

Games with informational externalities

Dinah Rosenberg¹, Eilon Solan² and Nicolas Vieille³

¹Paris XIII

²Tel Aviv University

³HEC Paris

November 24-26th 2008,

Roscoff

- We are interested in discrete time dynamic games with incomplete information.
- Issues: acquire / transmit / hide information at equilibrium?
- In **Zero-sum repeated games with incomplete information** (Aumann Maschler), information disclosure is a by-product of exploitation of information. Dilemma exploitation/transmission
- *Generally agents may or may not want others to acquire pieces of information. Strategic transmission.*
- In **bandit problems**, dilemma between exproation/exploitation of information. Aquisition is from nature.
- *Generally aquisition is either from nature or from other strategic agents. Information may be an externality or a trading asset.*

- Simple games in terms of payoffs: no payoff-interaction, collection of one-agent problems. General signals.
- Payoffs depend only on the state and one's own action: no direct care about transmission.
- Information is the only punishment or reward, trading asset.
- No cheap talk, costly communication (discounted games)
- Issues
 - Learning of the state as a function of the information structure
 - Speed of learning, equilibrium payoffs, impact of observation
 - information transmission, costly communication, strategies

Three papers with E. Solan and N. Vieille:

- (i) long term learning in a general model with general information
- (ii) interaction between exploration of the state of nature - observation of others and exploitation in a bandit model
- (iii) possibility of exchanging information when communication is costly in a game without interaction and independent states in a specific repeated game model

Social learning in one arm bandit problems

Background: Bandit problems

One player dynamic allocation problem

Basic version:

- Two arms are given. The Safe yields a payoff of 0 and the Risky a random stream of iid payoffs, X_n , given a state of nature $\theta \in \{\underline{\theta}, \bar{\theta}\}$.
- θ initially drawn with probability p and unknown.
- At each state the DM chooses one of the two arms.
Maximize the expected discounted reward

Result: The optimal policy is to pull the risky arm until the conditional probability that $\theta = \bar{\theta}$ falls below some π_* .

Social learning in one arm bandit problems

On the impact of the observation of others on exploration/exploitation dilemma. Properties of equilibrium strategies.

- Collection of bandit problems with common θ and independent draws.
- $N(= 2)$ players face OABP with a *common* $\theta \in \{\bar{\theta}, \underline{\theta}\}$, with prior p_0 . Assume $\underline{\theta} < 0 < \bar{\theta}$.
- Decisions to switch to the safe arm are irreversible.
- Actions are public information (information from the other player), Payoffs are privately observed (direct further information through own action and nature).
- Remarks
 - Payoffs X_n^i and X_n^j are correlated: Player j 's decisions matter to i (only) since they contain information on θ .
 - No Common conditional probability, no state variable to serve as a posterior belief.

Social learning in one arm bandit problems

Definition

$\mathcal{F}_n^i = (X_1^i, \dots, X_n^i)$: private information of i Cutoff strategy: Processes information in a simple way: (i) use private information to compute a belief $p_n^i = P(\bar{\theta} | \mathcal{F}_n^i)$, (ii) use public information to compute a cutoff $\pi_n^i(\alpha)$, (iii) drop out if $p_n^i \leq \pi_n^i(\alpha)$

Theorem

Under some assumptions,

- *There is a symmetric equilibrium. All equilibria are in cutoff strategies.*
- *Qualitative features: The cutoff sequences $\pi_n^i(*)$ are non increasing . (+ others)*
- *When $N \rightarrow \infty$, cutoffs in stage 1 converge to p_* (indifference). In stage 2 cutoffs converge to 1 (resp. 0) if the fraction of players who dropped out in stage 1 is below (resp. above) some ρ .*

Social learning in one arm bandit problems

- Information is processed in a simple way: private information is compared to a cutoff that depends on public information.
- Cutoffs depend on public information in two ways: (i) if j drops out, switch from $\pi^i(*)$ to $\pi^i(k)$; (ii) If j does not drop out switch from $\pi_n^i(*)$ to $\pi_{n+1}^i(*)$.
- If player j is still active at time $n + 1$, this is good news. But the decision of i depends on the continuation expected payoffs i.e. also on future learning perspectives.
- Partial learning: as players stop after finite time if the machine is bad.
- in Large games deterministic learning process, full learning after one stage. A non negligible fraction of players drops in stage 1 (see Salomon).

Strategic information exchange

On the possibility to exchange information in equilibrium when communication is costly, there is no incentive to disclose except trade. Characterization of equilibrium payoffs

- Two players, with action sets A and B .
- Two sets of states: S and T .
- Payoff functions $u : S \times A \rightarrow \mathbf{R}$ and $v : T \times B \rightarrow \mathbf{R}$.
- Stage 0:
 - States (\mathbf{s}, \mathbf{t}) are realized.
 - Players receive signals $\mathbf{l} \in L$ and $\mathbf{m} \in M$ respectively: no further direct information.
- Stage $n \geq 1$: players choose a_n and b_n , which are publicly disclosed.
- Only actions are observed: strategic exchange of information.
- Discount factor $\delta < 1$.

Assumption

- Information a pair of signals: $L = L^S \times L^T$, $M = M^S \times M^T$.
- The triples $(\mathbf{s}, \mathbf{l}^s, \mathbf{m}^s)$ and $(\mathbf{t}, \mathbf{l}^t, \mathbf{m}^t)$ are independent.

Strategic information exchange

Basic example: each player faces an independent decision problem with two states and actions. Each player knows the other's state.

Can they improve upon the autarky profile?

- To do so, some information has to be transmitted.
- Since there is no payoff interaction and states are independent the only reason to reveal information is to trade it: using does not reveal and there is no other trading asset.
- But communication takes place through actions playing the myopically suboptimal action is necessary.
- The cost of playing a myopically suboptimal action must be compensated in the future by a better continuation payoff, ie by transmission of valuable information later.
- Never full revelation.

Strategic information exchange

Given $\pi \in \Delta(S)$, $u_*(\pi)$ is the myopically optimal payoff. u_* is the expected payoff of the autarky profile and u_{**} the expected payoff with joint information.

Definition

The information held by player 2 is (interim) **valuable** for player 1 if

$$\mathbf{E}_p[u_*(\tilde{\mathbf{p}}_1) | \mathbf{I}^S] > u_*(\mathbf{p}_1), \text{ with proba. } 1.$$

Conditional on \mathbf{I}^S , optimal payoff would be strictly higher if also knew \mathbf{m}^S . Interim notion. Game-dependent notion.

Theorem

Assume that the information of each player is valuable to the other.

Then the limit set of sequential equilibrium payoffs, as $\delta \rightarrow 1$, is the set $[u_, u_{**}] \times [v_*, v_{**}]$.*

- All information can be disclosed, with a **negligible delay**.
- The cost of revealing information is the loss incurred by playing sub-optimally. This is independent of the amount of information revealed.
- First ask the players to reveal their signal about their own state: incentives to do so.
- Then one player transmits information. The other transmits information + compensates the cost of suboptimal play, and so on....
- Indeed, each player can compute the other's conditional probability and therefore his optimal action and the cost of revelation.

Optimal experimentation and emergence of consensus in games with informational externalities

- Sequential Bayesian decision problems. Many identical agents without payoff interaction.
- General information about states and actions. Networks, observe actions, communication, direct information from nature....
- Social learning, consensus among players on the true information, or on the true optimal action as a function of information? Consensus is a weak form of learning.
- Question about reaching consensus at equilibrium but in the long run: no equilibrium payoff analysis.

A general model of informational interaction

- n players with no payoff interaction and **same** payoff function $u(\theta, a)$
- Very general signalling function : depends on all past and present actions of other players, all past signals of other players and the true state.
- Let q_n^i be the belief over Θ of payer i given his information at stage n . By martingale convergence, define the limit belief q_∞^i
- A limit action is played infinitely often.

Definition

We say that a player j observes another player i if he can identify a subset of i 's limit actions and i knows which limit actions j identifies. This defines a graph of observation.

A general model of informational interaction

Theorem

- *If σ is an equilibrium then for each player i , a.s. any limit point of the sequence of actions of i is a myopic best reply to q_∞^i , i.e. is an action a^* maximizes over a $\mathbf{E}_{q_\infty^i}[u(\cdot, a)]$*
- *If the graph of observation is connected, for any two players i and j , $\mathbf{E}[u(q_\infty^i)] = \mathbf{E}[u(q_\infty^j)]$.*
- *If player i observes player j then a.s. any of j 's limit actions is myopically optimal in i 's view ie maximizes in a the payoff $\mathbf{E}_{q_\infty^i}[u(\cdot, a)]$.*

A general model of informational interaction

- Any two players eventually perform equally well in expected terms (imitation).
- If in addition each player observes his own payoffs $u(q_\infty^i) = u(q_\infty^j)$
- This result fails for neighbors of neighbors (indifference between two actions).
- This is a weak learning result: the players do not share all information but agree on its consequences.
- Uses an auxiliary result that bounds the total amount of experimentation in one agent problems with general signals.
- no result on equilibrium payoffs.

- Learning in social networks
 - Players are organized on a graph and observe their neighbors' actions.
 - Each player eventually thinks that his neighbors play myopically optimally, and all players end up with the same expected payoffs (and the same payoff if each player observes his own payoffs).
 - However, a player may not play optimally in the eyes of another player who does not see him.
- Strategic experimentation.
 - Each player operates a bandit machine with the same types. Then iid payoffs.
 - Player i observes payoffs and actions // only the actions // and nothing .
 - If the players are organized along a connected graph and observe the actions of their neighbors, all players have the same expected payoff.
 - Hence use the same arm.