Olivier Gossner

Game Theory

Lecture notes. Value of information

1 One person decision problems

We start with a (finite) set K of states of nature, $k \in K$ is what the agent is unknown about, we assume a probability distribution p on K.

1.1 Representations of information

1.1.1 Partition representation of information

There is a set Ω of *states of the world*. To each state is associated one $k \in K$ via a mapping $\kappa\colon \Omega \to K$. The probability distribution over Ω is P. The agent's information is represented by a partition $\mathscr{P}$ of Ω. Given an element π of the partition, the probability of a state k is given by Bayes's rule through:

$$P(\kappa(\omega) = k | \omega \in \pi)$$

1.1.2 Signals representation of information

A *transition probability* from a finite set A to a finite set B is a family of probabilities over B, one for each element of A.

Definition 1.1. A *statistical experiment* is given by a set of signals X, and a distribution over signals $\alpha_k \in \Delta(X)$ for every signal k. Upon receiving the signal $x \in X$, the conditional probability of $k \in K$ is given by Bayes's rule through

$$P(k|x) = \frac{P(x,k)}{P(x)} = \frac{p(k)\alpha_k(x)}{\sum_{k' \in K} p(k')\alpha_{k'}(x)}$$

Exercise 1.2. Can we find a partition representation from a signals representation? Can we do the converse?

In each state of nature k, $\alpha_k \in \Delta(X)$ is the probability over signals to the decision maker, and $\alpha_k(x)$ is the probability of the particular signal x.

Example 1.3. There are 2 states of nature, Rain or Shine. I can ask for a weather forecast to my aunt Alice, or to my uncle Bob.

Before a rainy day, my aunt Alice predicts rain 60% of the time, a shiny day 20% of the time, and is indecisive 20% of the time. Before a sunny day, Alice predicts rain 20% of the time, is indecisive 10% of the time, and predicts sun 70% of the time.

Alice's statistical experiment can be represented by a matrix, where each row corresponds so a state of nature (k_R for rain, k_S for shine), each column corresponds to a signal (x_R for rain, x_S for shine, and x_I for indecisive) and the number in the cell corresponds to the probability of the signal if the given state of nature realizes. The matrix is as follows:

	x_R	x_I	x_S
k_R	0.6	0.2	0.2
k_S	0.2	0.1	0.7

Alice's statistical experiment

Before a rainy day, Bob predicts rain 70% of the time, and sun 30% of the time. Before a sunny day, he predicts rain 25% of the time, and sun 75% of the time. The statistical experiment corresponding to Bob is as follows:

	y_R	y_S
k_R	0.7	0.3
k_S	0.25	0.75

Bob's statistical experiment

1.1.3 Beliefs representation

A *beliefs representation* of information is given by a distribution over $\Delta(K)$, it is an element of $\Delta(\Delta(K))$. Element q has probability $p(q)$ of realising. The interpretation is that q is the agent's posterior belief with probability $p(q)$. The distribution of posterior beliefs satisfies the martingale property:

$$\sum_q p(q)q = p$$

Exercise 1.4.
- Consider the different posteriors arising from a statistical experiment. Show that they satisfy the martingale property.
- Consider any beliefs representation of information that satisfies the martingale property, show that there exists a statistical experiment that induces this distribution of posteriors.

Exercise 1.5. Assume that rain and shine are equally likely, and find the beliefs representation arising from the Alice's and Bob's statistical experiments.

1.2 Comparison of statistical experiments

1.2.1 Decision problems

A *decision problem* is given by a set of actions (choices) A and a payoff function $g: A \times K \to \mathbb{R}$. $g(a,k)$ is the payoff of action a in state k.

Agent's optimal strategy: The agent needs to find a rule: $f^*: X \to K$ that maximizes the total payoff

$$V_{\alpha,g} = \mathbb{E}_{x,k} g(f^*(x),k)$$

The agent's problem is to find, for each x, an action $f^*(x)$ that maximizes:

$$\mathbb{E}_{P(k|x)} g(f^*(x),k)$$

We introduce the function $v: \Delta(X) \to \mathbb{R}$ given by:

$$v(q) = \arg\max_a \mathbb{E}_q g(a,k)$$

This is the maximum expected payoff to an agent of belief q.

Using the beliefs representation: We have

$$\mathbb{E}_{x,k} g(f^*(x),k) = \mathbb{E}_q v(q)$$

The value of information is read directly from v and the distribution of posterior beliefs.

Observe that the mapping v is convex. Because of the martingale property:

$$\mathbb{E}_p v(q) \geq v(p)$$

Showing that the value of information is not negative.

Example 1.6. Assume that rain and sun are equally likely, and I need to decide if I take my umbrella or not. My payoff is 0 if I take my umbrella, -1 if I do not take it on a rainy day, and 1 if I do not take it on a sunny day.

Before deciding if I take my umbrella or not, should I rather listen to my uncle Bob, or to my aunt Alice?

The payoffs, depending on the state of nature and action, are as follows (U for taking the umbrella, N for not taking it):

	U	N
k_R	0	-1
k_S	0	1

The decision problem

Let p be the belief (conditional on the signal) that the day is sunny. The expected payoff by taking the umbrella is 0 independently on p, and it is $2p-1$ for not taking it. It is then optimal to take the umbrella whenever $p \geq \frac{1}{2}$, and optimal not to take it when $p \leq \frac{1}{2}$.

Let's look at Alice's statistical experiment. The probability that the announces x_R is $\frac{1}{2}0.6+\frac{1}{2}0.2 = .4$, the probability she announces x_I is .15, and the probability she announces X_S is .45. Conditional on her announcing x_R, the probability of a sunny day is $p(k_S|x_U) = \frac{.5*.2}{.5*.6+.5*.2} = .25$ by Bayes's rule. Thus when she announces x_R, I take my umbrella and obtain an expected payoff of 0. When she announces x_I, I compute $p(k_S|x_I) = \frac{1}{3}$, so I take my umbrella for an expected payoff of 0. Finally, $p(k_S|x_S) = \frac{7}{9} > \frac{1}{2}$, so following x_S I do not take my umbrella and my expected payoff is $\frac{5}{9}$. In total, my expected payoff is I listen to Alice and follow this optimal strategy is:

$$.4*0+.15*0+.45*\frac{5}{9} = .25$$

Now let's look at information from Bob. We compute $p(y_R) = .475$, and $p(y_S) = .525$, $p(k_S|y_R) = \frac{.125}{.475} < \frac{1}{2}$, $p(k_S|y_S) = \frac{.375}{.525} > \frac{1}{2}$. So my optimal strategy is to take the umbrella following y_R, and not to take it following y_S. My expected payoff is then:

$$.475*0+.525*(2*\frac{.375}{.525}-1) = .225$$

Overall, I am better off by listening to Alice (although she can be indecisive sometimes) rather than to Bob.

1.2.2 Information garblings

Definition 1.7. Let X and Y be sets of signals, a garbling from X to Y is a transition probability from X to Y.

Example 1.8. I sometimes ask my nephew Quentin to report to me what Alice says about the weather. Quentin doesn't reproduce perfectly Alice's information. When Alice predicts Sun or Rain, he conveys the message accordingly, but, when Alice is indecisive, Quentin reports a prediction for Sun or for Rain with equal probabilities.

Before a sunny day and before a rainy day, what are the probabilities of Quentin's reports?

Let z_R or z_S represent Quentin's forecast. In the state of nature k_R, we compute the probability that Quentin announces z_R as follows:

$$\begin{aligned} p(z_R|k_R) &= p(z_R, x_R|k_R) + p(z_R, x_I|k_R) + p(z_R, x_S|k_R) \\ &= p(x_R|k_R)p(z_R|k_R, x_R) + p(x_I|k_R)p(z_R|k_R, x_I) + p(x_S|k_R)p(z_R|k_R, x_S) \\ &= .6*1+.2*.5+.2*0 = .7 \end{aligned}$$

And $p(z_S|k_R) = .3$. I also compute $p(z_R|k_S) = 0.2*0+.1*.5*.7*1 = .75$ and $p(z_S|k_S) = .25$. I can thus represent the information from Quentin as follows:

	z_R	z_S
k_R	0.7	0.3
k_S	0.25	0.75

Quentin's statistical experiment

In general (meaning, independently of the decision problem that I face), am I better off asking Alice directly, or asking Quentin to report Alice's forecast? Whenever I have information from Alice, I can reproduce what Quentin does by flipping a coin if Alice is indecisive. Therefore, Alice is always providing at least as valuable information as Quentin, and this is independent of the decision problem that I face.

Looking at numbers, we see that Quentin and Bob's information structures are the same. Hence, I am indifferent to listening to Bob or to Quentin. Since Alice is better than Quentin, Alice's information is more valuable that Bob's information no matter the decision problem.

Definition 1.9. A statistical experiment $\alpha = (\alpha_k)$ with set of signals X is more informative than the statistical experiment $\beta = (\beta_k)$ with set of signals Y if there exists a garbling $Q = (Q_x)_x$ that mimics the latter from the former:

$$\forall k, y \quad \beta_k(y) = \sum_x \alpha_k(x) Q_x(y)$$

What is a better experiment than another when having to make decisions? Given a decision problem G and a statistical experiment α, and assuming the uniform probability over states of nature, G_α represents the induced Bayesian game with one player. We let $V(\alpha, G)$ denote the maximal expected payoff to the decision maker over all possible strategies in G_α.

1.2.3 Comparison of statistical experiments: usefulness

Definition 1.10. A statistical experiment α is more useful than a statistical experiment β if, for every decision problem G,

$$V(\alpha, G) \geq V(\beta, G)$$

Is Alice more useful than Bob? If we want to answer this question, we must show that the payoff from listening to Alice is as high as from listening to Bob in **every** decision problem! Is there a simpler way to do this?

1.2.4 Blackwell's Theorem

When is a statistical experiment more useful than another? The answer is given by Blackwell's theorem on comparison of statistical experiments.

Theorem 1.11. *α is more useful than β if only if α is more informative than β.*

Proof. We first prove that more informative experiments are always more useful. Let Q be a garbling from X to Y that mimics β from α. Let G be any decision problem, with action set A and payoff function g. For any strategy σ in G extended by β, define the strategy σ' in D extended by α by:

$$\sigma'_x(a) = \sum_y Q_x(y)\sigma_y(a)$$

The strategy σ' corresponds to 1) using Q to compute random signals in Y and 2) follow σ in the decision problem according to the garbled signal obtained. We verify that σ and σ' give the same expected payoff in every state of nature k:

$$\begin{aligned}\sum_x \alpha_k(x)\sigma'_x(a)g(a,k) &= \sum_x \alpha_k(x)\sum_y Q_x(y)\sigma_y(a)g(a,k)\\ &= \sum_y \beta_k(y)\sigma_y(a)g(a,k)\end{aligned}$$

Therefore, for every strategy in G extended by β, there exists a strategy in G extended by α which gives at least the same payoff. Thus, α is more useful than β.

Now we prove the converse by an application of the min max theorem. Let D the set of decision problems where $A = Y$, and with payoff function bounded by 0 and 1:

$$D = \{g \colon Y \times K \mapsto [0,1]\}$$

Let C be the set of (behavioral) strategies in the decision problem D with statistical experiment. An element of C is thus a transition probability σ from X to Y. For every g,

$$V(\alpha,g) \geq V(\beta,g)$$

In particular, for every $g \in D$, there exists σ in C such that the expected payoff given by σ in g extended by α is no less than the payoff of the identity strategy in g extended by β:

$$\sum_{x,k}\sum_y \alpha_k(x)\sigma_x(y)g(y,k) \geq \sum_k\sum_y \beta_k(y)g(y,k) \tag{1}$$

Consider the auxiliary game between players I and II in which player I chooses an element of C, player II chooses an element of D, and the payoff to player I is:

$$\gamma(\sigma,g) = \sum_{x,k}\sum_y \alpha_k(x)\sigma_x(y)g(y,k) - \sum_k\sum_y \beta_k(y)g(y,k)$$

Player I and II's strategy sets are convex and compact, and the payoff function γ is linear in each player's strategy. We can thus apply the min max theorem:

$$\min_D \max_\sigma \gamma(\sigma,g) = \max_\sigma \min_D \gamma(\sigma,g)$$

Equation (1) shows that $\max_\sigma \min_D \gamma(\sigma,g) \geq 0$. Thus, there exists $\sigma^* \in C$ such that, for every $g \in D$,

$$\gamma(\sigma^*,g) \geq 0 \tag{2}$$

In particular, for any pair k,y, considering the decision problem for which $g(k',y')=1$ if $(k',y')=(k,y)$ and $g(k',y')=0$ otherwise, equation (2) shows:

$$\sum_x \alpha_k(x)\sigma_x(y) \geq \beta_k(y) \tag{3}$$

For any k, summing the left side or the right side of (3) gives a total of 1. So we conclude that for every k,y:

$$\sum_x \alpha_k(x)\sigma_x(y) = \beta_k(y)$$

which shows that σ, seen as a garbling from X from Y, mimics β from α. $\square$

1.3 Rational inattention

Now we let the agent choose which information to receive. We do not want to put an artificial limit on what are the experiments she may choose from. Rather, we let her choose any experiment she wishes! The constraint is that information extraction is costly. More precisely, the cost of information we consider is proportional to the decrease in *entropy* between her prior and posterior beliefs. This is known as *rational inattention*.

1.3.1 Entropy: definition

Let p be a probability distribution in $\Delta(K)$

- The entropy of p is defined as $H(p) = -\sum_k p(k)\log p(k)$
- $\log = \log_2$ by convention, $0\log 0 = 0$ by continuity
- Concave, maximal for the uniform distribution, 0 iff $p(k)=1$ for some k

Justifications: originally, information theory

- Axiomatizations
- Shortest coding of $k \in K$, example $p = (1/2, 1/4, 1/8, 1/8)$
- Coding of an iid. sequence in K^n (Shannon 48)

Reference: great book by Cover and Thomas (2006).

1.3.2 Rational inattention

Growing literature, important papers are Sims 00, 02, Matejka McKay 15, Caplin Dean 16.

Consider a situation in which the agent decides how to allocate her information.

- Her initial belief is $p \in \Delta(K)$
- She chooses any information structure, represented by a distribution $p(q)$ over posterior beliefs q
- The cost of information is $C(H(p) - \mathbb{E}_q H(q))$ for $C > 0$.
- After receiving information, she makes decisions in a decision problem with action set A and payoff $g \colon A \times K \to \mathbb{R}$

The agent chooses both *what information* to receive (and to pay for), and *how to use that information*. Entropy captures the cost of processing information and of attention.

1.3.3 Solving for the rational inattention model

We know that, with belief q, optimal decisions lead to an expected payoff of $v(q)$. The problem becomes to maximize

$$E_q\left(v(q) + CH(q)\right) - H(p)$$

over all distributions of beliefs q such that $\mathbb{E}q = p$ (beliefs representation of information).

Consider the function $w_C(q) = v(q) + CH(q)$, it is in general neither concave nor convex. Let $\operatorname{cav} w$ the lowest concave function that is above w. The value of the decision problem for the agent is:

$$\operatorname{cav} w(p) - H(p)$$

Analysis shows that either the agent stays at p and does not extract information, or locally, a small change in p does not change the terminal beliefs of the agent.

1.4 Bayesian persuasion

Consider the situation in which a *sender* is informed of a state of nature, and chooses which experiment is used by the *receiver*. Once she receives her information, the receiver chooses some action, which affects both her payoff and the payoff of the sender. For instance, the sender can be a pharmaceutical firm and the receiver a regulator. The objective of the regulator is to approve a new drug if it is effective enough, while the objective of the pharmaceutical firm is to have her drug approved by the regulator. What is the best way possible choice of information by the sender?

1.4.1 Model

There is a set K of states of nature. The sender and receiver's prior is a distribution p. The receiver has a finite set of actions, A, and her payoff function is $g_R \colon A \times K \to \mathbb{R}$. The sender's payoff function is $g_S \colon A \times \mathbb{R}$.

We use the beliefs representation of information. The sender chooses a *statistical experiment*, given by a probability distribution over $\Delta(K)$ such that $\mathbb{E}q = p$.

Given a belief $q \in \Delta(K)$, the received chooses an action

$$a^*(q) \in A^*(q) = \arg\max_a \mathbb{E}_q g_R(a,q).$$

1.4.2 Solving Bayesian persuasion

The sender manipulates the receiver's beliefs. In order to find his optimal strategy, we need to associate a payoff to each belief of the receiver. Since there may be several optimal actions for the receiver, we choose the one that gives the best payoff to the sender (to avoid selection problems).

We thus let:

$$v(q) = \max_{a \in A^*(q)} \mathbb{E} g_S(a,q).$$

The objective of the sender is then to find a distribution of q that averages to p, that maximizes

$$\mathbb{E}_q v(q).$$

The solution to this problem is given by

$$\operatorname{cav} v(p).$$

2 Bayesian Games: Definition

A Bayesian game is given by an information structure and a payoff structure. The set of players I and a set of "states of nature", K are common to both.

The information structure describes the uncertainty about the state of nature and the player's information about it. It is given by:

- Ω is the set of states of the world
- $p \in \Delta(\Omega)$ is the common prior probability distribution over Ω
- P_i is the information partition of player i

- $\kappa : \Omega \mapsto K$ describes the state of nature as a function of the state of the world.

The payoff structure represents the player's strategic interaction. It is given by:

- A set of actions A_i for each player i
- A payoff function g_i from $A := \Pi_i A_i \times K$ to $\mathbb{R}$ for each player i

When Ω has the structure of a product space $\Omega = \Pi_i T_i$, and each player i is informed of the component t_i of $\omega = (t_j)_j$, we say that t_i is player i's type.

An information structure $\mathscr{I}$ and a payoff structure G define a Bayesian game $G_{\mathscr{I}}$ in which:

- A state of the world ω is drawn according to p, the state of nature being $k = \kappa(\omega)$
- Each player i is informed of the element of P_i containing ω
- Actions in A_i are chosen, defining an action profile $a = (a_i)_i$.
- The payoff for player i is $g_i(a,k)$.

A pure strategy for player i in the Bayesian game is a mapping f_i from Ω to A_i that depends only on the information partition element of player i: for every ω and ω' such that $\omega' \in P_i(\omega)$, $\omega' = f_i(\omega)$. We let Σ_i be the set of such strategies.

A behavioral strategy for player i is a P_i-measurable mapping σ_i from Ω to $\Delta(A_i)$.

When the state of nature is ω and the profile of strategies is $f = (f_i)$, the chosen profile of actions is $f(\omega) = (f_i(\omega))_i$, and the state of nature is $\kappa(\omega)$. The corresponding payoff to player i in the game is $g_i(f(\omega), \kappa(\omega))$. According to the probability distribution p, the expected payoff for player i corresponding to the profile of strategies $f = (f_i)$ is

$$\gamma_i(f) = \mathbb{E}_p g_i(f(\omega), \kappa(\omega))$$

Definition 2.1. A Bayesian-Nash equilibrium of the game $G_{\mathscr{I}}$ is a Nash equilibrium of the game where player i's strategy set is Σ_i and player i's payoff function is γ_i.

Proposition 2.2 (Existence)**.** *Assume that Ω and the actions sets A_i are all finite. Then the Bayesian Game $G_{\mathscr{I}}$ has a Nash equilibrium in behavioral strategies.*

Proof. $G_{\mathscr{I}}$ has a finite number of strategies per player. So, by Nash's existence theorem, it has a Nash equilibrium in *mixed* strategies, where a mixed strategy is a randomization over pure strategies. By Kuhn's theorem, mixed strategies are equivalent to behavioral strategies, hence the existence of a Nash equilibrium in behavioral strategies. □

3 Example: Do you really want to defect?

Let us consider the two following two games which differ when player 2 plays D.

	C	D
C	$0,-2,$	$-10,-1$
D	$-1,-10$	$-5,-5$

G_L

	C	D
C	$0-2,$	$-10,-7$
D	$-1,-10$	$-5,-11$

G_R

G_L is played with probability p, and G_R is played with probability $1-p$.

1. **Assume both players know which game is played.** Then, the information structure can be represented as follows: $\Omega = \{\omega_L, \omega_R\}$, $K = \{L, R\}$, $p(L) = p$, $p(R) = 1-p$, $P_1(\omega) = P_2(\omega) = \{\omega\}$ for each ω, $\kappa(\omega_L) = L$, $\kappa(\omega_R) = R$. The payoff structure is given by the payoff matrices above. In the Bayesian game, player 1 has two strategies C and D, and player 2 has two strategies L and R. The Bayesian game in strategic form is:

	C	D
C	$0,-2,$	$-10,6p-7$
D	$-1,-10$	$-5,6p-11$

   No player informed

   We only solve for Bayesian equilibria in pure strategies. D for player 1 is a best-response to D, and D for player 2 is a best response to D if $p \geq \frac{1}{6}$. So for $p \geq \frac{1}{6}$, (D,D) is a Bayesian Nash equilibrium. C for player 1 is a best-response to C, and C is a best-response for player 2 to C if $p \leq \frac{5}{6}$, so for $p \leq \frac{5}{6}$, C,C is a Bayesian Nash equilibrium.

2. **Assume that only player 2 knows what game is played**. The information structure is as before, except that $P_1(\omega_L) = P_1(\omega_R) = \{\omega_L, \omega_R\}$. The payoff structure is as before. In the normal form, player 1 has the same strategy set as before. Player 2 has now 4 strategies, since a different action can be chosen in G_L and in G_R. We label these strategies as CC, CD, DC, DD, where the first coordinate indicates the action chosen in G_L, while the second coordinates indicates the action chosen in G_R. The payoff matrix is now:

	CC	CD	DC	DD
C	$0,-2,$	$10p-10,5p-7$	$-10p,p-2$	$-10,6p-7$
D	$-1,-10$	$4p-5,p-11$	$-4p-1,5p-10$	$-5,6p-11$

   Player 2 informed

   Player 2 has a dominant strategy: DC. This can be checked directly on the payoff matrix, or with a simple reasoning, D is dominant in G_L, and C is dominant in G_R, since the payoff in the combined game is a convex combination of the payoff in G_L and in G_R, DC is dominant is the Bayesian game. We check that C is a best-response to DC if $p \leq \frac{1}{6}$, and D is a best-response if $p \geq \frac{1}{6}$. Thus, if $p > \frac{1}{6}$, (D, DC) is the unique Bayesian Nash equilibrium, and if $p < \frac{1}{6}$, (C, DC) is the unique Bayesian Nash equilibrium. If $p = \frac{1}{6}$, there is a continuum of Bayesian Nash equilibria, in which player 1 randomizes between C and D, and player 2 plays DC.

Remark that the strategy C of player 2 in the game in which no player is informed corresponds to the strategy CC in which the game in which player 2 is informed, and the same holds for D versus DD. More generally, a strategy in a game with less information can always be played if the player has more information, since the extra information can always be ignored.

4 Example: Adverse selection in car sales

A car may be of good quality (G), or bad quality (B), with equal probabilities ($1/2$ each).

The seller may decide to sell (action S), or not not to sell (action NS). The buyer may accept to buy (B), or refuse to buy (NB). The price for the car is fixed, and is equal to some value p.

The utility for a car of good quality is 3 for the seller, and 4 for the buyer (units in thousand british pounds). A car of bad quality is worth 0, both to the buyer and to the seller.

Both Informed Assume that both players are informed of the quality of the car. Describe the situation as a Bayesian Game. Solve, depending on the value of p.

None Informed Assume that no player knows the quality of the car. Describe the situation as a Bayesian Game. Solve, depending on the value of p.

Information asymmetry Assume that only the seller is informed of the car's quality. Describe the situation as a Bayesian Game. Solve, depending on the value of p.

5 Betting

5.1 A betting example

There are three states of nature, $\omega_1, \omega_2, \omega_3$. Each state is equally likely. Player 1's information partition is

$$P_1 = \{\{\omega_1, \omega_2\}\{\omega_3\}\}$$

And Player 2's information partition is

$$P_2 = \{\{\omega_1\}, \{\omega_2, \omega_3\}\}$$

Players 1 and 2 each decide if they accept a bet (action B) or refuse the bet (action R). If both accept the bet, a transfer takes place depending on which state of nature is realized: player 1 pays 1 £ to player 2 in state ω_1, receives 2 £ from player 2 in state ω_2, and pays 5 £ to player 2 in state ω_3.

What are the Bayesian Nash equilibria of this game?

5.2 No-Trade Theorem

The previous example shows that in a betting game where each player can refuse the bet, at equilibrium the sum of expected payoffs is the same as if betting never takes place. This result extends to a wide range of situations that can model betting, poker playing, or trading in the stock market.

The set of player I is arbitrary. Consider an arbitrary payoff structure $\mathscr{I}$, and a payoff structure G. The set of actions of player i is A_i the set of states of nature is K, and player i's payoff function is g_i. We make two assumptions on the payoff structure.

The first assumption is that *every player can opt out*: There exists a strategy $n_i \in A_i$ such that, for every $k \in K$ and $a_{-i} \in A_{-i}$ (A_{-i} represents $\prod_{j \neq i} A_j$), $g_i(n_i, a_{-i}, k) = 0$. By playing the strategy n_i, player i guarantees a payoff of 0 no matter k and the actions of the other players.

The second assumption is *no gains from trade*: For every k, a, $\sum_i g_i(a,k) \leq 0$. This assumption distinguishes the betting example from, for instance, the car sale example. According to this assumption, any gain by a player must correspond to an at least equivalent loss by the other players.

Theorem 5.1 (No Trade Theorem)**.** *If a Bayesian game is such that every player can opt out and there are no gains from trade, then, at every Bayesian Nash equilibrium, the expected payoff to every player is 0.*

Proof. Let f be a Bayesian Nash equilibrium, and let $u = (u_i)_i$ be the corresponding vector of expected payoffs. The no gains from trade condition implies that

$$\sum_i u_i = \sum_\omega p(\omega) \sum_i g_i((f_i(\omega))_i, k) \leq 0.$$

On the other hand, if a player i plays the constant strategy n_i (for every ω, play n_i at ω), then this player obtains 0 in the Bayesian game no matter what the other's strategies are. Therefore, the expected equilibrium payoff of each player i is at least 0:

$$\text{for every } i,\ u_i \geq 0.$$

The two conditions imply together that $u_i = 0$ for every i. □

6 Multi player case

This section provides insights on the value of information in the multi-player case. Since not all the material is covered in lectures, this is meant for your own reading and thinking if you're interested. Section 6.1 provides an example in which the value of information is negative, Section 6.2 shows that the value of information is always non-negative in a zero-sum game, and Section 6.3 offers to understand the value of information as the value of possessing more strategies.

6.1 General Games

We first define what a more informative information structure is, using the partition approach.

Definition 6.1. The information structure $\mathscr{I}$ is more informative than the information structure $\mathscr{J}$ for player i if $\mathscr{J}$ can be obtained from $\mathscr{I}$ by replacing i's information partition by a coarser information partition.

In general games, is a player better off when his information partition is finer?

Example 6.2. Consider an investment game. There are two investment opportunities, A and B. Only one of them is good, and the probabilities are 1/2 on each one being good. Two firms, I and II, can invest in either A, or B. Firm I invests, and firm II invests after observing firm I's choice. The payoff for a firm investing in the good project is 3 if the other firm doesn't invest in the good project. If both invest in the good project, the payoff to each firm is 1. A firm not investing in the good project gets a payoff of 0.

No information No firm is informed of which project is the good one. Solve the game. What are the equilibrium payoffs?

Firm I informed Assume firm I, but not firm II, is informed of which project is good. Solve the game. What are the equilibrium payoffs?

Secret information Assume that with probability p, firm I is not informed of which project is good. With the remaining probability $1-p$, firm I knows (with certainty) which project is good. Firm II doesn't know whether firm I is informed or not. Represent the information structure, and firm I's corresponding statistical experiment. Inside of this framework, is firm I better off when informed, or not informed?

6.2 Zero-Sum Games

A payoff structure with two players $1,2$ is *zero-sum* if for every state of nature k and action pairs a_1,a_2, we have $g_1(k,a_1,a_2)+g_2(k,a_1,a_2)=0$.

An information structure $\mathscr{I}$ and a zero-sum payoff structure G define a zero-sum game with incomplete information $G_{\mathscr{I}}$. We let $Val(\mathscr{I},G)$ denote the value of this game.

The following proposition shows that the value of information is always non-negative in zero-sum games.

Proposition 6.3. *Let $\mathscr{I}$ be more informative for player 1 than $\mathscr{J}$, or less informative for player 2, then, for every zero-sum payoff structure G:*

$$Val(\mathscr{I},G) \geq Val(\mathscr{J},G)$$

Proof. Application of the $\min\max$ theorem. □

6.3 More Information as More Strategies

Let $\mathscr{I}$ and $\mathscr{J}$ be two information structures such that player i is more informed in $\mathscr{I}$ than in $\mathscr{J}$, and let G be any payoff structure.

For $j \neq i$, player j's strategy sets are the same in $G_{\mathscr{I}}$ and in $G_{\mathscr{J}}$. For player i, any strategy in $G_{\mathscr{I}}$ can be identified to a strategy in $G_{\mathscr{J}}$.

Hence, more information implies more strategies.

Can we find a converse to this result?

Example 6.4. Consider the following games G and G'

	C	D
C	$3,3$	$-1,2$
D	$4,-1$	$0,0$

G

	C	D
C	$3,3,$	$-1,2$

G'

Player 1 has more strategies in G than in G'. Can we construct two information structures $\mathscr{I}$ and $\mathscr{J}$, and a payoff structure G, such that player 1 is more informed in $\mathscr{I}$ than in $\mathscr{J}$, $G_{\mathscr{I}}$ can be represented by G, and $G_{\mathscr{J}}$ can be represented by G'?

Hint: Use a continuum of states of nature. The state of nature is a "key" needed for player 1 to play the D strategy.

7 Repeated zero-sum games, optimal information revelation

In this section we study zero-sum repeated games and incomplete information on one side. The study is through examples, for the general theory, we refer the reader to the books:

- Repeated Games with Incomplete Information, by Aumann and Maschler (1995)
- A First Course on Zero-Sum Repeated Games, by Sorin (2002)
- Repeated Games, by Mertens, Sorin and Zamir (2015)

Consider two equally likely states of nature, $k \in \{k_1, k_2\}$. P1, the maximizer, is informed of the k. P2, the minimizer, isn't.

A game takes place in stages between P1 and P2. At the end of each stage, P2 observes P1's action in the previous stage (Not her own payoff, which is kept in a secret bank account). We study optimal information revelation by P1.

7.1 First example

Assume the payoff functions to P1 in states k_1 and k_2 are given by:

	L	R
T	1	0
B	0	0

k_1

	L	R
T	0	0
B	0	1

k_2

We analyse the situation through the following questions:

1. Assume the game with incomplete information is played only once. What are the strategies for P1? For P2?
2. What is the optimal strategy for P1 in the one-shot Bayesian game? What is the corresponding $\min\max$ payoff?
3. Assume now that P1 is using a completely revealing strategy: a strategy that plays T in state k_1, and B in state k_2. After observing action T, what is the posterior of P2 on the state of nature? Same question after observing B.
4. If P1 uses a completely revealing strategy and P2 plays a best response to it, what is the best payoff that P1 can expect in subsequent repetitions of the game? (in stage 2, stage 3, and so on).
5. Now consider non-revealing strategies of P1, which uses the same mixed strategy in states k_1 and k_2. What is the best non-revealing strategy for P1? What is P1's expected payoff in stage 1 if he uses this non-revealing strategy and P2 plays a best response? What about stage 2 onwards (P1 uses the same non-revealing strategy at every stage)?
6. Is it better for P1 to use a completely revealing strategy, or a non-revealing one?

7.2 Second example

Assume now the payoffs are given by:

k_1:

	L	R
T	-1	0
B	0	0

k_2:

	L	R
T	0	0
B	0	-1

Same questions as before.

7.3 Third example

Finally consider the payoff functions to P1 in states k_1 and k_2 given by:

k_1:

	L	M	R
T	4	0	2
B	4	0	-2

k_2:

	L	M	R
T	0	4	-2
B	0	4	2

- Answer the same questions as before.
- Consider the following partially revealing strategies of P1: If the state is k_1, with probability 3/4 play T forever, and with probability 1/4, play B forever. If the state is k_2, with probability 3/4 play B forever, and with probability 1/4, play T forever. Conditional on observing T in the first stage, what is P2's posterior belief on the states of nature? What is P2's best response to P1's strategy from stage 2 on? Same questions when P2 observes P1 playing B in the first stage.
- What expected average payoff per stage does the partially revealing strategy guarantees to P1 in the long-run?
- What is the best strategy for P1, a non revealing strategy, a completely revealing strategy, or a partially revealing strategy?

7.4 A common analysis of all 3 cases

For each of the payoff structures:

- Assume that the state of nature is k_1 with probability p and k_2 with probability $1-p$, and that P1 plays a non-revealing strategy.
- Graph the value $u(p)$ of the one-shot Bayesian game as a function of p. Draw the function $(\text{cav}\, u)(p)$, the smallest concave function that is larger or equal than u.

Comment on the optimal revelation of information.

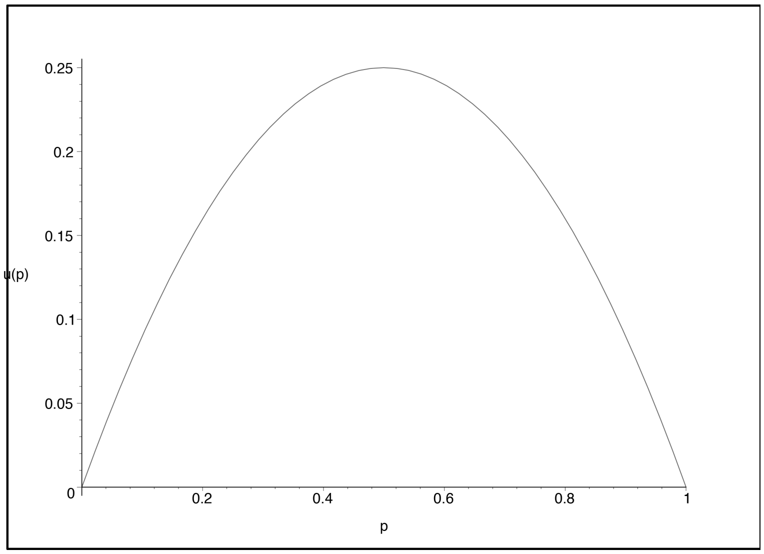

Figure 1: First example revisited

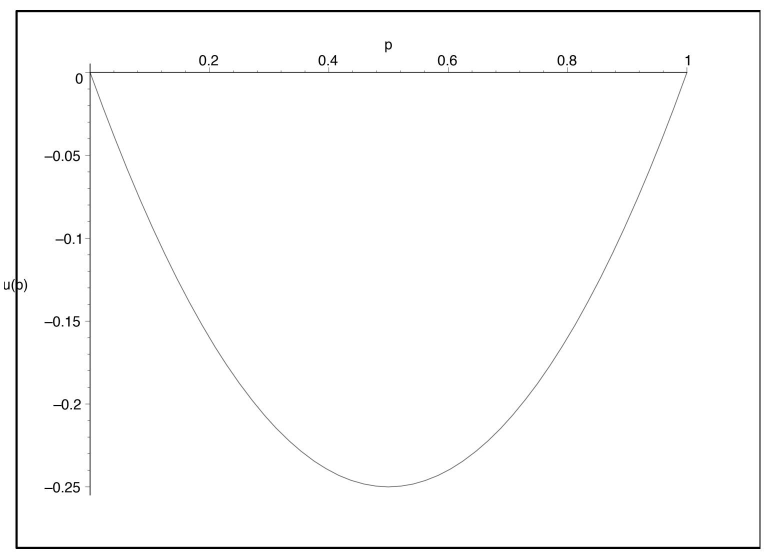

Figure 2: Second example revisited

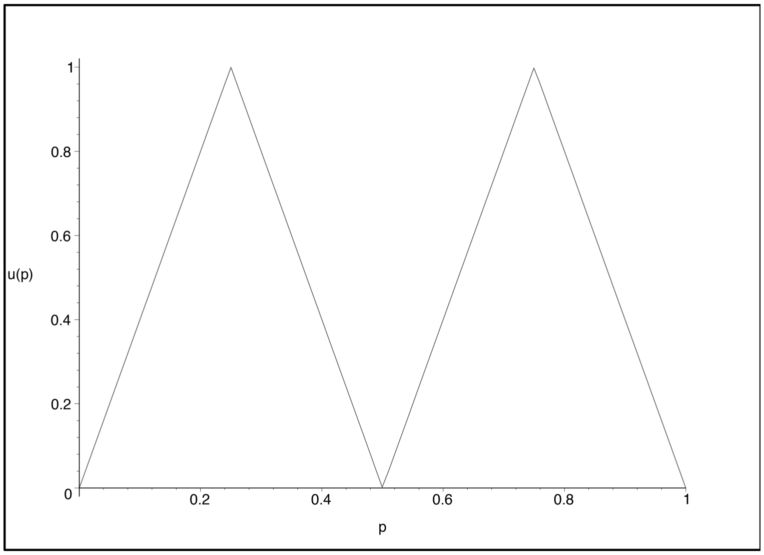

Figure 3: Third example revisited

The optimal information revelation corresponds to a *splitting* of the initial belief p into a random belief q with the constraint that $\mathbb{E}q = p$.