Computational Game Theory

Lecture 6

P2/2023

György Dán

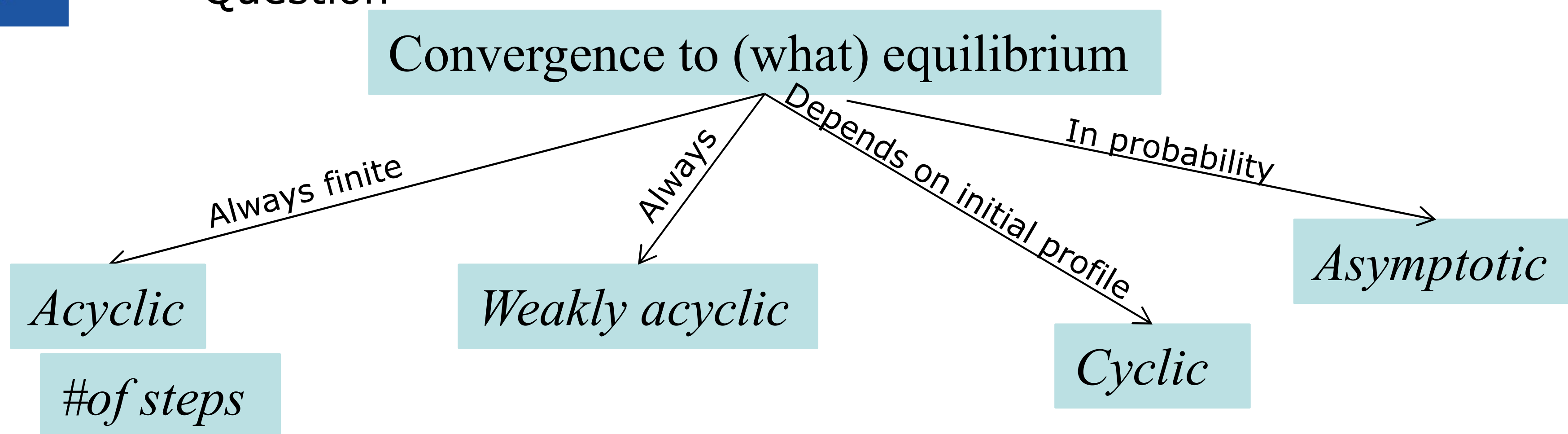
Division of Network and Systems Engineering

A bird's eye view of learning

- Strategic game G=<N,(A_i),(u_i)>
- Revision rule determines

$$a_i^{(k+1)} = f_i(a^{(0)},\ldots,a^{(k)})$$

- Sequence of strategy profiles
  - $a^{(0)}, a^{(1)},\ldots, a^{(k)}$
- Question



Taxonomy of Learning models

- Revision rule
  - Myopic learning
    - Best/better reply
    - Fictitious play
    - Adaptive play
    - Regret-minimization
    - Experimentation dynamics
  - Coordinated Bayesian learning
    - Rational learning

E. Kalai, E. Lehrer, "Rational learning leads to Nash equilibrium", Econometrica, 61(5), 1993

- Revision opportunity (Revision process)
  - Synchronous
  - Asynchronous
  - Independent
  - Plesiochronous/Conflict-free (on graphs)

Revision process

- Revision process is a probability measure q on the subsets of N, $P(N)=2^N$, such that
  $$\forall i \in N\ \exists J \subseteq N\ s.t.\ i \in J \wedge q_J > 0,$$
  where q_J is the probability that exactly players $i \in J$ receive a revision opportunity (independently across periods)
- Revising set for revision process q: $R^q=\{J \subseteq N | q_J>0\}$

- Special cases:
  - Asynchronous learning: $R^q=\{\{i\}|i \in N\}$
  - Synchronous learning: $R^q=\{N\}$
  - Independent learning: $R^q=\{P(N)\}$
  - Regular learning: $\{\{i\}|i \in N\} \subseteq R^q$

Best and Better Reply Dynamic

- Better reply dynamic

  $$\exists i \in N : u_i(a_i^{(k+1)}, a_{-i}^{(k)}) > u_i(a^{(k)}) \quad \text{and} \quad a_{-i}^{(k+1)} = a_{-i}^{(k)}$$

- Best reply dynamic

  $$\exists i \in N : a_i^{(k+1)} \in B_i(a_{-i}^{(k)}) \quad \text{and} \quad a_{-i}^{(k+1)} = a_{-i}^{(k)}$$

- Improvement graph
  - Digraph *G(A,E)*

    $$(a,a') = E \Leftrightarrow \exists i \in N : u_i(a_i', a_{-i}') > u_i(a) \quad \text{and} \quad a_{-i}' = a_{-i}$$

  - Sink: NE
- Best-improvement graph
  - Digraph *G(A,E)*

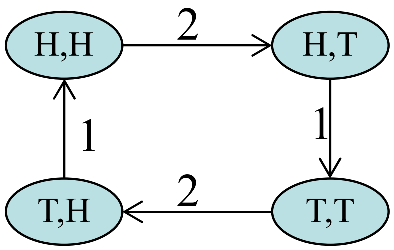
    $$(a,a') = E \Leftrightarrow a_i' \in B_i(a_{-i}), u_i(a') > u_i(a) \quad \text{and} \quad a_{-i}' = a_{-i}$$

  - Sink: NE

Examples

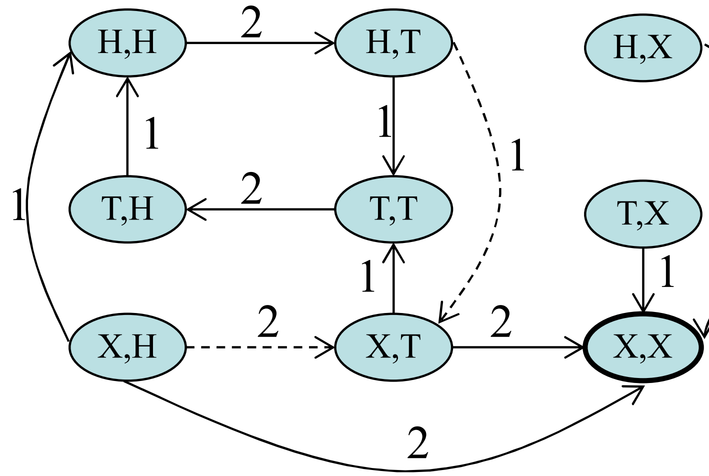
- Matching pennies

	H	T
H	2,0	0,2
T	0,2	2,0



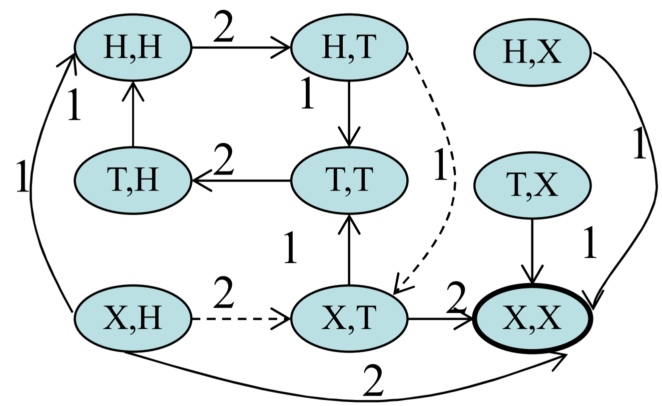
- Modified MP

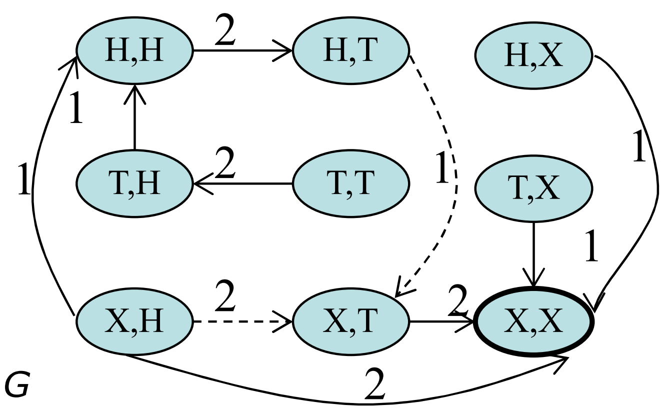
	H	T	X
H	2,0	0,2	0,0
T	0,2	2,0	0,0
X	0,0	1,0	3,3



Schedulers and x-Cyclicity

- Scheduler for game Γ
  - Subgraph G′(A,E′) of *G(A,E)*
    - $E' \subseteq E, s.t. (a,a') \in E \rightarrow \exists(a,a'') \in E'$
  - If E″⊆E′ then *G″* weaker than *G′*
- Default scheduler: G
- Path in scheduler *G(A,E)*
  - Sequence of edges *e*∈*E*
- Cyclicity
  - Scheduler *G*
    - Acyclic: no closed path in *G*
    - Weakly acyclic: ∃*G′* acyclic, weaker than *G*
  - Game Γ ⇔ default scheduler
    - I-acyclic ⇒ BI-acyclic ⇒ BI-weakly acyclic ⇒ I-weakly acyclic



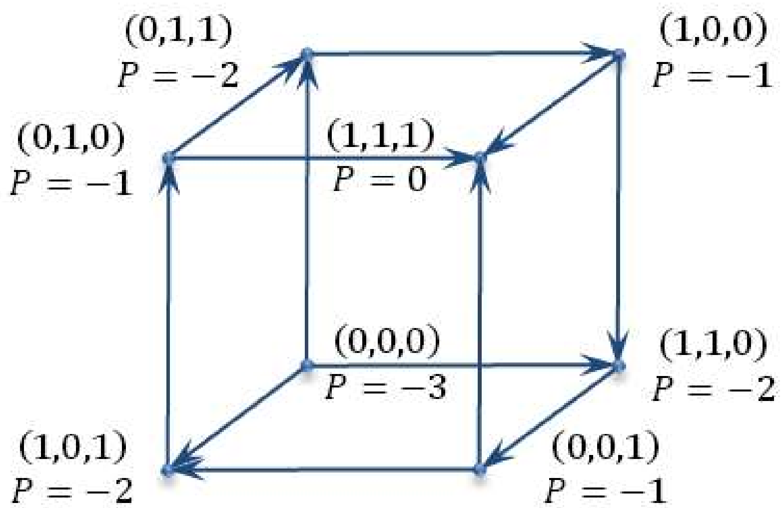


Milchtaich, "Schedulers, Potentials and Weak Potentials in Weakly Acyclic Games," Tech. Rep, May 2013
Apt, K.R. and Simon, S. "A classification of weakly acyclic games" Theory and Decision, v.78, 2015

Potential and Weak Potential

- Generalized ordinal potential Ψ for game Γ
  - $(a,a')\in E \Rightarrow \Psi(a')>\Psi(a)$
- Weak generalized ordinal potential Ψ for game Γ
  - $\exists G'$ weaker scheduler, s.t. $(a,a')\in E' \Rightarrow \Psi(a')>\Psi(a)$
- Construction of the (weak) potential

$$P(a) = -\max\{m \geq 0 \mid \exists \text{path of length } m \text{ starting at } a\}$$



I. Milchtaich, “Schedulers, Potentials and Weak Potentials in Weakly Acyclic Games,” Tech. Rep, May 2013

Sufficient Conditions for WA

- Game Γ is solvable by iterated elimination of never best response strategies (IENBS) if there is a sequence of games $\Gamma^{(0)}(=\Gamma),\Gamma^{(0)},\ldots,\Gamma^{(m)}$, s.t., all strategy profiles in $\Gamma^{(m)}$ are equilibria.
- Let Γ be a finite game. If Γ is solvable by IENBS then Γ is weakly BI-acyclic.
  - *Proof:* Assume Γ is solvable, consider $\Gamma^{(0)}(=\Gamma),\Gamma^{(0)},\ldots,\Gamma^{(m)}$

    Define height $h(a_i)=\max_l a_i \in \Gamma^{(l)}$

    Define height $h(a)$ as average height of strategies

    Claim: $\Psi(a)=h(a)$ is a weak potential for Γ

    if $h(a)=m \Rightarrow a$ is equilibrium

    if $h(a)<m \Rightarrow$ let $l_{ik}=a_{ik}=\min_i(h(a_i))$

    $a_{ik} \notin \Gamma(l_{ik}+1) \Rightarrow$ never best response

    pick a_{ij} best response, $a'=(a_{ij},a_{-i})$

    $h(a')>h(a)$

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

Sufficient Conditions for WA

- Subgame Γ'<$N,(A'_i),(u_i)$> of game Γ<$N,(A_i),(u_i)$>
  - $A'_i \subset A_i$ *for some* $i \in N$
- Let Γ be a finite game. If every subgame has a *unique* equilibrium then Γ is weakly BI-acyclic.

- Example: Matching pennies v2

	H	T	X
H	2,0	0,2	0,0
T	0,2	2,0	0,1
X	0,0	1,0	3,3

- Equilibria?
- (Best)-improvement graph cyclic/wa/acyclic?

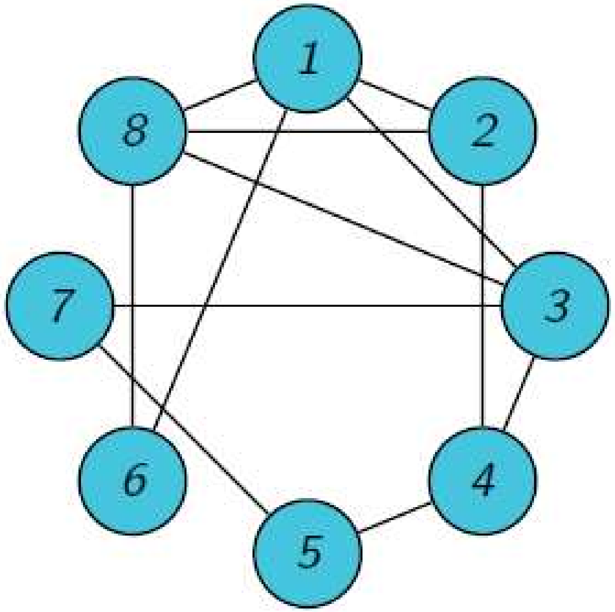
Fabrikant, A., Jaggard, A. D. and Schapira, M. “On the structure of weakly acyclic games,” LNCS 6386, 2010, pp.126–137

Example: Resource Allocation

- Digraph $G=(N,E)$
- Players N
  - $(i,j)\in E \Leftrightarrow (j,i)\in E$
  - $\delta_{ij}=\delta_i \ \forall (i,j)\in E$
- Action sets $A_i \subseteq \{a,b,c,d\}$
  - $|a_i|=K_i$
- Utility

$$U_i^r(a_i,a_{-i})|_{r\in a_i} = \begin{cases} c_{ir} & \text{if } \forall j\in \mathrm{N}(i) \quad r\notin a_j \\ \delta_i c_{ir} & \text{if } \exists j\in \mathrm{N}(i)\ s.t.\ r\in a_j \end{cases}$$

$$U_i(a)=\sum_r U_i^r(a)$$



V. Pacifici, G. Dán, "Convergence in Player-Specific Graphical Resource Allocation Games," IEEE JSAC, 30(11), 2012

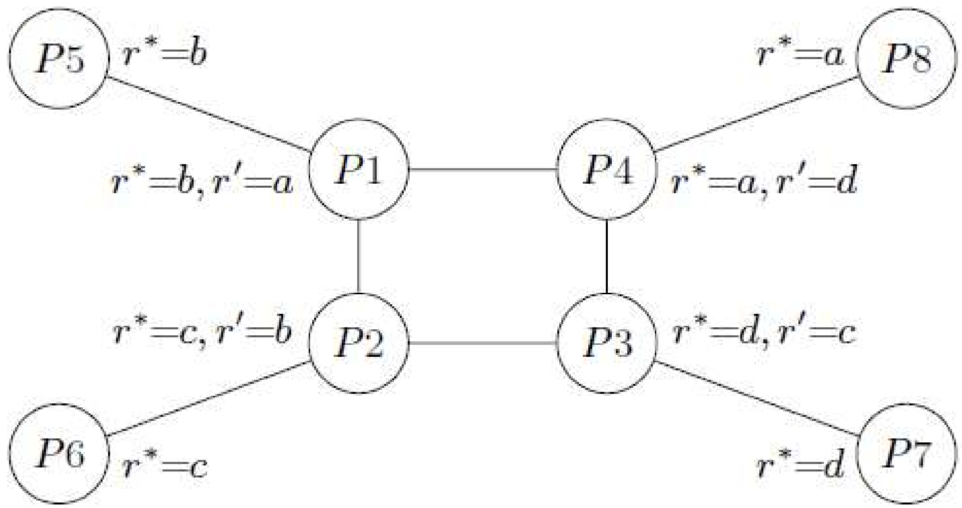
Resource Allocation

- G complete: BI-acyclic
- G non-complete: BI-weakly acyclic

$P1-P8: c_{ir*} > c_{ir}$

$P1-P4: \exists r' \; s.t. \; c_{ir'} > \delta_i c_{ir*}$

$P5-P8: \delta_i c_{ir*} > c_{ir}$



- Cycle: $(\boldsymbol{a}, b, \boldsymbol{d}, \boldsymbol{a}) \underset{3}{\rightarrow} (\boldsymbol{a}, b, c, \boldsymbol{a}) \underset{1}{\rightarrow} (\boldsymbol{b}, \boldsymbol{b}, c, \boldsymbol{a}) \underset{4}{\rightarrow} (\boldsymbol{b}, \boldsymbol{b}, c, d)$

$$\underset{2}{\rightarrow} (\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{c}, d) \underset{1}{\rightarrow} (a, \boldsymbol{c}, \boldsymbol{c}, d) \underset{3}{\rightarrow} (a, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{d})$$

$$\underset{2}{\rightarrow} (a, b, \boldsymbol{d}, \boldsymbol{d}) \underset{4}{\rightarrow} (\boldsymbol{a}, b, \boldsymbol{d}, \boldsymbol{a})$$

- NE: *(a,b,c,d)*

V. Pacifici, G. Dán, “Convergence in Player-Specific Graphical Resource Allocation Games,“ IEEE JSAC, 30(11), 2012

Synchronous Best and Better Reply

- Best reply - Synchronous

$$a_i^{(k+1)} \in B_i(a_{-i}^{(k)}) \quad \forall i \in N$$

- Convergence problematic
  - Coordination games
  - Anticoordination games

- Super/submodular games
  - A_i compact,
  - u_i upper/lower-semicontinuous on $A_i(a_{-i})$ $\forall a_{-i} \in A_i$
  - $a^{(0)}$ is least/greatest element of A

    $\Rightarrow$ convergence to equilibrium
  - Due to ascending/descending property of the best response correspondences $B_i(a_{-i})$

	L	R
U	1,1	0,0
D	0,0	1,1

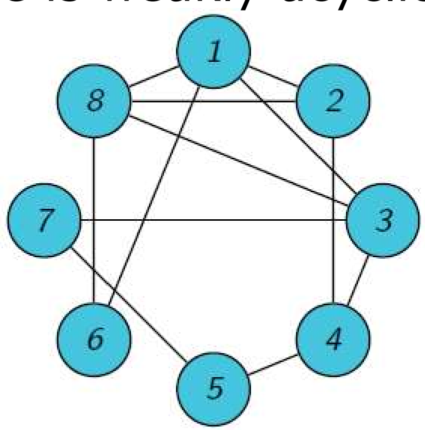
	L	R
U	0,0	1,1
D	1,1	0,0

	L	R
U	6,6	0,9
D	9,0	1,1

Regular Best and Better reply

- Let Γ be (B)I-weakly acyclic. Under a regular revision process (best)better-reply dynamic converges to equilibrium with probability 1.
  - Proof: If players play asynchronously long enough it converges.
    P(players play asynchronously long enough)>0

- Let Γ be (B)I-acyclic graphical game. Under a plesiochronous (conflict-free) revision process the game is weakly acyclic.



V. Pacifici, G. Dán, "Convergence in Player-Specific Graphical Resource Allocation Games," IEEE JSAC, 30(11), 2012
V. Pacifici, G.Dán, "Content-peering Dynamics of Autonomous Caches in a Content-centric Network," in Proc. of IEEE Infocom, Apr. 2013.

Fictitious Play - Discrete

- Infinite memory of past strategies
- Belief μ_i of player j's play
  - Empirical distribution of past strategies

$$\mu_i^t(a_j) = \sum_{\tau=1}^{t} I_{\{a_j(\tau)=a_j\}}$$

- Other players' play assumed independent

$$\mu_i^t(a_{-i}) = \prod_{j \neq i} \mu_i^t(a_j)$$

- Revision rule: best reply to belief

$$U_i(a_i, \mu_{-i}^t) = \sum_{a_{-i} \in A_{-i}} U_i(a_i, a_{-i}) \prod_{a_j \in a_{-i}} \mu_{-i}^t(a_j)$$

$$a_i^{t+1} \in B_i(\mu_{-i}^t)$$

G.W.Brown,"Iterative Solution of Games by Fictitious Play", Activity Analysis of Production, Wiley, New York, 1951
D. Monderer, L.S.Shapley, "Fictitious Play Property for Games with Identical Interests", Journal of Economic Theory, 68, pp.258-265,1996
U. Berger, "Brown's Original Fictitious Play", Journal of Ecomonic Theory 135(1), pp. 572-578, 2007

Asynchronous Fictitious Play

- 2 players
- If an AFP process for a non-degenerate game contains a switch from (a_{1i},a_{2j}) to (a_{1k},a_{2l}) then there is an improvement path from (a_{1i},a_{2j}) to (a_{1k},a_{2l})
  - Proof: Switch $i \rightarrow k$ $(A\mu_2^{t-1})_i - (A\mu_2^{t-1})_k \geq 0$ and $(A\mu_2^{t})_i - (A\mu_2^{t})_k \leq 0$

    But: $$\mu_2^t(a_{2j}) = \frac{t-1}{t}\mu_2^{t-1}(a_{2j}) + \frac{1}{t} I_{\{a_2(t)=a_{2j}\}}$$

    thus $(A)_{ij} - (A)_{kj} \leq 0$

- Consequence: Every AFP process converges to a pure NE in a non-degenerate 2 player ordinal potential game.

G.W.Brown,"Iterative Solution of Games by Fictitious Play", Activity Analysis of Production, Wiley, New York, 1951
D. Monderer, L.S.Shapley, "Fictitious Play Property for Games with Identical Interests", Journal of Economic Theory, 68, pp.258-265,1996
U. Berger, "Brown's Original Fictitious Play", Journal of Ecomonic Theory 135(1), pp. 572-578, 2007

Synchronous Fictitious Play

- Example

	NC	C
NC	3,3	0,4
C	4,0	1,1

  - Equilibrium?
  - Does SFP converge to NE? In what sense?
- Convergence:
  - Zero-sum games
  - 2x2 games
  - Games with identical interests (weighted potential)
  - Dominance solvable games

G.W.Brown,“Iterative Solution of Games by Fictitious Play”, Activity Analysis of Production, Wiley, New York, 1951
D. Monderer, L.S.Shapley, “Fictitious Play Property for Games with Identical Interests”, Journal of Economic Theory, 68, pp.258-265,1996
U. Berger, “Brown’s Original Fictitious Play”, Journal of Ecomonic Theory 135(1), pp. 572-578, 2007

Fictitious Play - Observations

- Example

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

  - Equilibria? Convergence?

- Pure Equilibrium FP: If $a^t=a$ for $t \geq T$ then a is a pure NE of G.
- Mixed Equilibrium FP: If $\lim_{t\to\infty} \mu^t = \mu$ and μ is a mixed NE of G then FP converges to equilibrium.
- FP converges in beliefs to equilibrium $\Leftrightarrow$ it converges in beliefs to equilibrium in the Cesaro mean.

$$\lim_{T\to\infty} \frac{\#\{1 \leq t \leq T : \mu(t) \notin B_\delta(\mathrm{M})\}}{T} = 0 \quad \forall \delta > 0$$

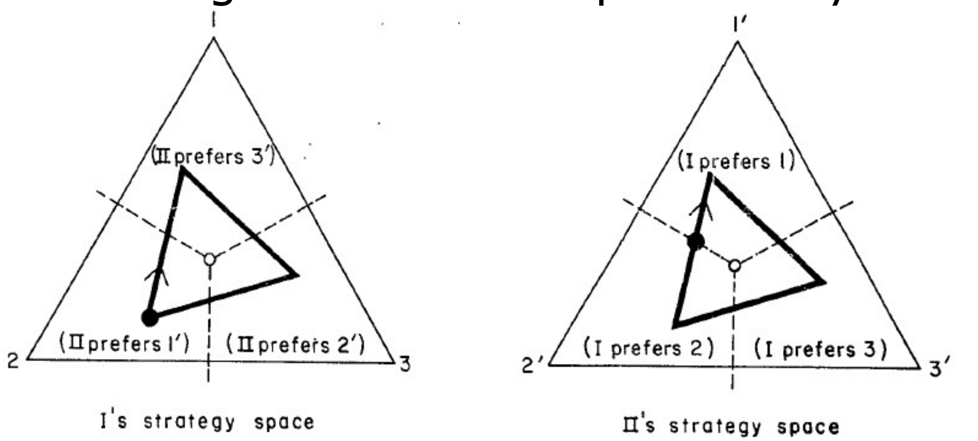
D. Monderer, L.S.Shapley, "Fictitious Play Property for Games with Identical Interests", Journal of Economic Theory, 68, pp.258-265,1996
D. Monderer, A. Sela, "Fictitious Play and No-Cycling Conditions", mimeo, 1992

Fictitious Play - Cycling

- Generalized RSP
- NE?

	L	M	R
T	1,0	0,0	0,1
C	0,1	1,0	1,0
D	0,0	0,1	1,0

- Synchronous FP
- Cycle if starting from (T,M)
  - Run lengths increase exponentially



Shapley L. (1964) "Some Topics in Two-Person Games" in Advances in Game Theory M. Dresher, L.S. Shapley, and A.W. Tucker (Eds.), Princeton University Press

Fictitious Play - Continuous

- Infinite memory of past strategies
- Belief μ_i of player i's play
  - Empirical distribution of past strategies
    - $\mu_i{}^t(a_j) = \int_{s=0}^{t} I_{\{a_j(s)\}} ds \;\; t \geq t_1$
- Other players' play assumed independent

$$\mu_i^t(a_{-i}) = \prod_{j \neq i} \mu_i^t(a_j)$$

- Revision rule: Drift towards an element of best response

$$\frac{d\alpha_i^t}{dt} \in B_i(\mu_{-i}^t) - \alpha_i^t$$

- If $\lim_{t\to\infty} \mu^t = \mu$ and μ is a mixed NE of *G* then CFP converges to equilibrium
- Convergence
  - 2x3 games (non-degenerate)

D. Monderer, A. Sela, "Fictitious Play and No-Cycling Conditions", mimeo, 1992

Joint Strategy Fictitious Play

- Infinite memory of past strategies
- Belief μ_i of other players' play
  - Empirical distribution of past strategies

$$\mu_i^t(a_{-i}) = \sum_{\tau=1}^{t} I_{\{a_{-i}(\tau)=a_{-i}\}}$$

- Other players' play assumed correlated

$$\mu_i^t(a_{-i}) \neq \prod_{j \neq i} \mu_i^t(a_j)$$

- Revision: best reply to belief

$$a_i^{t+1} \in B_i(\mu_{-i}^t)$$

$$U_i(a_i, \mu_{-i}^t) = \sum_{a_{-i} \in A_{-i}} U_i(a_i, a_{-i}) \mu_i^t(a_{-i}) = \frac{t-1}{t} U_i(a_i, \mu_{-i}^{t-1}) + \frac{1}{t} U_i(a_i, a_{-i}^{t-1})$$

D. Monderer, A. Sela, "Fictitious Play and No-Cycling Conditions", mimeo, 1992

Observation

- Let Γ be finite n-person game. If for $t>0$ the action profile a^t generated by JSFP is a strict NE then $a^{t+\tau}=a^t$, $\tau>0$
  - Proof:

$$\left.\begin{aligned} &U_i(a_i^t,\mu_{-i}^t)\geq U_i(a_i',\mu_{-i}^t) \quad \forall i,a_i' \\ &U_i(a_i^t,a_{-i}^t)> U_i(a_i',a_{-i}^t) \quad \forall i,a_i'\neq a_i^t \end{aligned}\right]$$

$$\begin{aligned} U_i(a_i^t,\mu_{-i}^{t+1}) &= \frac{t}{t+1}U_i(a_i^t,\mu_{-i}^t)+\frac{1}{t+1}U_i(a_i^t,a_{-i}^t) \\ &> \frac{t}{t+1}U_i(a_i',\mu_{-i}^t)+\frac{1}{t+1}U_i(a_i',a_{-i}^t) \end{aligned}$$

- Example:
  - $N=\{1,2,3\}$
  - $A_i=\{a,b\}$
  - Under FP: $a^0=(a,A,\alpha)\text{->}(b,B,\beta)\text{->}(a,B,\alpha)\text{->}\ldots$
  - Under JFP: $a^0=(a,A,\alpha)\text{->}(b,B,\beta)$

α	A	B
a	0,0,0	1,1,1
b	1,1,1	0,0,0

β	A	B
a	1,1,1	-1,-1,-1
b	-100,-100,-100	1,1,1

J.R. Marden, G. Arslan, J. Shamma, “Joint strategy fictitious play with inertia for potential games,” IEEE TAC, 54(2), 2009

Adaptive Play

- Finite memory *m* of past strategies with sample size *S*
  - Belief based on sample

$$\mu_i^k(a_{jl}) = \sum_{\tau=k_1}^{k_s} I_{\{a_j(\tau)=a_{jl}\}} \qquad \mu_i^k(a_{-i}) = \prod_{a_{jl}\in a_{-i}} \mu_i^k(a_{jl})$$

$$j_s \in [k-m,k], \quad 1 \le s \le S, \quad S \le m$$

- Revision rule
  - Play best reply to belief with probability *1-*ε

$$a_i^{(k+1)} \in B_i(\mu_i^k)$$

  - Play random strategy with probability ε
- Transitions for $\epsilon = 0$ described by Markov chain P^0
  - Strict equilibrium played *m* times= absorbing state
- Best-reply graph Γ weakly acyclic $\Rightarrow$
  if $s \le m/(\max_a L(a)+2)$ then AP(ε=0) converges to strict NE w.p.1
  - *L(a)* is length of shortest path from *a* to a strict NE in the best reply graph *G*

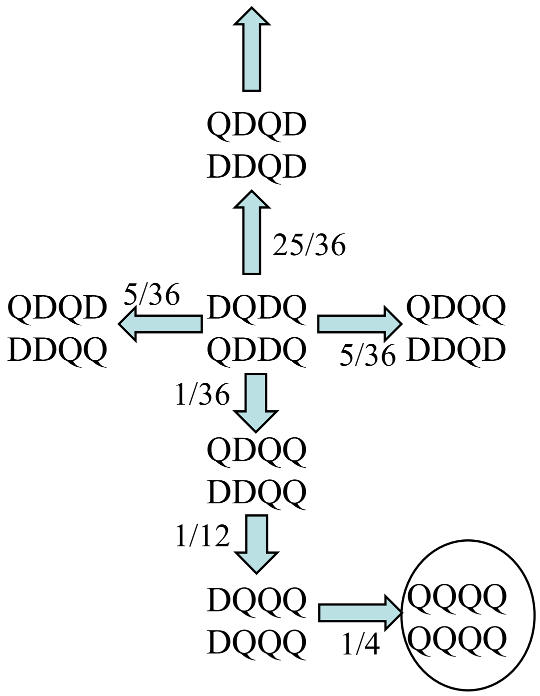
H. P. Young, "The Evolution of Conventions", Econometrica 61(1), pp. 57-84, 1993

Example: Adaptive Play

- Typewriter game

	D	Q
D	5,5	0,0
Q	0,0	4,4

- Let $m=4, S=2, \varepsilon=0$
  - $Sample=\{QQ\} \Rightarrow$ *best reply* $= Q$
  - $Sample \neq \{QQ\} \Rightarrow$ *best reply* $= D$
- Absorbing states
  - {DDDD}
  - {QQQQ}



H. P. Young, "The Evolution of Conventions", Econometrica 61(1), pp. 57-84, 1993
H.P. Young, "Individual strategy and social structure," Princeton Univ. Press, 2001

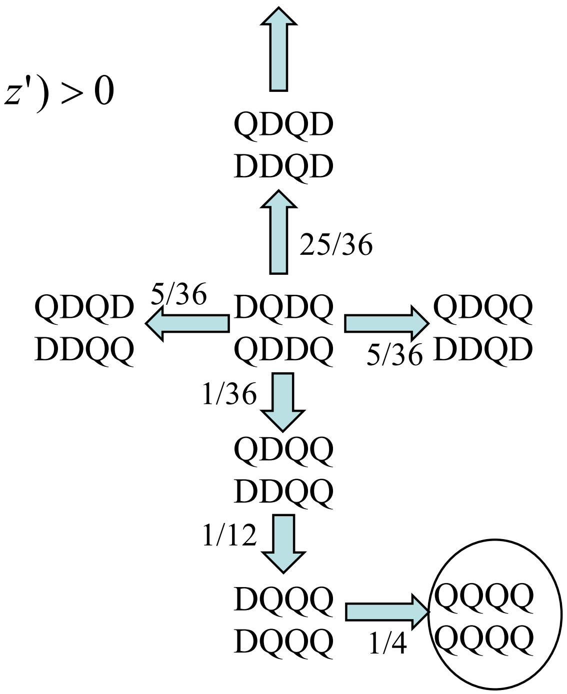
Perturbed Markov Process

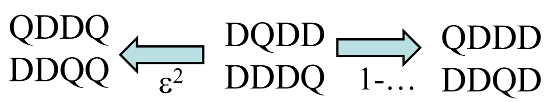
- P^ε is a regular perturbed Markov process for P^0 if P^ε is irreducible for every $\varepsilon\in(0,\varepsilon^*]$, and

$$\lim_{\varepsilon\to 0} P^{\varepsilon}_{zz'} = P^{0}_{zz'} \quad \forall z,z'\in Z$$

$$P^{\varepsilon}_{zz'} > 0 \Rightarrow 0 < \lim_{\varepsilon\to 0} P^{\varepsilon}_{zz'} / \varepsilon^{r(z,z')} < \infty, \quad r(z,z') > 0$$

- Resistance: $r(z,z')\geq 0$
- P^ε irreducible $\Rightarrow$ μ^ε exists
  - Stationary distribution
- Example:
  - Typewriter: $0\leq r(z,z')\leq 2$
  - General: #errors



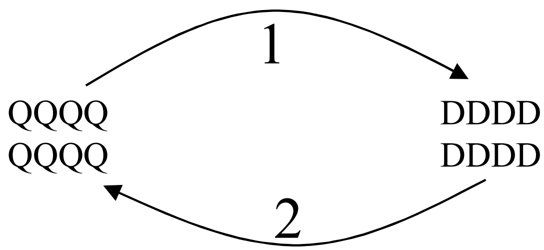


H. P. Young, "The Evolution of Conventions", Econometrica 61(1), pp. 57-84, 1993
H.P. Young, "Individual strategy and social structure," Princeton Univ. Press, 2001

Stochastic Potential

- State z' accessible from state z if $P(z,z')^k>0$ for some $k>0$
- States z' and z communicate if mutually accessible
  - Communication class
- Recurrence class E_i: communication class s.t. no state $z \notin E_i$ is accessible
- ij-path: $\gamma=(z_1, \ldots, z_n)$, $z_1 \in E_i$, $z_n \in E_j$
- Resistance r
  - ij-path: $r(\xi)=r(z_1,z_2)+\ldots+r(z_{n-1},z_n)$
  - $r_{ij}=\min_\xi r(\xi)>0$
- Digraph of recurrence classes
  - Rooted trees T_{it} for E_i
    - Resistance $r(T_{it})$
- *Stochastic potential of* E_i

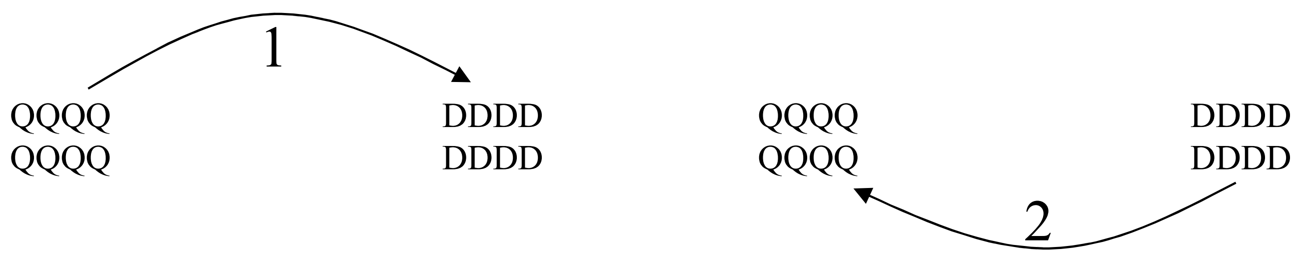
$$\gamma(i)=\min_t r(T_{it})$$



Stochastic Stability

- Let P^{ε} be a regular perturbed Markov process for P^0, and let μ^{ε} be the unique stationary distribution of P^{ε} for each $\varepsilon>0$. Then
  - $\exists \lim_{\varepsilon \to 0} \mu^{\varepsilon} = \mu^0$
  - μ^0 is a stationary distribution of P^0.
  - $z \in E_i$ s.t. $\gamma(i)=\min_i \gamma(i) \Rightarrow z$ stochastically stable

- Rooted trees for the Typewriter game



  - DDDD is stochastically stable

Logit-response Dynamic

- Logit choice function

$$p(a_i, a_{-i}) = \frac{e^{\beta u_i(a_i, a_{-i})}}{\sum_{a_i' \in A_i} e^{\beta u_i(a_i', a_{-i})}} \quad \beta > 0$$

- Let Γ be a potential game with potential $\Psi(a)$. Under asynchronous learning (with prob q_i) the invariant distribution is (Gibbs-Boltzmann)

$$\mu^\beta(a) = \frac{e^{\beta \Psi(a)}}{\sum_{a' \in A} e^{\beta \Psi(a')}}$$

  - Proof: it satisfies the detailed balance equations

$$\mu^\beta(a) P^\beta_{a,a'} = \frac{e^{\beta \psi(a)}}{\sum_{a'' \in A} e^{\beta \psi(a'')}} q_i \frac{e^{\beta \psi(a_i', a_{-i})}}{\sum_{a_i'' \in A_i} e^{\beta \psi(a_i'', a_{-i})}} = \mu^\beta(a') P^\beta_{a',a}$$

C.A. Ferrer, N. Netzer, “The logit-response dynamics,” Games and Economic Behavior, 68(2), 2010, pp. 413-427

Logit-response dynamics

- *Asynchronous learning:* Let Γ be a potential game with potential $\Psi(a)$. The set of stochastically stable states of the logit-response dynamics is equal to the set of maximizers of Ψ.
  - Proof: Limit of previous result for $\beta \to \infty$

- *Synchronous regular learning:* Let Γ be BI-weakly acyclic. The stochastically stable states are a subset of the NE.

- *Independent learning:* Let Γ be a potential game with potential $\Psi(a)$. Let player *i* use logit-response with probability $\omega = e^{-\beta m} > 0$. For large *m* the stochastically stable states are the maximizers of Ψ.

C.A. Ferrer, N. Netzer, "The logit-response dynamics," Games and Economic Behavior, 68(2), 2010, pp. 413-427
J.R. Marden, J. Shamma, "Revisiting Log-Linear Learning: Asynchrony, Completeness and Payoff-Based Implementation," Games and Economic Behavior, 75(2), 2012

Regret-matching

- Regret of player i for action a_i

$$R_i^{a_i}(t) = \left[\frac{1}{t} \sum_{\tau=0}^{t-1} (U_i(a_i, a_{-i}^{\tau}) - U_i(a^{\tau})) \right]^{+}$$

- Regret matching

$$P(a_i^t = a_i) = \frac{R_i^{a_i}(t)}{\sum_{a_j \in A_i} R_i^{a_j}(t)}$$

- If every player follows regret matching then the empirical distribution of play μ_t converges almost surely as $t \to \infty$ to the set of correlated equilibrium distributions of the game.

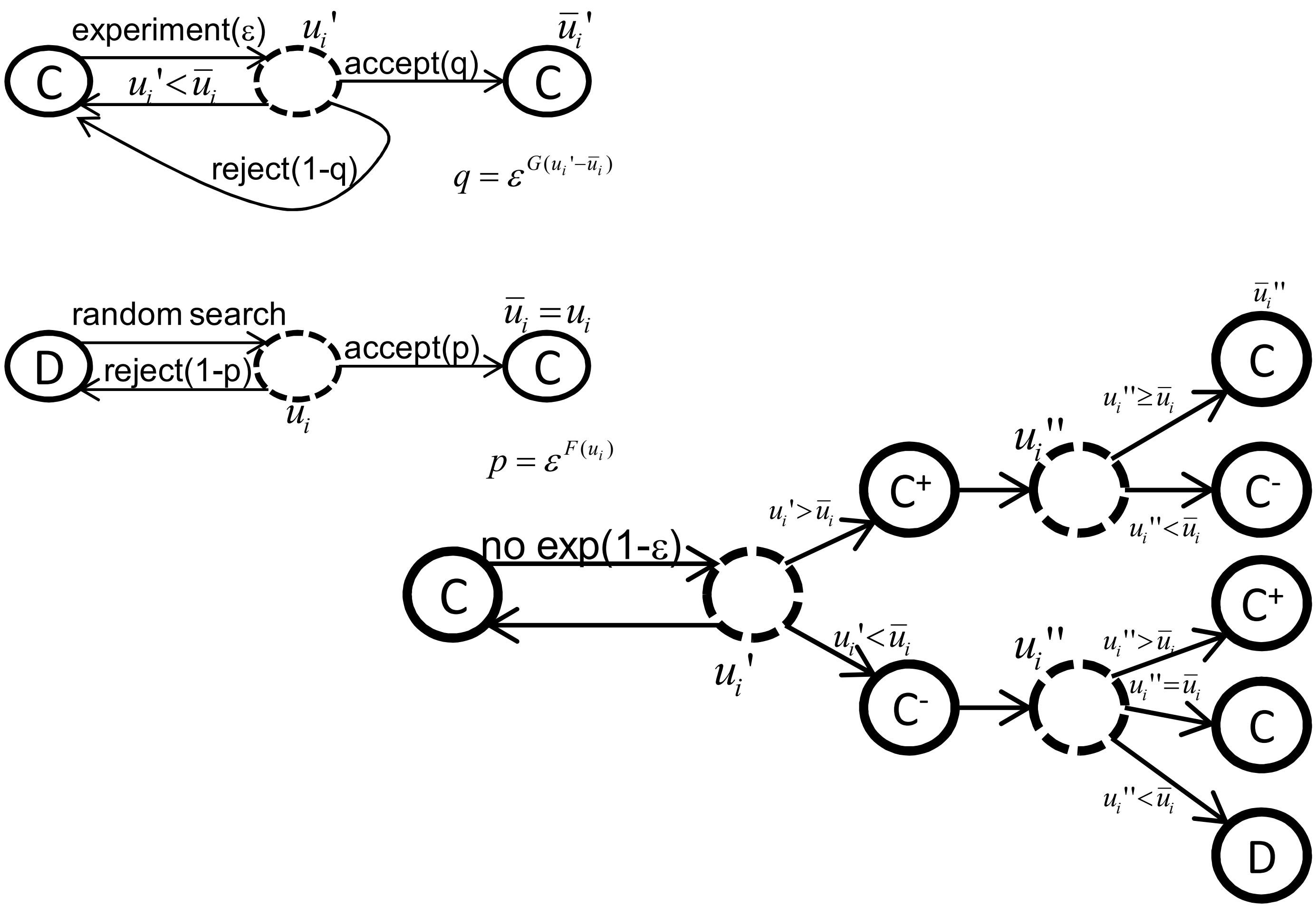
S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," Econometrica, vol. 68, no. 5, pp. 1127–1150, 2000.

Experimentation Dynamics

- State of player $z_i = (m_i, \bar{a}_i, \bar{u}_i)$
  - m_i → mood
  - $\bar{u}_i$ → benchmark payoff
  - $\bar{a}_i$ → benchmark action
- Moods
  - **c**ontent, **d**iscontent, watchful (c^-), hopeful (c^+)
- Experimentation probability ε
- Acceptance functions
  - $F(u)=-f_1*u+f_2$, $f_1>0$, $0<F(u)<1/2n$
  - $G(\Delta u)=-g_1*\Delta u+g_2$, $g_1>0$, $0<G(\Delta u)<1/2$

B.S.R. Pradelski, H.P.Young, "Learning efficient equilibria in distributed systems," Games and Economic Behavior, 75(2), 2012

State transitions



B.S.R. Pradelski, H.P.Young, "Learning efficient equilibria in distributed systems," Games and Economic Behavior, 75(2), 2012

Convergence

- Welfare of state: $W(z)=\sum_{i=1}^{n} u_i(\overline{a})$
- ε-equilibrium:

$$u_i(a_i,\overline{a}_{-i})-u_i(\overline{a})\leq\varepsilon \quad \forall i,\forall a_i\in A_i$$

- Instability of state z: $S(z)=\min\{\varepsilon : z \text{ is a } \varepsilon \text{ equilibrium}\}$
- Interdependence:

  Game $\Gamma=\langle N,(A_i),(u_i)\rangle$ is interdependent if for every $a\in A$ and every $\emptyset\subset J\subset N$ there exists some player $i\notin J$ and an action profile a_J' such that $u_i(a_J',a_{N\setminus J})\neq u_i(a_J,a_{N\setminus J})$.

- Let Γ be interdependent game and log-linear trial and error learning with experimentation probability ε.
  - *If there is a pure NE then every stochastically stable state is a NE that maximizes W(z)*
  - *If there is no pure NE then every stochastically stable state maximizes* $f_1W(z)-g_1S(z)$

B.S.R. Pradelski, H.P.Young, "Learning efficient equilibria in distributed systems," Games and Economic Behavior, 75(2), 2012

Example

- Coordination game

	L	R
U	a,a	c,d
D	d,c	b,b

	L	R
U	6,6	4,1
D	1,4	8,8

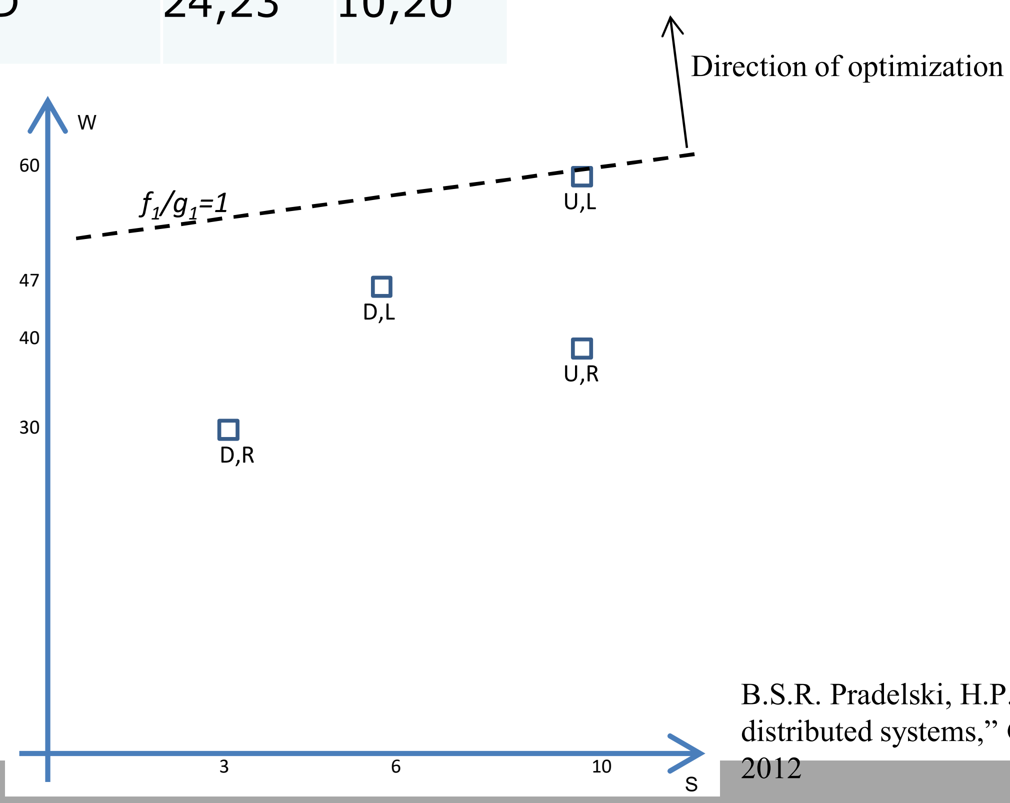
- Assume *a-d > b–c > 0, and b>a*
  - *(D,R)* Pareto optimal
  - *(U,L)* risk dominant
- *(D,R)* selected most of the time as $\varepsilon \rightarrow 0$

B.S.R. Pradelski, H.P.Young, "Learning efficient equilibria in distributed systems," Games and Economic Behavior, 75(2), 2012

Example – No NE

- Game w/o pure NE

	L	R
U	30,30	0,40
D	24,23	10,20



B.S.R. Pradelski, H.P.Young, "Learning efficient equilibria in distributed systems," Games and Economic Behavior, 75(2), 2012

Literature

- M.Osborne, A Rubinstein, “A course in game theory”, MIT press, 1994
- D. Fudenberg, J. Tirole, “Game Theory”, MIT press, 1991

- D.M. Topkis, “Equilibrium points in nonzero-sum n-person submodular games”, SIAM J. Control and Optimization 17(6), pp.773-787, 1979.
- D.M. Topkis, “Supermodularity and complementarity”, Princeton Univ. Press., 1998
- G.W.Brown,“Iterative Solution of Games by Fictitious Play”, Activity Analysis of Production, Wiley, New York, 1951
- D. Monderer, L.S.Shapley, “Fictitious Play Property for Games with Identical Interests”, Journal of Economic Theory, 68, pp.258-265,1996
- U. Berger, “Brown’s Original Fictitious Play”, Journal of Ecomonic Theory 135(1), pp. 572-578, 2007
- H. P. Young, “The Evolution of Conventions”, Econometrica, vol. 61., num. 1., pp. 57-84, 1993