Outline of Part IV

- Games and equilibria
- Nash dynamics
- Fictitious play
- No-regret dynamics
- Trail and error learning

Learning by trials and errors (Young, 2008)

Algorithm

- **Marden et al.**: experiment rarely, and compare with the average payoff received over long periods. Adopt the new action when it leads to significantly better payoff.
  - Work for weakly acyclic games (convergence to NEs)

- New idea: experimentations triggered by decreases in payoffs
  - Convergence to NEs in games where a pure NE exists
  - Proof of convergence: uses Freidlin-Wentzell perturbation theory
  - Hereafter, synchronous moves

Algorithm

Idea: enrich the *state* of agent

$$x_i(t) = (\bar{s}_i(t), \bar{u}_i(t), m_i(t))$$

Baseline action

Baseline utility

Mood $\in \{C, D, H, W\}$

Algorithm: content

At the beginning of each time period t: if $m_i(t) = C$

- Play benchmark action w.p. 1-ε
  - If $u_i(a) > \overline{u_i}$, become hopeful
  - If $u_i(a) = \overline{u_i}$, be content
  - If $u_i(a) < \overline{u_i}$, become watchful

- Explore and play a_i randomly chosen
  - If $u_i(a) > \overline{u_i}$, adopt a_i and update your benchmarks
  - If $u_i(a) \leq \overline{u_i}$, don't change anything

Algorithm: watchful

At the beginning of each time period t: if $m_i(t) = W$, play benchmark action

- If $u_i(a) > \overline{u_i}$, become hopeful
- If $u_i(a) = \overline{u_i}$, be content
- If $u_i(a) < \overline{u_i}$, become discontent

Don’t change the benchmarks

Algorithm: hopeful

At the beginning of each time period t: if $m_i(t) = H$, play benchmark action

- If $u_i(a) > \overline{u_i}$, become content, update $\overline{u_i} = u_i(a)$
- If $u_i(a) = \overline{u_i}$, become content
- If $u_i(a) < \overline{u_i}$, become watchful

Algorithm: discontent

At the beginning of each time period t: if $m_i(t) = D$, play a random action a_i

- Become content; adopt the new action and update the benchmarks with probability $\phi(u_i(a), \overline{u_i})$
- Remain discontent with probability $1 - \phi(u_i(a), \overline{u_i})$

Convergence

- Assume that the game as at least one pure NE, and denote by $\Omega^\star$ the set of pure NEs.

Theorem *For any $\delta > 0$, there exists ϵ such that:*

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{i=0}^{t-1} 1_{\{s(i)\in\Omega^\star\}} \geq 1-\delta$$

Perturbed Markov chains

Idea from **Young**, *The evolution of conventions*, Econometrica 1993

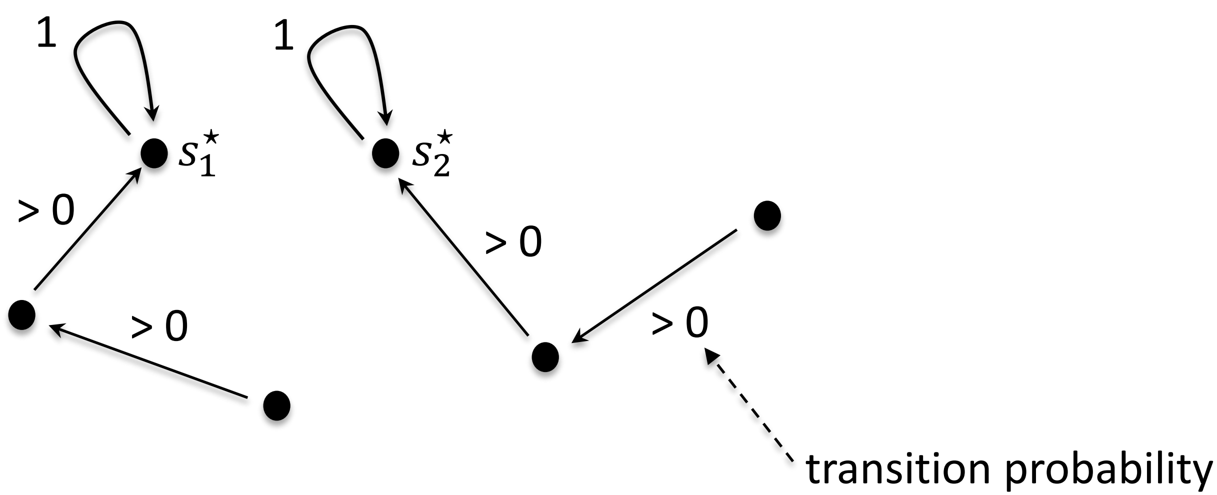
Step 1. Construct a Markov chain absorbed in states maximizing social welfare

Step 2. Perturb the Markov chain to achieve irreducibility

Step 3. Show that in steady-state, the perturbed Markov chain concentrates on socially optimal states

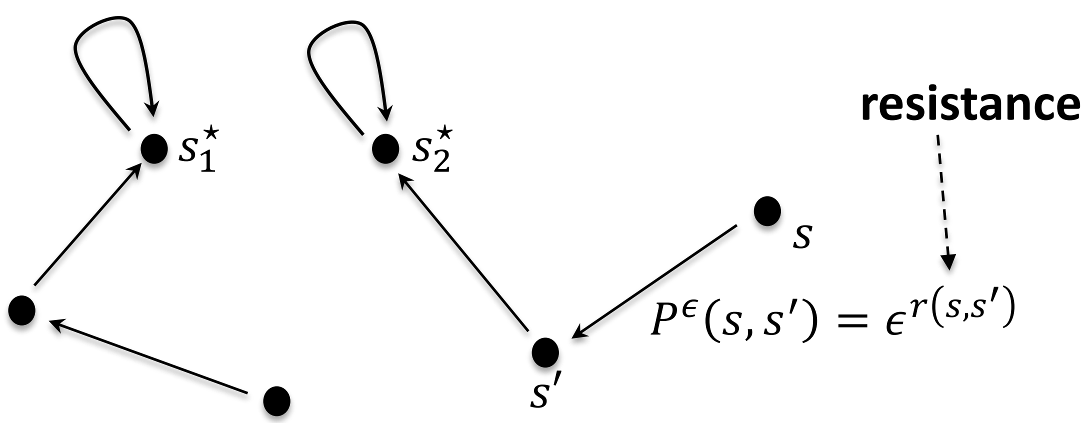
Transient Markov chain

Let $\Omega^\star$ be the set of socially optimal states.



Resistance, rooted trees, potential

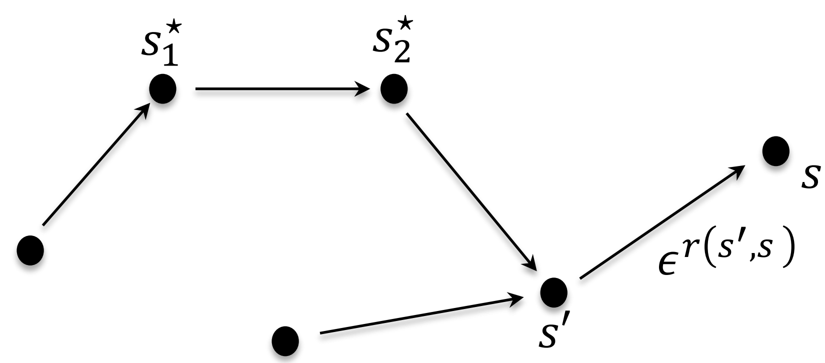
Step 2. Irreducible perturbed Markov chain



Resistance, rooted trees, potential

Step 2. Irreducible perturbed Markov chain

Tree rooted in *s*



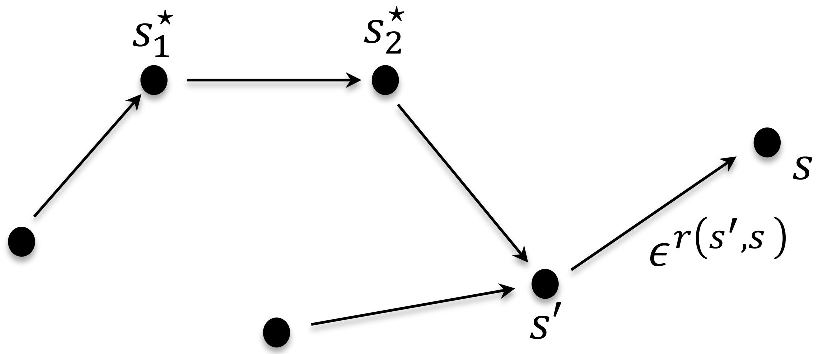
Steady-state distribution: $\pi^\epsilon(s) \sim \sum_{T \in \mathrm{Trees_s}} \epsilon^{\sum_{(s_1,s_2)\in T} r(s_1,s_2)}$

Potential of s: $\gamma(s) = \min_{T \in \mathrm{Trees_s}} \sum_{(s_1,s_2)\in T} r(s_1, s_2)$

Resistance, rooted trees, potential

Lemma *When $\epsilon \to 0$, π^{ϵ} concentrates on states with minimal potential.*

Tree rooted in *s*



Steady-state distribution: $$\pi^{\epsilon}(s) \sim \sum_{T \in \mathrm{Trees}_s} \epsilon^{\sum_{(s_1,s_2)\in T} r(s_1,s_2)}$$

Potential of s: $$\gamma(s) = \min_{T \in \mathrm{Trees}_s} \sum_{(s_1,s_2)\in T} r(s_1, s_2)$$