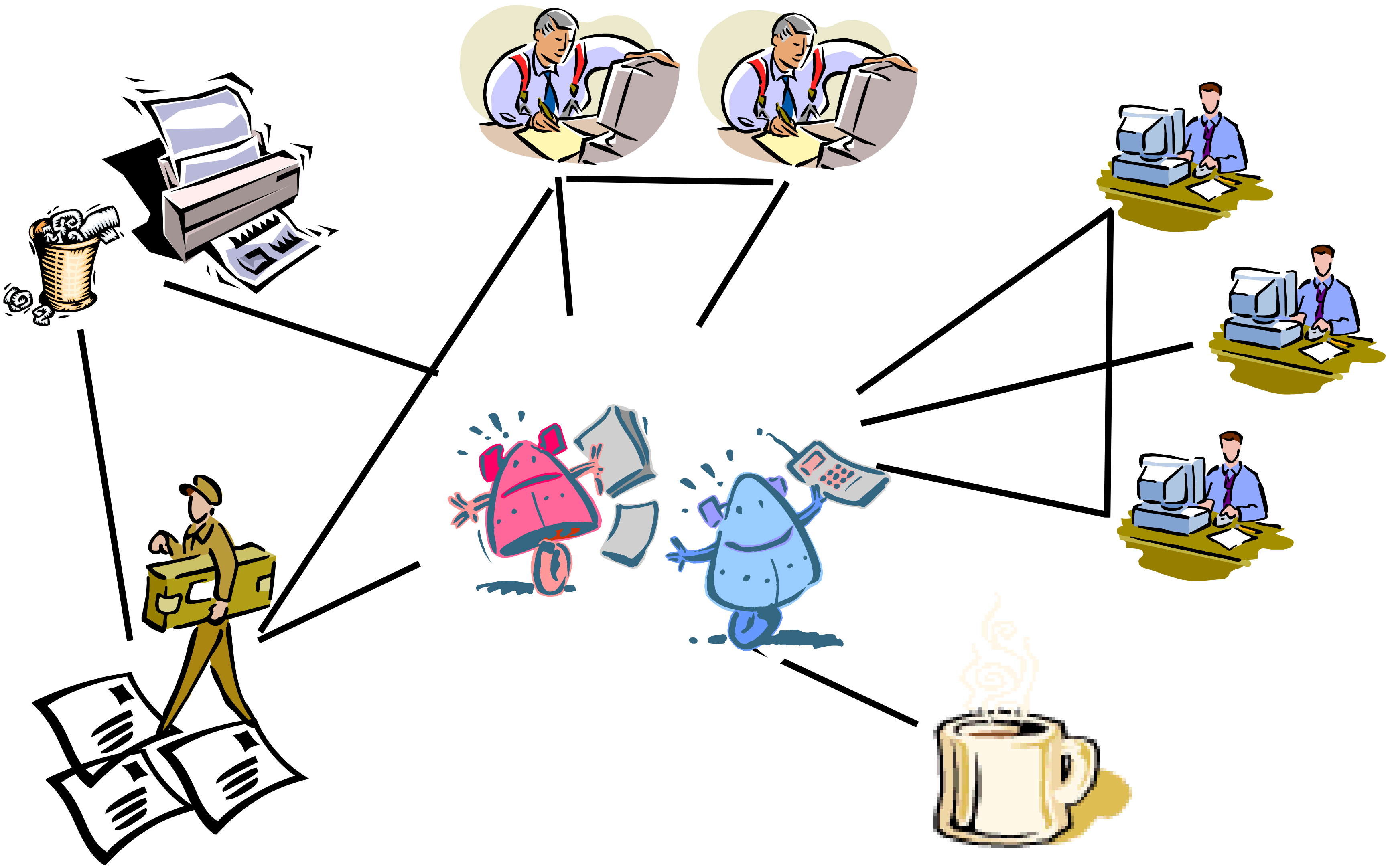
2534 Lecture 8: Intro to Game Theory

- Game Theory today
  - will focus on Shoham and Leyton-Brown for next two classes
  - today: main parts of Ch.3 and Ch.5
- Announcements
  - Problem Set 1 handed back (see web page for sample solutions)
  - Project pre-proposals back today
    - If sufficiently detailed and clear, my comments will indicate no need to submit a "final proposal" next week
  - Problem Set 2 will be posted on web site tonight or tomorrow AM (due in two weeks on Nov.11)

Multiagent Systems



Multiagent Decision Making

- So far we've considered single-agent decision making
  - focus on one-shot problems
  - will return to sequential problems (MDPs, POMDPs) later
- Multiagent decision making presents some unique challenges and complications
- Let's consider an example to illustrate: **Battlebots**!

Game 1: Battlebots

		Blue Robot: Coffee	Blue Robot: Tea
Red Robot	Coffee	0 / 0	10 / 8
	Tea	8 / 10	0 / 0

red payoff/blue payoff

- What should the robots do?
  - both go for coffee?
  - red coffee, blue tea? blue coffee, red tea?
  - what about each choosing coffee of tea randomly?
    - e.g., choose coffee with $p=5/9$, tea with $p=4/9$

Strategic Considerations

- Normally, we'd propose using MEU; but…
- Missing ingredient: *what will the other guy do?*
  - my own action not enough to determine outcome/utility
  - if I had a distribution (*beliefs*) over other's actions, I could adjust my strategy appropriately…
  - …but then, why couldn't he do the same thing?
    - dependence of decisions is, in some sense, circular!
- Game theory provides models and solution concepts for general multiagent interactions
  - generally, a solution is a *strategy profile* that is in some form of *equilibrium*
  - e.g., *<Red gets coffee; Blue gets tea>*

Normal Form Games

- A *normal (or strategic) form game*
  - set of agents (or players) $i = \{1, \ldots N\}$
  - action sets A_i for each agent
    - *joint action* set $A = X A_i$
  - utility functions $u_i : A \rightarrow \mathbf{R}$
    - $u_i(\mathbf{a})$ is utility of joint action $\mathbf{a}$ to agent i
- Assumptions *(more to come, plus we'll relax **these ones** later)*
  - players choose their actions simultaneously and independently
  - they all know the structure of the game:
    - moves available to all players
    - all player's utilities
  - … and the structure of the game is common knowledge

Normal Form Games

- Matrix notation (2Player, 2Action example):

		II (column)	
		C	D
I (row)	A	u1 / u2	u1 / u2
	B	u1 / u2	u1 / u2

- Notes:
  - actions often called *pure strategies*
  - joint actions called *pure strategy profiles*
  - utilities usually called *payoffs*

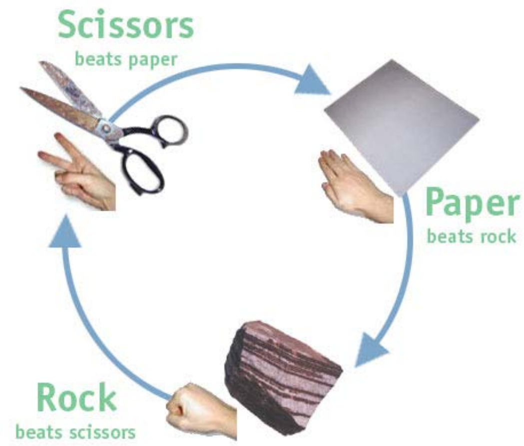
Example: Matching Pennies

	Heads	Tails
Heads	1 / -1	-1 / 1
Tails	-1 / 1	1 / -1

- What should the players do?
  - if blue plays heads, red wants to play heads
  - if red plays heads, blue wants to play tails
  - if blue plays tails, read wants to play tails
  - if read plays tails, blue wants to play heads…
- What about random choice? 50-50?

Example: Rock-Paper-Scissors

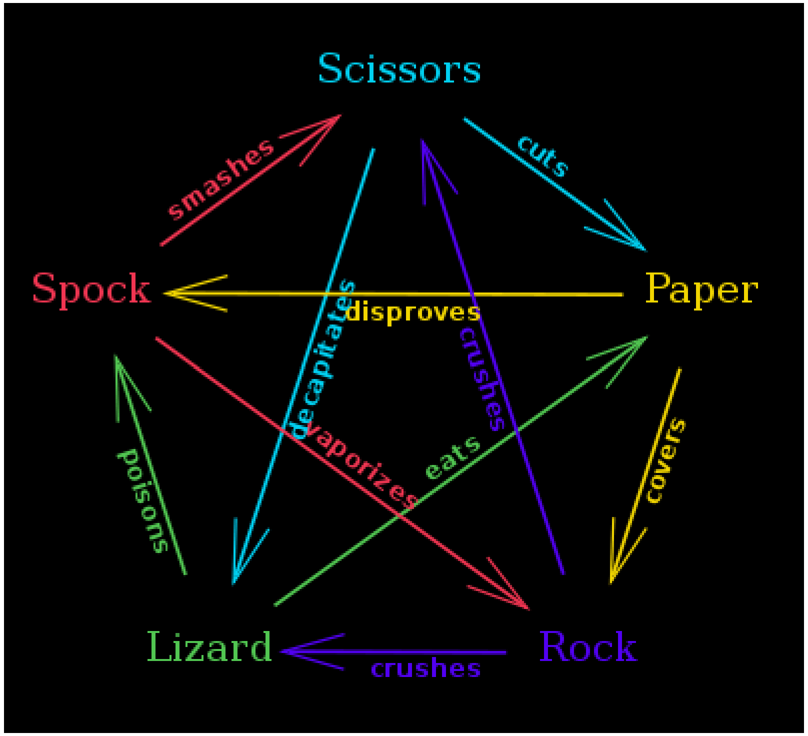
	R	P	S
R	0 / 0	-1 / 1	1 / -1
P	1 / -1	0 / 0	-1 / 1
S	-1 / 1	1 / -1	0 / 0

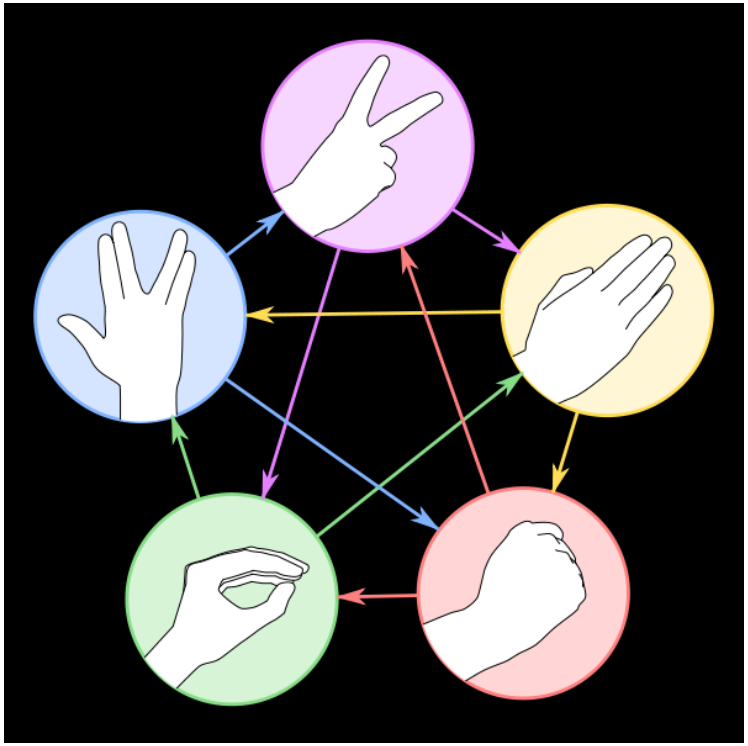


- Just matching pennies with more moves

Year	World Champion	Country
2002	Peter Lovering	Canada
2003	Rob Krueger	Canada
2004	Lee Rammage	Canada
2005	Andrew Bergel	Canada
2006	Bob Cooper	United Kingdom
2007	Andrea Farina	USA
2008	Monica Martinez	Canada
2009	Tim Conrad	USA

or… Rock Paper Scissors Lizard Spock





Example: Prisoner's Dilemma

	Cooperate	Defect
Cooperate	3 / 3	0 / 4
Defect	4 / 0	1 / 1

- What should the players do?
  - only "rational" thing to do is for both to defect
  - but both are worse off

Example: Coordination Game (I)

	A	B
A	1 / 1	0 / 0
B	0 / 0	1 / 1

- An example of a *common interest game*
  - each agent receives the same payoff (sometimes we write only one payoff in each cell)
    - e.g., want to meet at Starbucks or Second Cup
  - no competition, only issue is *coordination*

Example: A “Random” 2x3 Game

	A	B	C
A	10 / 0	0 / 10	2 / 6
B	4 / 7	5 / 6	8 / 8
C	9 / 10	3 / 8	6 / 2

Best Responses and Equilibria

- Notation: $\boldsymbol{a}_{-i}$ refers to pure strategy profile $\boldsymbol{a}$ with action for agent i removed
- Action $a_i \in A_i$ is a *best response* to $\boldsymbol{a}_{-i}$ iff

$$u_i(a_i \cdot \boldsymbol{a}_{-i}) \geq u_i(b \cdot \boldsymbol{a}_{-i}) \text{ for all } b \in A_i$$

- Let $BR_i(\boldsymbol{a}_{-i})$ denote the set of best responses
- A *pure strategy Nash equilibrium* is any pure strategy profile $\boldsymbol{a}$ s.t. $a_i \in BR_i(\boldsymbol{a}_{-i})$ for all i
  - intuitively, an equilibrium is stable in the sense that no agent has any incentive to *unilaterally deviate*

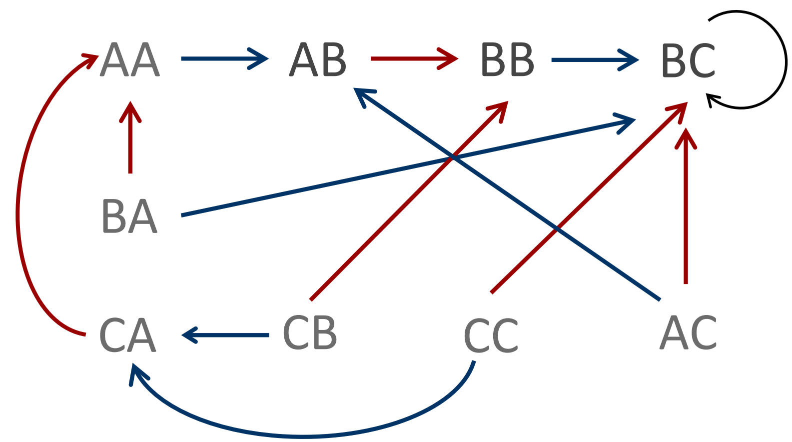
Pure Equilibria in Examples

		Blue Robot: Coffee	Blue Robot: Tea
Red Robot	Coffee	0 / 0	10 / 8
	Tea	8 / 10	0 / 0

- Battlebots:
  - <*C*,*T*> and <*T*,*C*> are both pure equilibria
  - which is better? which would each prefer? how to decide?
  - *equilibrium selection* a critical issue in predicting behavior

Pure Equilibria in Examples

	A	B	C
A	10 / 0	0 / 10	2 / 6
B	4 / 7	5 / 6	8 / 8
C	9 / 10	3 / 8	6 / 2



- Random 3x3 game:
  - can follow best response paths
    - *from any profile consider move one agent would to maximally improve payoff; if no player moves, then it's a pure NE*
  - only pure equilibrium is <*B,C*> (payoff 8,8)

Pure Equilibria in Examples

	Cooperate	Defect
Cooperate	3 / 3	0 / 4
Defect	4 / 0	1 / 1

- Prisoners Dilemma
  - *<Defect, Defect>* is only equilibrium
  - note: NE is *not even Pareto optimal* (let alone socially optimal)
  - is this good? bad? how would you fix it?

Pure Coordination Game

	Starbucks	SecondCup
Starbucks	4 / 4	0 / 0
SecondCup	0 / 0	4 / 4

- Common interest game
  - two pure NE
  - where do we meet? Any suggestions on how you might resolve the issue in practice?

Unbalanced Coordination Game

	Starbucks	SecondCup
Starbucks	4 / 4	0 / 0
SecondCup	0 / 0	3 / 3

- Common interest game, but unbalanced
  - two pure NE
  - choice “seems” easier now

Battle of the Sexes

	Sports Arena	Opera
Sports Arena	2 / 1	0 / 0
Opera	0 / 0	1 / 2

- Famous game (Luce and Raiffa 1957)
  - two pure NE
  - where do we meet?
  - equilibrium selection paramount

Stag Hunt (Rousseau)

	Hunt Stag	Hunt Hare
Hunt Stag	4 / 4	0 / 3
Hunt Hare	3 / 0	3 / 3

- Stag Hunt a classic example of equilibrium selection
  - If two hunter's team up, they will bag a stag; but if one tries alone, sure to fail
  - Safer option: hunt alone for a hare (no cooperation needed)
  - two (deterministic) NE: *(stag,stag)* and *(hare,hare)*
    - *(s,s)* gives better payoff, but poor outcome if partners opts out
    - *(h,h)* gives lesser payoff, but guaranteed
    - which would you predict? why? can you justify either prediction?

Example: Coordination Game (II)

	A	B	C
A	10	0	-40
B	0	2	0
C	-40	0	10

- Three pure NE
  - two are very attractive: (A,A) and (C,C)
  - but penalties make miscoordination “painful”
  - should you take the safe alternative (B,B)?

Three-Player Example: Voting Game

For (M1)

	For (L)	Against (R)
For (U)	1 / 1 / 1	1 / 2 / 1
Against (D)	2 / 1 / 1	0 / 0 / -1

Against (M2)

	For (L)	Against (R)
For (U)	1 / 1 / 2	-1 / 0 / 0
Against (D)	0 / -1 / 0	0 / 0 / 0

- ***Example:*** three MLAs: (closed) voting on salary increase. Vote passes if at least two vote for it. All want the raise, but would prefer to vote against it (and have the other two vote for it!)
  - P_k if raise passes, k votes against: 2
  - P_k if raise passes, k votes for: 1
  - P_k if raise fails, k votes against: 0
  - P_k if raise fails, k votes for: -1
- Four pure NE:
  - any pair vote for a raise, third votes against (3)
  - all three vote against (1)

Equilibrium Selection

- A very subtle topic, and important in practical settings
  - e.g., studied extensively in political science, behavioral economics, social psychology, etc.
- Focal points one “heuristic” means of resolving issue
  - factors external to the game may influence choices
    - *past history (met at Starbuck’s last week)*
    - *social conventions (generally agree to pass on the right)*
  - factors internal to game structure may matter
    - *as in the unbalanced coordination game, or penalty game*
    - *or risk attitude (Stag Hunt)*
  - learning often proposed as explanation for good eq. selection

Pure Equilibria in Examples

	Heads	Tails
Heads	1 / -1	-1 / 1
Tails	-1 / 1	1 / -1

- Matching Pennies:
  - no pure equilibrium
  - why? (consider the “infinite regress”)
  - rock-paper-scissors-lizard-Spock: same thing…

Mixed Strategies

- A *mixed strategy* σ_i for *i* is a distribution over A_i
  - a randomized choice (uncorrelated with other agents)
- A *mixed strategy profile*: $\sigma = <\sigma_1 \ldots \sigma_n>$
  - pure strategy profile a special case (singleton support for each *i*)
- Utility of strategy profile to player *i* : $u_i(\sigma) = E[\, u_i(\mathbf{a}) \mid \sigma \,]$
  - randomization of each agent is independent
- Define $BR_i(\sigma_{-i})$ in the standard way:

$$u_i(\sigma_i \cdot \sigma_{-i}) \geq u_i(\tilde{\sigma}_i \cdot \sigma_{-i}) \text{ for all } \tilde{\sigma}_i \in \Delta(A_i)$$

- A *Nash equilibrium* is any (mixed) strategy profile σ

  s.t. $\sigma_i \in BR_i(\sigma_{-i})$ for all *i*

Mixed Equilibria: Matching Pennies

- In MP: only NE is *<{0.5H,0.5T}, {0.5H,0.5T}>*
- Why is this an equlibrium?
  - If Even plays *0.5/0.5,* then for Odd we have

    EU(H) = EU(T) = EU(any mixture)
  - So no incentive for Odd to deviate from *0.5/0.5*
  - Reasoning for Even is symmetric
- Why is this a unique equilibrium?
  - If Even plays *p:H, (1-p):T* for any $p > 0.5$, best response for Odd is T with probability 1.0 *(and analogous switch if $p < 0.5$)*
  - But then Even would not want to stay with $p(H) > 0.5$
  - Intuitively, if Even does anything but 50-50, Odd will exploit it and get a higher payoff (which means Even gets a lower payoff)

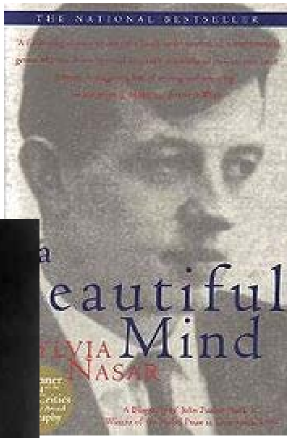
Mixed Equilibria: BattleBots

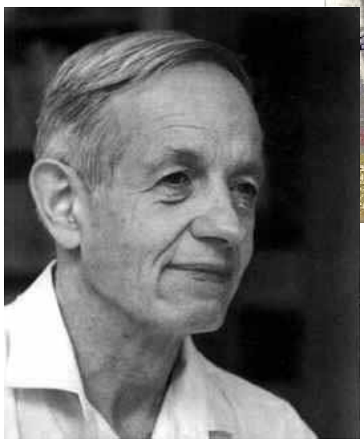
- Battlebots has a mixed NE
  - Both robots use same strategy
    - *Pr(coffee) = 5/9*
    - *Pr(tea) = 4/9*
- Why is this an equilibrium?
  - If Red plays (5/9; 4/9)
    - For Blue: EV(coffee) = 5/9(0) + 4/9(10) = 40/9
    - For Blue: EV(tea)= 5/9(8) + 4/9(0) = 40/9
  - No (positive) incentive for Blue to deviate

	Coffee	Tea
Coffee	0 / 0	10 / 8
Tea	8 / 10	0 / 0

Mixed Equilibria: Properties

- If mixed σ_i is a BR for i, then each a_i s.t. $Pr(a) > 0$ must also be a BR (otherwise EU would be higher without it)
  - this provides a way to compute equilibria in some cases
  - exercise: do it for BattleBots
- Some games have only *mixed* equilibria
  - e.g., Matching Pennies
- All (finite) games have some NE
  - famous result due to John Nash





Existence of Nash Equilibria

- **Nash's Celebrated Theorem:** Any finite normal form game has a (mixed strategy) equilibrium
- Proof intuition:
  - strategy profiles *S* form a compact (closed, bounded), convex subspace of R^m (where *m* is number of agent-action pairs)
  - *best response correspondence* (mapping) $BR: S \rightarrow SS$ (subsets of S)
    - Let $BR(\sigma)$ is set of all best response profiles for σ
  - $BR(\sigma)$ is a convex subset of *S* (recall that any linear combination of BRs must also be a BR)
  - The correspondence BR is also *upper-hemicontinuous*
    - This is a simple generalization of continuity for correspondences
  - Kakutani fixed-point theorem: BR must have a fixed point
    - in other words, there is some σ s.t. $\sigma \in BR(\sigma)$
  - By definition: any fixed point of the BR correspondence is an equilibrium

Existence of Equilibria

- The proof is entirely nonconstructive
- Computation of NE is generally quite difficult
  - we'll return to this
- Special cases: can be solved easily
  - identical interest games
  - two-person, zero-sum games

Interpretation of Mixed Equilibria

- *Genuine randomization* of choices to prevent exploitation
  - soccer penalty kicks, bluffing in poker, many contests and sports, ...
- Conditions on *stable beliefs* (or expectations, conjectures) about how opponent will play
  - if I believe you will play q = 0.5, and I believe that you believe I will play p = 0.5, and I believe that you believe that I believe that you will play q = 0.5, and …
- Learning interpretations: beliefs based on experience with opponent (or related opponents)
  - I've (P2) seen you (P1) play heads 47 times ,tails 53 times: I assume strategy is p = 0.47
  - My best response is heads (q = 1.0)
  - But you learn in similar fashion and *also* adopt best response to what you've observed
  - In some cases, observed frequencies converge to NE
  - But actually play may often be deterministic (and quite cyclic, not random at all)
- Evolutionary interpretations (evolutionary stable strategies)

Two-Person, Zero-Sum (2PZS) Games

- 2PZS games embody “pure competition”
  - two players (1,2); for all joint actions $\mathbf{a}$, $u_1(\mathbf{a}) = -u_2(\mathbf{a})$
  - usually write $u(\mathbf{a})$ assuming 1 maximizes, 2 minimizes
  - more generally, “constant sum”
  - MP, chess, soccer, etc.
- An agent *maxminimizes* if it selects a strategy whose guaranteed payoff is best
  - i.e., you choose best *a* assuming other guy does worst *a* for you
  - Player 1 plays $\sigma^*_1 \in \mathrm{argmax}_{\sigma 1}\ \min_{\sigma 2}\ u(\sigma_1,\sigma_2)$
  - Player 2 plays $\sigma^*_2 \in \mathrm{argmin}_{\sigma 2}\ \max_{\sigma 1}\ u(\sigma_1,\sigma_2)$

Two-Person, Zero-Sum Games

- **Theorem (von Neumann):**
  - In a 2PZS game, a strategy profile is a Nash equilibrium iff both players maxminimize in that profile
  - Thus if $<\sigma_1, \sigma_2>$ is an equilibrium, then
    $u(\sigma_1, \sigma_2) = \max \min u(\sigma_1, \sigma_2) = \min \max u(\sigma_1, \sigma_2)$
- So, while there may be multiple equilibria for a 2PZS game, they all have the same value to each player
- Can be solved by linear programming (next slide)
- In MP: $\max \min u(\sigma_1, \sigma_2) \rightarrow <0.5,0.5>$
  - *if Even does more than 0.5 heads, Odd picks tails and Even loses more than ½ the time*

LP for 2PZS Game (basic & slack formulations)

$$\min U_1^*$$
$$s.t. \sum_{a_2^k \in A_2} u_1(a_1^j, a_2^k) s_2^k \leq U_1^* \;\; \text{for all } a_1^j \in A_1$$
$$\sum_{a_2^k \in A_2} s_2^k = 1; \quad \text{and all} \quad s_2^k \geq 0$$

$$\min U_1^*$$
$$s.t. \sum_{a_2^k \in A_2} u_1(a_1^j, a_2^k) s_2^k + r_1^j = U_1^* \;\; \textit{for all } a_1^j \in A_1$$
$$\sum\nolimits_{a_2^k \in A_2} s_2^k = 1; \; \textit{all} \;\; s_2^k \geq 0; \textit{all} \;\; r_1^j \geq 0$$

- Vars:
  - Player 2's mixed strategy (probs s_k)
  - Player 1's minimax value
- Aim: min 1's payoff s.t. constraint that it's at least as great as the expected value of his best action
  - Slack formulation (useful in LCP to follow)
  - Slack for 1's action a: how much worse than a best response

Computation of Equilibria

- Computation of NE is generally quite difficult
- Two-player games, classic algorithm is Lemke-Howson, relies on formulation as a *linear complementarity* problem
  - Looks like LP for 2PZS, but with complementarity constraints to ensure that each player only mixes over pure best responses given strategy of other player (nonlinear constraint)
- In general, no great strategies other than to rely on some form of enumeration (e.g., use LCPs as subroutines)
  - simple enumerative algorithm: identify support sets and see if some mixture of these is a NE; finding probabilities for *fixed support sets* is easy: a linear (feasibility) program for 2 players (more generally, polynomial feasibility program)
  - see Chapter 4 of S&LB for more discussion of this

LCP for 2-player General Sum Game

$$\text{find } U_1^* \, U_2^*$$

$$s.t. \sum_{a_2^k \in A_2} u_1(a_1^j, a_2^k) s_2^k + r_1^j = U_1^* \;\; \textit{for all}\, a_1^j \in A_1$$

$$\sum_{a_1^j \in A_1} u_2(a_1^j, a_2^k) s_1^j + r_2^k = U_2^* \;\; \textit{for all}\, a_2^k \in A_2$$

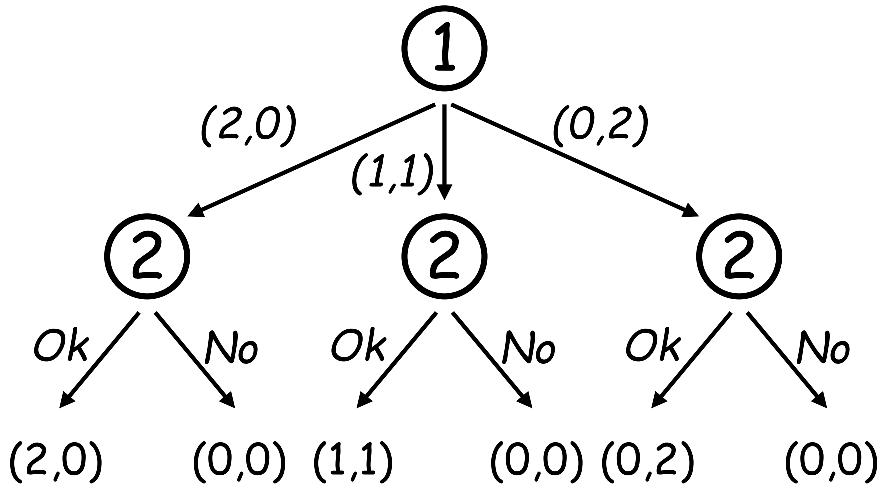
$$\textstyle\sum_{a_2^k \in A_2} s_2^k = 1; \;\; \textit{all} \;\; s_2^k \geq 0; \textit{all} \;\; r_1^j \geq 0 \;\; \textit{(and for the other player)}$$

$$\boldsymbol{r_1^j s_1^j \geq 0 \;\; \text{and}\; r_2^k s_2^k \geq 0 \;\; \textit{for all}\, a_1^j \in A_1, a_2^k \in A_2}$$

- Much like the LP for 2PZS, but with an objective, and the nonlinear complementarty constraints: agent can only assign an action positive probability if it has zero slack (i.e., is a best response)

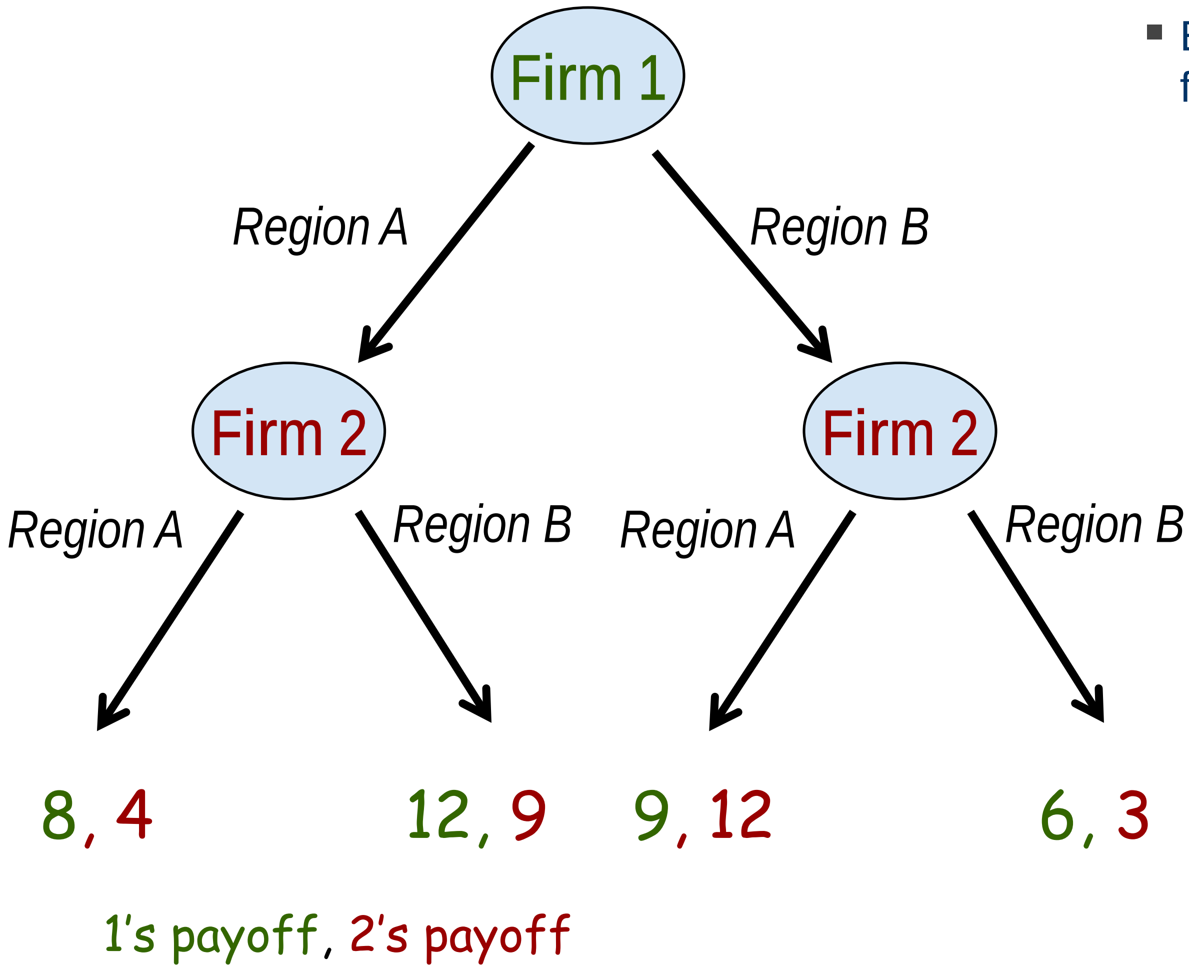
Extensive Form Games

- Normal form (matrix) very general, but games seem somewhat limited
  - Players move simultaneously and outcome determined at once
  - No observation, reaction, etc.
- Most games have a *dynamic* structure (turn taking)
  - Chess, tic-tac-toe, cards games, soccer, corporate decisions…
  - See what "opponent" does before making move



Sharing Game (S&LB):
-Player 1 offers a split of 2 goods
-Player 2 accepts or rejects

A More Realistic Sharing Game



- Example: two firms competing for market in two areas
  - Each firm, 1 and 2, can tackle one area only
  - Total revenue ($M) in Area A: 12, Area B: 9
  - If firm is alone in one area, get all of that area's revenue
  - If both firms target same area, "first mover" gets 2/3, second 1/3
  - Firm 1 makes first move, Firm 2 chooses *after* Firm 1
  - Critical: Firm 2 observers Firm 1's choice before moving!

Perfect Information Games

- Looks exactly like a decision tree (as we will discuss when we look at sequential problems), except:
  - each node *controlled by a player*, and corresponds to a complete history of the game
  - edges correspond to actions
  - payoff vector at each *terminal* node
  - often no *chance nodes* internal to tree, but easy enough to do
  - see formal definition in S&LB
- Since player knows entire history, this is called a *perfect information, extensive form (PIEF) game*
  - knows all prior moves, stochastic outcomes
  - much like *full observability* in MDPs

Strategies and Equilibria

- A *pure strategy* for i in a PIEF-game is any function that assigns an action $a_i \in A(h)$ to any node/history h owned by i
  - like policies in decision trees: each node must be assigned, whether it is realizable or not
  - unlike (seemingly) strategic/normal form games, my decision can depend on what you did in the past
- A *Nash equilibrium* is any strategy profile σ s.t. $u_i(Term(\sigma_i, \sigma_{-i})) \geq u_i(Term(\sigma'_i, \sigma_{-i}))$ for all i, σ'_i
  - note: we'll focus on deterministic strategies (why?)

Determining Nash Equilibria

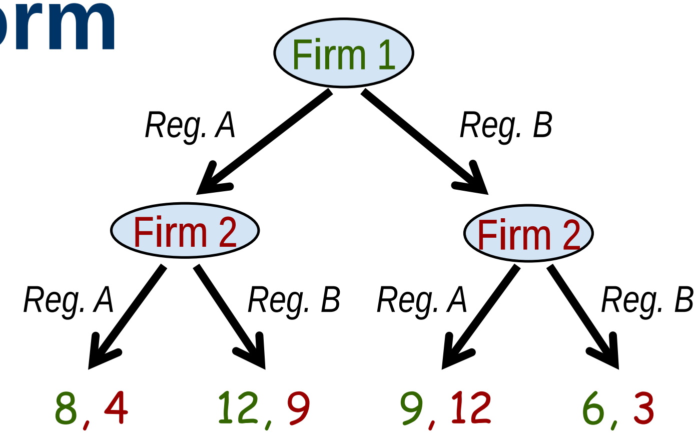
- Backward induction of game tree (like in a decision tree) will determine some, but not all NE (we'll see this in a moment)
- But PIEF games easily converted to normal form
  - consider each pure strategy in tree to be an action
    - e.g., strategy [ONO]: "will accept if offered [2,0], reject if [1,1], accept if [0,2]"
  - NE in normal form identical to NE in extensive form (by defn)

	[2]	[1]	[0]
[000]	2,0	1,1	0,2
[00N]	2,0	1,1	0,0
[0N0]	2,0	0,0	0,2
[0NN]	2,0	0,0	0,0
[N00]	0,0	1,1	0,2
[N0N]	0,0	1,1	0,0
[NN0]	0,0	0,0	0,2
[NNN]	0,0	0,0	0,0

First player strategy: make offer (column) 2=[2,0], 1=[1,1], 0=[0,2]
Second player strategy: OK or NO for each of the three possible offers (row)

Identify two different (classes of) pure strategy equilibria in this game. Which is most natural?

Conversion to Normal Form



Firm 1 : market in region A or B

Firm 2: four strategies

- AA/AB (i.e., always A): A if Firm 1 does A, A if 1 does B
- BA/AB: B if Firm 1 does A, A if 1 does B
- AA/BB: ditto
- BA/BB (always B): ditto

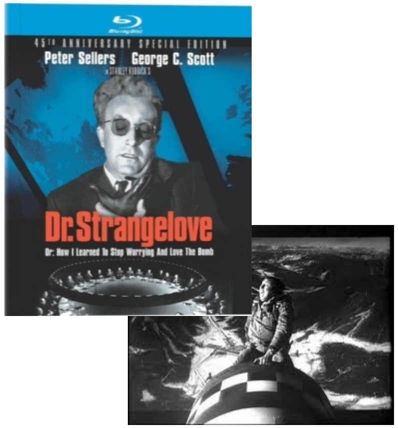
	AA/AB	AA/BB	BA/AB	BA/BB
A	8,4	8,4	12,9	12,9
B	9,12	6,3	9,12	6,3

Game has three NE:

- A/BA,AB and A/BA,BB(equivalent outcomes)
  - fact that 1 chooses A makes 2's choice "when 1 picks B" irrelevant
- B/AA,AB: *Firm 2 threatens to move into A if Firm1 does*
  - Threat sufficient to make Firm 1 move to B (alone): full amount (9) of smaller payoff is better than 2/3-share (8) of larger payoff
  - Is this threat credible?

Subgame Perfect Equilibria

- A *subgame* is the game corresponding to a specific subtree (written *SG(h)* for node *h*)
- A *subgame perfect equilibrium (SGPE)* is a profile σ such that $\sigma[h]$ (i.e., σ restricted to *h*) is a NE for *SG(h), for every node h*
- Subgame perfection rules out NE *B/AA,AB* in the previous game
  - It is not an equilibrium at “node 2 subgame” (2’s choice after 1 picks A)
  - SGP ignores history (hence ability to “punish” for past actions)
  - intuitively, punishment not rational/threats not credible (in one-shot game)
- If one could precommit to a strategy, threats could be credible
  - e.g., computer program, legal contract, …
  - however, such precommitment is *a move* in a more involved game
  - retaliation in future interactions possible, but also a different game (repeated, or Markov game)
  - SGPE is the most natural form of equilibrium for EF



The Ambassador reveals that his side has installed a doomsday device that will automatically destroy life on Earth if there is a nuclear attack against the Soviet Union. The American President expresses amazement that anyone would build such a device. But Dr. Strangelove… admits that it would be "an effective deterrent... credible and convincing.“

Strangelove explains the technology behind the Doomsday Machine and why it is essential that not only should it destroy the world in the event of a nuclear attack but also be fully automated and incapable of being deactivated. He further points out that the "whole point of the Doomsday Machine is lost if you keep it a secret".

When asked why the Soviets did not publicize this, Ambassador de Sadeski sheepishly answers it was supposed to be announced the following Monday at the (Communist) Party Congress because "the Premier loves surprises."

From en.wikipedia.org/wiki/Dr._Strangelove#Plot

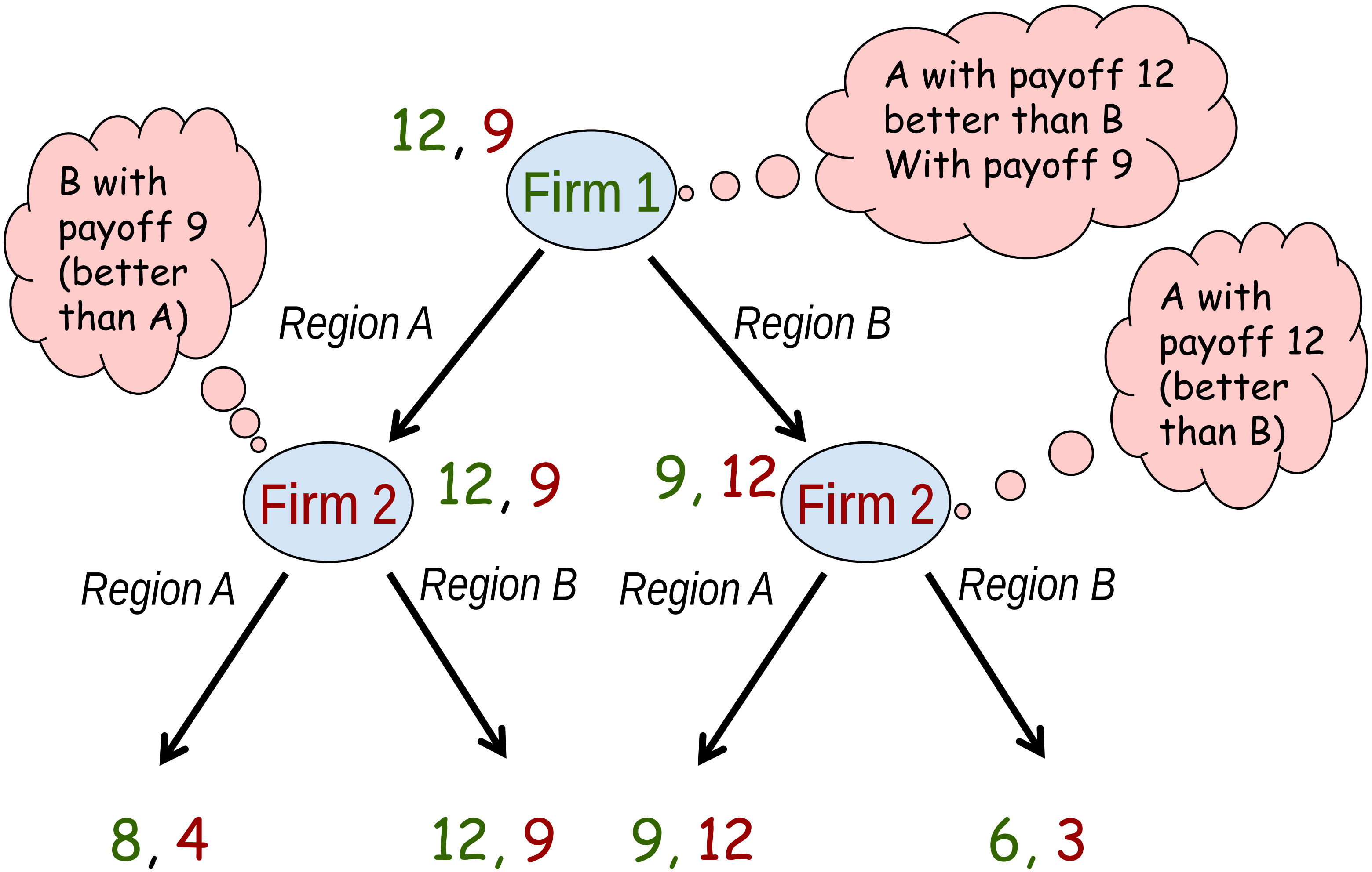
SGPE: Notes

- Chance nodes easy to incorporate into picture
  - utility of strategy profile given by expectation
  - history must still be observable (like decision trees)
- Randomization comes in two forms
  - randomize over the strategies σ_i (i.e., mixture within normal form)
  - randomize choices at specific nodes (independently); called *behavioral strategies* (choices at different nodes are uncorrelated)
- **Theorem (Kuhn):** There is a SGPE for any finite extensive form game
  - fairly obvious: backward induction (DP) to calculate it
  - unique (except for ties); no role for randomization

Backward Induction

- Simple way to compute a SGPE
  - back values up the tree exactly as in a decision tree
  - only differences
    - you back up entire payoff vectors
    - at choice node controlled by player *i*, action is the one maximizing that player's value
  - leaves no role for threats, action choice made at nodes independent of path taken to get to that node

Backward Induction: Toy Illustration



Criticisms of Backward Induction

- Centipede game show difficulties of NE (even subgame perfect NE) in predicting behavior
  - players alternate decisions: across (continue), down (terminate)
- Only NE (hence only SGPE) is (D,D,D,D,D)
  - seems more problematic as game goes on longer!

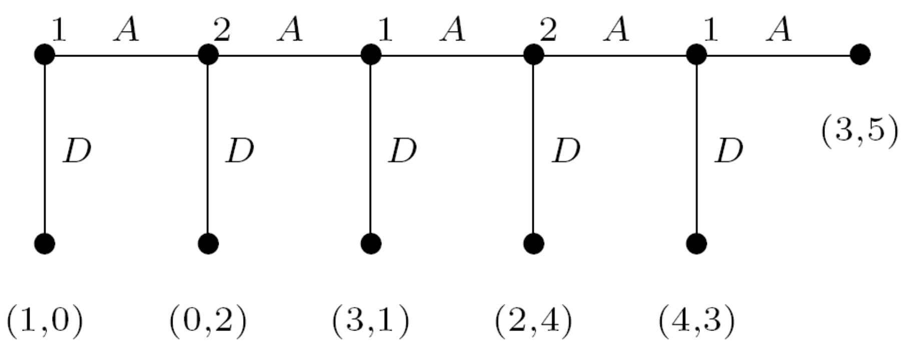


Figure 5.9: The Centipede game.

Fig from Multiagent Systems, Shoham and Leyton-Brown, 2009

Different (Related) Criticism of NE

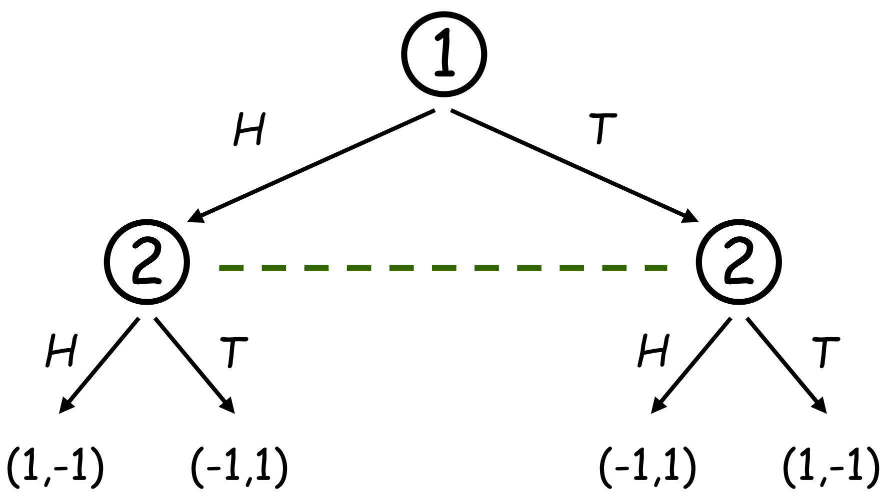
- Traveller's dilemma
  - backdrop: two travellers lose suitcase with *identical items* (e.g., valuable souvenirs) on airline, airline customer service rep trying to assess ("elicit") true value
  - two players, must write a dollar value between 2 and 100
  - if numbers are the same, each get number they wrote
  - if different, each gets lowest number, but bonus of $2 to person with lower number, *penalty* of $2 to person with higher number
  - only NE (reasoning similar to SGPE) is for each to write down $2
    - 100 dominated by 99 (no matter what other guy does, your payoff is better if you write down 99 instead of 100)
    - repeating this reasoning, driven down to only NE of (2,2)
  - experimentally, people rarely do this: claims are much higher in the lab; but they are influenced by the penalty: higher penalties cause people to make lower claims

Imperfect Information EF Games

- Matching pennies can't be modeled as a perfect info EFG
  - the simultaneity of moves can't be captured
  - many other truly dynamic games have partial obervability, partial simultaneity, etc. (like move from MDPs to POMDPs)
- *Imperfect information games* capture this
  - *information set* for a player *i*: a subset of nodes owned by player *i* in an EFG; intuitively, this set reflects a set of histories that *can't be distinguished* by *i*
  - *the set of moves at each node in an info set must be the same*
  - *a player's strategy is restricted to choosing same move (or mixed move) at each node in a given information set*

Imperfect Information

- See S&LB for formal defn; somewhat informally
  - each player’s nodes must be *partitioned* into info sets
  - strategy maps info sets into action choices
  - we can no longer ensure pure strategy NE exist; in general mixed strategies needed (since the full observability of moves no longer holds)
  - NE defined in this strategy space in usual way



A dashed line (sometimes an oval) connects nodes in the same info set

Imperfect Info Games: Notes

- Every NF game can be converted to an imperfect info EF game
  - what is the obvious way?
- A natural way to incorporate some lack of information about states of nature, exogenous events
- *Sequential equilibria:* analog of SGP when we have incomplete info
  - players need to form beliefs over possible histories at each information set and reason about other player's beliefs at *their* information sets
  - *beliefs and strategies both* must be equilibrium
- Backward induction no longer makes sense
  - can resort to normal form computation (or use "sequence representation")
- Usually we assume *perfect recall*
  - history of i's *own* actions at nodes in any information set must be identical (*i* remembers everything it did)
  - we can restrict randomization to randomizing action choices independently at each information set (no extra power in randomizing over complete strategies)
  - we call such strategies *behavioral strategies* (exactly as in perfect info EFGs)

Imperfect vs. Incomplete Info Games

- Don't confuse *imperfect* information games with *incomplete* information games
  - incomplete info games are games where full game structure unknown
  - we'll discuss these later; vital in mechanism design

Repeated and Markov Games

- Extensive form games capture some sequential reasoning, but are "hand-crafted" like decision trees
- MDPs and POMDPs are more "generic"
- Repeated games, Markov games extend normal form games by allowing agents to repeatedly interact, possibly in different contexts
- *Repeated games:* matrix game played over finite or infinite horizon
  - payoffs are sum of payoffs received over all games
  - strategies map history of play into (randomized) action choice
  - e.g., repeated Prisoner's dilemma: much richer strategy space, allows threats of retaliation, promises of cooperation, etc.
- *Markov games:* similar, but different "games" played at each stage
  - at each stage, agents are in a state (as in an MDP)
  - (randomized) action choices at each stage determine immediate payoffs
  - action choices also determine (stochastic) state transition: next stage
  - goal is to maximize total or discounted sum of payoffs