$. When x^i is i 's signal in \mathcal{I} , the interpreted

signal in \mathcal{J} is y^i with probability $\phi^i(x^i)(y^i)$. An **interpretation** from \mathcal{I} to \mathcal{J} is a family $\phi = (\phi^i)_{i \in I}$ of interpretation mappings for all the players. ϕ and $\phi^{-i} = (\phi^j)_{j \neq i}$ define mappings from X to $\Delta(Y)$ and from X^{-i} to $\Delta(Y^{-i})$ when $\Pi_i \Delta(Y^i)$ and $\Pi_{j \neq i} \Delta(Y^j)$ are identified to subsets of $\Delta(Y)$ and $\Delta(Y^{-i})$.

Definition 3.2 *A consistent interpretation from \mathcal{I} to \mathcal{J} is an interpretation $\phi = (\phi^i)_i$ such that the image of μ by ϕ is ν , namely such that for every $y \in Y$, $\mathbf{E}_\mu \phi(x)(y) = \nu(y)$. \mathcal{J} is **reproducible** from \mathcal{I} when there exists a consistent interpretation from \mathcal{I} to \mathcal{J} .*

For the remainder of the section ϕ represents a compatible interpretation from \mathcal{I} to \mathcal{J} . Let P be the probability induced on $X \times Y$ by μ and the transition law ϕ . Explicitly $P(x, y) = \mu(x)\phi(x)(y)$. The marginals of P on X and Y are μ and ν , respectively.

We shall say that ϕ is faithful whenever no player loses information about the interpreted signal of the others by relying on his interpreted signal and forgetting his original one. We shall express it in different ways that will turn out to be equivalent.

We can compare the conditional probabilities $q(y^i)(y^{-i}) = P(y^{-i}/y^i)$ and $r(x^i)(y^{-i}) = P(y^{-i}/x^i)$ for $\phi(x^i)(y^i) > 0$. These can be seen as “ex ante” and “ex post” conditional probabilities over the interpreted signals of players others than i when i 's signal is x^i and his interpreted signal is y^i . We shall view $q(y^i)$ and $r(x^i)$ as random variables with values in $\Delta(Y^{-i})$. Note that $r(x^i) = \mathbf{E}_{p(x^i)} \phi^{-i}(x^{-i})$.

Definition 3.3 *An interpretation ϕ from \mathcal{I} to \mathcal{J} is **faithful** whenever it is compatible and for every i and $x^i \in X^i$, $\mathbf{E}_{\phi^i(x^i)} q(y^i) = r(x^i)$. When there exists a faithful interpretation from \mathcal{I} to \mathcal{J} , we say that \mathcal{J} is **faithfully reproducible** from \mathcal{I} .*

Another way of comparing information is to use the notion of **comparison of experiments** due to Blackwell (see [Bl51] [Bl53]). Recall that an **experiment** is a collection $\alpha = (u_1, \dots, u_n)$ of probability measures over some (finite) space \tilde{X} . A point $\tilde{x} \in \tilde{X}$ is selected according to one of the distributions

(u_1, \dots, u_n) , and is observed by the statistician. Given two experiments α and $\beta = (v_1, \dots, v_n)$ with $v_k \in \Delta(\tilde{Y})$ for $k \in \{1, \dots, n\}$, we say that α is **sufficient** for β whenever there exists a **stochastic transformation** from α to β , that is a family $(Q(\tilde{x}))_{\tilde{x} \in \tilde{X}}$ of probability measures over \tilde{Y} such that for all $1 \leq k \leq n$ and $\tilde{y} \in \tilde{Y}$, $v_k(\tilde{y}) = \mathbf{E}_{u_k} Q(\tilde{x})(\tilde{y})$. When α is sufficient for β and β is sufficient for α , α and β are called **equivalent**.

In our case, the statistician is player i , and two experiments are given by the families $\alpha^i = (u_{y^{-i}})_{\{y^{-i}, \nu(y^{-i}) > 0\}}$ and $\beta^i = (v_{y^{-i}})_{\{y^{-i}, \nu(y^{-i}) > 0\}}$ of probabilities over X^i and Y^i defined by $u_{y^{-i}}(x^i) = P(x^i/y^{-i})$ and $v_{y^{-i}}(y^i) = P(y^i/y^{-i})$. Since $\mathbf{E}_{m_{y^{-i}}} \phi^i(x^i)(y^i) = \sum_{x^i} P(x^i/y^{-i})P(y^i/x^i) = P(y^i/y^{-i})$, ϕ^i defines a stochastic transformation from α^i to β^i . Therefore α^i is sufficient for β^i .

Theorem 3.1 *For all i , the following statements are equivalent:*

- (i) $\forall x^i \in X^i$, $\mathbf{E}_{\phi^i(x^i)} q(y^i) = r(x^i)$
- (ii) $\forall x^i \in X^i$, $\phi^i(x^i)(y^i) > 0 \Rightarrow q(y^i) = r(x^i)$
- (iii) $\forall x_0^i, x_1^i \in X^i$, $\phi^i(x_0^i)(y^i)\phi^i(x_1^i)(y^i) > 0 \Rightarrow r(x_1^i) = r(x_0^i)$
- (iv) α^i and β^i are equivalent
- (v) *The distributions of $r(x^i)$ and $q(y^i)$ on $\Delta(Y^{-i})$ are equal. Namely $\forall q_0 \in \Delta(Y^{-i})$, $\mu(\{r(x^i) = q_0\}) = \nu(\{q(y^i) = q_0\})$.*

In particular, ϕ is faithful if and only if any of these conditions is true for all player i .

Proof of the theorem:

We start the proof with a lemma :

Lemma 3.1 *Writing $\psi^i(y^i)(x^i) = P(x^i/y^i)$, we have $q(y^i) = \mathbf{E}_{\psi^i(y^i)} r(x^i)$ for all $y^i \in Y^i$.*

Proof of the lemma: $q(y^i)(y^{-i}) = \sum_{x^i} P(x^i/y^i)P(y^{-i}/x^i) = \mathbf{E}_{\psi^i(y^i)} r(x^i)$. ■

(ii) \Rightarrow (i) is straightforward.

(iii) \Rightarrow (ii) See that for $y^i \in Y^i$, $\psi(y^i)(x^i) > 0$ if and only if $\phi^i(x^i)(y^i) > 0$.

If for all these x^i , $r(x^i)$ takes the same value, then $q(y^i)$ also takes this value.

(i) \Rightarrow (iii) Consider a connected component C of the graph \mathcal{G} on $X^i \cup Y^i$ where x^i and y^i are connected whenever $\phi(x^i)(y^i) > 0$. Now let q_0 be an

extremal point of $\text{co}(\{r(x^i), x^i \in C \cap X^i\} \cup \{q(y^i), y^i \in C \cap Y^i\})$. Since $r(x^i) = \mathbf{E}_{\phi(x^i)} q(y^i)$, any $x^i \in C \cap X^i$ such that $r(x^i) = q_0$ is only connected to points y^i such that $r(y^i) = q_0$. Conversely, as $q(y^i) = \mathbf{E}_{\psi(y^i)} r(x^i)$, for $y^i \in C \cap Y^i$, y^i connected to x^i and $q(y^i) = q_0$ imply that $r(x^i) = q_0$. Therefore, if $x_0^i, x_1^i \in X^i$ are elements of the same connected component of \mathcal{G} , $r(x_0^i) = r(x_1^i)$, which implies (iii).

(iv) is equivalent to (v) In the case where the marginal of P on $\{y^{-i} \in Y^{-i}, P(y^{-i}) > 0\}$ is uniform, the distributions of $r(x^i)$ and of $q(y^i)$ are the standard measures associated with α^i and β^i . Theorem 4 of [Bl51] shows that the standard measures associated to α^i and β^i are equal if and only if α^i and β^i are equivalent. Blackwell's result easily extends to the case where the marginal of P on $\{y^{-i} \in Y^{-i}, P(y^{-i}) > 0\}$ is not uniform.

(v) \Rightarrow (ii) Assume that for all $q_0 \in \Delta(Y^{-i})$, $\mu(\{r(x^i) = q_0\}) = \nu(\{q(y^i) = q_0\}) = \nu(\{\mathbf{E}_{\psi(y^i)} r(x^i) = q_0\})$. Let q_0 be extremal in $\text{co}\{r(x^i), x^i \in X^i\}$. We get $\sum_{r(x^i)=q_0} \mu(x^i) \leq \sum_{r(x^i)=q_0} \mu(x^i) \phi^i(\{q(y^i) = q_0\})$. This can hold only if $r(x^i) = q_0$ implies $\phi^i(\{q(y^i) = q_0\}) = 1$. Therefore (ii) holds if $r(x^i) = q_0$. We get an equivalent system of equations if $r(x^i) \neq q_0$, and the proof is completed by induction on the cardinality of X^i .

(ii) \Rightarrow (v) For $q_0 \in \Delta(Y^{-i})$, $\nu(\{q(y^i) = q_0\}) = \mathbf{E}_{\mu} \phi^i(x^i)(\{q(y^i) = q_0\}) = \mu(\{r(x^i) = q_0\})$. ■

We also need to define an ε -**faithful** interpretation. First we define a metric on $\Delta(Z)$ for Z finite by $d(\rho_1, \rho_2) = \max_{z \in Z} |\rho_1(z) - \rho_2(z)|$.

Definition 3.4 For $\varepsilon > 0$, a consistent interpretation ϕ is a ε -faithful interpretation when for all i , $x^i \in X^i$ and $y^i \in Y^i$, $\phi^i(x^i)(y^i) > 0$ implies $d(r(x^i), q(y^i)) \leq \varepsilon$.

Proposition 3.1 \mathcal{J} is faithfully reproducible from \mathcal{I} if and only if there exists an ε -faithful interpretation for all $\varepsilon > 0$.

Proof: The direct proof is obvious since a faithful interpretation is also a ε -faithful interpretation. For all $\varepsilon > 0$, the set of ε -faithful interpretations from

\mathcal{I} to \mathcal{J} is compact in the set of interpretations from \mathcal{I} to \mathcal{J} endowed with the metric $D(\phi_1, \phi_2) = \max_{i, x^i} d(\phi_1^i(x^i), \phi_2^i(x^i))$. If these sets are non-empty, their intersection is also non-empty. ■

4 Equivalence theorem

Theorem 4.1 (Main Theorem) *\mathcal{I} is richer than \mathcal{J} if and only if \mathcal{J} is faithfully reproducible from \mathcal{I} .*

The remaining of this section is devoted to a proof of the theorem and to a corollary.

4.1 Construction of strategies in $\Gamma(\mathcal{J}, G)$ from strategies in $\Gamma(\mathcal{I}, G)$ and from ϕ

Assume that ϕ is an interpretation from \mathcal{I} to \mathcal{J} . Let G be a compact game, and f a I -tuple of strategies in $\Gamma(\mathcal{J}, G)$. The I -tuple of strategies e in $\Gamma(\mathcal{I}, G)$ is defined by $e^i(x^i)(B^i) = \sum_{y^i} \phi^i(x^i)(y^i) f^i(y^i)(B^i)$ for B^i Borel subset of S^i .

Lemma 4.1 *If ϕ is a compatible interpretation, e and f induce the same distribution on S .*

Proof: Let $B = B^1 \times \dots \times B^I$ be a product of Borel subsets of S^1, \dots, S^I . We get $\mathbf{E}_\mu e(x)(B) = \mathbf{E}_\mu \Pi_i e^i(x^i)(B^i) = \mathbf{E}_\mu \Pi_i \sum_{y^i} \phi^i(x^i)(y^i) f^i(y^i)(B^i) = \mathbf{E}_\nu \Pi_i f^i(y^i)(B^i) = \mathbf{E}_\nu f(y)(B)$. ■

Lemma 4.2 *If ϕ is faithful and f is a Nash equilibrium of $\Gamma(\mathcal{J}, G)$, e is a Nash equilibrium of $\Gamma(\mathcal{I}, G)$*

Proof: We use the remark 2.1. For $x^i \in X^i$ and $\sigma^i \in \Sigma^i$,

$$\begin{aligned} \mathbf{E}_{p(x^i)} g^i(\sigma^i, e^{-i}(x^{-i})) &= \mathbf{E}_{p(x^i)} \mathbf{E}_{\phi^{-i}(x^{-i})} g^i(\sigma^i, f^{-i}(y^{-i})) = \mathbf{E}_{r(x^i)} g^i(\sigma^i, f^{-i}(y^{-i})) \\ &= \sum_{y^i} \phi^i(x^i)(y^i) \mathbf{E}_{q(y^i)} g^i(\sigma^i, f^{-i}(y^{-i})) \end{aligned}$$

Since f is a Nash equilibrium of $\Gamma(\mathcal{J}, G)$, this is at most

$$\sum_{y^i} \phi^i(x^i)(y^i) \mathbf{E}_{q(y^i)} g^i(f^i(y^i), f^{-i}(y^{-i})) = \mathbf{E}_{r(x^i)} g^i(e^i(x^i), f^{-i}(y^{-i}))$$

$$= \mathbf{E}_{p(x^i)} g^i(e^i(x^i), e^{-i}(x^{-i}))$$

■

This completes the first part of the proof of the main theorem.

4.2 Construction of a faithful interpretation if \mathcal{I} is richer than \mathcal{J}

First, we assume that the inclusion $D(\mathcal{I}, G) \supset D(\mathcal{J}, G)$ is also satisfied when G is an upper semi-continuous game, and prove the existence of a faithful interpretation under this assumption. Then we complete the proof of the main theorem using approximations of upper semi-continuous games by continuous games.

4.2.1 Case where the payoff function may be upper semi-continuous

An **upper semi-continuous (or u.s.c.) game** is given by $((S^i)_i, g)$, where the sets S^i 's are compact, and $g : S \rightarrow (\mathbb{R} \cup \{-\infty\})^I$ is an upper semi-continuous payoff function. For G an u.s.c. game and an information structure \mathcal{I} , $\Gamma(\mathcal{I}, G)$ and $D(\mathcal{I}, G)$ are defined as in the case of a compact game.

Given an information structure \mathcal{J} , we construct an u.s.c. game $G = G(\mathcal{J})$ as follows: An element of $S^i = Y^i \times \Delta(Y^{-i})$ is a couple (s_Y^i, s_Δ^i) , and the payoff function is defined by $g^i(s_Y, s_\Delta) = \ln s_\Delta^i(s_Y^{-i})$ if $s_\Delta^i(s_Y^{-i}) > 0$, and $g^i(s_Y, s_\Delta) = -\infty$ otherwise. The payoff of i does not depend on s_Δ^{-i} nor on s_Y^i . We write it $\tilde{g}^i(s_\Delta^i, s_Y^{-i})$.

Now consider an information structure \mathcal{I} , and a I -tuple e of strategies in $\Gamma(\mathcal{I}, G)$. For $x^i \in X^i$, $e^i(x^i)$ is a probability measure over $Y^i \times \Delta(Y^{-i})$. We denote by $e_Y^i(x^i)$ and $e_\Delta^i(x^i)$ its marginals on Y^i and $\Delta(Y^{-i})$ respectively.

In G , each player announces a signal and a probability over the signals of the others. The payoff function is designed such that each player has incentive to announce as probability his conditional probability over the signals announced by the others. More precisely, e induces with μ a probability P_e on $X \times Y \times \prod_i \Delta(Y^{-i})$, and $P_e(s_Y^{-i}/x^i) = \mathbf{E}_{p^i(x^i)} e_Y^{-i}(x^{-i})(s_Y^{-i})$ is the conditional probability on x^i that the other players announce the signal s_Y^{-i} . (Recall that $p^i(x^i)(x^{-i}) = P_e(x^{-i}/x^i) = \frac{\mu(x^i, x^{-i})}{\mu(x^i)}$.)

Lemma 4.3 e is a Nash equilibrium of $\Gamma(\mathcal{I}, G)$ if and only if for all i and $x^i \in X^i$, $e_{\Delta}^i(x^i)$ is the Dirac mass at $\mathbf{E}_{p^i(x^i)} e_Y^{-i}$.

Proof: From remark 2.1, e is a Nash equilibrium if and only if for all i and $x^i \in X^i$, $e_{\Delta}^i(x^i) \in \arg \max_{\sigma_{\Delta}^i \in \Delta(\Delta(Y^{-i}))} \mathbf{E}_{p^i(x^i)} \tilde{g}^i(\sigma_{\Delta}^i, e_Y^{-i}(x^{-i}))$. Note that for $\gamma, \delta \in \Delta(Y^{-i})$,

$$\mathbf{E}_{\gamma} \tilde{g}^i(\gamma, y^{-i}) - \mathbf{E}_{\delta} \tilde{g}^i(\delta, y^{-i}) = \sum_{y^{-i} \in Y^{-i}} \gamma(y^{-i}) \ln \frac{\gamma(y^{-i})}{\delta(y^{-i})} = D(\gamma \parallel \delta)$$

Where $D(\gamma \parallel \delta)$ is the relative entropy (or Kullback Leibler distance [CT91]) between γ and δ . By property of the relative entropy, $D(\gamma \parallel \delta) \geq 0$, with $D(\gamma \parallel \delta) = 0$ if and only if $\delta = \gamma$. This proves the lemma. ■

Proposition 4.1 Assume that $D(\mathcal{I}, G) \supset D(\mathcal{J}, G)$ for the previously defined u.s.c. game G , then \mathcal{J} is faithfully reproducible from \mathcal{I} .

Proof: Let the (pure) strategies in $\Gamma(\mathcal{J}, G)$ be defined by $f^i(y^i) = (y^i, q^i(y^i))$ ($q^i(y^i)(y^{-i}) = \frac{\nu(y^{-i}, y^i)}{\nu(y^i)}$). By lemma 4.3, f is a Nash equilibrium. Let e be a Nash equilibrium inducing the same distribution on the actions of G . An interpretation ϕ is given by $\phi^i(x^i)(y^i) = e_Y^i(y^i)$. The proposition is a consequence of the two next lemmas.

Lemma 4.4 ϕ is a compatible interpretation from \mathcal{I} to \mathcal{J} .

Proof: Just note that the marginals on Y of the distributions induced by e and f on S are ν and the image of μ by ϕ respectively. ■

Lemma 4.5 ϕ is faithful.

Proof: Take $x^i \in X^i$ and $y^i \in Y^i$ such that $\phi^i(x^i)(y^i) > 0$. As e is a Nash equilibrium, $e_{\Delta}^i(x^i)$ is the Dirac mass at $\mathbf{E}_{p^i(x^i)} e_Y^{-i}(x^{-i}) = \mathbf{E}_{p^i(x^i)} \phi^{-i}(x^{-i})$. Then $e^i(x^i)(y^i, \mathbf{E}_{p^i(x^i)} \phi^{-i}(x^{-i})) > 0$. As e and f induce the same distribution on the actions of G , there exists $y'^i \in Y^i$ such that $f^i(x^i)(y'^i, \mathbf{E}_{p^i(x^i)} \phi^{-i}(x^{-i})) > 0$. By definition of f we have $y'^i = y^i$, and $q^i(y^i) = \mathbf{E}_{p^i(x^i)} \phi^{-i}(x^{-i})$. ■

4.2.2 Case where the payoff function is continuous

Here we approximate the game $G = G(\mathcal{J})$ by a family of continuous games $G_K = G_K(\mathcal{J})$, for $K < 0$. We study the best response correspondence of $\Gamma(\mathcal{I}, G_K)$, then we construct an interpretation from \mathcal{I} to \mathcal{J} from a Nash equilibrium of $\Gamma(\mathcal{J}, G)$ and prove that it is an ε -faithful interpretation.

For some K , $G_K = G_K(\mathcal{J})$ is the compact game whose spaces of strategies are S^i , and with payoff function given by $g_K^i(s_Y, s_\Delta) = \max\{g(s_Y, s_\Delta), K\}$. We again write $g_K^i(s_Y, s_\Delta) = \tilde{g}_K^i(s_\Delta^i, s_Y^{-i})$.

Lemma 4.6 *Let $\gamma \in \Delta(Y^{-i})$, and $\delta \in \arg \max_{\delta' \in \Delta(Y^{-i})} \mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i})$. There exists a subset J of Y^{-i} such that $\delta(y^{-i}) = \frac{\gamma(y^{-i})}{\sum_{j \in J} \gamma(j)}$ if $y^{-i} \in J$, $\delta(y^{-i}) = 0$ if $y^{-i} \notin J$, and $y^{-i} \in J$ if $\gamma(y^{-i}) > -\frac{1}{K}$*

Proof: Take $\delta \in \arg \max_{\delta' \in \Delta(Y^{-i})} \mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i})$, and $\delta' \in \Delta(Y^{-i})$. Put $p_m = \exp K$.

$$\mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i}) = \sum_{\delta'(y^{-i}) < p_m} \gamma(y^{-i})K + \sum_{\delta'(y^{-i}) \geq p_m} \gamma(y^{-i}) \ln \delta'(y^{-i})$$

Then if $\delta(y^{-i}) < p_m$, $\delta(y^{-i}) = 0$. Maximizing $\sum_{\delta'(y^{-i}) \geq p_m} \gamma(y^{-i}) \ln \delta'(y^{-i})$ shows that if $\delta(y^{-i}) \neq 0$,

$$\delta(y^{-i}) = \frac{\gamma(y^{-i})}{\sum_{\delta(j) \geq p_m} \gamma(j)}$$

Put $J = \{y^{-i}, \delta(y^{-i}) > 0\}$, we have to prove that $y^{-i} \in J$ if $\gamma(y^{-i}) > -\frac{1}{K}$. Take $j_0 \notin J$, and denote $J_1 = J \cup \{j_0\}$. Now define δ' by $\delta'(y^{-i}) = \frac{\gamma(y^{-i})}{\sum_{j \in J_1} \gamma(j)}$ if $y^{-i} \in J_1$, and $\delta'(y^{-i}) = 0$ otherwise.

$$\mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i}) \geq \sum_{y^{-i} \notin J_1} \gamma(y^{-i})K + \sum_{y^{-i} \in J} \gamma(y^{-i}) \ln \left(\frac{\gamma(y^{-i})}{\sum_{j \in J_1} \gamma(j)} \right)$$

Write $a = \sum_{j \in J} \gamma(j)$ and $b = \gamma(j_0)$. We get

$$\mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i}) - \mathbf{E}_\gamma \tilde{g}_K^i(\delta, y^{-i}) \geq b \ln b - Kb + a \ln a - (a + b) \ln(a + b)$$

Since $a + b \leq 1$,

$$\mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i}) - \mathbf{E}_\gamma \tilde{g}_K^i(\delta, y^{-i}) \geq b \ln b - Kb + (1 - b) \ln(1 - b) \geq \ln \frac{1}{2} - Kb$$

Therefore $\mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i}) > \mathbf{E}_\gamma \tilde{g}_K^i(\delta, y^{-i})$ if $\gamma(j_0) > -\frac{1}{K}$, which proves that if $\gamma(j_0) > -\frac{1}{K}$, $j_0 \in J$. \blacksquare

Lemma 4.7 *For $\varepsilon > 0$, there exists $K < 0$ such that for , $\gamma \in \Delta(Y^{-i})$, $\delta \in \arg \max_{\delta' \in \Delta(Y^{-i})} \mathbf{E}_\gamma \tilde{g}_K^i(\delta', \gamma)$ implies $d(\delta, \gamma) < \varepsilon$.*

Proof: Take $0 < \varepsilon < \frac{1}{2}$. As Y^{-i} is finite, we can take $K < -\frac{1}{\varepsilon}$ such that for all $\gamma \in \Delta(Y^{-i})$, $\sum_{j \in J} \gamma(j) > 1 - \frac{\varepsilon}{2}$ with $J = \{y^{-i}, \gamma(y^{-i}) > -\frac{1}{K}\}$. Take $\delta \in \arg \max_{\delta'} \mathbf{E}_\gamma \tilde{g}_K^i(\delta', y^{-i})$. If $\delta(y^{-i}) > 0$, $\gamma(y^{-i}) \leq -\frac{1}{K} \leq \varepsilon$, and if $\delta(y^{-i}) \neq 0$:

$$|\delta(y^{-i}) - \gamma(y^{-i})| = \gamma(y^{-i}) \left(\frac{1}{\sum_{j \in J} \gamma(j)} - 1 \right) \leq \frac{1}{1 - \frac{\varepsilon}{2}} - 1 \leq \varepsilon$$

\blacksquare

Now we can construct ε -faithful interpretations.

Proposition 4.2 *If \mathcal{I} is richer than \mathcal{J} , there exists an ε -faithful interpretation from \mathcal{I} to \mathcal{J} for all $\varepsilon > 0$.*

Proof: For $\varepsilon > 0$, let K be chosen such that $\delta \in \arg \max_{\delta'} \mathbf{E}_\gamma \tilde{g}_K^i(\delta, y^{-i})$ implies $d(\delta, \gamma) < \frac{\varepsilon}{2}$. To $y^i \in Y^i$ we associate $f_\Delta(y^i)$ that maximizes $\mathbf{E}_{q^i(y^i)} \tilde{g}_K^i(\cdot, y^{-i})$. The I -tuple of (pure) strategies f in $\Gamma(\mathcal{J}, G_K)$ defined by $f^i(y^i) = (y^i, f_\Delta(y^i))$ is a Nash equilibrium. Take a Nash equilibrium e of $\Gamma(\mathcal{I}, G_K)$ inducing the same distribution on S as f , and let $e_Y^i(x^i)$ and $e_\Delta^i(x^i)$ be the marginals of $e^i(x^i)$ on Y^i and $\Delta(Y^{-i})$. An interpretation from \mathcal{I} to \mathcal{J} is again defined by $\phi^i(x^i) = e^i(x^i)$. We see, as in the case where G is u.s.c., that ϕ is a compatible interpretation. We have to prove that ϕ is ε -faithful.

Consider $x^i \in X^i$ and y^i such that $\phi^i(x^i)(y^i) > 0$. Let $U \subset Y^i \times \Delta(Y^{-i})$ be the support of the image of μ by f^i . The support T of $e^i(x^i)$ is included in U . By definition of f , the section of U by $\{y^i\} \times \Delta(Y^{-i})$ is $\{y^i, f_\Delta^i(y^i)\}$. The section of T by $\{y^i\} \times \Delta(Y^{-i})$ is not empty since $e_Y^i(x^i)(y^i) > 0$, therefore it is also $\{y^i, f_\Delta^i(y^i)\}$. Then $f_\Delta^i(y^i) \in \arg \max_\delta \mathbf{E}_{p^i(x^i)} \tilde{g}_K^i(\delta, e^{-i}(x^{-i})(y^{-i})) = \arg \max_\delta \mathbf{E}_{r^i(x^i)} \tilde{g}_K^i(\delta, y^{-i})$, with $r^i(x^i) = \mathbf{E}_{p^i(x^i)} e_Y^{-i}(x^{-i}) = \mathbf{E}_{p^i(x^i)} \phi^{-i}(x^{-i})$. From the choice of K , $d(f_\Delta^i(y^i), r^i(x^i)) < \frac{\varepsilon}{2}$, but also $d(f_\Delta^i(y^i), q^i(y^i)) < \frac{\varepsilon}{2}$ since f is a Nash equilibrium. This finally proves that $d(r^i(x^i), q^i(y^i)) < \varepsilon$.

Proof of Theorem 4.1: As seen in proposition 3.1, the existence for all $\varepsilon > 0$

of an ε -faithful interpretation from \mathcal{I} to \mathcal{J} implies the existence of a faithful interpretation.

Corollary 4.1 *A consistent interpretation ϕ from \mathcal{I} to \mathcal{J} is a faithful interpretation if and only if for every compact game G and every Nash equilibrium f of $\Gamma(\mathcal{J}, G)$, the strategies defined by $e(x^i) = \sum_{y^i} \phi(x^i)(y^i)f(y^i)$ form a Nash equilibrium of $\Gamma(\mathcal{I}, G)$.*

Proof: The *only if* sense is a consequence of the lemmas 4.1 and 4.2. For the *if* proof, take $\varepsilon > 0$, then choose K and construct an equilibrium f_0 of $\Gamma(\mathcal{I}, G_K)$ as in the proof of proposition 4.2. Also let $x^i \in X^i$ and $y^i \in Y^i$ be such that $\phi^i(x^i)(y^i) > 0$. We still have $d(f_\Delta(y^i), q(y^i)) < \frac{\varepsilon}{2}$. On the other side, $(y^i, f_\Delta(y^i))$ is in the support of $e^i(x^i)$. Therefore $f_\Delta(y^i) \in \arg \max_\delta \mathbf{E}_{e_{Y^i}^{-i}} \tilde{g}_K^i(\delta, y^{-i})$, which implies that $d(f_\Delta(y^i), \mathbf{E}_{p^i(x^i)} \phi^{-i}(x^{-i})) < \frac{\varepsilon}{2}$. This proves that ϕ is ε -faithful for all ε , and therefore ϕ is faithful. ■

5 Equivalence classes – Examples

5.1 examples

Example 5.1

With two players we represent information structures by matrices. Each cell contains its probability to be drawn, player 1 is informed about the row, and player 2 about the column.

	a_1	b_1
x_1	1/3	1/3
y_1	1/3	0

	a_2	b_2	c_2
x_2	1/6	0	1/6
y_2	0	1/6	1/6
z_2	1/6	1/6	0

\mathcal{I}_1

\mathcal{I}_2

Let ϕ be the interpretation from \mathcal{I}_2 to \mathcal{I}_1 defined by $\phi^1(x_2) = \phi^1(y_2) = x_1$, $\phi^1(z_2) = y_1$ for player 1, and $\phi^2(a_2) = \phi^2(b_2) = a_1$, $\phi^2(c_2) = b_1$ for player 2. ϕ is a faithful interpretation from \mathcal{I}_2 to \mathcal{I}_1 .

5.2 Equivalence relation

Definition 5.1 \mathcal{I} and \mathcal{J} are **equivalent** information structures whenever \mathcal{I} is richer than \mathcal{J} and \mathcal{J} is richer than \mathcal{I} .

Example 5.2

	a_3	b_3	c_3
x_3	1/6	1/12	1/12
y_3	1/6	1/12	1/12
z_3	1/3	0	0

\mathcal{I}_3

We see that in \mathcal{I}_3 the second and third columns are the same, as are the first and second rows. If we set $\phi^1(x_3) = \phi^1(y_3) = x_1$, $\phi^1(z_3) = y_3$, and $\phi^2(a_3) = a_1$, $\phi^2(b_3) = \phi^2(c_3) = b_1$, we see that \mathcal{I}_3 is richer than \mathcal{I}_1 . Conversely, a faithful interpretation from \mathcal{I}_1 to \mathcal{I}_3 is defined by $\phi^1(x_1) = \frac{1}{2}x_3 + \frac{1}{2}y_3$, $\phi^1(y_1) = z_3$ and $\phi^2(a_1) = a_3$, $\phi^2(b_1) = \frac{1}{2}b_3 + \frac{1}{2}c_3$. Therefore \mathcal{I}_3 and \mathcal{I}_1 are equivalent.

5.3 Deterministic interpretations

Definition 5.2 A interpretation ϕ from \mathcal{I} to \mathcal{J} is said to be **deterministic** when for every $x^i \in X^i$, the support of $\phi^i(x^i)$ is a singleton, that is when ϕ^i is a mapping from X^i to Y^i .

Example 5.3

This is the case of the interpretation from \mathcal{I}_2 and \mathcal{I}_3 to \mathcal{I}_1 , but not of the one from \mathcal{I}_1 to \mathcal{I}_3 .

Definition 5.3 An information structure $\tilde{\mathcal{J}}$ is called **minimal** when any faithful interpretation from an information structure \mathcal{I} to $\tilde{\mathcal{J}}$ is deterministic.

Proposition 5.1 For any information structure \mathcal{I} , there exists an information structure $\tilde{\mathcal{J}}$ that is equivalent to \mathcal{J} and minimal.

Proof: On Y^i we define an equivalence relation by $y^i \mathcal{R} y'^i$ when $q(y^i) = q(y'^i)$. Let \tilde{Y}^i be the set of equivalence classes for \mathcal{R} on Y^i , and let ψ be defined by

$\psi^i(y^i)(\tilde{y}^i) = 1$ if $y^i \in \tilde{y}^i$, and $\psi^i(y^i)(\tilde{y}^i) = 0$ if $y^i \notin \tilde{y}^i$. $\tilde{\nu}$ represents the image by ψ of ν , so that ψ is a compatible interpretation from \mathcal{J} to $\tilde{\mathcal{J}} = ((\tilde{Y}^i)_i, \tilde{\nu})$.

Lemma 5.1 *ψ is a faithful interpretation from \mathcal{J} to $\tilde{\mathcal{J}}$.*

For $y^i, y'^i \in Y^i$ and $\tilde{y}^i \in \tilde{Y}^i$ such that $\psi^i(y^i)(\tilde{y}^i)\psi^i(y'^i)(\tilde{y}^i) > 0$, $y^i \in \tilde{y}^i$ and $y'^i \in \tilde{y}^i$. Then $q(y^i) = q(y'^i)$, therefore $\mathbf{E}_{q(y^i)}\psi^{-i}(y^{-i}) = \mathbf{E}_{q(y'^i)}\psi^{-i}(y^{-i})$. ■

On the other hand, an interpretation $\tilde{\psi}$ from $\tilde{\mathcal{J}}$ to \mathcal{J} is given by $\tilde{\psi}^i(\tilde{y}^i)(y^i) = \frac{\nu(y^i)}{\tilde{\nu}(\tilde{y}^i)}$ if $y^i \in \tilde{y}^i$, and $\tilde{\psi}^i(\tilde{y}^i)(y^i) = 0$ if $y^i \notin \tilde{y}^i$.

Lemma 5.2 *$\tilde{\psi}$ is a faithful interpretation from $\tilde{\mathcal{J}}$ to \mathcal{J} .*

See that for $\tilde{y} \in \tilde{Y}$, the restriction of ν to the cartesian product of the sets \tilde{y}^i is a product measure. If $y \in \tilde{y}$, $\nu(y) = \tilde{\nu}(y)\prod_i \frac{\nu(y^i)}{\tilde{\nu}(\tilde{y}^i)}$, and thus $\nu(y) = \tilde{\nu}(\tilde{y})\tilde{\psi}(\tilde{y})(y)$. Then $\tilde{\psi}$ is a consistent interpretation from $\tilde{\mathcal{J}}$ to \mathcal{J} . If $\tilde{\psi}^i(\tilde{y}^i)(y^i)\tilde{\psi}^i(\tilde{y}^i)(y'^i) > 0$, $y^i, y'^i \in \tilde{Y}^i$, and $q(y^i) = q(y'^i)$. ■

To get $\tilde{\mathcal{J}}$ from \mathcal{J} , we identify the signals of player i in \mathcal{J} that lead to the same conditional probability on the signals of the other players. The next lemma shows that if we do this a second time, we still get $\tilde{\mathcal{J}}$. Let $\tilde{q}(\tilde{y}^i)$ be the conditional probability of $\tilde{\nu}$ over \tilde{Y}^{-i} given \tilde{y}^i .

Lemma 5.3 *For $\tilde{y}^i, \tilde{y}'^i \in \tilde{Y}^i$, $\tilde{q}(\tilde{y}^i) = \tilde{q}(\tilde{y}'^i)$ implies $\tilde{y}^i = \tilde{y}'^i$.*

Proof: Take $y^i \in \tilde{y}^i$ and $y'^i \in \tilde{y}'^i$, since $\tilde{\psi}$ is faithful $q(y^i) = \mathbf{E}_{\tilde{q}(\tilde{y}^i)}\tilde{\psi}(y^{-i}) = \mathbf{E}_{\tilde{q}(\tilde{y}'^i)}\tilde{\psi}(y^{-i}) = q(y'^i)$. Therefore $\tilde{y}^i = \tilde{y}'^i$.

Proof of the proposition: Consider a faithful interpretation ϕ from \mathcal{I} to $\tilde{\mathcal{J}}$. Take $x^i \in Y^i$, and $\tilde{y}^i, \tilde{y}'^i \in \tilde{X}^i$ such that $\phi^i(x^i)(\tilde{y}^i)\phi^i(x^i)(\tilde{y}'^i) > 0$. $\tilde{q}(\tilde{y}^i) = \tilde{p}(\tilde{y}'^i)$, therefore $y^i = y'^i$. ■

6 Concluding remarks.

We have introduced a dual approach to the classical approach of correlated equilibria. We considered normal form games extended by information structures, but rather than making the information structure vary to get all the correlated equilibrium distributions of the normal form game, we compared two information structures by making the normal form game vary. We then obtained a

characterization of “ \mathcal{I} is richer than \mathcal{J} ” where the normal form game does not appear.

The relation “ \mathcal{I} is richer than \mathcal{J} ” defines a partial order relation on the information structures. The smallest element is the trivial information structure where each player has only one signal that appears with probability one.

In this framework, the normal form games are supposed compact and the information structures finite. One may wonder if the main theorem holds with a smaller class of games (finite games for instance) and with a bigger class of information structures.

References

- [Au74] Aumann, R. J. (1974) Subjectivity and Correlation in Randomized Strategies. *Journal of Mathematical Economics* **1**, 67-95.
- [Au87] Aumann, R. J. (1987) Correlated Equilibrium as an Expression of Bayesian Rationality. *Econometrica* **55**, 1-18.
- [Bl51] Blackwell, D. (1951) Comparison of Experiments. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 93-102.
- [Bl53] Blackwell, D. (1953) Equivalent Comparison of Experiments. *Annals of Mathematical Statistics* **24**, 265-272.
- [CT91] Cover, T.M., and J.A. Thomas (1991) Elements of Information Theory. *Wiley Series in Telecommunications*.
- [Me94] Mertens, J-F. (1994) Correlated- and Communication Equilibria. J.F. Mertens and S. Sorin Eds., *Game-Theoretic Methods in General Equilibrium Analysis*, NATO ASI Series, 243-248.