Computational Game Theory

Lecture 7

P2/2023

György Dán

Division of Network and Systems Engineering

Dynamic games

- Players make decisions at different points in time
- Extensive game
  - Players make decisions one by one (approx)
  - Can learn about the environment and others’ choices

- Repeated game
  - Players play multiple strategic games
  - Decision is influenced by the history
  - Extension of extensive game

- Other forms of dynamic games
  - Stochastic game
  - Differential game

Extensive game w. perfect inf.

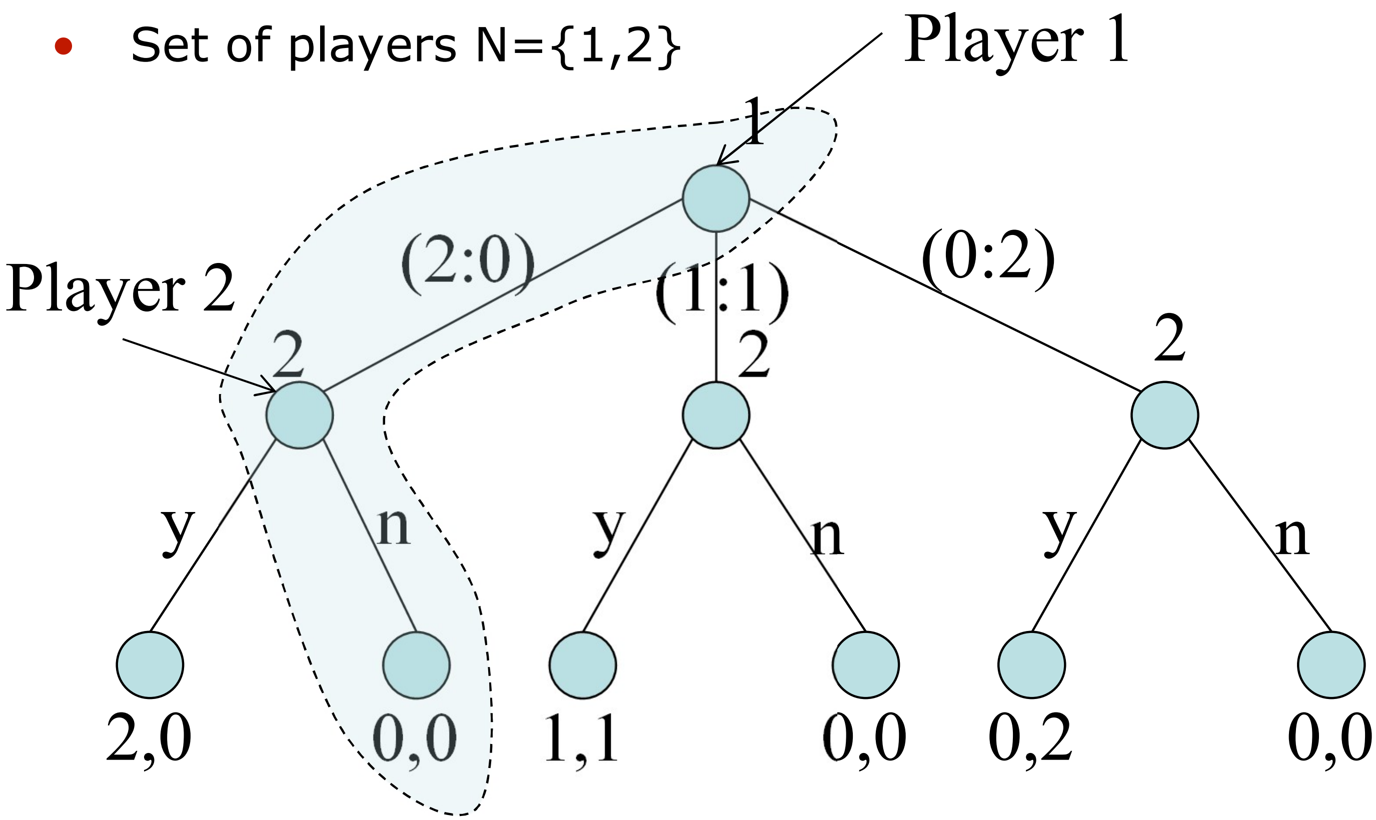
- A set of players *N*
- A set of actions for each player *A*
- A predefined sequence of choosing actions
  - Previous choices are known to all players

- Sequence *h* of actions called history
  - $(a^k)_{k=1\ldots K} \in Z \subseteq H$ terminal history if
    - *K* is infinite
    - $\neg\exists a^{K+1}\ s.t.\ (a^k)_{k=1..K+1} \in H$

- The history is
  - finite if $|H|<\infty$
  - finite horizon if longest $h \in H$ is finite

A 2-Player Extensive Game

- Set of players N={1,2}



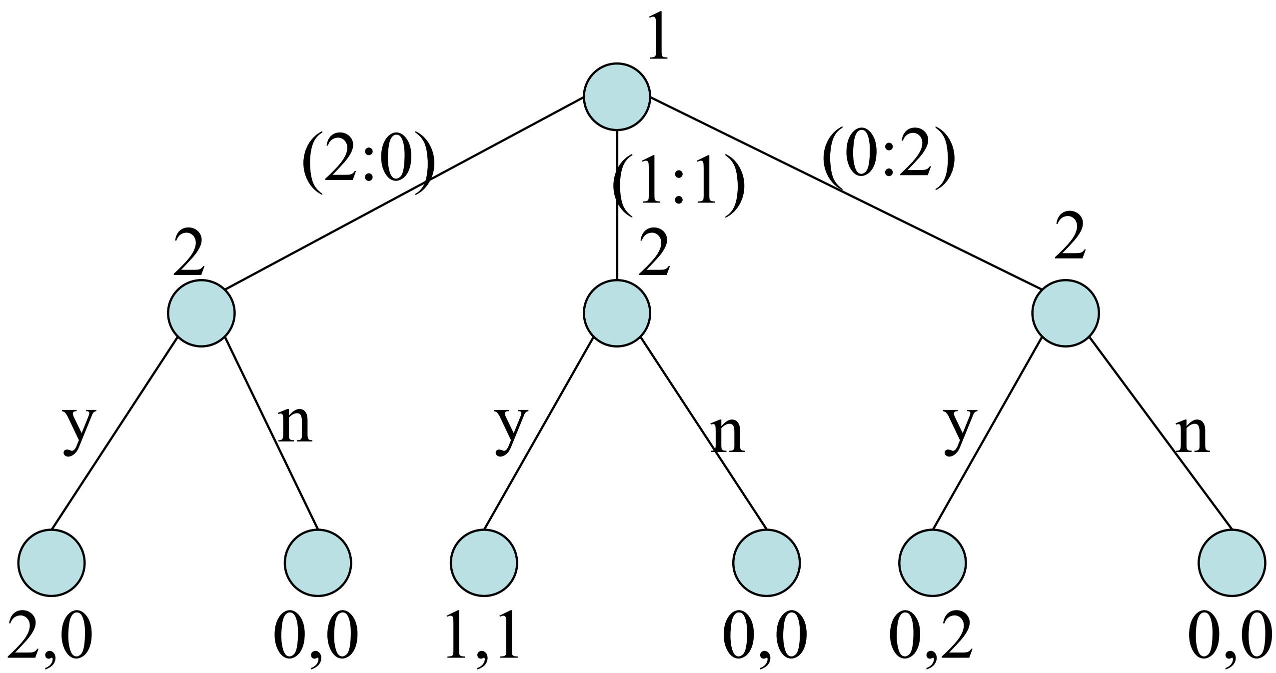
Extensive game - definition

- An extensive game with perfect information $G=<N,H,P,\succcurlyeq_i>$ consists of
  - A set N of players
  - A set H of sequences (histories) that satisfies
    - $\varnothing \in H$
    - if $(a^k)_{k=1...K} \in H$ and $L<K \Rightarrow (a^k)_{k=1...L} \in H$
    - if $(a^k)_{k=1}^{\infty}$ satisfies $(a^k)_{k=1...L} \in H$ for $\forall L>0 \Rightarrow (a^k)_{k=1}^{\infty} \in H$
  - A function $P: H \setminus Z \rightarrow N$ (player function)
  - A preference relation $\succcurlyeq_i$ on Z for $\forall i \in N$
- Similar to strategic games, $\succcurlyeq_i$ may be represented by $u_i: Z \rightarrow R$
- Set of actions implicitly defined

$$A(h) = \{a : (h,a) \in H\}$$

Example I - Definition

- Set of players N={1,2}
- Player function P(∅)=1, P((2:0))=P((1:1))=P((0:2))=2
- Set of histories H={∅,(2:0),(1:1),(0:2),((2:0),y), ((2:0),n),((1:1),y),((1:1),n), ((0:2),y),((0:2),n)}

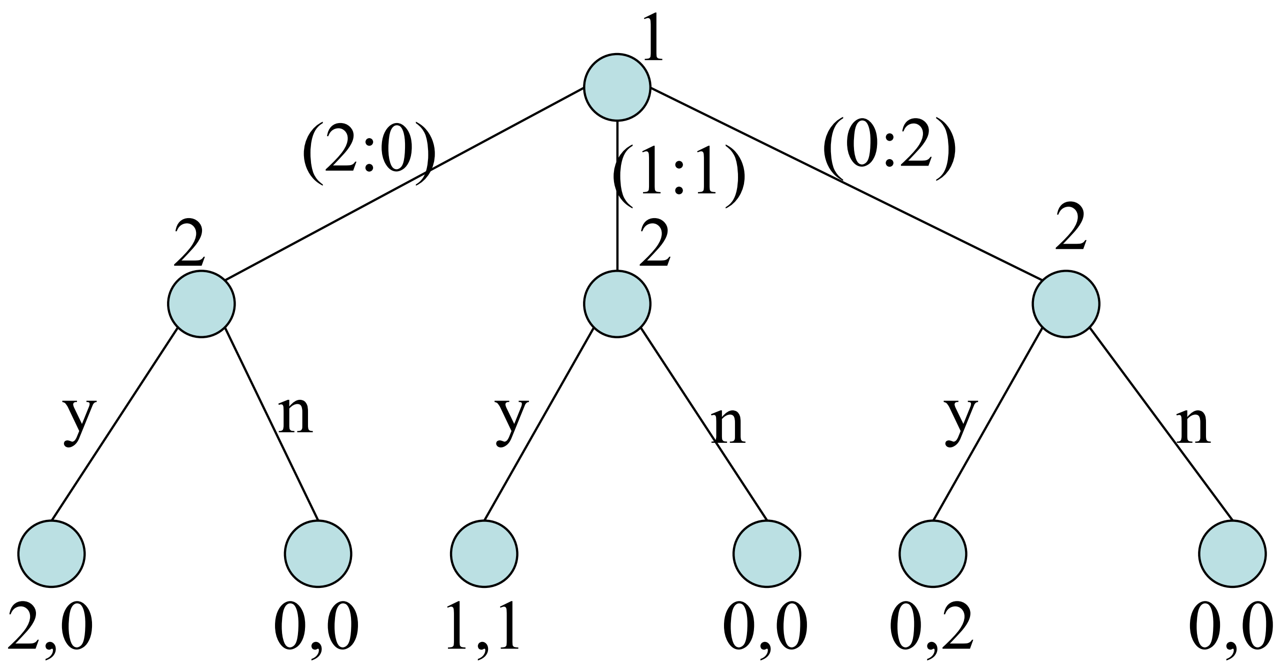


Example I – Definition contd.

- Preference relations

$$((2:0),y) \succ_1 ((1;1),y) \succ_1 ((0:2),y) \sim_1 ((2:0),n) \sim_1 ((1:1),n) \sim_1 ((0:2),n)$$

$$((0:2),y) \succ_2 ((1:1),y) \succ_2 ((2:0),y) \sim_2 ((2:0),n) \sim_2 ((1:1),n) \sim_2 ((0:2),n)$$



Strategies

- A strategy of player $i \in N$ in the extensive game with perfect information $G=<N,H,P, \succcurlyeq_i>$ is a function that *assigns* an *action* in *A(h) to every* history in $\{h \in H \backslash Z: P(h)=i\}$
  - Strategy depends on *N,H,P*

- Example strategies:
  - Player 1: (2:0), (1:1), (0:2)
  - Player 2: (y,y,y), (y,y,n), (y,n,n), (y,n,y), (n,y,n), (n,y,y), (n,n,y), (n,n,n)

Outcomes

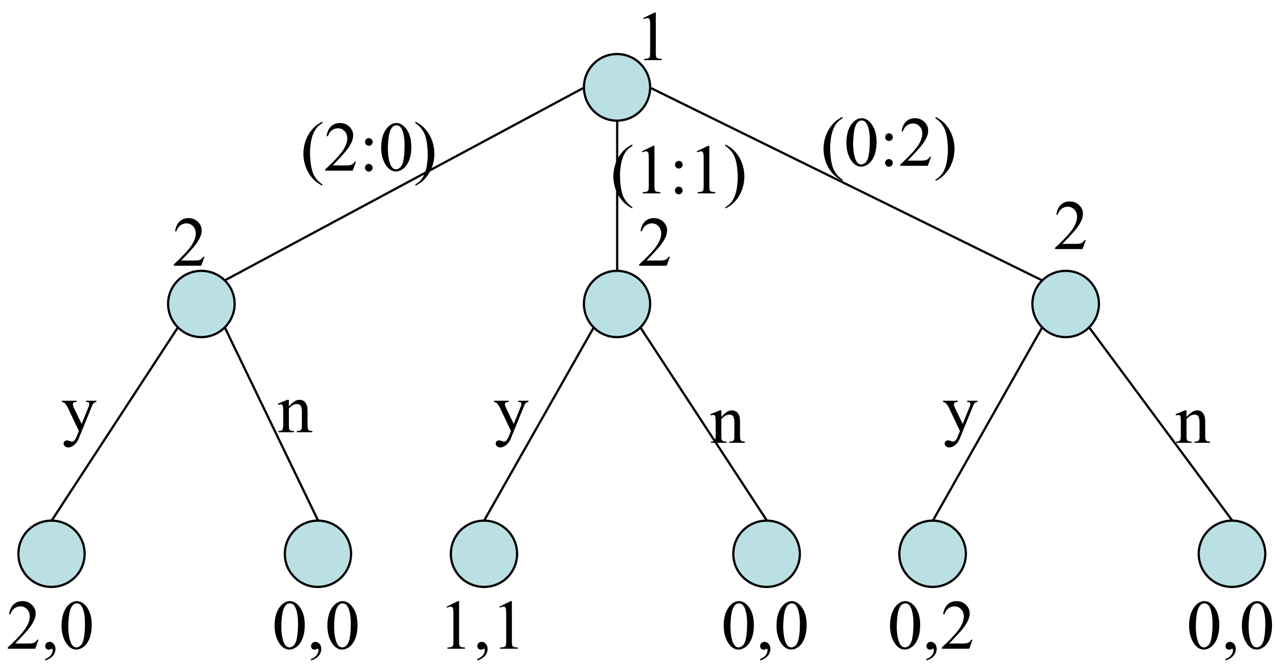
- The outcome $O(s)$ of a strategy profile $(s_i)_{i \in N}$ in the extensive game with perfect information $G=<N,H,P, \succcurlyeq_i>$ is the terminal history $h \in Z$ that results if every player follows its strategy s_i.

- $O(s)=(a^1,a^2,\ldots a^K) \in Z$ such that

$$s_{P(a^1,\ldots,a^k)}(a^1,\ldots,a^k) = a^{k+1} \quad 0 \le k \le K$$

Example I contd.

- What is the solution of the game?



Nash equilibrium

- A Nash equilibrium of an extensive game with perfect information $G = \langle N,H,P, \succcurlyeq_i \rangle$ is a strategy profile s^* such that for $\forall i \in N$

$$O(s_{-i}^*, s_i^*) \succsim_i O(s_{-i}^*, s_i) \quad \forall s_i$$

- A Nash equilibrium of an extensive game with perfect information $G = \langle N,H,P, \succcurlyeq_i \rangle$ is the Nash equilibrium of the strategic game $G^* = \langle N,(A_i),(\succcurlyeq_i') \rangle$ given as
  - $A_i = S_i$
  - $a \succcurlyeq_i' a' \quad \Leftrightarrow \quad O(s_i, s_{-i}) \succsim_i O(s_i', s_{-i}) \quad \forall s, s' \in S = \times_{i \in N} S_i$

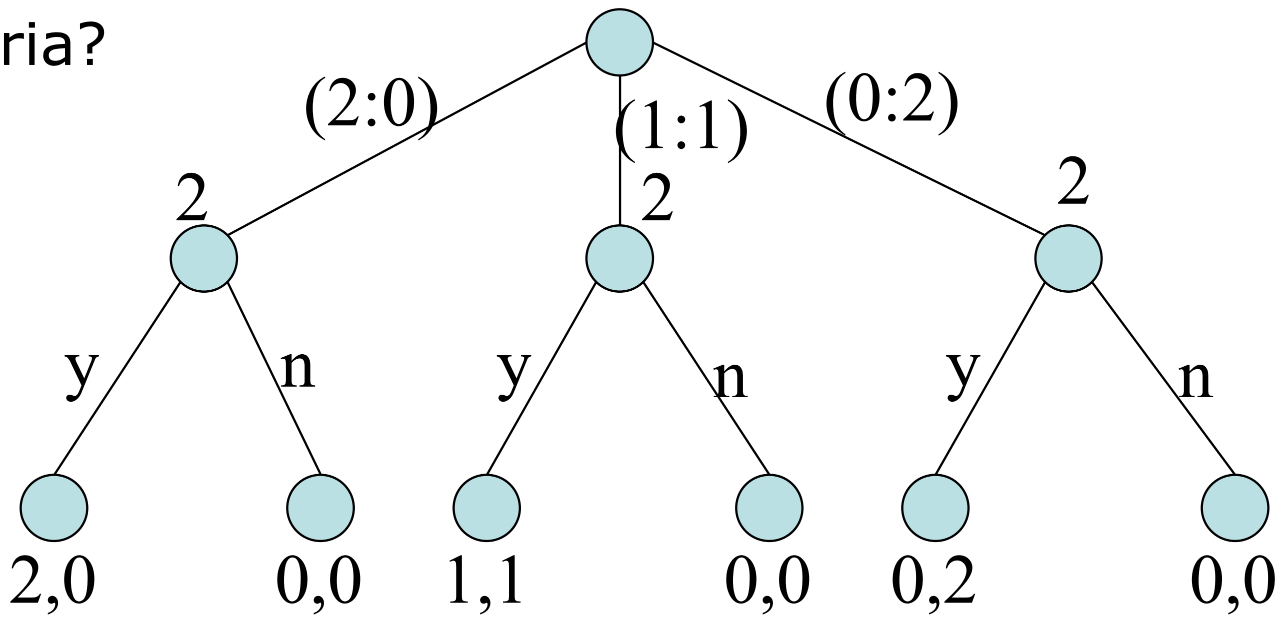
Example I revisited

	(y,y,y)	(y,y,n)	(y,n,n)	(y,n,y)	(n,y,n)	(n,y,y)	(n,n,y)	(n,n,n)
(2:0)	2,0	2,0	2,0	2,0	0,0	0,0	0,0	0,0
(1:1)	1,1	1,1	0,0	0,0	1,1	1,1	0,0	0,0
(0:2)	0,2	0,0	0,0	0,2	0,0	0,2	0,2	0,0

Nash equilibria?

- Plausible?

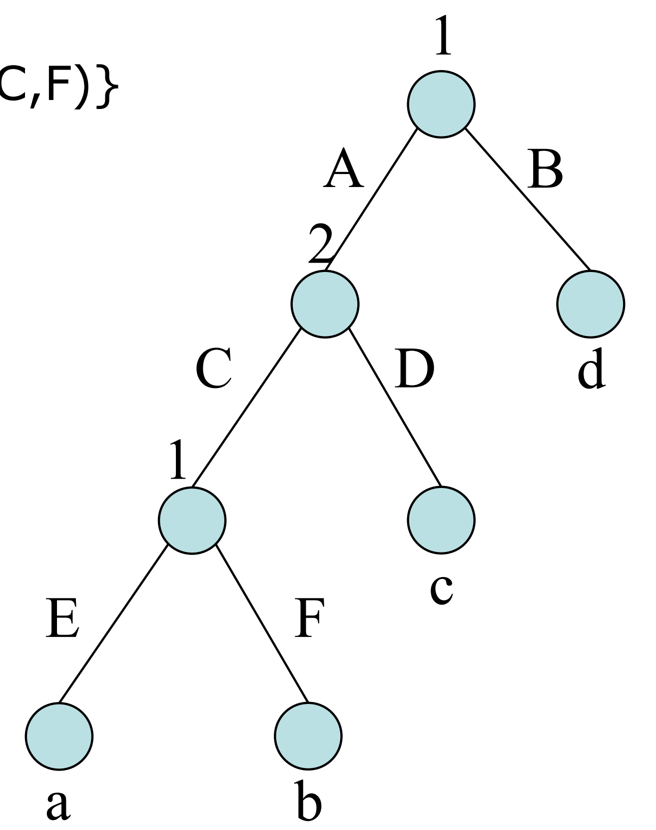
((2:0),yyy), ((2:0),yyn), ((2:0),ynn), ((2:0),yny), ((2:0),nnn), ((2:0),nny), ((1:1),nyy), ((1:1),nyn), ((0:2),nny)



Another example (II)

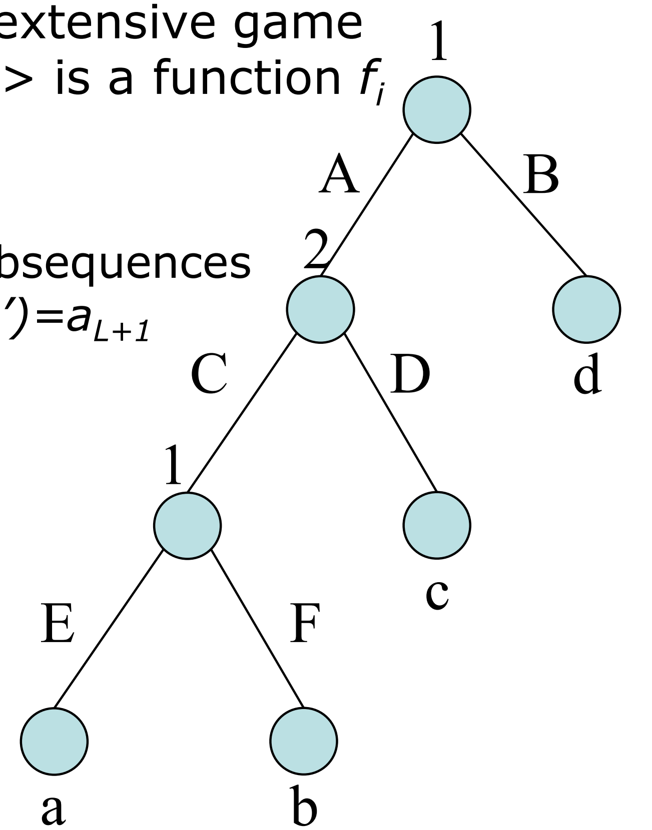
- N={1,2}
- H={∅, (A),(B),(A,C),(A,D),(A,C,E),(A,C,F)}
- P(∅)=1,P(A)=2,P((A,C))=1
- Strategies
  - S_1={(A,E),(A,F),(B,E),(B,F)}
  - S_2={(C),(D)}
- Strategy is not necessarily consistent
  - Outcomes are indifferent
- Corresponding strategic game

	C	D
AE	a	c
AF	b	c
BE	d	d
BF	d	d



Reduced strategy

- The reduced strategy of player *i* in an extensive game with perfect information $G=<N,H,P,\succcurlyeq_i>$ is a function f_i such that
  - its domain is $dom(f_i) \subseteq \{h \in H : P(h)=i\}$
  - $h \in dom(f_i) \Leftrightarrow h=(a^k)$ and for all its subsequences $h'=(a^k)_{k=1...L}$ with $P(h')=i$ we have $f_i(h')=a_{L+1}$
- Example II reduced strategies
  - Player 1
    - $f_1(\varnothing)=B$
    - $f_1(\varnothing)=A$ and $f_1((A,C))=E$
    - $f_1(\varnothing)=A$ and $f_1((A,C))=F$
  - Player 2
    - $f_2(A)=C$
    - $f_2(A)=D$



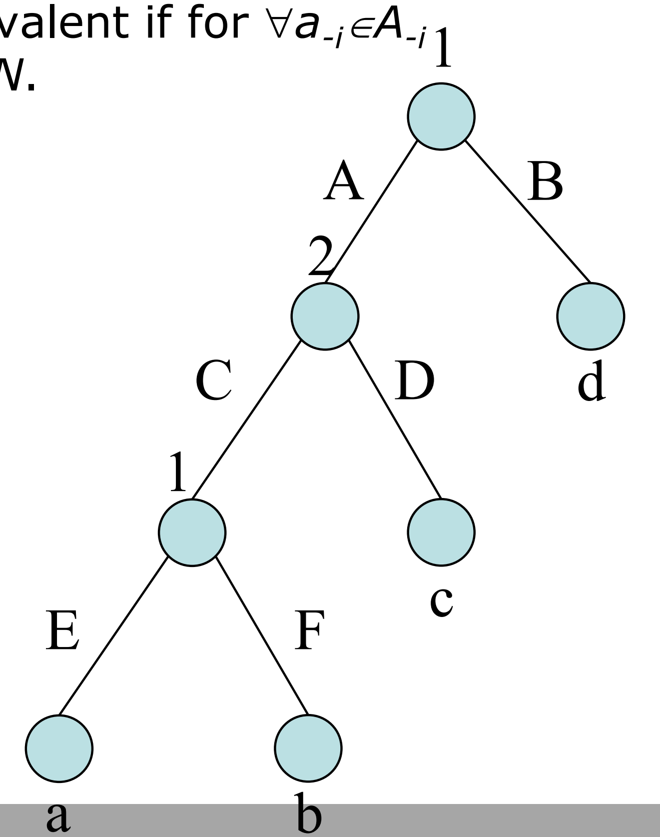
Reduced strategic form

- Let $G=<N,H,P,\succcurlyeq_i>$ be an extensive game with perfect information and $<N,(A_i),(\succcurlyeq_i)>$ be its strategic form.

  For $i\in N$ actions $a_i\in A_i$ and $a'_i\in A_i$ are equivalent if for $\forall a_{-i}\in A_{-i}$ we have $(a_{-i},a_i)\sim_j(a_{-i},a'_i)$ for every $j\in N$.

- The reduced strategic form of G is the strategic game $<N,(A'_i),(\succcurlyeq''_i)>$ in which A'_i contains only one of the equivalent strategies $a_i\in A_i$ and $\succcurlyeq''_i$ is the preference ordering over $\times_{j\in N}A'_j$ induced by $\succcurlyeq'_i$.

	C	D
AE	a	c
AF	b	c
B	d	d



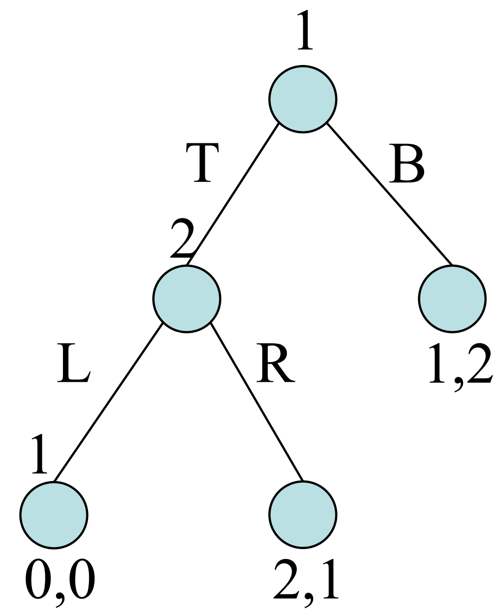
A similar example (III)

- N={1,2}
- H={∅,B,T,(T,L),(T,R)}
- P(∅)=1, P(T)=2
- Nash equilibria?
  - Strategic form

(T,R)

(B,L)

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



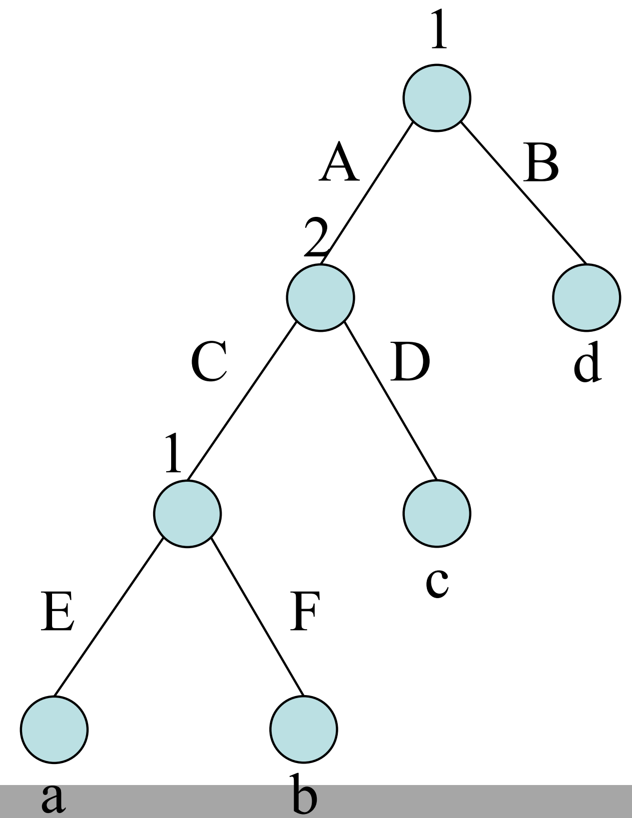
  - Reduced strategic form

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

More suitable equilibrium concept?

Subgame of a game

- The subgame of the extensive game with perfect information $G=<N,H,P, \succcurlyeq_i>$ that follows the history h is the extensive game with perfect information $G=<N,H|_h,P|_h, \succcurlyeq_{i|h}>$, where
  - $H|_h=\{h':(h,h')\in H\}$,
  - $P|_h(h')=P(h,h')$ for $h'\in H|_h$,
  - $h' \succcurlyeq_{i|h} h'' \Leftrightarrow (h,h') \succcurlyeq_i (h,h'')$



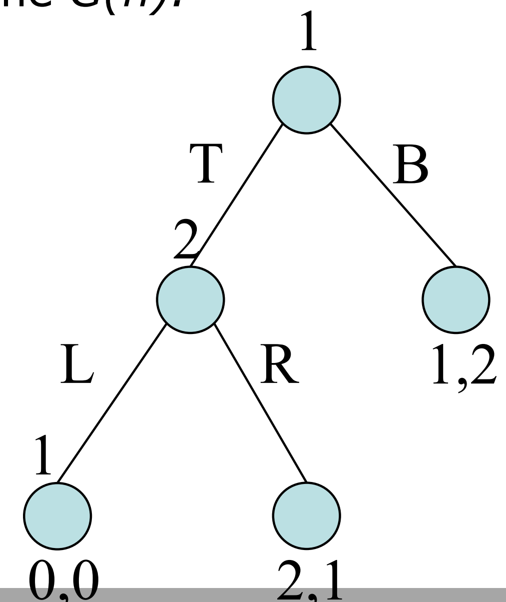
Subgame perfect equilibrium

- A subgame perfect equilibrium of an extensive game with perfect information G=<*N,H,P,* $\succcurlyeq_i$> is a strategy profile s^* such that for every player $i \in N$ and every nonterminal history $h \in H \setminus Z$ for which $P(h)=i$

$$O(s^*_{-i}|_h, s^*_i|_h) \succsim_i|_h O(s^*_{-i}|_h, s_i) \quad \forall s_i$$

for every strategy s_i of player *i* in the subgame *G(h).*

- Example:
  - The NE of the game were
    - (B,L)
    - (T,R)
  - What are the SPE of the game?
    - what are the nonterminal histories?



One deviation principle

- Let $G=<N,H,P, \succcurlyeq_i>$ be a finite horizon extensive game with perfect information. The strategy profile s^* is a SPE of G iff for every player i and every history $h \in H$ for which $P(h)=i$ we have

$$O(s^*_{-i}|_h, s^*_i|_h) \succsim_i|_h O(s^*_{-i}|_h, s_i) \quad \forall s_i$$

  for every strategy s_i of player i in the subgame $G(h)$ that differs from $s^*_i|_h$ only in the action it prescribes after the initial history of $G(h)$.

- Consequence
  - Can find the SPE of a finite horizon game with backwards induction (and some patience)

Existence and uniqueness of SPE

- Every finite extensive game with perfect information has a SPE.

- Proof

  Use the one deviation principle to construct a SPE from every terminal history $h \in Z$

- If none of the players is indifferent between any two outcomes then the SPE is unique.

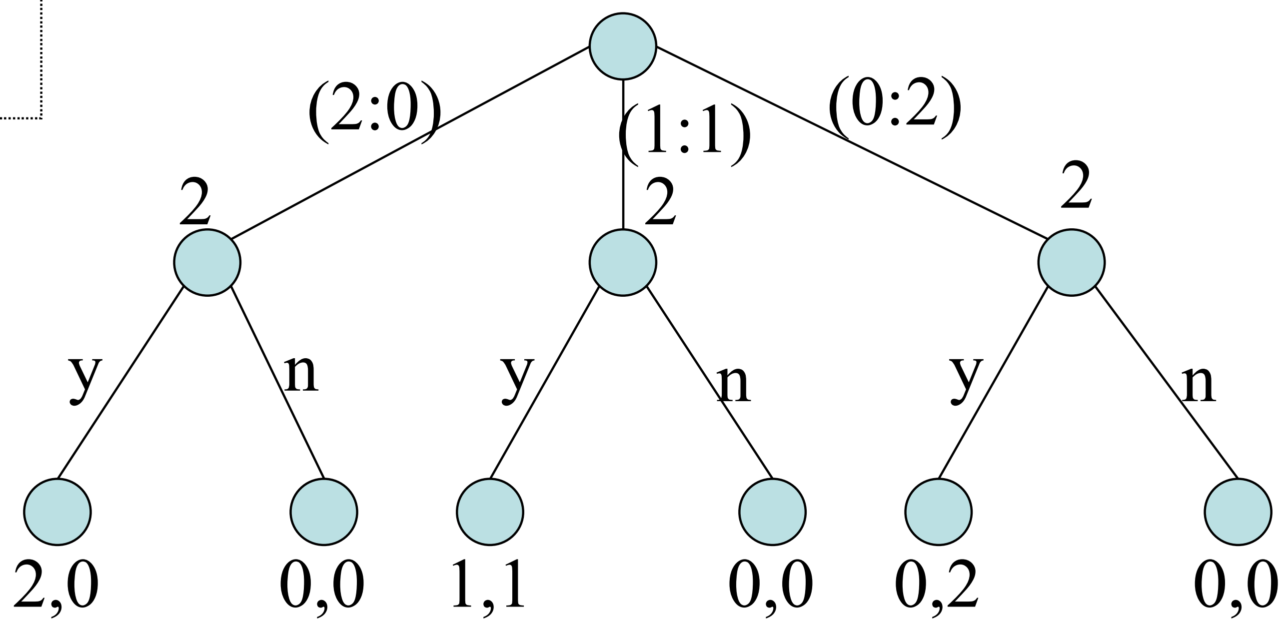
- Q: What about finite/infinite horizon?

Example I again

- The NE of the game were
  ((2:0),yyy), ((2:0),yyn), ((2:0),ynn), ((2:0),yny), ((2:0),nnn), ((2:0),nny), ((1:1),nyy), ((1:1),nyn), ((0:2),nny)
- What are the SPE of the game?

((2:0),yyy)

((1:1),nyy)

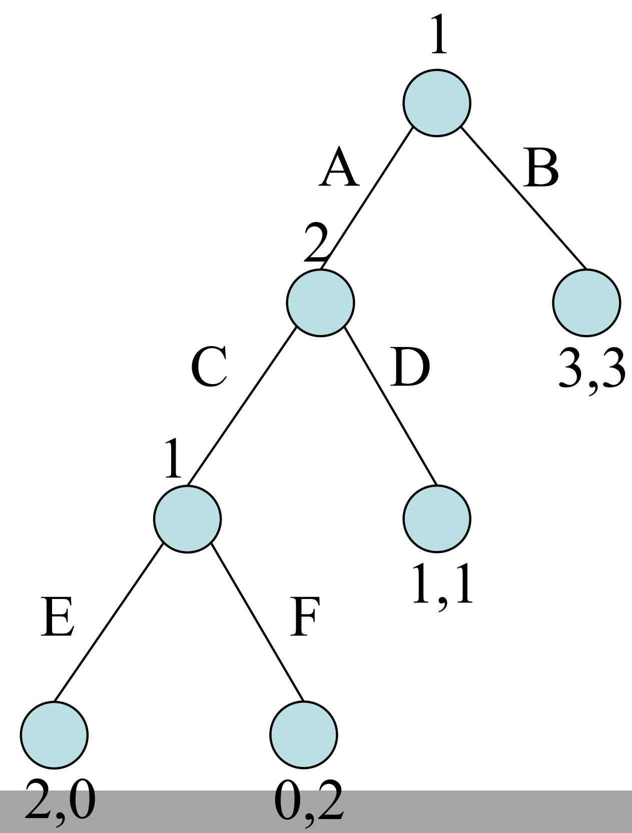


Iterated elimination of weakly dominated actions and SPE

- For a finite extensive game with perfect information and no indifferent outcomes the IEWDA in the strategic form of the game *can* lead to the unique SPE
  - depends on the order of elimination

- Example
  - What is the SPE?
  - What is the order of IEWDA?

	C	D
AE	2,0	1,1
AF	0,2	1,1
BE	3,3	3,3
BF	3,3	3,3



Some extensions

- Introduce an "environment" player *c*
  - $P(h)=c$ for some $h \in H \setminus Z$
  - c picks action from $A_c(h)$ at random (with density $f_c(h)$)
  - preferences interpreted over lotteries
  - called chance moves

- Imperfect information
  - Players may not know other players' past actions
  - Notion of *information set*

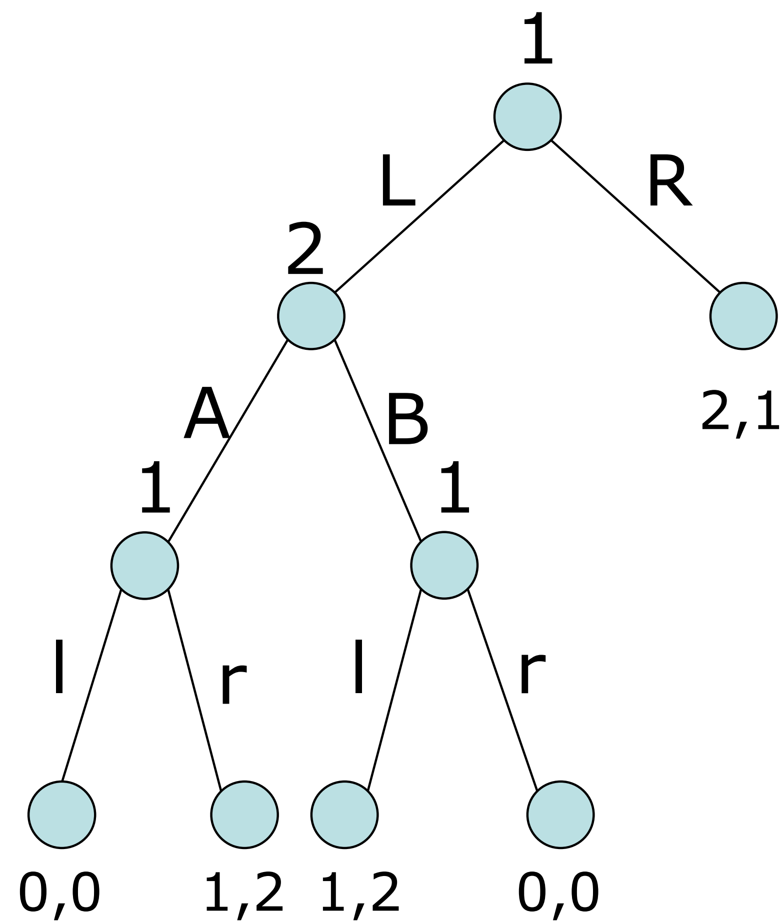
- Introduce simultaneous moves
  - $P(h) \subseteq N$
  - History $h \in H$ is a sequence of vectors

Mixed vs. Behavioral strategies

- Mixed strategy not used in extensive games with perfect information w/o simultaneous moves
  - Does not lead to new solutions
- Mixed strategy of player *i*
  - Probability measure over the set of player *i*'s pure strategies
- Behavioral strategy of player *i*
  - Collection of independent probability measures over the sets of possible actions for each non-terminal history
- Kuhn's theorem: In an extensive game of perfect recall for every mixed strategy there is a behavioral strategy that yields the same payoff to every player.

Example

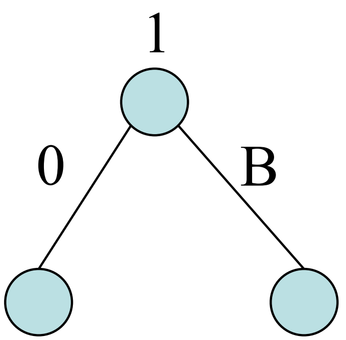
- Player 1's pure strategies
  - (R,l,l), (R,l,r),(R,r,l),(R,r,r) (L,l,l), (L,l,r),(L,r,l),(L,r,r)
- Player 2's pure strategies
  - (A), (B)
- Player 1's mixed strategies
  - $\alpha_{11},\dots,\alpha_{18}$
- Player 2's mixed strategies
  - α_{21}, α_{22}
- Player 1's behavioral strategies
  - α_{111}, α_{112}
  - α_{121}, α_{122}
  - α_{131}, α_{132}
- Player 2's behavioral strategies
  - α_{21}, α_{22}



A last example

- Slightly modified BoS game
- Player 1 can burn a dollar bill before

- What is the SPE?



	B	S
B	3,1	0,0
S	0,0	1,3

	B	S
B	2,1	-1,0
S	-1,0	0,3

$u_1(h) \geq 3/4$

	BB	BS	SB	SS
0B	3,1	3,1	0,0	0,0
0S	0,0	0,0	1,3	1,3
BB	2,1	-1,0	2,1	-1,0
BS	-1,0	0,3	-1,0	0,3

Repeated games

- A set of players *N*
- A set of actions for each player *A*
- Players play the "constituent" strategic game repeatedly
- Number of times the game is played can be
  - infinite
  - finite

- Objective vs. subjective number of repetitions

- Formally
  - Extensive game with simultaneous moves

Infinitely Repeated Game

- Let G=<N,(A_i), $\succsim_i$> be a strategic game, A_i is compact, and $\succsim_i$ is continuous. An infinitely repeated game of G is an extensive game with perfect information and simultaneous moves G=<*N,H,P,* $\succsim_i^*$> in which
  - $H = \{\emptyset\} \cup \{\cup_{t=1}^{\infty} A^t\} \cup A^{\infty}$
  - *P(h)=N* $\forall t$
  - $\succsim_i^*$ is a preference relation on A^{∞} that satisfies the condition of weak separability, i.e., if $(a^t) \in \mathrm{A}^{\infty}, a, a' \in A$, and $a \succsim_i a'$

    $(a^1, \ldots, a^{t-1}, a, a^{t+1}, \ldots) \succsim_i (a^1, \ldots, a^{t-1}, a', a^{t+1}, \ldots)$
- Strategy of player *i* assigns an action to every *h* $\in$*H**Z* (*Z*=A^{∞})

Preference relations

- Preference relation $\succcurlyeq_i^*$ based on the payoff u_i in G
  - assume u_i is bounded

- Payoff profile of G

$$v = \underline{u}(a) = \left(u_1(a),....,u_{|N|}(a)\right) \quad for \quad a \in A$$

- v is a *feasible payoff profile* of G if

$$v = \sum_a \lambda_a \underline{u}(a), \quad \sum_a \lambda_a = 1$$

- How can strategies be compared?
  - Payoffs have “time” dimension
    - (0,0,1,0,0,0,….) (0,1,0,0,0,0,….) ???
  - Model different forms of “human” preferences
    - compare sequences of payoffs

δ-discounted criterion

- Payoff profile in the repeated game

$$\sum_{t=1}^{\infty} \delta^{t-1} v_i^t \quad \delta \in (0,1)$$

- Preference relation defined as

$$(v^t) \succsim_i^* (w^t) \Leftrightarrow \sum_{t=1}^{\infty} \delta^{t-1} (v_i^t - w_i^t) \geq 0 \quad \delta \in (0,1)$$

- δ-discounted infinitely repeated game of $G=<N,(A_i),(u_i)>$

$$(1,1,1,0,0,0,...) \succ (0,0,0,2,2,2,2,\ldots) \qquad \delta < \sqrt[3]{\frac{1}{3}}$$

$$(0,0,0,2,2,2,2,\ldots) \succ (1,1,1,0,0,0,...) \qquad \delta > \sqrt[3]{\frac{1}{3}}$$

Limit of means criterion

- Payoff profile in the repeated game

$$\lim_{T\to\infty}\frac{1}{T}\sum\nolimits_{t=1}^{T} v_i^t$$

- Preference relation defined as

$$(v^t) \succ_i^* (w^t) \Leftrightarrow \liminf_{T\to\infty}\frac{1}{T}\sum\nolimits_{t=1}^{T}(v_i^t - w_i^t) > 0$$

- Limit of means infinitely repeated game of $G=\langle N,(A_i),(u_i)\rangle$

$$(0,\ldots,0,2,2,2,2,\ldots) \succ (1,1,1,\ldots,1,0,0,0,\ldots)$$

$$(-1,2,0,\ldots) \sim (0,\ldots)$$

Overtaking criterion

- Payoff profile in the repeated game

$$\sum_{t=1}^{\infty}(v_i^t)$$

- Preference relation defined as

$$(v^t) \succ_i^* (w^t) \Leftrightarrow \liminf_{T\to\infty} \sum_{t=1}^{T}(v_i^t - w_i^t) > 0$$

- Overtaking infinitely repeated game of $G=<N,(A_i),(u_i)>$

$$(1,\text{-}1,0,\ldots) \sim (0,\ldots)$$

$$(\text{-}1,2,0,\ldots) \succ (0,\ldots)$$

Famous example

- Infinitely repeated prisoner's dilemma
- Constituent game

	Do not confess	Confess
Do not confess	3,3	0,4
Confess	4,0	1,1

- Should the players play the NE of the constituent game?
  - Is that a NE?
- What is a subgame perfect equilibrium?
- What payoff profiles should we expect?

Folk theorems

- Characterize the set of payoff profiles of the repeated game
  - Nash equilibrium
  - Subgame perfect equilibrium

- Proofs constructive
  - Strategies that lead to the profile
  - Strategies often described as state machines
    - finite
    - infinite

- Not strong results
  - depend on the criterion used

The worst outcome: Minmax

- Player i's minmax payoff: The lowest payoff that other players can force upon player i

$$v_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_{-i}, a_i)$$

- Use it as a threat
  - p_{-i} is the most severe punishment
  - $B_i(p_{-i})$ are the best responses to the punishment

- Enforceable payoff profile (and corresponding outcome a)

$$w_i \geq v_i \quad i \in N$$

- Strictly enforceable payoff profile (and outcome a)

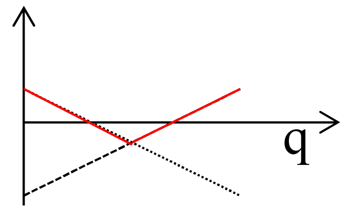
$$w_i > v_i \quad i \in N$$

Example (mixed vs. pure)

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

- Pure strategies
  - $v_1=1,\ v_2=1$
- Mixed strategies
  - Player 1's minmax payoff
    - $q=\alpha_2(L)$

$$v_T(q) = -3q+1$$
$$v_M(q) = 3q-2$$
$$v_B(q) = 0$$



    - Minimize $\max(v_T,v_M,v_B)$
      - $q=0.5 \rightarrow v_T=v_M=-0.5,\ v_1=v_B=0$
  - Player 2's minmax payoff
    - $p_T=\alpha_1(T),\ p_M=\alpha_1(M)$

$$v_L(p_T,p_M) = 2(p_T - p_M) + (1 - p_T - p_M)$$
$$v_R(p_T,p_M) = -2(p_T - p_M) + (1 - p_T - p_M)$$

    - Minimize $\max(v_L,v_R)$
      - $p_T=0.5,\ p_M=0.5 \rightarrow v_2=v_L=v_R=0$

The worst outcome

- Every Nash equilibrium payoff profile of the repeated game of $G=<N,(A_i),(u_i)>$ is an <u>enforceable payoff profile</u> of G
  - for the limit of means criterion
  - for the δ-discounting criterion ($\delta \in (0,1)$)

- Proof:

  Assume s^* is NE and $w_i<v_i$ for player i (i.e., not enforcable)

  Then s^*_i can be improved

  $s'_i(h) \in B_i(s_{-i}(h)) \quad \Rightarrow \quad w_i \geq v_i \quad \Rightarrow$ s* is not a NE

Nash folk theorems

- *Limit of means:* Every feasible enforceable payoff profile of $G=<N,(A_i),(u_i)>$ is a NE payoff profile for the limit of means infinitely repeated game of G.
  - play each outcome *a* for β_a number of times in every cycle of rounds

$$w=\sum_{a\in A}\frac{\beta_a}{\gamma}u(a), \quad \text{where} \quad \gamma=\sum_{a\in A}\beta_a$$

  - players $j \neq i$ punish player *i* who first deviates from this strategy by playing $(p_{-i})_j$ forever
    - player *i* loses by deviating $\Rightarrow$ *NE*

- *δ-discounted:* Let *w* be a feasible strictly enforceable payoff profile of $G=<N,(A_i),(u_i)>$.
  Then $\forall\varepsilon>0$ $\exists\delta^*<1$ s.t. if $\delta>\delta^*$ then the δ-discounted infinitely repeated game of G has a NE with payoff profile w', $|w-w'|<\varepsilon$.

Plausibility

- Consider these two constituent games

G_1	D	C
D	3,3	0,4
C	4,0	1,1

minmax

G_2	D	C
D	2,3	1,5
C	0,1	0,1

- Threat is not credible
  - Punishes the punisher

Perfect folk theorems

- Punishment phase should not punish the punisher
  - Punish deviation for a limited amount of time
    - Just enough to cancel out the gain of the deviation
  - Compensate the punisher if needed
- PFT for limit of means criterion
  - Every <u>strictly enforceable feasible</u> payoff profile
  - Punish for a limited length of time
- PFT for overtaking criterion
  - Any <u>strictly enforceable</u> outcome a^*
  - Punish for a limited length of time and punish misbehaving punishers

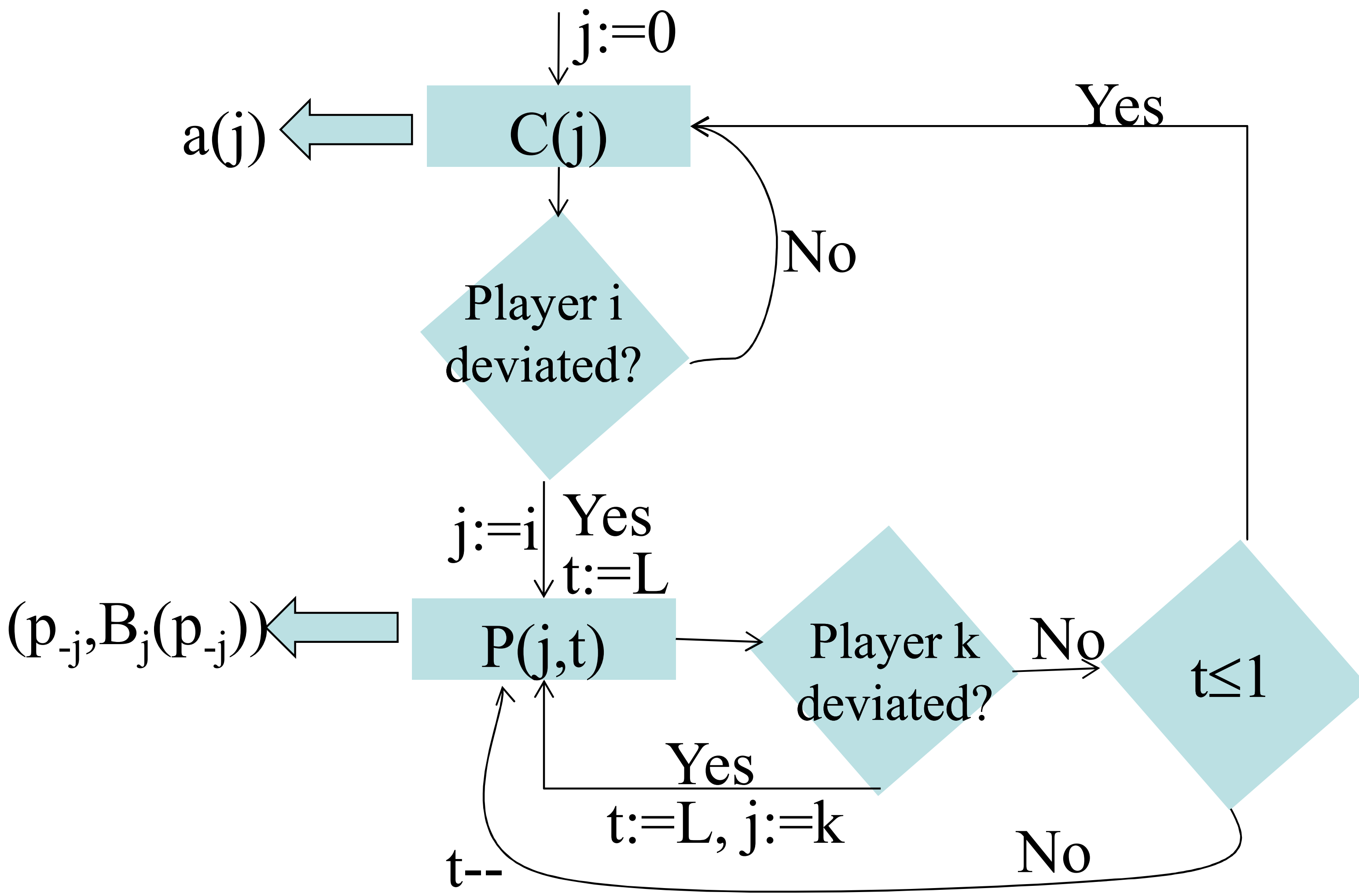
PFT for the discounting criterion

- Let a^* be a <u>strictly enforceable</u> outcome of $G=<N,(A_i),(u_i)>$. Assume that there is a collection $(a(i))_{i\in N}$ of strictly enforceable outcomes of G s.t.
  - $a^* \succ_i a(i)$
  - $a(j) \succ_i a(i)$

  for all $j \in N\setminus\{i\}$. Then $\exists\delta^*<1$ s.t. $\forall\delta>\delta^*$ there is a subgame perfect equilibrium of the δ-discounted infinitely repeated game of G that generates the path (a^t) in which $a^t=a^*$ for $\forall t$
- Proof:
  - Start with profile a^*
  - Punish deviation of player j
    - Play $(p_{-j},B_j(p_{-j}))$ for a period L large enough
    - Then choose outcome $a(j)$
    - Unless a punisher k misbehaves
      - choose $a(k)$ for period L to punish the misbehaving punisher

D. Fudenberg, E.S.Maskin, "The folk theorem in repeated games with discounting or with incomplete information", Econometrica, vol. 54, pp. 533-554, 1986

State Machine for the PFT



PFT for the discounting criterion

- Deter player *i* from deviating from outcome *a(j)*
  - Choose *L* large enough

$$\underbrace{M - u_i(a(j))}_{\text{Gain from deviation}} < L\underbrace{(u_i(a(j)) - v_i)}_{\text{Non discounted loss of payoff during punishment}} \quad \forall i \in N, j \in \{0\} \cup N$$

  - Choose $\delta' < 1$ s.t. for $\delta > \delta'$

$$\underbrace{M - u_i(a(j))}_{\text{Gain from deviation}} < \underbrace{\sum_{k=2}^{L+1} \delta^{k-1}(u_i(a(j)) - v_i)}_{\text{Discounted loss of payoff during punishment}}$$

- Deter punisher from deviating from the punishment rule
  - Choose $\delta^* > \delta'$ s.t. for $\delta > \delta^*$

$$\underbrace{\sum_{k=1}^{L} \delta^{k-1}\left(M - u_i(p_{-j}, b_j(p_{-j}))\right)}_{\text{Deviation gain for the punisher}} < \underbrace{\sum_{k=L+1}^{\infty} \delta^{k-1}(u_i(a(j)) - u_i(a(i))}_{\text{Potential punishment of the punisher}}$$

D. Fudenberg, E.S.Maskin, "The folk theorem in repeated games with discounting or with incomplete information", Econometrica, vol. 54, pp. 533-554, 1986

Some extensions to infinitely repeated games

- Long run and short run players
- Overlapping generations of players
- Randomly matched opponents

Finitely repeated games

- Let $G=<N,(A_i),\succcurlyeq_i>$ be a strategic game, A_i is compact, and $\succcurlyeq_i$ is continuous. A repeated game of G is an extensive game with perfect information and simultaneous moves $G=<N,H,P,\succcurlyeq_i^*>$ in which
  - $H=\{\varnothing\}\cup\left\{\bigcup_{t=1}^{T} A^t\right\}$
  - $P(h)=N$
  - $\succcurlyeq_i^*$ is a preference relation on A^T that satisfies the condition of weak separability, i.e., for $\forall t$

$$\left(a^t\right)\in A^T, a\in A, a'\in A, a\succsim_i^* a' \Rightarrow (a^1,\ldots,a^{t-1},a,a^{t+1},\ldots)\succsim_i^* (a^1,\ldots,a^{t-1},a',a^{t+1},\ldots)$$

- Strategy of player i assigns an action to every $h\in H\setminus Z$
- Preference relation (similar to limit of means)

$$(v^t)\succ_i^* (w^t) \Leftrightarrow \frac{1}{T}\sum_{t=1}^{T}(v_i^t - w_i^t) > 0$$

- T period finitely repeated game

Example

- Finitely repeated PD

	Do not confess	Confess
Do not confess	3,3	0,4
Confess	4,0	1,1

- Should the players play the NE of the constituent game?

Another example

- Modified PD

	L	M	R
T	3,3	0,4	0,0
C	4,0	1,1	0,0
B	0,0	0,0	0.5,0.5

- Should the players play the NE of the constituent game?

Minmax payoffs in all NE

- If the **payoff profile in every NE** of the constituent game *G* is the profile (v_i) of **minmax payoffs** in *G* then for any value of T the outcome (a^1,…,a^T) of every NE of the T-period repeated game of *G* is such that ***a^t* is a NE of G** for t=1,…,T.
  - Proof: by contradiction. If not all actions are NE, player *i* can improve by exchanging the last non-NE action to the NE, and then play $B_i(p_{-i})$.

- If the constituent game G has a **unique NE payoff profile** then for any T the **action profile** chosen after any history in any SPE of the T-period finitely repeated game of G **is a NE of G**.
  - Proof: by induction, the last period has to be a NE, etc.

Nash folk theorem

- If the constituent game G has a NE a^* s.t. $u_i(a^*)>v_i$ then for any <u>strictly enforceable outcome a'</u> of G and $\varepsilon>0$ $\exists T^*$ s.t. the T period repeated game of G has a NE $(a^1,\ldots,a^T)$ for which

$$\left|\frac{1}{T}\sum_{t=1}^{T} u_i(a^t) - u_i(a')\right| < \varepsilon \quad \forall T > T^*$$

  - Proof sketch:
    - Play a' until period T-L
    - Play a^* after period T-L
    - Punish player j by playing $(p_{-j})_i$
    - Choose L to cancel gain of deviation
      $$\max_{a_i\in A_i} u_i(a'_{-i}, a_i) - u_i(a') \le L(u_i(a^*) - v_i)$$
    - Choose T^* big enough to be within ε
      $$\left|\frac{1}{T^*}[(T^* - L)u_i(a') + Lu_i(a^*)] - u_i(a')\right| < \varepsilon$$

Perfect folk theorem

- Let a^* be a strictly enforceable outcome of the constituent game G. Let G be s.t.
  - $\forall i \in N$ there are two NE of G that differ in their payoffs for player i
  - there is a collection $(a(i))_{i \in N}$ of strictly enforceable outcomes of G such that
    - $a^* \succ_i a(i) \quad \forall i \in N$

      $a(j) \succ_i a(i) \quad \forall j \in N \setminus \{i\}$

Then $\forall \varepsilon > 0$ $\exists T^*$ s.t. the T-period repeated game of G has a SPE $(a^1,\ldots,a^T)$ in which

$$\left| \frac{1}{T} \sum_{t=1}^{T} u_i(a^t) - u_i(a^*) \right| < \varepsilon \quad \forall T > T^*$$

Dynamic games

- Players make decisions at different points in time
- Extensive game
  - Players make decisions one by one
  - Can learn about the environment and others' choices
- Repeated game
  - Players play multiple strategic games
  - Decision is influenced by the history
  - Extension of extensive game
- Other forms of dynamic games
  - Stochastic game
  - Differential game

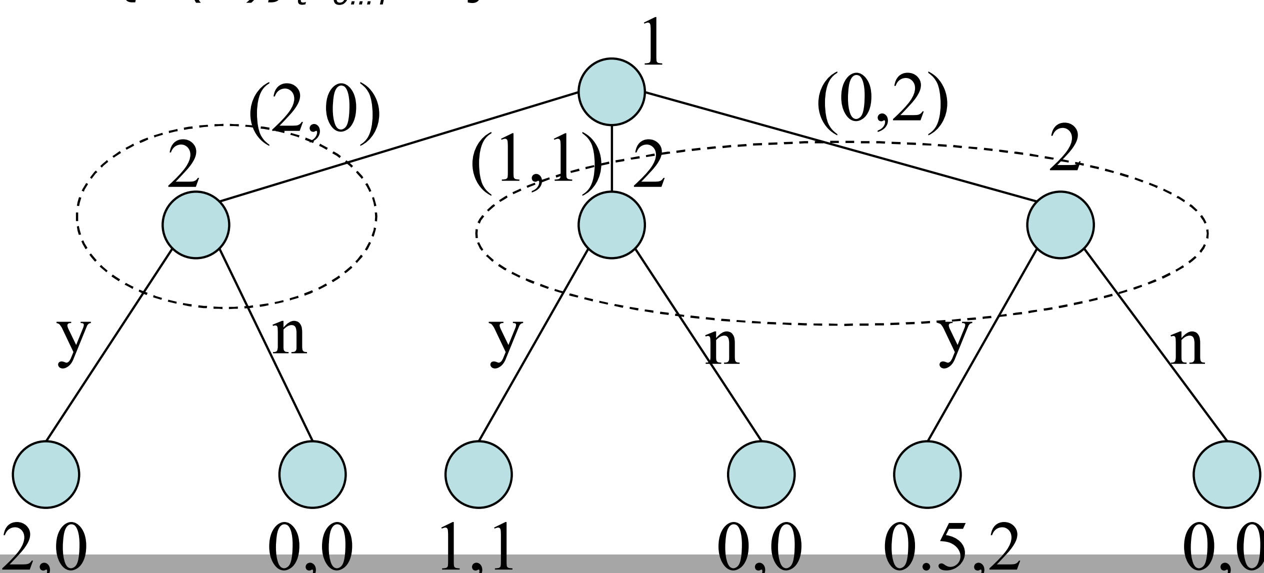
Reduction of the history set

- Consider an extensive game $G=<N,H,P,(u_i)>$
  - For all t we can write

    $$u_i(a^0,\ldots,a^T)=u_i(h^t,f^t)$$

    (f^t is future)

- For each t partition the set of histories
  - $\{H^t(h^t)\}_{t=0\ldots T}$ disjoint and exhaustive



Sufficient partition

- A partition $\{H^t(h^t)\}_{t=0\ldots T}$ is <u>sufficient</u> if, for all t, h^t and $h^{t'}$ such that $H^t(h^t)=H^t(h^{t'})$, the subgames starting at date t after histories h^t and $h^{t'}$ are equivalent
  - identical action spaces

$$A_i^{t+\tau}(h^t, a^t, \ldots, a^{t+\tau-1}) = A_i^{t+\tau}(h^{t'}, a^t, \ldots, a^{t+\tau-1}) \quad \forall i, \forall \tau \geq 0$$

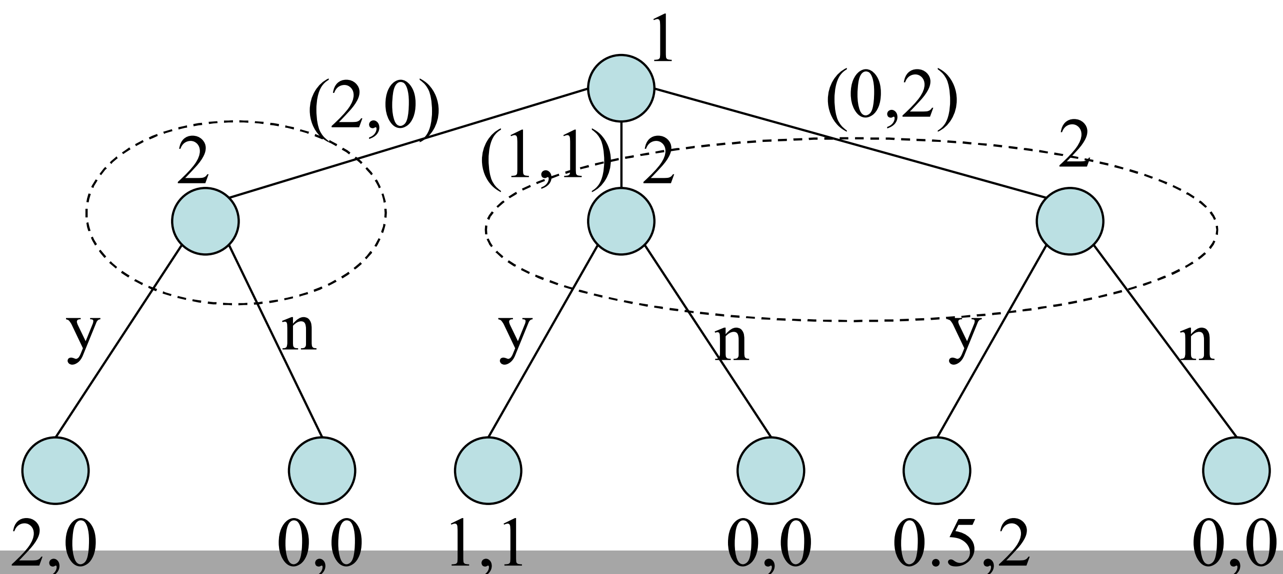
  - utility functions represent the same preferences
    - uniqueness of the utility function to an affine transformation

$$u_i(h^t, f^t) = \lambda_i(h^t, f^t) u_i(h^{t'}, f^t) + \mu_i(h^t, h^{t'}, f_{-i}^t)$$

- Trivial sufficient partition
  - $H^t(h^t)=\{h^t\}$

Payoff relevant history

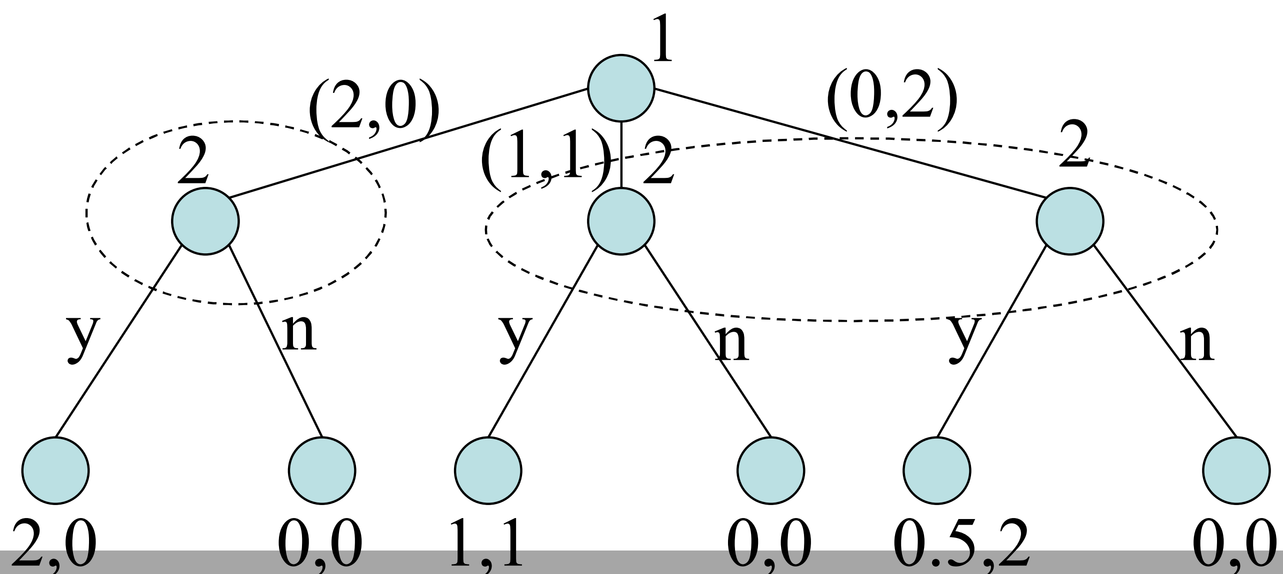
- Payoff relevant history is the minimal sufficient partition
  - the coarsest sufficient partition



Markov strategy

- Markov strategy is a strategy that is measurable with respect to the payoff relevant history
$$H^t(h^t) = H^t(h^{t'}) \Rightarrow \sigma_i^t(h^t) = \sigma_i^t(h^{t'}) \quad \forall i$$
  - consistent with rationality – no coarser history would give equally good payoffs

- No need to know the entire history



Markov perfect equilibrium

- Markov perfect equilibrium is a profile of Markov strategies σ that are subgame perfect equilibrium.

- Sufficient condition for existence (mixed case)
  - finite-horizon extensive game
  - infinite-horizon extensive game with continuous payoff profile at ∞ $\lim_{t\to\infty} \sup_{h,\tilde{h}\ s.t.\ h^t=\tilde{h}^t} |u_i(h) - u_i(\tilde{h})| = 0$
    - δ-discounted criterion (δ<1), per-period payoffs are bounded

E. Maskin, J. Tirole, "Markov Perfect Equilibrium, I," Journal of Economic Theory, vol. 100, pp. 191-219, 2001

Stochastic games

- History summarized in "state"
  - Available actions depend on the state
  - Current payoffs depend on the state and the actions

- A stochastic game $G=<N,K, (\Delta A_i(k)),Q, \succcurlyeq_i>$ consists of
  - Set N of players
  - Set K of states
  - Sets of mixed action profiles on $A_i(k)$
  - Transition function $Q=(q(k^{t+1}|k^t,a^t))$
  - Preference relation on the sequence of outcomes and states (objective function)
    - δ-discounted $u_i = \sum_{t=0}^{\infty} \delta^t g_i(k^t, a^t)$
    - Limit of means

Markov (stationary) strategy in Stochastic Games

- Assume players other than i play Markov strategies
  - h' and h two histories both leading to state k
  - a_i and $a_{i'}$ actions chosen by player i after h and h' resp.
  - value $V_i(k,s_{-i})$ highest expected payoff i can achieve starting from state k

- Value function V_i

$$V_i(k;s_{-i}) = \max_{a_i \in A_i(k)} \mathrm{E}\left[g_i(k, s_{-i}(k), a_i) + \delta \sum_{k' \in \mathrm{K}} q(k' | k, s_{-i}(k), a_i) V_i(k', s_{-i}) \right]$$

  - Maximizers form Markov best response

Existence of MPE

- Markov perfect equilibria always exist in stochastic games with a finite number of states and actions.
  - Proof:
    - Markov strategic form
      - Agent (i,k) has u_i of player i starting from state k

$$u_{i,k}(a) = \mathrm{E}\left[\sum_{t\geq 0} \delta^t g_i(k^t, a(1,k^t), \ldots a(N,k^t)) \mid k^0 = k\right]$$

    - Finite states $\Rightarrow$ finite # of agents and actions
      - There is a mixed strategy NE $(\sigma^*_{i,k})$
    - Markov strategy of player i is $\sigma^*_i(k)=\sigma^*_{i,k}$
      - Depends on the state only
    - By construction it is subgame perfect
      - agents optimize in each state

- Other existence results
  - Countably infinite state space
  - etc.

T. Parthasarathy, "Existence of Equilibrium Stationary Strategies in Discounted Stochastic Games", Sankhya Series A, vol 44, pp. 114-127, 1982

Differential games

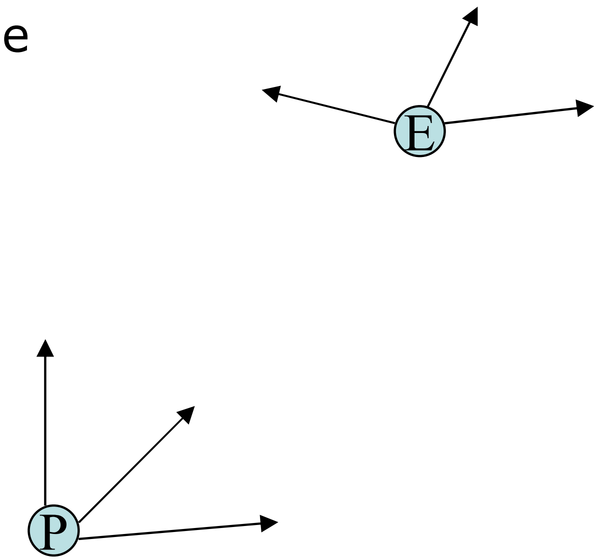
- Continuous time stochastic games

- A differential game G=<N,(k^t),(h_j^t), (u_i)> consists of
  - Set N of players - often $|N|=2$
  - State vector $k^t = (k_1^t, \ldots, k_n^t) \in \mathrm{R}^n$
  - Sets of actions $A_i(k^t) \in R_i^a$
  - Transition functions
    $$\frac{dk_j^t}{dt} = h_j^t(k^t, a^t)$$
  - Payoff functions
    $$u_i = \int_0^T g_i^t(k^t, a^t)dt + v_i^T(k^T)$$
  - Initial condition
    $$k^0 = k(0) \in \mathrm{R}^n$$

Example

- Simple pursuit game in the plane
  - Two players: P and E
    - P has speed W
    - E has speed w
      - W>w
  - State variable
    - Position
  - Action space
    - Angle
- Objective
  - Time of capture

- Markov perfect equilibrium?
  - Direct fleeing



R. Isaacs, "Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization", Courier Dover, 1999

Literature

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

- T. Parthasarathy, "Existence of Equilibrium Stationary Strategies in Discounted Stochastic Games", Sankhya Series A, vol 44, pp. 114-127, 1982
- D. Levhari, L.Mirman, "The great fish war", Bell Journal of Economics, 1980, pp.322-344
- J.H. Case, "Towards a theory of many player differential games", SIAM Journal of Control, vol. 7, 1969, pp 179-197.
- A. Starr and Y.C. Ho, "Nonzero-sum differential games", Journal of Optimization Theory and Applications, vol 3., 1969., pp. 183-206