Intelligent Autonomous Agents:

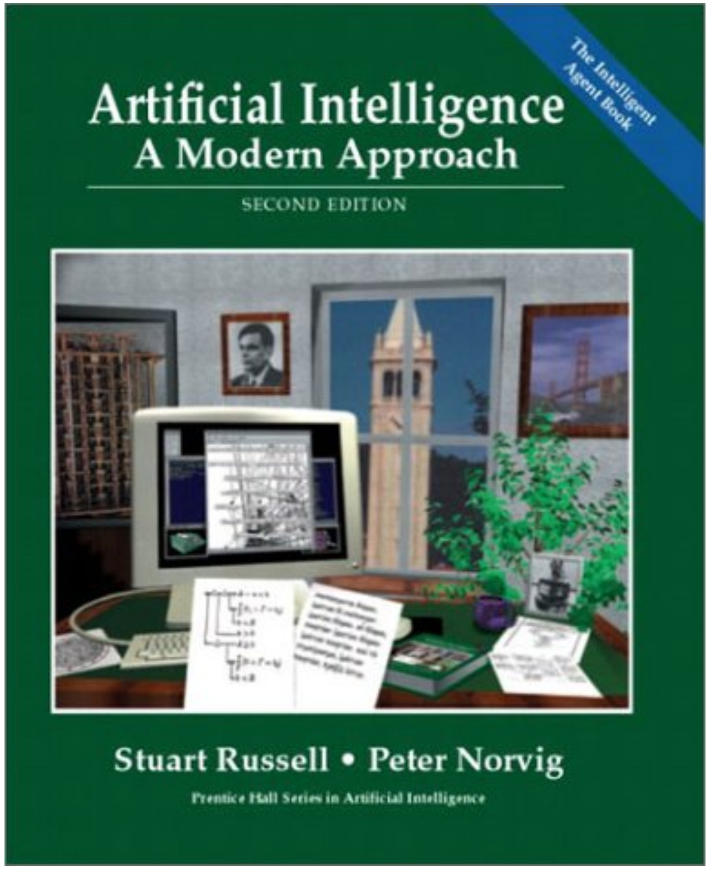
Agents and Rational Behavior

Lecture 10: Multiple Agents and Game Theory

Ralf Möller

Hamburg University of Technology

Literature

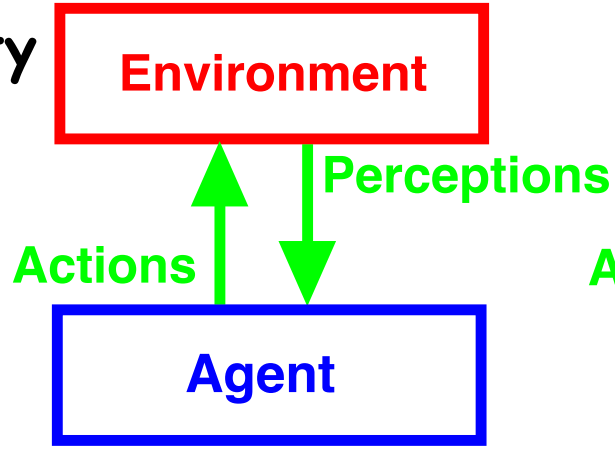


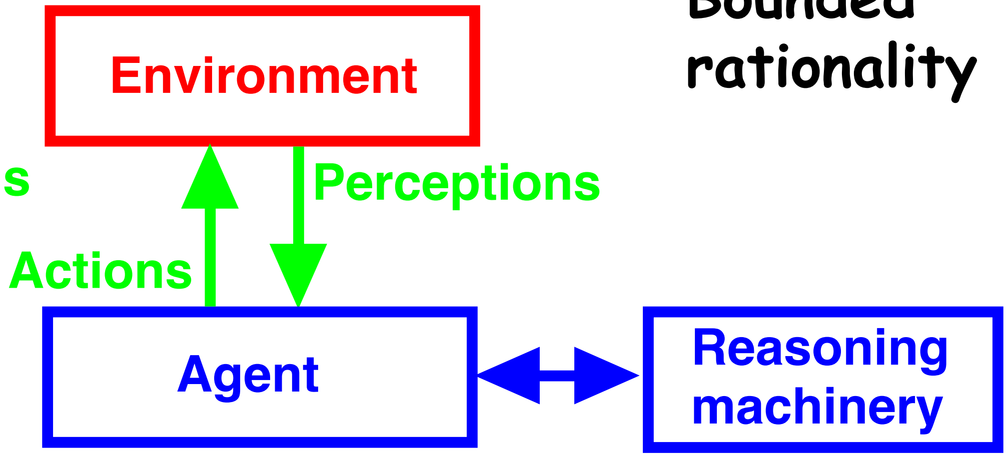
- Chapter 17

Presentations from CS 886
Advanced Topics in AI
Electronic Market Design
Kate Larson
Waterloo Univ.

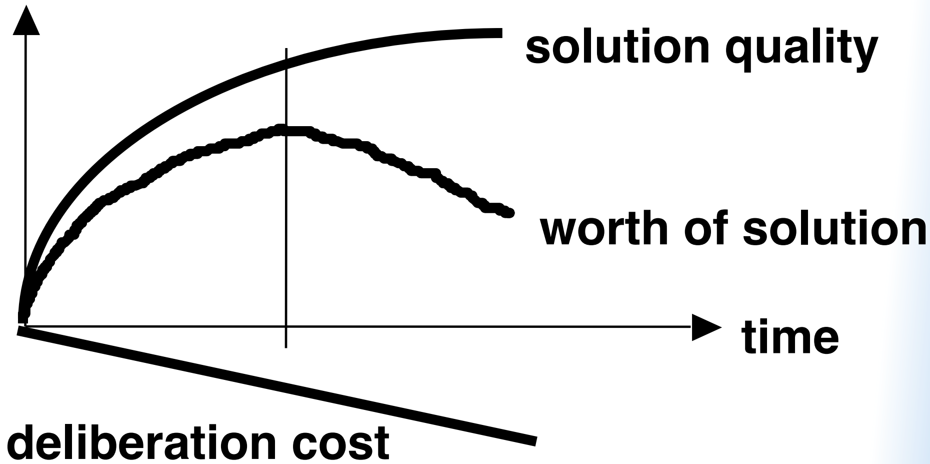
Full vs bounded rationality

Full rationality





Bounded rationality



Descriptive vs. prescriptive theories of bounded rationality

Multiagent Systems: Criteria

- Social welfare: $\max_{outcome} \sum_i u_i(outcome)$
- Surplus: social welfare of outcome – social welfare of status quo
  - Constant sum games have 0 surplus.
  - Markets are not constant sum
- Pareto efficiency: An outcome o is Pareto efficient if there exists no other outcome o' s.t. some agent has higher utility in o' than in o and no agent has lower
  - Implied by social welfare maximization
- Individual rationality: Participating in the negotiation (or individual deal) is no worse than not participating
- Stability: No agents can increase their utility by changing their strategies
- Symmetry: No agent should be inherently preferred, e.g. dictator

Game Theory: The Basics

- **A game:** Formal representation of a situation of strategic interdependence
  - Set of agents, I (|I|=n)
    - AKA players
  - Each agent, j, has a set of actions, A_j
    - AKA moves
  - Actions define outcomes
    - For each possible action there is an outcome.
  - Outcomes define payoffs
    - Agents' derive utility from different outcomes

Normal form game*

(matching pennies)

		Agent 2	
		H	T
Agent 1	H	-1, 1	1, -1
	T	1, -1	-1, 1

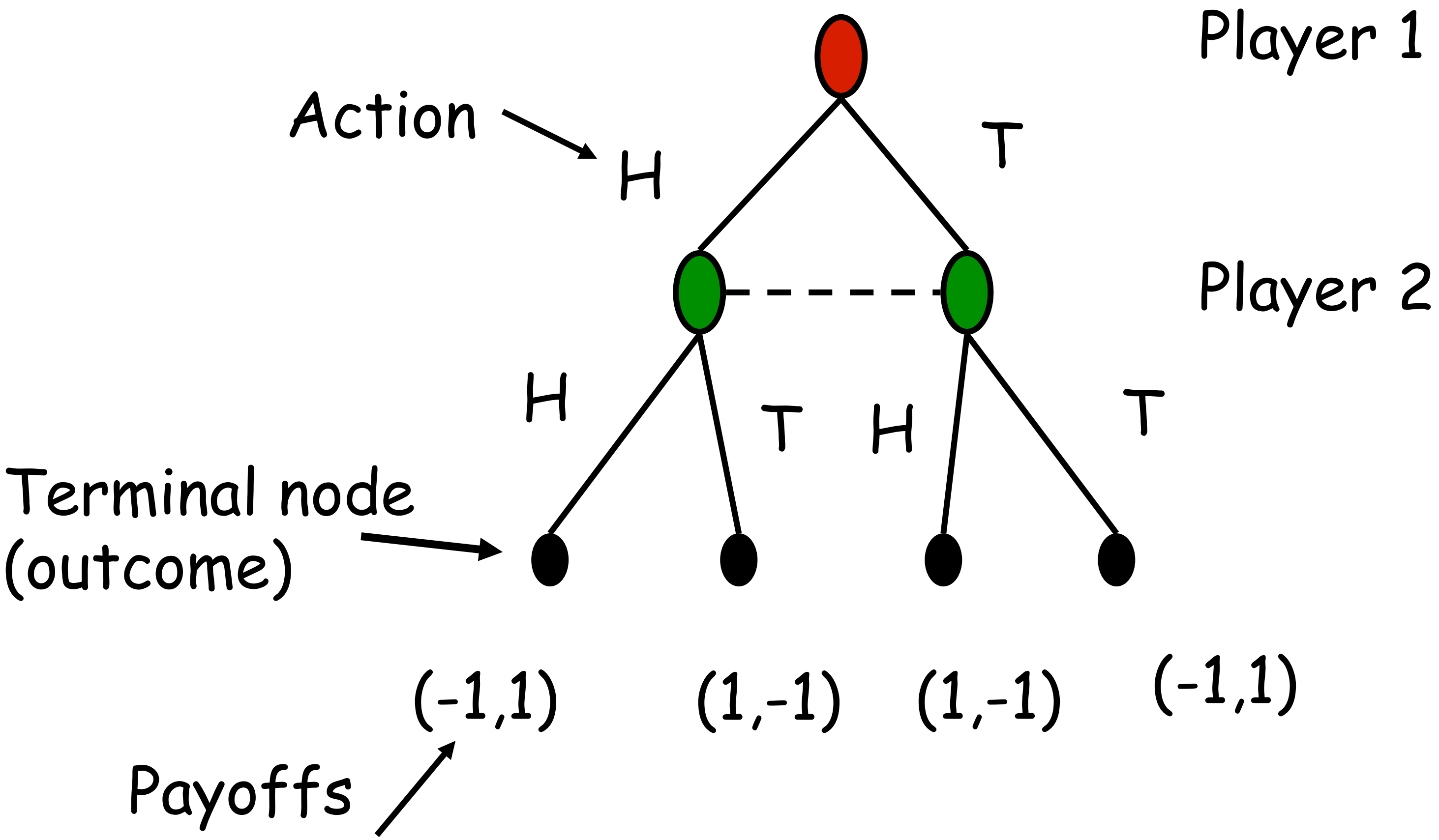
Action

Outcome

Payoffs

*aka strategic form, matrix form

Extensive form game
(matching pennies)



Strategies (aka Policies)

- Strategy:
  - A strategy, s_j, is a **complete contingency plan**; defines actions agent j should take for all possible states of the world
- Strategy profile: $s=(s_1,\ldots,s_n)$
  - $\mathbf{s_{-i} = (s_1,\ldots,s_{i-1},s_{i+1},\ldots,s_n)}$
- Utility function: $u_i(s)$
  - Note that the utility of an agent depends on the strategy profile, not just its own strategy
  - We assume agents are **expected utility maximizers**

Normal form game*
(matching pennies)

Agent 1 \ Agent 2	H	T
H	-1, 1	1, -1
T	1, -1	-1, 1

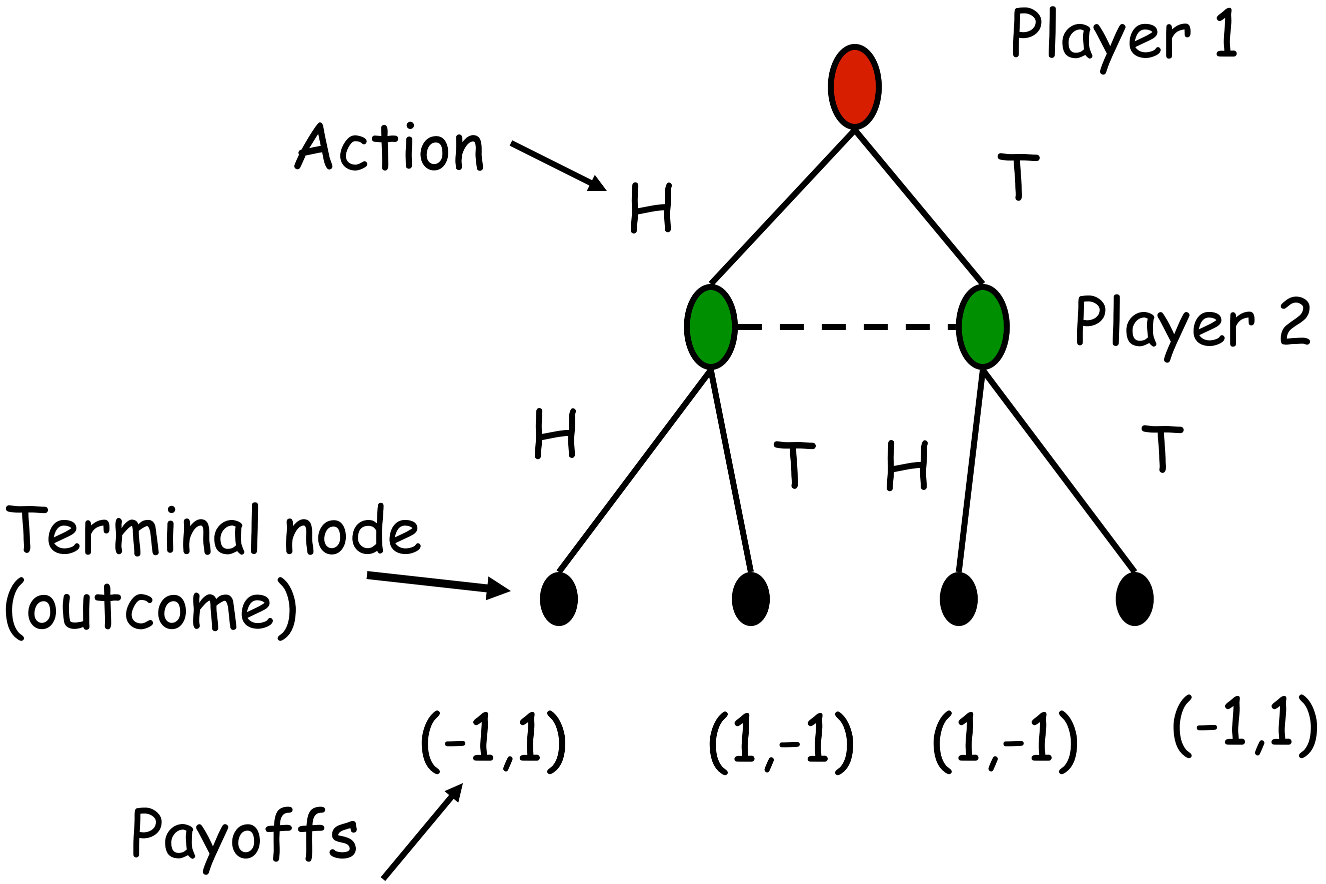
Strategy for agent 1: H

Strategy profile (H,T)

U1((H,T))=1
U2((H,T))=-1

*aka strategic form, matrix form

Extensive form game
(matching pennies)



Strategy for agent 1: T

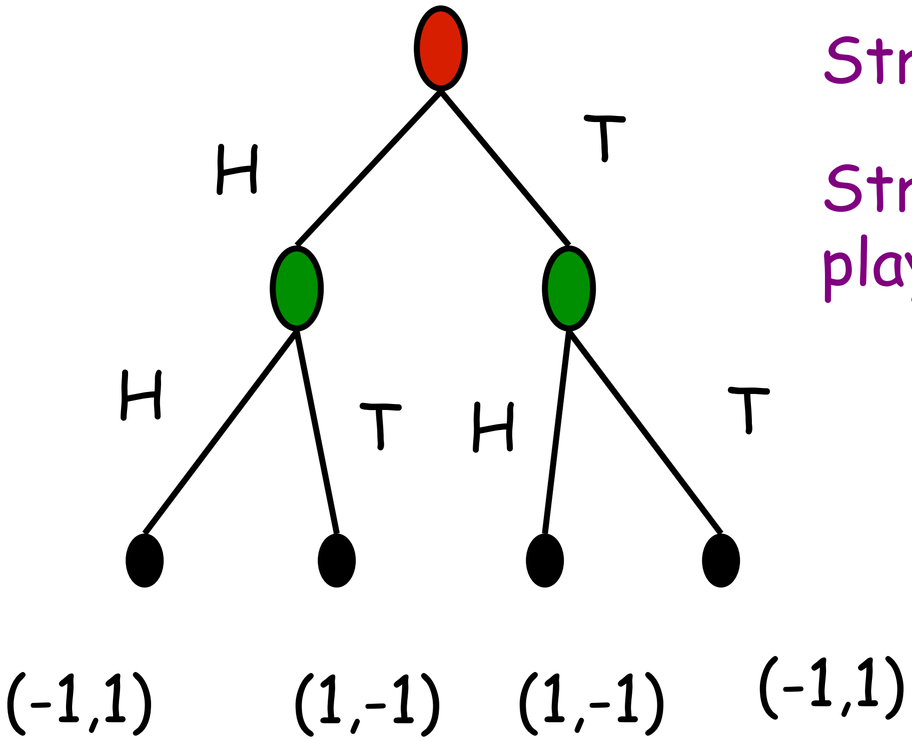
Strategy profile: (T,T)

U1((T,T))=-1

U2((T,T))=1

Extensive form game
(matching pennies, seq moves)

Recall: A strategy is a contingency plan for all states of the game



Strategy for agent 1: T

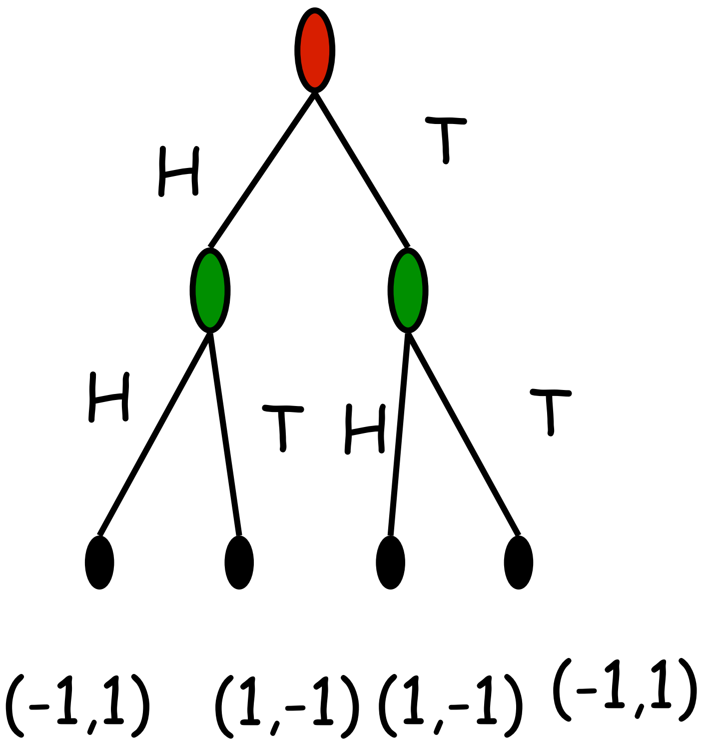
Strategy for agent 2: H if 1 plays H, T if 1 plays T (H,T)

Strategy profile: (T,(H,T))

U1((T,(H,T)))=-1

U2((T,(H,T)))=1

Game Representation



	H,H	H,T	T,H	T,T
H	-1,1	-1,1	1,-1	1,-1
T	1,-1	-1,1	1,-1	-1,1

Potential combinatorial explosion →

Example: Ascending Auction

- State of the world is defined by (x,p)
  - $x \in \{0,1\}$ indicates if the agent has the object
  - p is the current next price
- Strategy $s_i((x,p))$

$$s_i((x,p)) = \begin{cases} p, \text{ if } v_i >= p \text{ and } x=0 \\ \text{No bid otherwise} \end{cases}$$

Dominant Strategies

- Recall that
  - Agents' utilities depend on what strategies other agents are playing
  - Agents' are expected utility maximizers
- Agents' will play best-response strategies

s_i^* is a best response if $u_i(s_i^*, s_{-i}) \geq u_i(s_i', s_{-i})$ for all s_i'

- A dominant strategy is a best-response for all s_{-i}
  - They do not always exist
  - Inferior strategies are called dominated

Dominant Strategy Equilibrium

- A dominant strategy equilibrium is a strategy profile where the strategy for each player is dominant
  - $s^*=(s_1^*,\ldots,s_n^*)$
  - $u_i(s_i^*,s_{-i}) \geq u_i(s_i',s_{-i})$ for all i, for all s_i', for all s_{-i}
- **GOOD**: Agents do not need to counterspeculate!

Example: Prisoner's Dilemma

- Two people are arrested for a crime. If neither suspect confesses, both are released. If both confess then they get sent to jail. If one confesses and the other does not, then the confessor gets a light sentence and the other gets a heavy sentence.

	A: Confess	A: Don't Confess
B:Confess	B=-5, A=-5	B=-1, A=-10
B:Don't Confess	B=-10, A=-1	B=-2, A=-2

Dom. Str. Eq

Pareto Optimal Outcome

Example: Split or Steal

Does communication help?

	A: Steal	A: Split
B:Steal	B=0, A=0	B=100, A=−10
B:Split	B=−10, A=100	B=50, A=50

Dom. Str. Eq

Pareto Optimal Outcome

Example: Vickrey Auction (2nd price sealed bid)

- Each agent i has value v_i
- Strategy $b_i(v_i) \in [0, \infty)$

$$u_i(b_i, b_{-i}) = \begin{cases} v_i - \max\{b_j\} \text{ where } j \neq i \text{ if } b_i > b_j \text{ for all } j \\ 0 \text{ otherwise} \end{cases}$$

Given value v_i, $b_i(v_i) = v_i$ is (weakly) dominant.

Let $b' = \max_{j \neq i} b_j$. If $b' < v_i$ then any bid $b_i(v_i) > b'$ is optimal. If $b' \geq v_i$, then any bid $b_i(v_i) \leq v_i$ is optimal. Bid $b_i(v_i) = v_i$ satisfies both constraints.

Example: Bach or Stravinsky

- A couple likes going to concerts together. One loves Bach but not Stravinsky. The other loves Stravinsky but not Bach. However, they prefer being together than being apart.

	B	S
B	2,1	0,0
S	0,0	1,2

No dom. str. equil.

Nash Equilibrium

- Sometimes an agent's best-response depends on the strategies other agents are playing
  - No dominant strategy equilibria
- A strategy profile is a **Nash equilibrium** if no player has incentive to deviate from his strategy given that others do not deviate:
  - for every agent i, $u_i(s_i^*, s_{-i}) \geq u_i(s_i', s_{-i})$ for all s_i'

	B	S
B	2,1	0,0
S	0,0	1,2

Iterated Elimination of Dominated Strategies

- Let Ri⊆Si be the set of removed strategies for agent i
- Initially Ri=Ø
- Choose agent i, and strategy s_i such that $s_i \in S_i \setminus R_i$ and there exists $s_i' \in S_i \setminus R_i$ such that

  $$u_i(s_i', s_{-i}) > u_i(s_i, s_{-i}) \text{ for all } s_{-i} \in S_{-i} \setminus R_{-i}$$
- Add s_i to R_i, continue

- **Thm:** (Soundness) If a unique strategy profile, s*, survives then it is a Nash Eq.
- **Thm:** (Completeness) If a profile, s*, is a Nash Eq then it must survive iterated elimination.

Example: Iterated Dominance

	r	l	c
U	3,−3	7,−7	15,−15
D	9,−9	8,−8	10,−10

Nash Equilibrium

- Interpretations:
  - Focal points, self-enforcing agreements, stable social convention, consequence of rational inference..
- Criticisms
  - They may not be unique (Bach or Stravinsky)
    - Ways of overcoming this
      - Refinements of equilibrium concept, Mediation, Learning
  - Do not exist in all games (in the form defined above)
  - They may be hard to find
  - People don't always behave based on what equilibria would predict (ultimatum games and notions of fairness,…)

Example: Matching Pennies

	H	T
H	-1, 1	1,-1
T	1,-1	-1, 1

So far we have talked only about **pure** strategy equilibria.

Not all games have pure strategy equilibria.
Some equilibria are **mixed** strategy equilibria.

Mixed strategy equilibria

- Mixed strategy:

  Let $\sum_i$ be the set of probability distributions over S_i

  We write σ_i for an element of $\sum_i$

- Strategy profile: $\sigma=(\sigma_1,\ldots, \sigma_n)$
- Expected utility: $u_i(\sigma)=\sum_{s\in Si} \sigma_i(s)u_i(s)$
- Nash Equilibrium:
  - σ^* is a (mixed) Nash equilibrium if

  $u_i(\sigma^*_i, \sigma^*_{-i})\geq u_i(\sigma_i, \sigma^*_{-i})$ for all $\sigma_i\in\sum_i$, for all i

Example: Matching Pennies

		q H	1-q T
p	H	−1, 1	1, −1
1-p	T	1, −1	−1, 1

Want to play each strategy with a certain probability so that the competitor is indifferent between its own strategies.

$1p+(-1)(1-p)=(-1)p+1(1-p) \Rightarrow p=1/2$

$q-(1-q)=-q+(1-q) \Rightarrow q=1/2$

Mixed Nash Equilibrium

- Thm (Nash 50):
  - Every game in which the strategy sets, S_1, …,S_n have a finite number of elements has a mixed strategy equilibrium.
- Finding Nash Equil is another problem
  - "Together with prime factoring, the complexity of finding a Nash Eq is, in my opinion, the most important concrete open question on the boundary of P today" (Papadimitriou)

Bayesian-Nash Equil

(Harsanyi 68)

- So far we have assumed that agents have complete information about each other (including payoffs)
  - Very strong assumption!
- Assume agent i has type $\theta_i \in \Theta_i$, which defines the payoff $u_i(s, \theta_i)$
- Agents have common prior over distribution of types $p(\theta)$
  - Conditional probability $p(\theta_{-i} | \theta_i)$ (obtained by Bayes Rule when possible)

Bayesian-Nash Equil

- Strategy: $\sigma_i(\theta_i)$ is the (mixed) strategy agent i plays if its type is θ_i
- Strategy profile: $\sigma=(\sigma_1,\ldots, \sigma_n)$
- Expected utility:
  - $U_i(\sigma_i(\theta_i),\sigma_{-i}(),\theta_i)=\sum_{\theta-i} p(\theta_{-i}|\theta_i)u_i(\sigma_i(\theta_i),\sigma_{-i}(\theta_{-i}),\theta_i)$

- Bayesian Nash Eq: Strategy profile σ^* is a Bayesian-Nash Eq if for all i, for all θ_i,

  $U_i(\sigma^*_i(\theta_i),\sigma^*_{-i}(),\theta_i) \geq U_i(\sigma_i(\theta_i),\sigma^*_{-i}(),\theta_i)$

  (best responding w.r.t. its beliefs about the types of the other agents, assuming they are also playing a best response)

Example: 1^{st} price sealed-bid auction

2 agents (1 and 2) with values v_1, v_2 drawn uniformly from [0,1].

Utility of agent i if it bids b_i and wins the item is $u_i = v_i - b_i$.

Assume agent 2's bidding strategy is $b_2(v_2) = v_2/2$

How should 1 bid? (i.e. what is $b_1(v_1) = z$?)

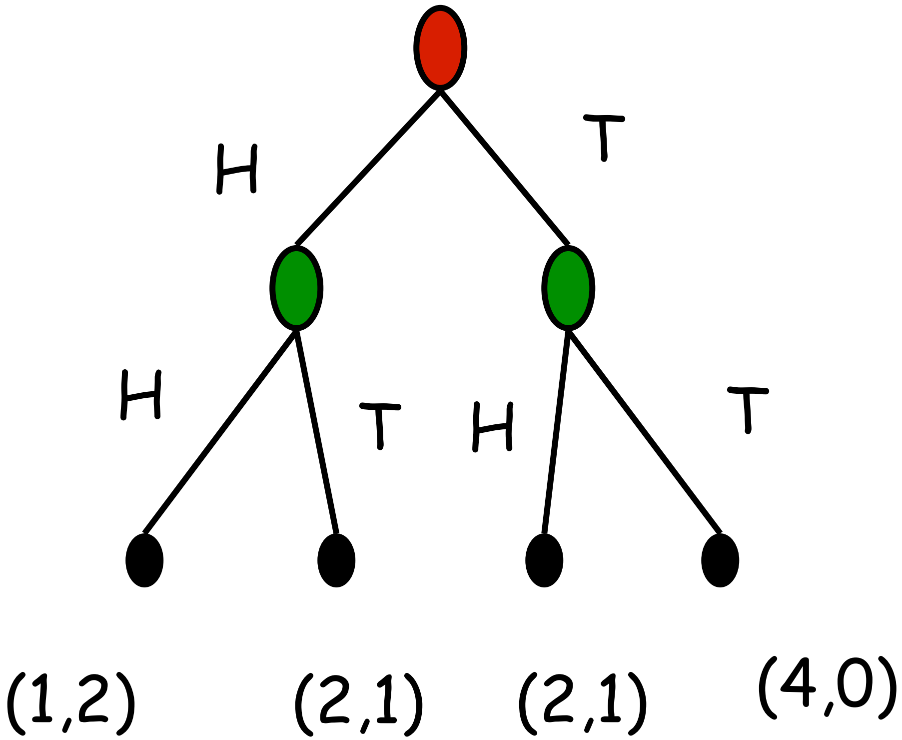
$$U_1 = \int_{z=0}^{2z} (v_1 - z) dz = (v_1 - z) 2z = 2zv_1 - 2z^2$$

Note: given $b_2(v_2) = v_2/2$, 1 only wins if $v_2 < 2z$

Therefore, $\text{Max}_z[2zv_1 - 2z^2]$ when $z = b_1(v_1) = v_1/2$

Similar argument for agent 2, assuming $b_1(v_1) = v_1/2$.
We have an equilibrium

Extensive Form Games



Any finite game of perfect information has a pure strategy Nash equilibrium. It can be found by backward induction.

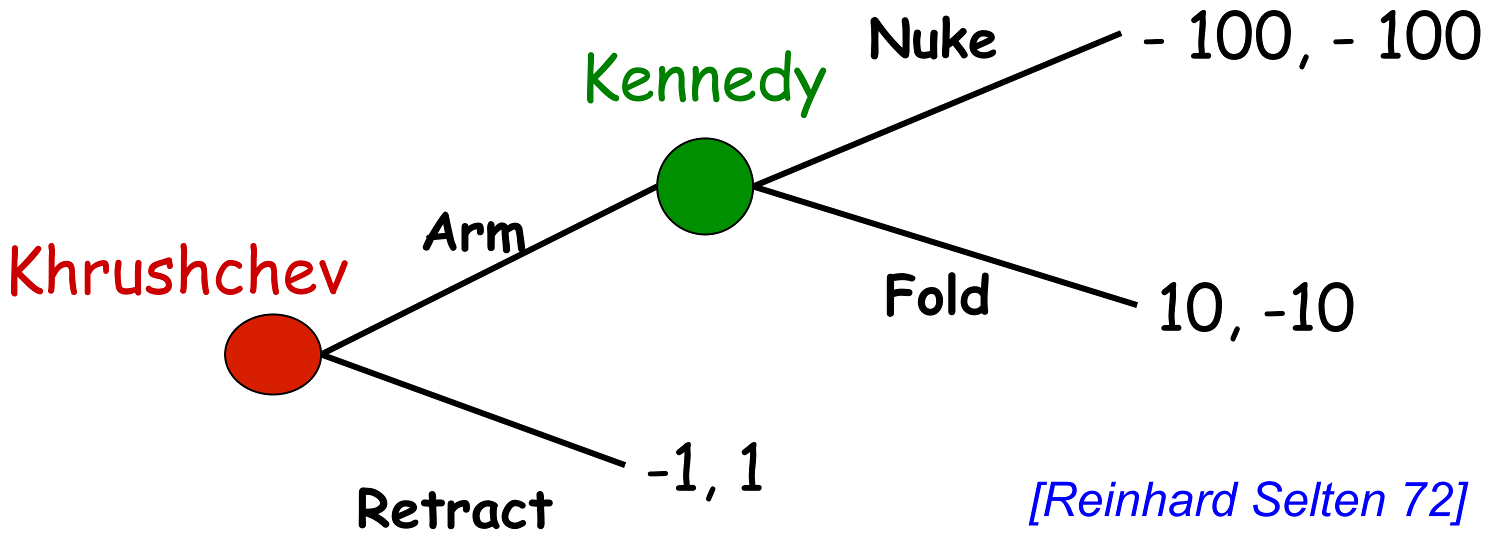
Chess is a finite game of perfect information. Therefore it is a “trivial” game from a game theoretic point of view.

Subgame perfect equilibrium & credible threats

- Proper subgame = subtree (of the game tree) whose root is alone in its information set

- Subgame perfect equilibrium
  - Strategy profile that is in Nash equilibrium in every proper subgame (including the root), **whether or not that subgame is reached** along the equilibrium path of play

Example: Cuban Missile Crisis



[Reinhard Selten 72]

Pure strategy Nash equilibria: (Arm, Fold) and (Retract, Nuke)

Pure strategy subgame perfect equilibria: (Arm, Fold)

Conclusion: Kennedy's Nuke threat was not credible.