Reasoning about reasonable strategy sets for games of conventions
Presentation, ILLC July 2, 2008
Pelle Guldborg Hansen
Section for Philosophy and Science studies, Roskilde University Denmark

---

Lewis' theory of convention

Convention: *A behavioural regularity the conformity to which instantiates one out of* *<u>multiple strict coordination equilibria,</u>* *made salient by precedent and operational by this being common knowledge*

Game 1: *A pure coordination game*

player 2

		l	r
player 1	l	**1, 1**	0, 0
	r	0, 0	**1, 1**

Game 2: *The telephone tag game*

player 2

		c	w
player 1	c	0, 0	**1, 1**
	w	**1, 1**	0, 0

- reformulating action descriptions
- equilibrium selection and the problem of salience:

(a) underdetermination, (b) lack of reason to chose the salient

Learning and convention

Best reply models of learning

• *maximize expected payoff given expectations*

• *learning rules for forming these expectations given observations*

- <u>*stationarity assumption:*</u> *agents learn on the assumption that other agents are playing some stationary (possibly mixed) strategy*

In this model, an agent k updates his probability distribution $\mu_k^n(\cdot)$ according to the following formula: If one of m possible outcomes $A_1, ..., A_m$ can occur in each of n rounds, and n_{A_i} is the number of times that A_i is observed to occur in the n rounds of the game, then

$$\mu_k^n(A_i) = \frac{n_{A_i} + \gamma_{A_i}}{n + \sum_j \gamma_{A_j}}, \tag{1}$$

where $\gamma_{A_i} \geq 0$, $1 \leq j \leq m$ and $\sum_j \gamma_{A_j} > 0$. Notice that the value of γ_{A_i} reflects the strength of k's initial expectation that A_i will occur in any given round. Hence before any rounds have occurred, that is, when $n = 0$, the quotient $\gamma_{A_i} / \sum_j \gamma_{A_j}$ determines an agents *prior probability* that A_i occurs.

Learning and convention

Best reply models of learning

• *maximize expected payoff given expectations*

• *learning rules for forming these expectations given observations*

- *stationarity assumption:* *agent learn on the assumption that other agents are playing some stationary (possibly mixed) strategy*

• **Successes and failures**

• **Convention as correlated equilibrium**

$$f(\omega) = \begin{cases} (A_{k_1}, A_{i_2}) \text{ if } \omega = \omega_1 \\ (A_{k_2}, A_{i_1}) \text{ if } \omega = \omega_2 \end{cases}.$$

Presumptuous learning

stationarity assumption:

learning under the assumption that other agents are playing some stationary (possibly mixed) strategy

minimal universe:

unique and salient partition of W which is common knowledge

common inductive standards:

unique projections of past observations into the future

Reasoning about reasonable strategy sets for games of conventions
Presentation, ILLC July 2, 2008
Pelle Guldborg Hansen
Section for Philosophy and Science studies, Roskilde University Denmark

Framing theory

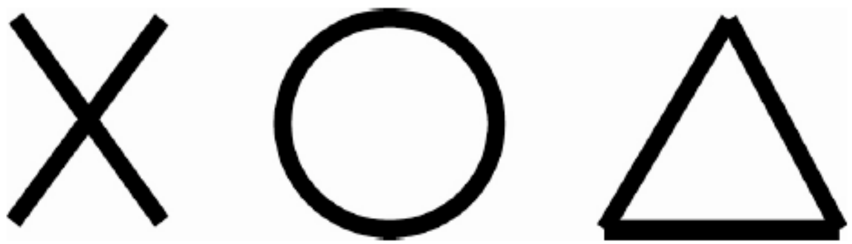

Figure 2: Three symbols

In general, let $S = \{x_1, ...x_n\}$ be a set of objects of choice and $P = \{\phi_1, ...\phi_n\}$ be a set of predicates. A frame $F \subseteq P$, then, is a set of predicates suitable for describing the objects in S, where if $\phi_i \in F$, then $E(\phi_i)$ is a function that denotes the *extension* of ϕ in S, that is the (possibly empty) subset of S which satisfy ϕ. We say that x and $y \in S$ are *F-equivalent*, written $x = Fy$, if

(1) $$y \in E(\phi) \quad \text{iff } x \in E(\phi)\ \forall \phi \in F$$

That is, equation (1) says that if x and y are F-equivalent, then the predicates of a frame do not suffice to discriminate x from y. In this way a frame F induces a partition, P_F, on S, whose cells are F-equivalence classes [Bacharach 2003, 63].