Richer Representations:
Beyond the Normal and Extensive Forms

Richer Representations

- **Infinite games** cannot be represented in normal or extensive form
  - e.g. what happens when a simple normal-form game such as the Prisoner's Dilemma is repeated infinitely?
- Games played by an **uncountably infinite** set of agents
- An interval of the **real numbers** as each player's action space

Richer Representations

- Assumption that agents have **perfect knowledge** of everyone's payoffs
  - Agents might have **private information** that affects their own payoffs
  - Agents might have only **probabilistic information** about each others' private information
- Big impact: agent's actions can depend on **what he knows** about another agent's payoffs

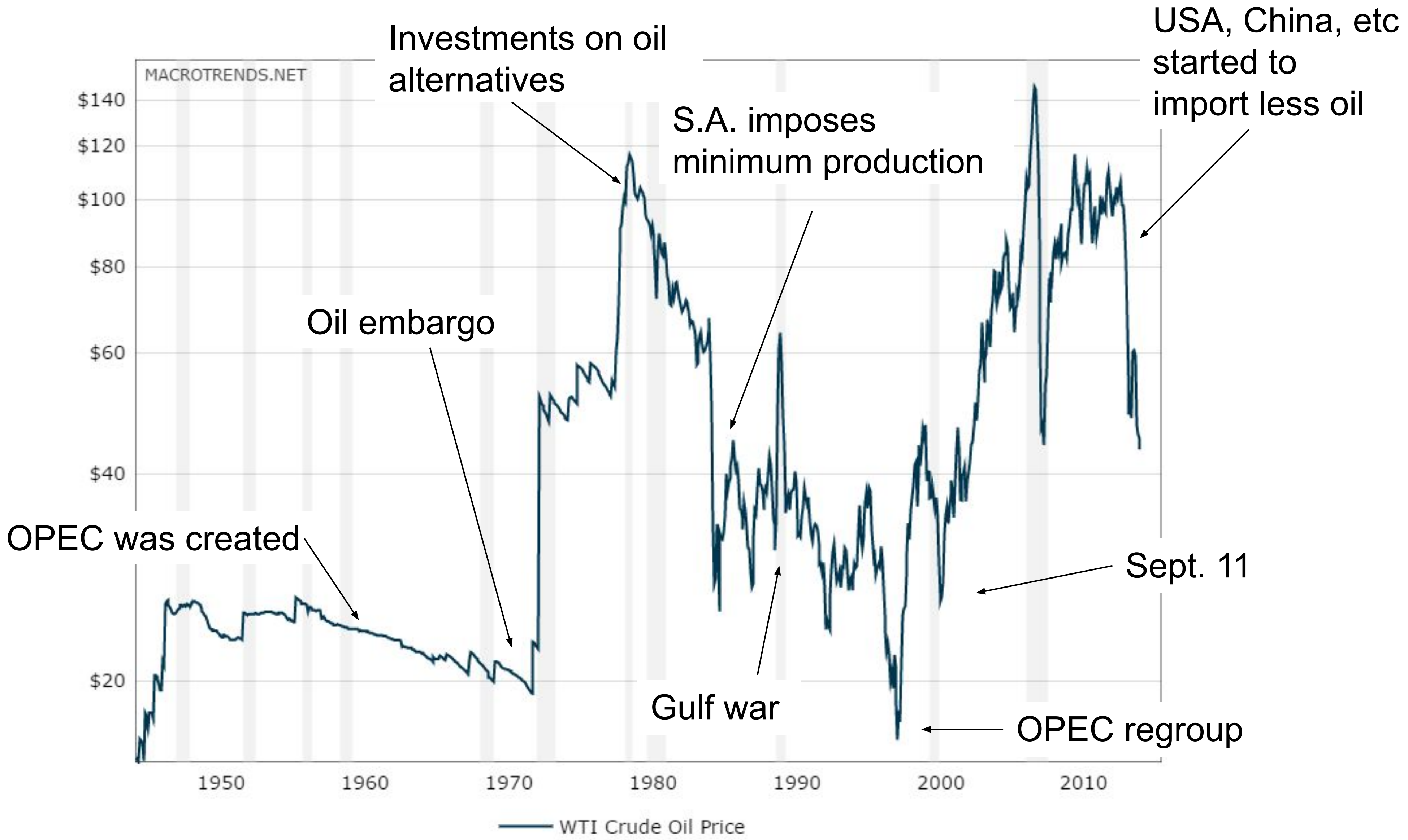
Richer Representations

- As the number of players and actions grow, games can be **too large** to reason about
- Settings that are most interesting in practice tend to involve **highly structured** payoffs
  - a large game actually corresponds to finitely repeated play of a **small game**
  - the number of agents who are able to directly affect any given agent's payoff is **small**

Repeated Games

- What happens when a simple normal-form game such as the **Prisoner's Dilemma** is repeated infinitely?

Repeated games



Repeated games

- Many (most?) interactions occur more than once:
  - Firms in a marketplace
  - Political alliances
  - Friends (favor exchange...)
  - Workers (team production...)
- Many are like a **repeated Prisoner’s Dilemma**
  - Need to easily observe each other’s plays and react (quickly) to punish undesired behavior

Repeated games

- Consider the following representation of a two-stage prisoner`s dilemma game

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

$\Rightarrow$

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Repeated games

- Do agents see what the other agents played earlier?
- Do they remember what they knew?
- What is the utility of the entire repeated game?

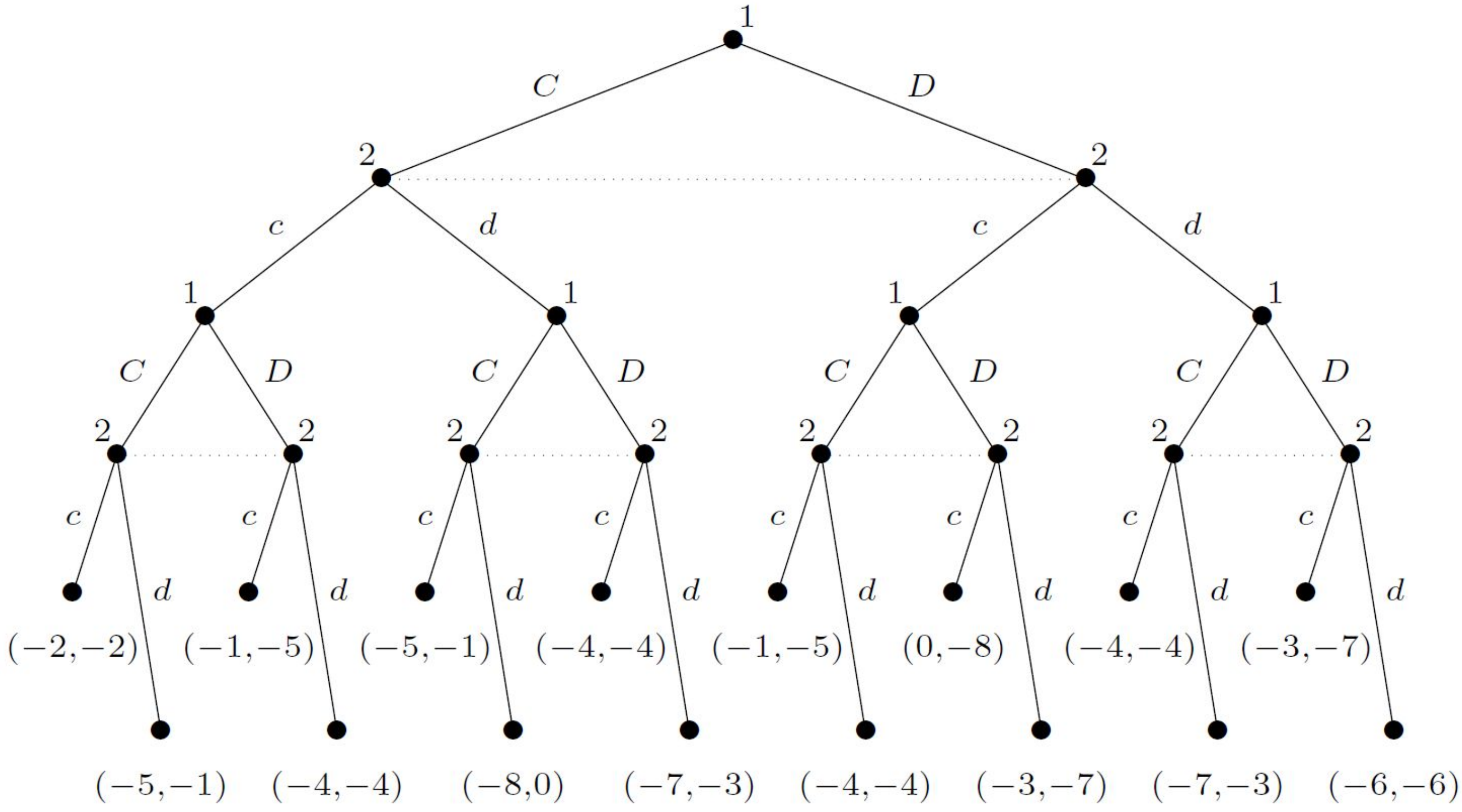
	C	D
C	$-1,-1$	$-4,0$
D	$0,-4$	$-3,-3$

$\Rightarrow$

	C	D
C	$-1,-1$	$-4,0$
D	$0,-4$	$-3,-3$

Finitely repeated games

- <u>Imperfect-information game</u> in extensive form to completely disambiguate the semantics of a **finitely repeated game**

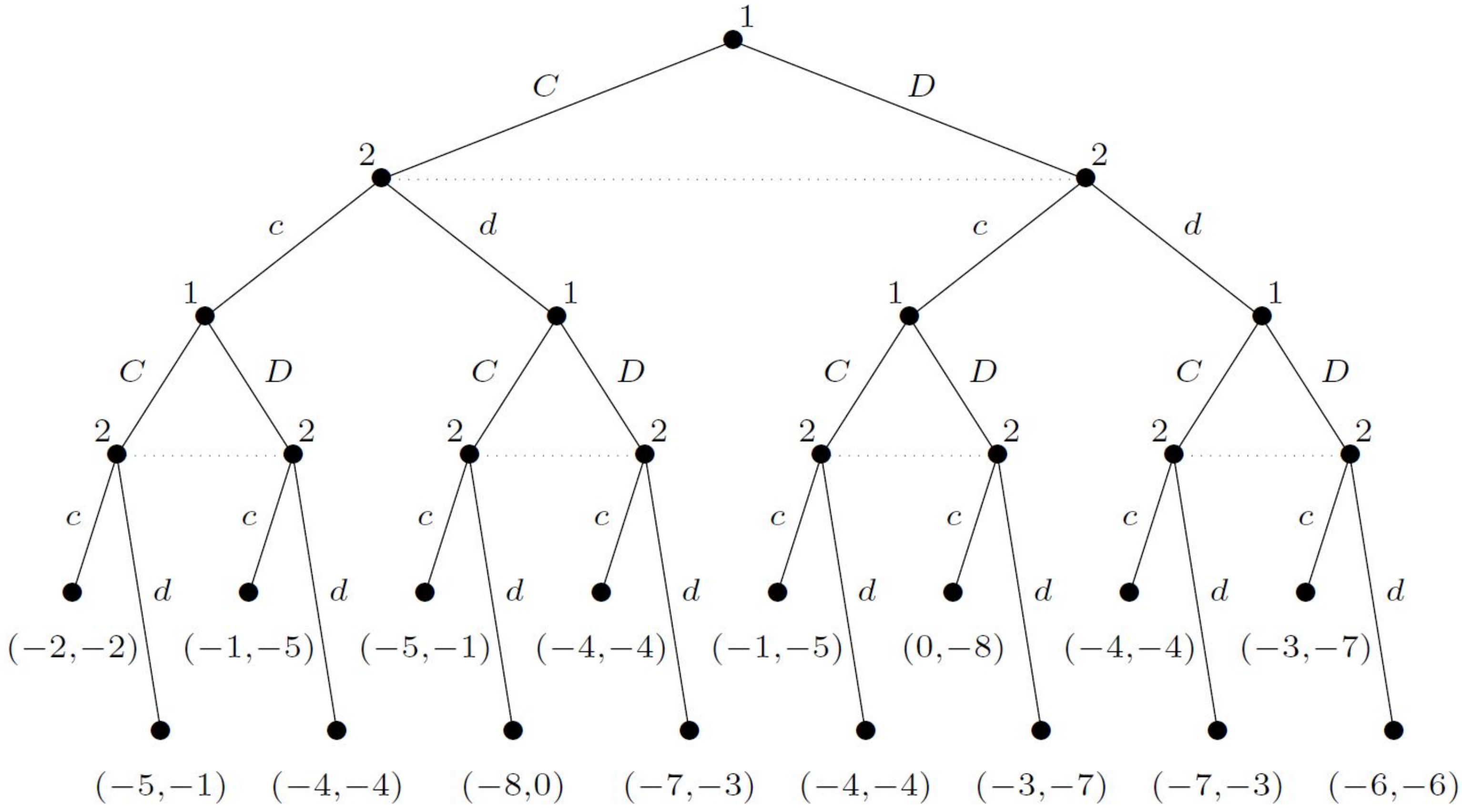


Finitely repeated games

- At each iteration the players do not know what the other player is playing but afterward they do
- The payoff function of each agent is **additive**

Finitely repeated games

- What is the equilibrium of this game?



Infinitely repeated games

- Does it mean that the game will go on forever?
  - No, it means that the players do not know when it will end

Infinitely repeated games

- When the infinitely repeated game is transformed into extensive form, the result is an **infinite tree**
  - Payoffs cannot be attached to any terminal nodes
  - Payoffs cannot be defined as the sum of the payoffs in the stage games
    - **Q:** why not?
    - **A:** in general, they will be infinite
- There are two common ways of defining a player's payoff to get around this problem

Infinitely repeated games

- Given an infinite sequence of payoffs $r_i^{(1)}, r_i^{(2)}, \ldots$ for player i, **the average reward** of i is

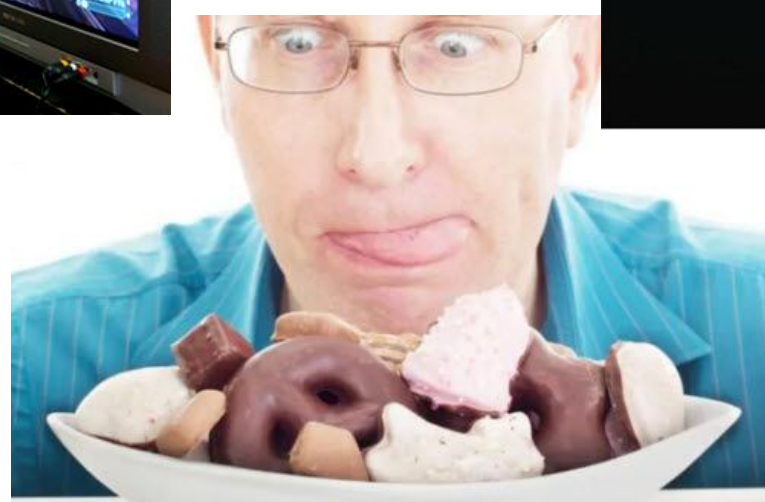
$$\lim_{k\to\infty} \frac{\sum_{j=1}^{k} r_i^{(j)}}{k}$$

Infinitely repeated games

- Two friends are in a running competition
- Every saturday they check who have run the most
- Both of them prefer not running to running
- They have to plan how much they should run each week
- Is the average reward model appropriate?

Infinitely repeated games

- Sometimes, the agent may care much more about the present than the future



Infinitely repeated games

- Given an infinite sequence of payoffs $r_i^{(1)}, r_i^{(2)}, \ldots$ for player i, and a discount factor β with $0 \le \beta \le 1$, the **future discounted reward** of i is

$$\sum_{j=1}^{\infty} \beta^j r_i^{(j)}$$

- e.g.: $U_i = 0.8^1\, r_i^{(1)} + 0.8^2\, r_i^{(2)} + 0.8^3\, r_i^{(3)} + \ldots$

Infinitely repeated games

- The discount factor β can be interpreted as
  - The agent cares more about his well-being in the **near term** than in the long term
  - The agent cares about the future just as much as he cares about the present, but with some probability $1 - \beta$ the **game will be stopped** any given round
- The analysis of the game is **not affected** by which perspective is adopted

Infinitely repeated games

nature human behaviour

ARTICLES

https://doi.org/10.1038/s41562-022-01392-w

Check for updates

OPEN

The globalizability of temporal discounting

Economic inequality is associated with preferences for smaller, immediate gains over larger, delayed ones. Such temporal discounting may feed into rising global inequality, yet it is unclear whether it is a function of choice preferences or norms, or rather the absence of sufficient resources for immediate needs. It is also not clear whether these reflect true differences in choice patterns between income groups. We tested temporal discounting and five intertemporal choice anomalies using local currencies and value standards in 61 countries ($N = 13{,}629$). Across a diverse sample, we found consistent, robust rates of choice anomalies. Lower-income groups were not significantly different, but economic inequality and broader financial circumstances were clearly correlated with population choice patterns.

Infinitely repeated games

- What is a **pure strategy** in an infinitely-repeated game?
  - a choice of action at every decision point
  - an action at every stage game, which is an infinite number of actions!
- Can be thought of an 'instruction' to a third person who will play the game for you

Infinitely repeated games

- What is a **pure strategy** in an infinitely-repeated game?
- Some famous strategies (repeated PD):
  - **Tit-for-tat:** Start out cooperating. If the opponent defected, defect in the next round. *Then go back to cooperation (atualizar)*
  - **Grim trigger:** Start out cooperating. If the opponent ever defects, defect forever

Infinitely repeated games

- What can we say about Nash equilibria?
  - Unable to construct an induced normal form :(
    - Nash's theorem only applies to finite games
  - With an infinite number of strategies, there could be an infinite number of pure-strategy equilibria!
- However, we can characterize a set of payoffs that are achievable under equilibrium
  - without having to enumerate the equilibria

Reminder: maxmin e minmax

What are the maxmin value $\boldsymbol{v'_1}$ and the minmax value $\boldsymbol{v_1}$?

	L	***R***
T	0,a	3,b
M	2,c	1,d
B	4,e	0,f

Reminder: maxmin e minmax

The maxmin value $\boldsymbol{v'_i}$ for a player is at most its minmax value $\boldsymbol{v_i}$
i.e., $\boldsymbol{v'_i \leq v_i}$

	L	***R***
T	0,a	3,b
M	2,c	1,d
B	4,e	0,f

3 is the minmax value for player ***1***

1 is the maxmin value for player ***1***

Infinitely repeated games

- Consider any n-player game ***G = (N, A, u)*** and any payoff profile $\boldsymbol{r = (r_1, r_2, \ldots, r_n)}$
- Let

$$v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_{-i}, s_i)$$

- $\boldsymbol{v_i}$ is player ***i***'s minmax value
  - his utility when the other players play minmax strategies against him, and he plays his best response

Infinitely repeated games

- **Definitions**
  - A payoff profile $\boldsymbol{r = (r_1, r_2, . . . , r_n)}$ is **enforceable** if $\boldsymbol{\forall i \in N, r_i \geq v_i}$
  - A payoff profile $\boldsymbol{r = (r_1, r_2, . . . , r_n)}$ is **feasible** if there exist <u>rational</u>, nonnegative values $\boldsymbol{\alpha_a}$ such that for all $\boldsymbol{i}$, we can express $\boldsymbol{r_i}$ as $\boldsymbol{\Sigma_{a \in A} \alpha_a u_i(a)}$, with $\boldsymbol{\Sigma_{a \in A} \alpha_a = 1}$

Infinitely repeated games

- Classify the following payoff profiles
  - (-1,1)
    - not enforceable, not feasible
  - (10,10)
    - enforceable, not feasible
  - (0,0)
    - feasible, not enforceable
  - (1,1)
    - feasible, enforceable
  - (2,2)
    - feasible, enforceable

	L	R
U	4,0	1,1
D	0,0	0,4

Infinitely repeated games

- **Theorem** (Folk Theorem)
- Consider any n-player game $\boldsymbol{G}$ and any payoff vector $\boldsymbol{r = (r_1, r_2, \ldots, r_n)}$
  - 1. If $\boldsymbol{r}$ is the payoff in any Nash equilibrium of the infinitely repeated $\boldsymbol{G}$ with average rewards, then for each player $\boldsymbol{i}$, $\boldsymbol{r_i}$ is enforceable
    - not feasible because sometimes it may yield irrational $\boldsymbol{\alpha}$'s
  - 2. If $\boldsymbol{r}$ is both feasible and enforceable, then $\boldsymbol{r}$ is the payoff in some Nash equilibrium of the infinitely repeated $\boldsymbol{G}$ with average rewards

Folk Theorem (Part 1)

- Payoff in Nash ⇒ enforceable
  - Suppose r is not enforceable, i.e. $r_i < v_i$ for some i
  - Then consider a deviation of player i to $b_i(s_{-i}(h))$ for any history h of the repeated game
    - b_i is any best-response action in the stage game
    - $s_{-i}(h)$ is the strategy of other players given h
  - By definition of a **minmax** strategy, player i will receive a payoff of at least v_i in every stage game if he adopts this strategy, and so i's average reward is also at least v_i
  - Thus i cannot receive the payoff $r_i < v_i$ in any Nash equilibrium

Folk Theorem (Part 2)

- Feasible and enforceable ⇒ Nash
  - Since $\boldsymbol{r}$ is a feasible enforceable payoff profile, we can write it as $r_i = \Sigma_{a \in A} (\beta_a / \gamma)\, u_i(a)$, where β_a and γ are nonnegative integers
    - Recall that α_a were required to be rational, so γ can be their common denominator
  - Since the combination was convex, we have $\gamma = \Sigma_{a \in A} \beta_a$

$\beta = (3, 1, 0, 3)$
$\gamma = 7$

	L	R
U	3/7	1/7
D	0/7	3/7

Folk Theorem (Part 2)

- Feasible and enforceable ⇒ Nash
  - We are going to construct a strategy profile that will cycle through all outcomes $\boldsymbol{a} \in \boldsymbol{A}$ of $\boldsymbol{G}$ with cycles of length $\boldsymbol{\gamma}$, each cycle repeating action $\boldsymbol{a}$ exactly $\boldsymbol{\beta_a}$ times
  - Let $\boldsymbol{(a^t)}$ be such a sequence of outcomes
  - Let us define a strategy $\boldsymbol{s_i}$ of player $\boldsymbol{i}$ to be a **trigger** version of playing $\boldsymbol{(a^t)}$
    - if nobody deviates, then $\boldsymbol{s_i}$ plays $\boldsymbol{a_i^t}$ in period $\boldsymbol{t}$
    - However, if there was a period $\boldsymbol{t'}$ in which some player $\boldsymbol{j \neq i}$ deviated, then $\boldsymbol{s_i}$ will play $\boldsymbol{(p_{-j})_i}$, where $\boldsymbol{(p_{-j})}$ is a solution to the minimization problem in the definition of $\boldsymbol{v_j}$

Folk Theorem (Part 2)

- Feasible and enforceable ⇒ Nash
  - If everybody plays according to s_i, then, by construction, player i receives average payoff of r_i
    - look at averages over periods of length γ
  - This strategy profile is a <u>Nash equilibrium</u>
    - Suppose everybody plays according to s_i, and player j deviates at some point
    - Then, forever after, player j will receive his **minmax** payoff $v_j \leq r_j$, rendering the deviation unprofitable

Discounted Repeated Games

- Stage game: (N, A, u)
- Discount factors: $\beta_1, \ldots, \beta_n$, $\beta_i \in [0, 1]$
- Assume a common discount factor for now
  - $\beta_i = \beta$ for all i
- Payoff from a play of actions $a_1, \ldots, a_t, \ldots$

$$\sum_t \beta_i^t u_i(a^t)$$

Discounted Repeated Games

- Histories of length $\boldsymbol{t}$: $\boldsymbol{H^t = \{h^t : h^t = (a_1, \ldots, a_t) \in A^t\}}$
- All finite histories: $\boldsymbol{H = U_t H^t}$
- A strategy: $\boldsymbol{s_i : H \rightarrow \Delta(A_i)}$
  - a map from every possible history into a possibly mixed strategy over what I can do in a given period given a history

Discounted Repeated Games

- Prisoners Dilemma
  - $A_i = (C, D)$
  - A history for ***3*** periods: $h^3 = ((C,C), (C,D), (D,D))$
  - A strategy for period ***4*** would specify what a player would do after seeing h^3

Discounted Repeated Games

- **Subgame perfection**
  - Profile of strategies that are Nash in every subgame
  - So, a Nash equilibrium following every possible history
  - Repeatedly playing a Nash equilibrium of the stage game is always a subgame perfect equilibrium of the repeated game

Discounted Repeated Games

- Repeated Prisoner's Dilemma
  - What happens if all players play the **Grim Trigger** strategy?
    - Cooperate as long as everyone has in the past
    - Both players defect forever after if anyone ever deviates

	C	D
C	3,3	0,5
D	5,0	1,1

Reminder

$$\left.\begin{aligned}
\sum_{k=0}^{n} x^k &= \frac{x^{n+1}-1}{x-1} \\
\sum_{k=0}^{\infty} x^k &= \frac{1}{1-x} \\
\sum_{k=1}^{\infty} x^k &= \frac{1}{1-k} - 1
\end{aligned}\right\} |x| < 1$$

$$\sum_{k=1}^{n} (ca_k + b_k) = c\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k$$

Discounted Repeated Games

- Let's check the payoff for deviating:
  - Cooperate: 
  - Defect: 

	C	D
C	3,3	0,5
D	5,0	1,1

Discounted Repeated Games

- Let's check the payoff for deviating:
  - Cooperate: 
  - Defect: 
  - Difference: 
  - Difference is nonnegative if:

 or  or 

Discounted Repeated Games

- Need to care about tomorrow **at least half as much** as today!

Discounted Repeated Games

- Basic logic:
  - Play something with relatively high payoffs, and if anyone deviates
  - Punish by resorting to something that
    - has lower payoffs (at least for that player)
    - and is credible: it is an equilibrium in the subgame

Folk Theorem

- Consider a finite normal form game $G = (N, A, u)$
- Let $a = (a_1, ..., a_n)$ be a **Nash equilibrium** of the stage game G
- If $a' = (a_1', ..., a_n')$ is such that $u_i(a') > u_i(a)$ for all i,
- then there exists a discount factor $\beta < 1$, such that
- if $\beta_i > \beta$ for all i, then there exists a subgame perfect equilibrium of the infinite repetition of G that has a' played in every period on the equilibrium path

Folk Theorem

- Outline of the Proof:
    - Play ***a'*** as long as everyone has in the past
    - If any player ever deviates, then play ***a*** forever after (Grim Trigger)
    - Check that this is a subgame perfect equilibrium for high enough discount factors

Folk Theorem

- Check that this is a subgame perfect equilibrium for high enough discount factors:
  - Playing ***a*** forever if anyone has deviated is a Nash equilibrium in any such subgame
  - Will **<u>someone</u>** gain by deviating from ***a'*** if nobody has in the past?
  - Maximum gain from deviating is

  - Minimum per-period loss from future punishment is

  - If deviate, the <u>maximum possible net gain</u> is 
  - Dev. not beneficial if: 

Folk Theorem

	C	D
C	3,3	0,5
D	5,0	1,1

- Maximum gain from deviating is

***M* = 5 – 3 = 2**

- Minimum per-period loss from future punishment is

***m* = 3 – 1 = 2**

- If deviate, the maximum possible net gain is 

or $2 - 2(\beta_i / (1 - \beta_i))$

- Dev. not beneficial if: 

$$\beta_i \geq 2 / 4 \geq ½$$

Folk Theorem

- More complicated play:

	C	D
C	3,3	0,10
D	10,0	1,1

Folk Theorem

- More complicated play:

10/02/2015 10h35 - Atualizado em 11/02/2015 09h32

Empreiteiras combinavam licitações desde o governo FHC, diz delator

Augusto Mendonça Neto prestou depoimento à Justiça nesta segunda.
Ele também disse que o 'clube' passou a ter 'efetividade' com Costa e Duque.

Lucas Salomão*
Do G1, em Brasília

FACEBOOK

O executivo Augusto Mendonça Neto, da Setal, em depoimento à Justiça

O executivo Augusto Ribeiro de Mendonça Neto, dono da Setal Engenharia e um dos delatores do esquema de corrupção na Petrobras, afirmou em depoimento à Justiça Federal do Paraná nesta segunda-feira (9) que o "clube" das empreiteiras passou a combinar resultados de licitações desde meados da década de 1990, época em que o país era presidido por Fernando Henrique Cardoso.

Segundo Mendonça Neto, as empresas e a **Petrobras** instituíram um grupo de trabalho

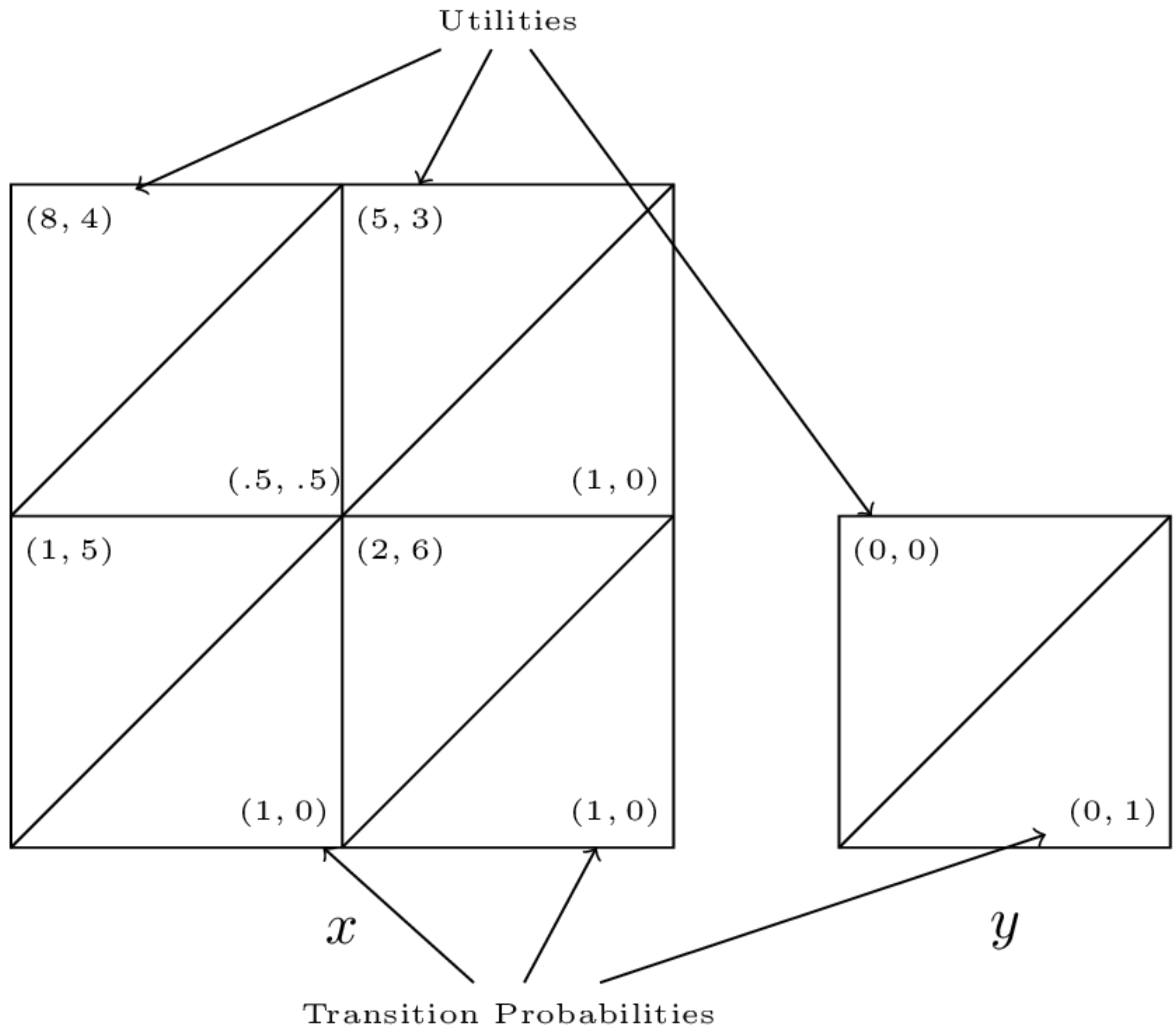
Repeated Games

- Players can condition future play on past actions
- Brings in many(!) equilibria: Folk Theorems
- Need key ingredients
  - Some (fast enough) observation about how others behave
  - Sufficient value to the future (limit of the means - extreme value) or high enough discount factor

Stochastic games

- Intuitively speaking, a **stochastic game** is a collection of normal-form games
  - the agents repeatedly play games from this **collection**
  - the particular game played at any given iteration **depends** probabilistically on the **previous game played** and on the **actions taken by all agents** in that game

Stochastic games



Bounded rationality

Bounded rationality

- The idea that in decision-making, rationality of individuals is **limited** by:
  - the information they have
  - the cognitive limitations of their minds
  - the finite amount of time they have to make a decision

Bounded rationality

- Until now, we have assumed that players are homogeneous and fully rational
- What happens when agents are not perfectly rational expected-utility maximizers?
- What happens when we impose specific computational limitations on them
- How can we represent such limitations?

Bounded rationality

- What happened in the games we played so far?
  - E.g. Prisoner's dilemma or Five Pirates' game

	C	D
C	$3, 3$	$0, 4$
D	$4, 0$	$1, 1$

Bounded rationality

- One approach: ε-equilibrium
  - Agents' rationality may be bounded when they are willing to settle for ε-lower payoffs
  - In the **finitely** repeated Prisoner's Dilemma game, the sets of cooperating ε-equilibria increases with size

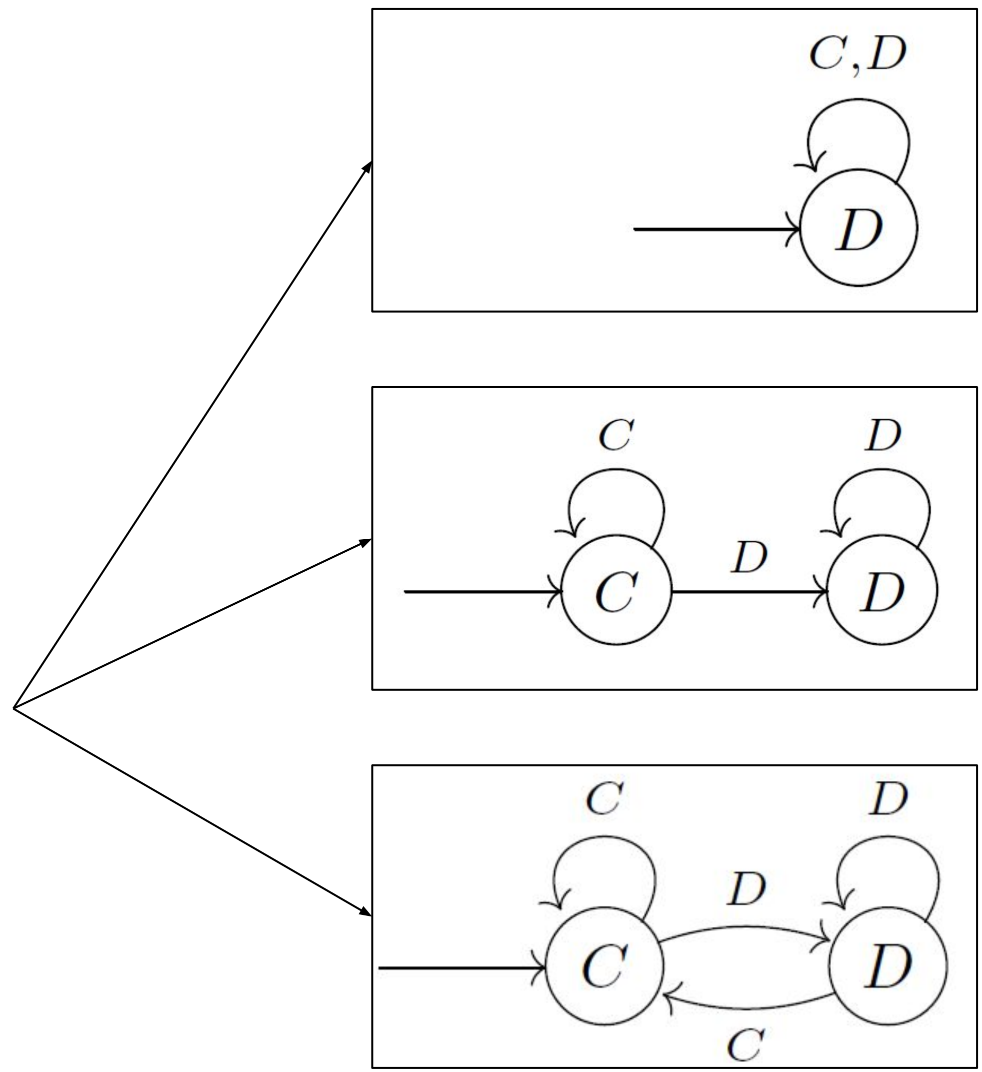
Bounded rationality

- Finitely repeated games (or imperfect-information extensive-form games)
  - A strategy for player ***i*** is a specification of an action for **every** information set belonging to ***i***
  - A strategy for ***k*** repetitions of an ***m***-action game is thus a specification of $(m^{k-1}) / (m-1)$ different actions
    - One action for every possible history
  - How to demand less rationality?

Bounded rationality

- How to demand less rationality?

	C	D
C	$3, 3$	$0, 4$
D	$4, 0$	$1, 1$



Finite-state automata

- A two-player **machine game** $G^M = (\{1, 2\}, M, G)$ of the k-period repeated game G is defined by:
  - a pair of players $\{1, 2\}$
  - $M = (M_1, M_2)$, where $M_i = \{m_1, m_2, \ldots\}$ is a set of available automata for player i
  - a normal-form game $G = (\{1, 2\}, A, u)$
- A pair $m_1 \in M_1$ and $m_2 \in M_2$ deterministically yield an outcome $o^t(m_1,m_2)$ at each iteration t
- Thus, G^M induces a normal-form game $(\{1, 2\}, M, U)$, in which each player i chooses an automaton $m_i \in M_i$, and obtains utility $U_i(m_1,m_2) = \Sigma^k_{t=1} u_i(o^t(m_1,m_2))$

Finite-state automata

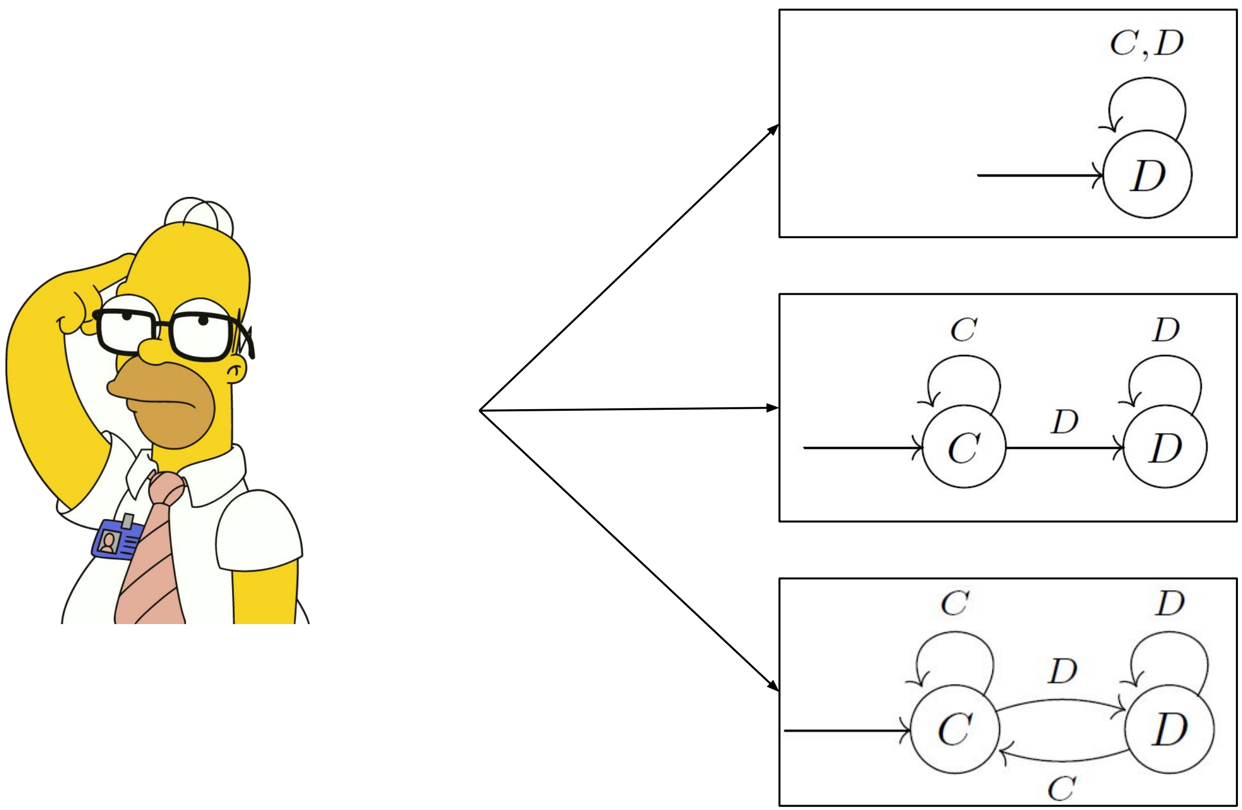
- A two-player **machine game** $G^M = (\{1, 2\}, M, G)$ of the k-period repeated game G is defined by:
  - a p
  - M =
    ava
  - a normal-form game $G = (\{1, 2\}, A, u)$

Can be easily replaced by the discounted utility

- A pair $m_1 \in M_1$ and $m_2 \in M_2$ deterministically yield an outcome $o^t(m_1,m_2)$ at each iteration t
- Thus, G^M induces a normal-form game $(\{1, 2\}, M, U)$, in which each player i chooses an automaton $m_i \in M_i$, and obtains utility $\boxed{U_i(m_1,m_2) = \Sigma^k_{t=1} u_i(o^t(m_1,m_2))}$

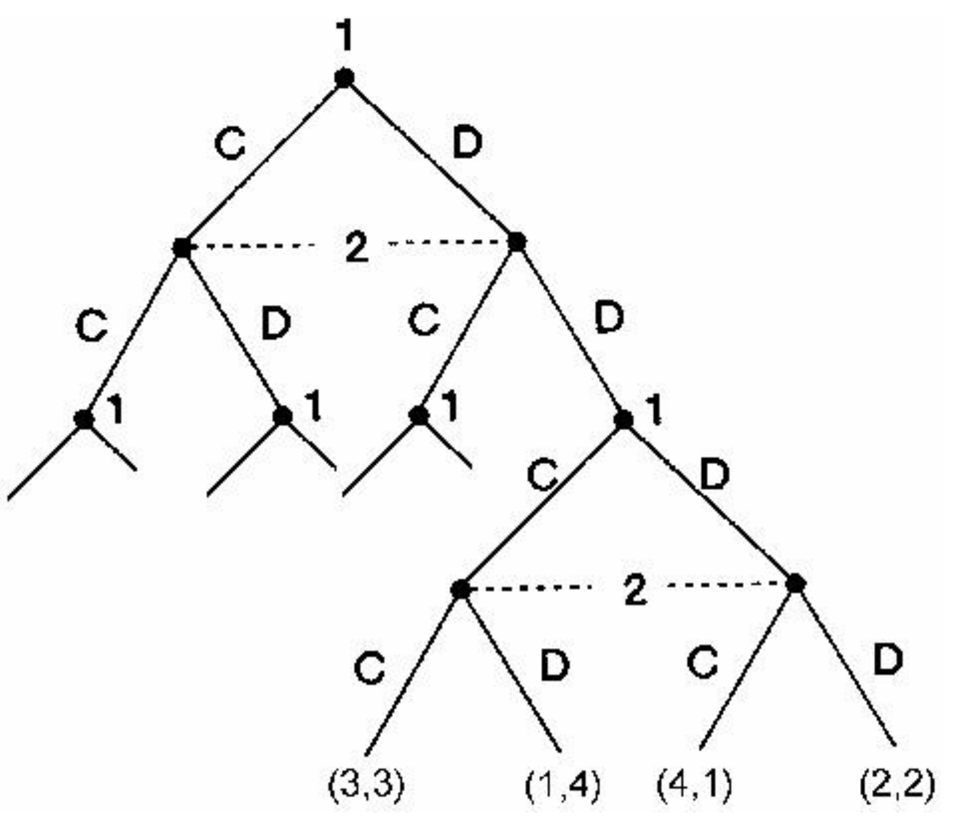
Automata of bounded size

- Intuitively, automata with **fewer** states represent simpler strategies
- Thus, one way to **bound the rationality** of the player is by limiting the number of states in the automaton



Automata of bounded size

- How to find the "always defect" equilibrium in the finite **RPD** via automatas?



Automata of bounded size

- Placing severe restrictions on the number of states not only induces an equilibrium in which cooperation always occurs, but also causes the always-defect equilibrium to disappear

Automata of bounded size

- How to find the "always defect" equilibrium in the finite **RPD** via automatas?
  - Each player has to use **backward induction** to find his dominant strategy
  - In order to perform backward induction in a ***k***-period repeated game, each player needs to **keep track** of at least ***k*** distinct states
    - one state to represent the choice of strategy in each repetition of the game

Automata of bounded size

- In what follows,
- the function $\boldsymbol{s : M \rightarrow Z}$ represents the number of states of an automaton $\boldsymbol{M}$
- the function $\boldsymbol{S(M_i) = \max_{M \in Mi} s(M)}$ represents the size of the largest automaton among a set of automata $\boldsymbol{M_i}$

Automata of bounded size

- In the Prisoner's Dilemma, it turns out that if

  $$2 < max(S(M_1), S(M_2)) < k ,$$

  then the constant-defect strategy does not yield a symmetric equilibrium, while the Tit-for-Tat automaton does

Automata of bounded size

- **Theorem**
  - For any integer x, there exists an integer k_0 such that
  - for all $k > k_0$, any machine game $G^M = (\{1, 2\}, M, G)$ of the k-period repeated PD game G,
  - in which $k^{1/x} \leq \min\{S(M_1), S(M_2)\} \leq \max\{S(M_1), S(M_2)\} \leq k^x$
  - holds has a Nash equilibrium in which the average payoffs to each player are at least $3 - 1/x$

Computing best-response automata

- **Theorem**

  Given a machine game $G^M = (N, M, G)$ of a limit average infinitely repeated game $G = (N, A, u)$

  with unknown N,

  and a choice of automata $m_1, \ldots, m_n$ for all players,

  there does not exist a polynomial time algorithm for verifying whether m_i is a best-response automaton for player i

  :) If we hold N fixed, than the problem belongs to P

Computing best-response automata

- **Theorem**

  Given a two-player machine game $G^M = (\{1, 2\}, M, G)$ of a limit average infinitely repeated two-player game $G = (N, A, u)$

  and a mixed strategy for player 2 in which the set of automata that are played with positive probability is **finite**,

  the problem of verifying that an automaton m_1 is a best-response automaton for player 1 is **NP-complete**

Computing best-response automata

- **Theorem**

  Given a two-player machine game $\boldsymbol{G^M = (\{1, 2\}, M, G)}$ of a limit average infinitely repeated Prisoner's Dilemma game $\boldsymbol{G}$

  an automaton $\boldsymbol{m_2}$, and an integer $\boldsymbol{k}$,

  the problem of computing a best-response automaton $\boldsymbol{m_1}$ for player $\boldsymbol{1}$,

  such that $\boldsymbol{size(m_1) \leq k}$,

  is **NP-complete**

  :(

From finite automata to Turing machines

- Turing machines are more powerful than finite-state automata due to their infinite memories
  - **Q:** Would game-theoretic results be preserved under this richer model?
  - **A:** No
  - For example, there is strong evidence that a PD game of two Turing machines can have equilibria that are arbitrarily close to the repeated C payoff

Machine Games

- For more information, please read Chapters 8 and 9 of "A Course in Game Theory", by Ariel Rubinstein