Multiagent decision making: Repeated Games

Outline

- **Introduction to Repeated games**
- Finitely-repeated games
- Infinitely-repeated games
- Folk Theorem
- Replicator dynamics

Repeated Games

- What happens if we play the same normal-form game over and over?

Repeated Games

- Questions we will need to answer in repeated games:
  - Can agents **observe** other agent’s actions?
  - Can agents **remember** the past?
  - What is the agent’s **utility** for the whole game?

Repeated Games

- Some of the questions will have **different answers** for:
  - **Finitely-repeated games**
  - **Infinitely-repeated games**
- The normal-form game that we play repeatedly is called the **stage game**

Outline

- Introduction to Repeated games
- **Finitely-repeated games**
- Infinitely-repeated games
- Folk Theorem
- Replicator dynamics

Finitely-Repeated Games

- How can we analyze a **repeated game with a finite number of repetitions**?
  - We can model this game as an **extensive-form game with imperfect information**
    - At each round, **players do not know what the other players have done, but afterward, they do**
    - **The overall payoff function is additive**: the sum of payoffs in stage games

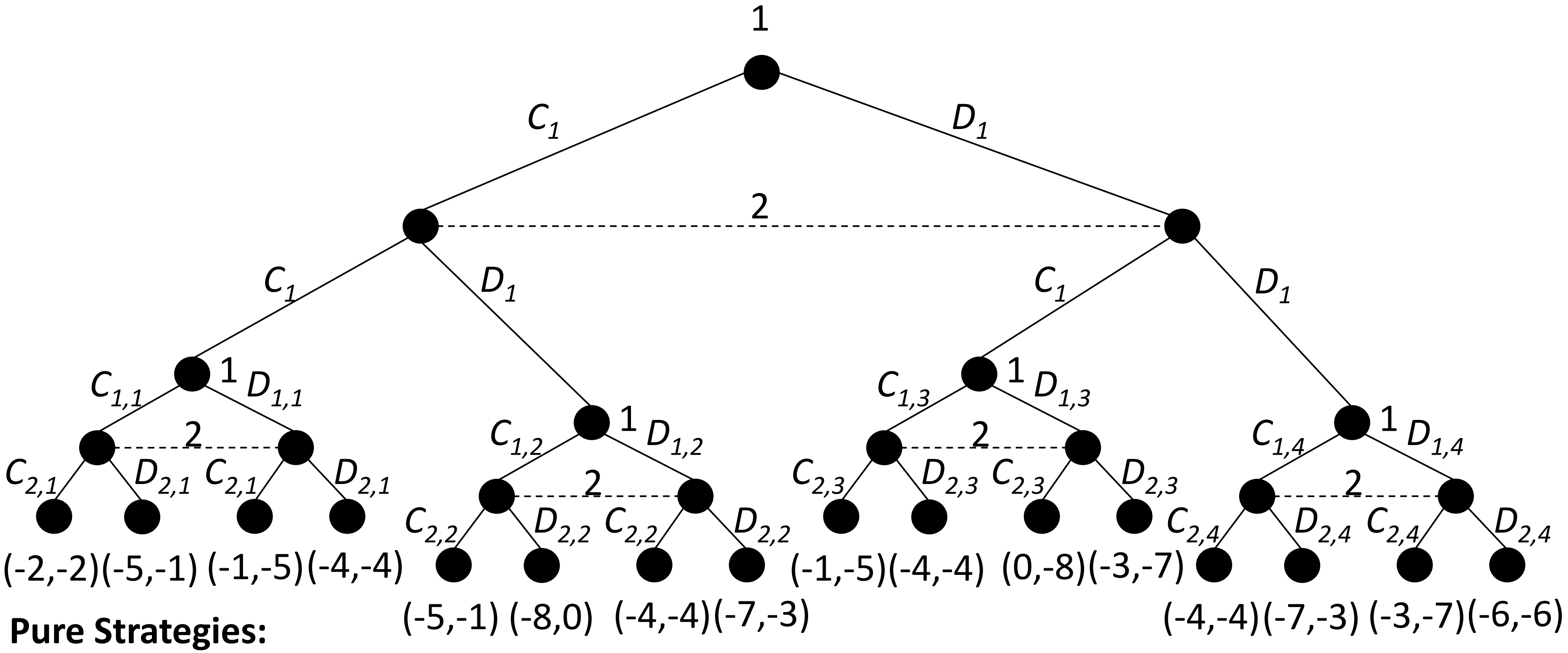
Finitely-Repeated Games

- Let us analyze the Prisoner's Dilemma in a two-stage game

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

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

Finitely-Repeated Games



Pure Strategies:

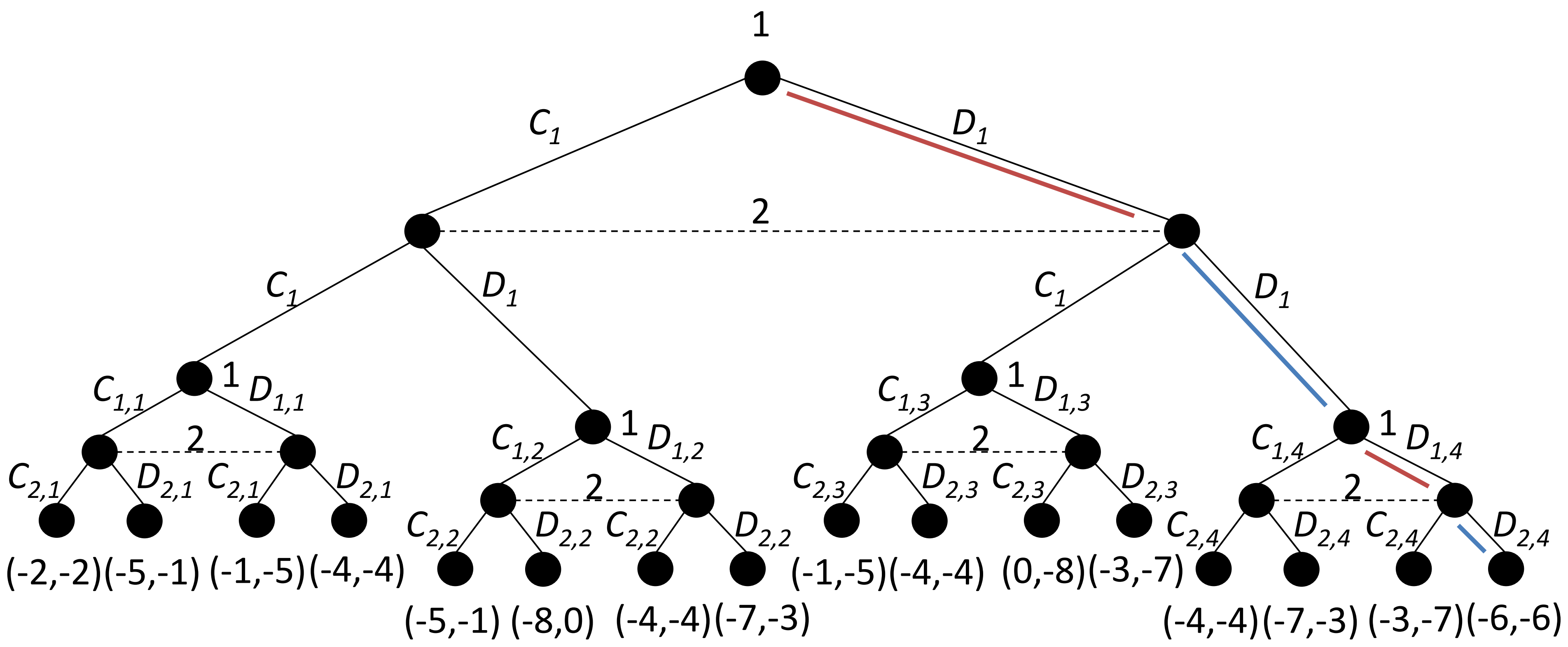
Agent 1:

{$(C_1, C_{1,1})$, $(C_1, D_{1,1})$, $(C_1, C_{1,2})$, $(C_1, D_{1,2})$, $(C_1, C_{1,3})$, $(C_1, D_{1,3})$, $(C_1, C_{1,4})$, $(C_1, D_{1,4})$,
$(D_1, C_{1,1})$, $(D_1, D_{1,1})$, $(D_1, C_{1,2})$, $(D_1, D_{1,2})$, $(D_1, C_{1,3})$, $(D_1, D_{1,3})$, $(D_1, C_{1,4})$, $(D_1, D_{1,4})$}

Agent 2:

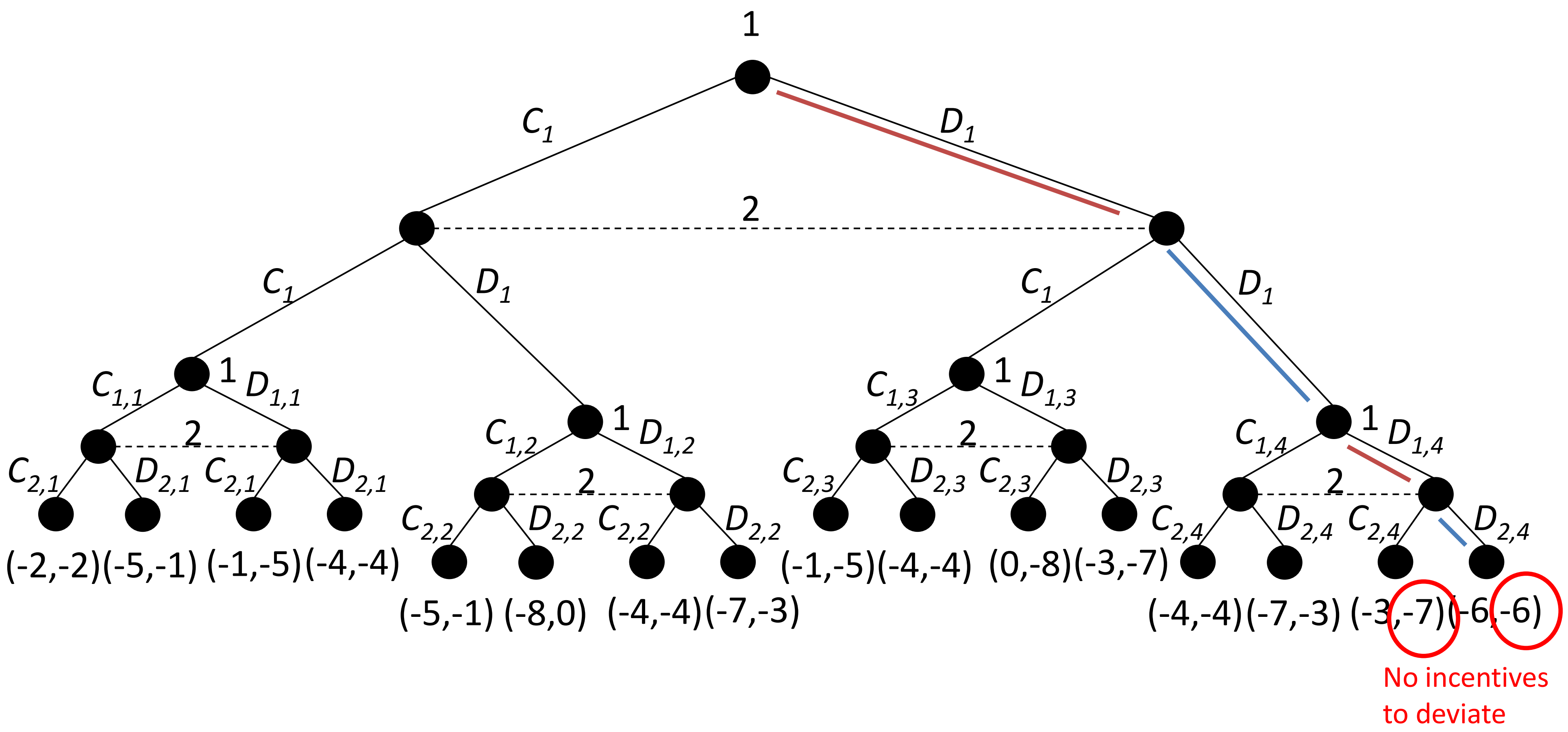
{$(C_1, C_{2,1})$, $(C_1, D_{2,1})$, $(C_1, C_{2,2})$, $(C_1, D_{2,2})$, $(C_1, C_{2,3})$, $(C_1, D_{2,3})$, $(C_1, C_{2,4})$, $(C_1, D_{2,4})$,
$(D_1, C_{2,1})$, $(D_1, D_{2,1})$, $(D_1, C_{2,2})$, $(D_1, D_{2,2})$, $(D_1, C_{2,3})$, $(D_1, D_{2,3})$, $(D_1, C_{2,4})$, $(D_1, D_{2,4})$,}

Finitely-Repeated Games



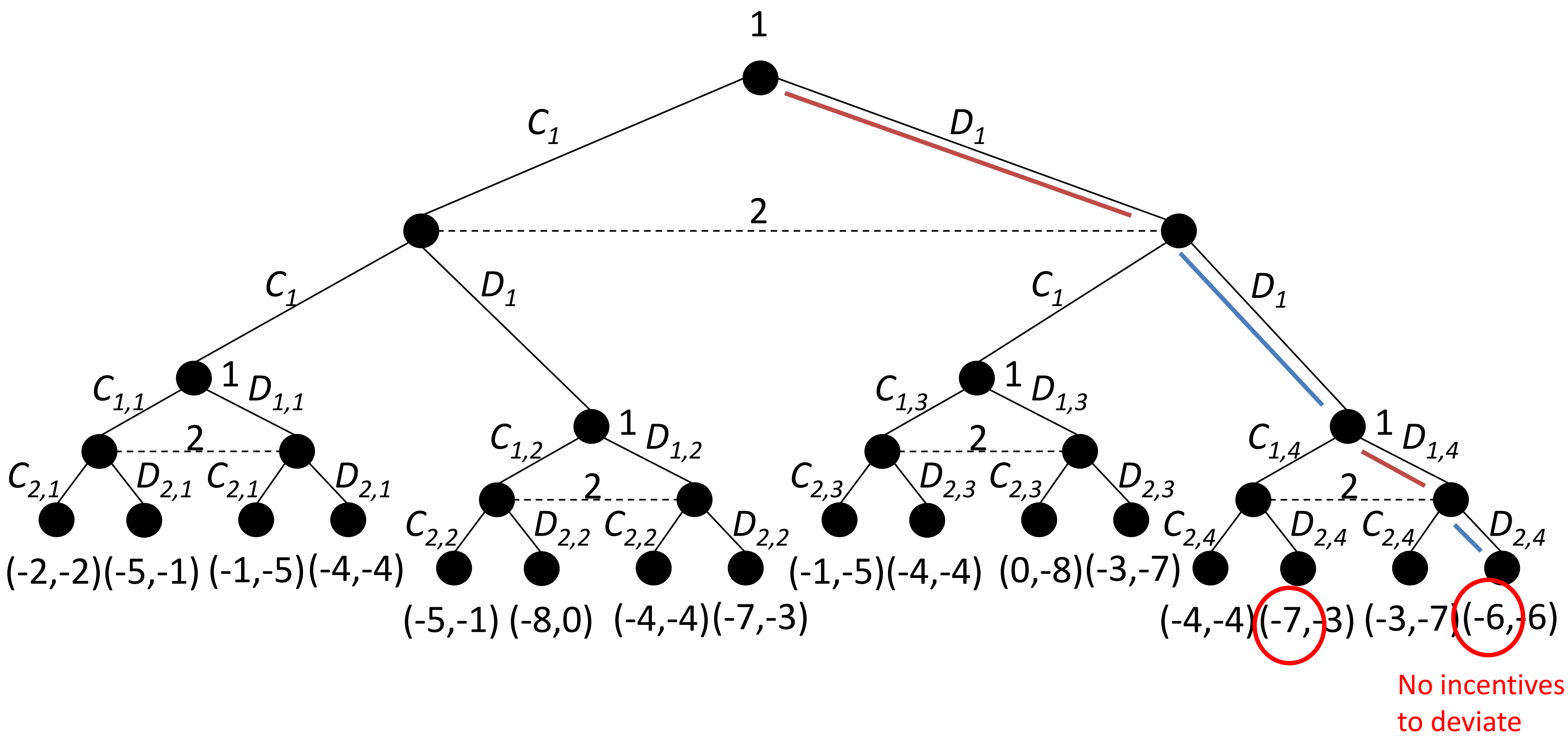
Is the strategy profile (D_1,$D_{1,4}$), (D_1,$D_{2,4}$) a Nash equilibrium?

Finitely-Repeated Games



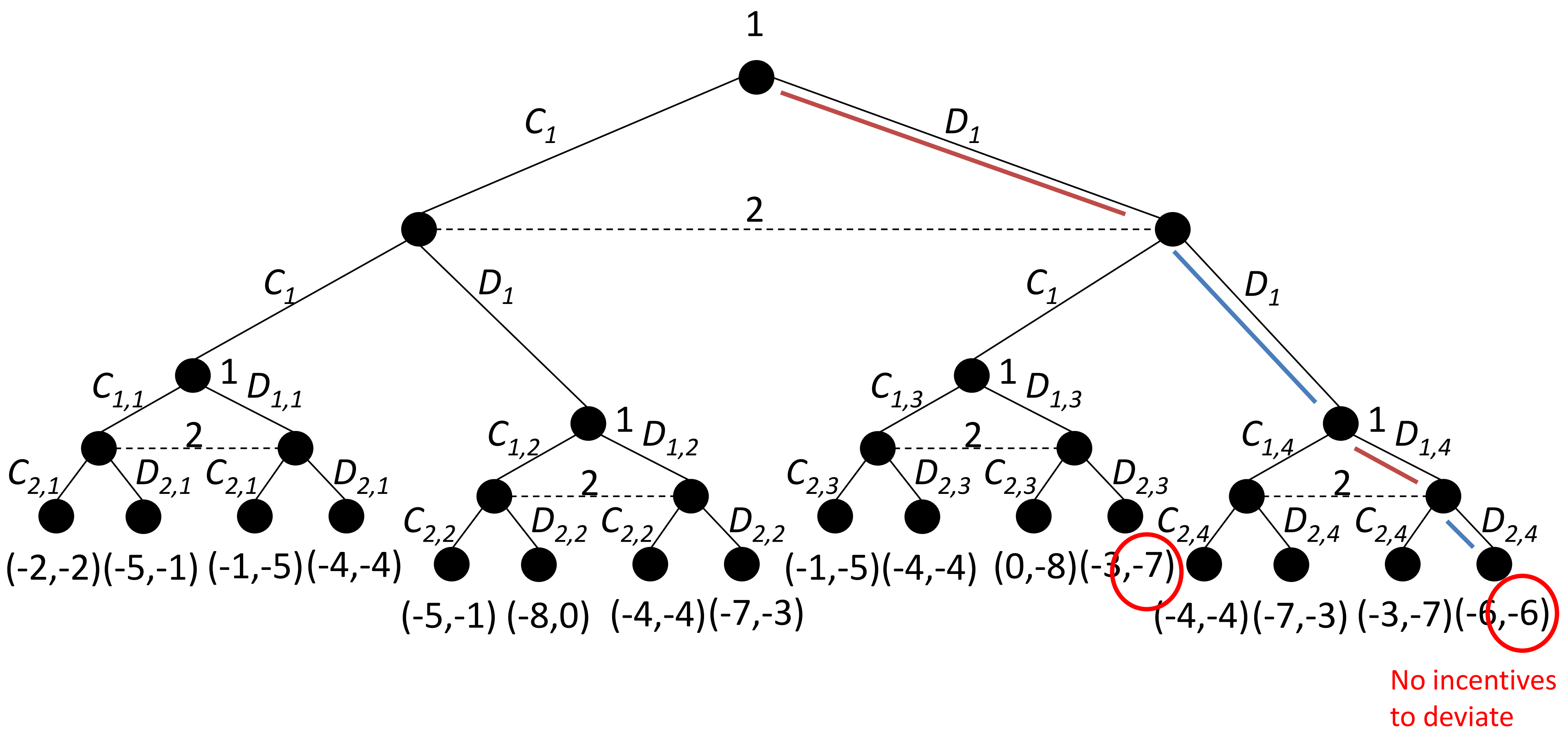
Is the strategy profile $(D_1, D_{1,4})$, $(D_1, D_{2,4})$ a Nash equilibrium?

Finitely-Repeated Games



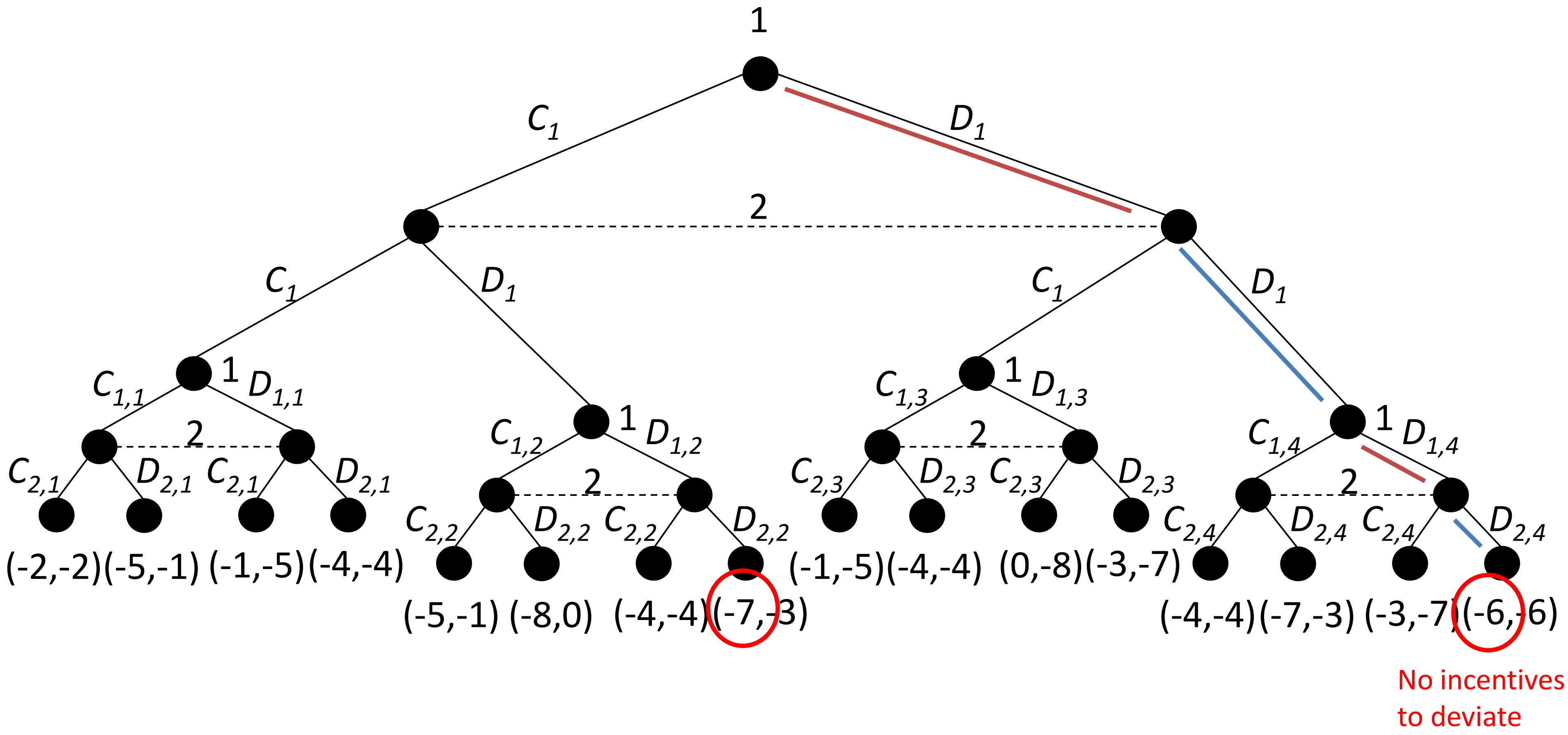
Is the strategy profile $(D_1, D_{1,4})$, $(D_1, D_{2,4})$ a Nash equilibrium?

Finitely-Repeated Games



Is the strategy profile $(D_1,D_{1,4})$, $(D_1,D_{2,4})$ a Nash equilibrium?

Finitely-Repeated Games



Is the strategy profile $(D_1, D_{1,4}), (D_1, D_{2,4})$ a Nash equilibrium?

YES

Finitely-Repeated Games

- **Proposition:** *If the stage game G has a unique Nash equilibrium then, for any finite T, the repeated game G(T) has a unique outcome: the Nash equilibrium of G is played in every subgame-perfect stage*

Outline

- Introduction to Repeated games
- Finitely-repeated games
- **Infinitely-repeated games**
- Folk Theorem
- Replicator dynamics

Infinitely-Repeated Games

- Let us now consider a **game is repeated an infinite number of times**

- Can we represent this as an **extensive-form game**?
  - An **infinite tree**!
  - Sum of payoffs is **infinity**!

Infinitely-Repeated Games

- **Definition**: Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for agent i, the **average reward** of agent i is:

$$\lim_{k \to \infty} \sum_{j=1}^{k} \frac{r_j}{k}$$

Infinitely-Repeated Games

- **Definition**: Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for agent i and discount factor β with $0 < \beta < 1$, the **future discounted reward** of agent i is:

$$\sum_{j=1}^{\infty} \beta^j r_j$$

- **Interpretation**: the agent cares more about her well-being in the near term than in the long term

Infinitely-Repeated Games

- What is a **pure strategy** in an infinitely-repeated game?
  - A **choice of action at every decision point**
    - In these games, this means **an action at every stage game** (for an infinite number of actions!)

Infinitely-Repeated Games

- More formally:
  - Histories of length t:
    - $H^t = \{h^t : h^t = (a^1, \ldots, a^t) \in A^t)$
    - $a^t = (a_1^t, \ldots, a_n^t)$ and $a_i^t \in A_i$
  - All finite histories: $H = \cup H^t$
  - A pure strategy: $s_i : H \to A_i$

Infinitely-Repeated Games

- Prisoner's Dilemma

  - $A_i = \{C, D\}$

  - A history of length $t = 3$:
    - $t = 1$: (C, C)
    - $t = 2$: (C, D)
    - $t = 3$: (D, D)

  - A strategy for period 4 would specify what an agent would do after seeing $(C, C), (C, D), (D, D)$

Infinitely-Repeated Games

- Some famous strategies (repeated Prisoner's Dilemma):
  - **Tit-for-tat**: Start out cooperating. If the opponent defected, defect in the next round. Then go back to cooperation.
    - **Agent 1: C, C, C, D, C, C,…**
    - Agent 2: C, C, D, C, C, C,…

Infinitely-Repeated Games

- Some famous strategies (repeated Prisoner's Dilemma):
  - **Trigger**: Start out cooperating. If the opponent ever defects, defect forever.
    - **Agent 1: C, C, C, D, D, D,…**
    - Agent 2: C, C, D, C, C, C,…

Infinitely-Repeated Games

- **Subgame-perfect equilibrium**
  - Profile of strategies that are a **Nash equilibrium in every subgame**
    - Hence, 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

Outline

- Introduction to Repeated games
- Finitely-repeated games
- Infinitely-repeated games
- **Folk Theorem**
- Replicator dynamics

Folk Theorem

- There are many **Folk Theorems** for infinitely-repeated games
  - **All are concerned with Nash equilibria** of an infinitely repeated game
  - This result was called the Folk Theorem because it **was widely known among game theorists** in the 1950s, even though **no one had published it**

Folk Theorem

- Consider a **finite normal form game $\boldsymbol{G}$** (stage game)
- Let $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ **be a Nash equilibrium** of the stage game G
- If $\boldsymbol{a}' = (\boldsymbol{a}'_1, \boldsymbol{a}'_2, \dots, \boldsymbol{a}'_n)$ **is such that $\boldsymbol{u}_i(\boldsymbol{a}') > \boldsymbol{u}_i(\boldsymbol{a})$** for all i, then:
  - there **exists a discount factor $\boldsymbol{\beta}$** $(0 < \beta < 1)$, such that $\beta_i \geq \beta$ for all i, and
  - there **exists a subgame perfect equilibrium of the infinite repetition of $\boldsymbol{G}$ that has $\boldsymbol{a}'$ played in every period on the equilibrium path**

Folk Theorem

- Outline of the Proof:
  - **Play a'** as long as **every agent is also playing**
  - If **any agent ever deviates**, then **play a forever after** (Trigger strategy)
  - 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 α forever** (if anyone deviated) is a **Nash equilibrium** in any subgame
  - Will **someone gain by deviating from α'** if nobody has in the past?

Folk Theorem

- Check that this is a **subgame-perfect equilibrium for high enough discount factors**
  - Will **someone gain by deviating from $\boldsymbol{a}$'** if nobody has in the past?
    - **Maximum gain from deviating** (over all agents) is

$$M = \max_{i,a''} u_i(a_i'', a'_{-i}) - u_i(a')$$

    - **Minimum per-period loss from future punishment** is

$$m = \min_i u_i(a') - u_i(a)$$

Folk Theorem

- Check that this is a **subgame-perfect equilibrium for high enough discount factors**
  - Will **someone gain by deviating from $\boldsymbol{a}$'** if nobody has in the past?
    - If an agent deviates, given the other agents' strategies, the **maximum possible net gain** is

$$M - m\frac{\beta_i}{1-\beta_i}$$

Folk Theorem

- Check that this is a **subgame-perfect equilibrium for high enough discount factors**
  - Will **someone gain by deviating from $\boldsymbol{a}$'** if nobody has in the past?
    - **Deviation is not beneficial** if

$$M - m\frac{\beta_i}{1-\beta_i} \leq 0$$

$$\frac{M}{m} \leq \frac{\beta_i}{1-\beta_i}$$

$$\beta_i \geq \frac{M}{M+m}\text{, for all } i$$

Infinitely-Repeated Games

- Example: **Prisoner's Dilemma – Can we sustain cooperation?**

We want to sustain cooperation

	C	***D***
C	$3, 3$	$0, 5$
D	$5, 0$	$1, 1$

Nash equilibrium of the stage game

Infinitely-Repeated Games

- Example: **Prisoner's Dilemma – Can we sustain cooperation?**
  - Agent cooperates as long as the other agent is cooperating in the past
  - Both agents defect if anyone deviates (**Trigger strategy**)

When is this an equilibrium? Is there a value of β that could sustain this equilibrium above?

	C	***D***
C	$3, 3$	$0, 5$
D	$5, 0$	$1, 1$

Infinitely-Repeated Games

- Example: **Prisoner's Dilemma – Can we sustain cooperation?**
  - **Always cooperate**: $3 + \beta 3 + \beta^2 3 + \beta^3 3 + \cdots = \frac{3}{1-\beta}$
  - **Always defect**: $5 + \beta 1 + \beta^2 1 + \beta^3 1 + \cdots = 5 + \beta \frac{1}{1-\beta}$

	C	***D***
C	$3, 3$	$0, 5$
D	$5, 0$	$1, 1$

Infinitely-Repeated Games

- Example: **Prisoner's Dilemma – Can we sustain cooperation?**

  - **Always cooperate**: $3 + \beta 3 + \beta^2 3 + \beta^3 3 + \cdots = 3 + \beta \frac{3}{1-\beta}$

  - **Always defect**: $5 + \beta 1 + \beta^2 1 + \beta^3 1 + \cdots = 5 + \beta \frac{1}{1-\beta}$

  - **Difference**: $-2 + \beta 2 + \beta^2 2 + \beta^3 2 + \cdots = -2 + \beta \frac{2}{1-\beta}$

Infinitely-Repeated Games

- Example: **Prisoner's Dilemma – Can we sustain cooperation?**
  - **Difference**: $-2 + \beta 2 + \beta^2 2 + \beta^3 2 + \cdots = \beta \frac{2}{1-\beta} - 2$
  - If we want the "**always cooperate**" to have a higher payoff:
    - $\beta \frac{2}{1-\beta} - 2 \geq 0$
    - $\beta \geq \frac{1}{2}$
    - **Interpretation**: if we want to sustain cooperation, the agent needs to care about tomorrow at least as half as much as it cares about today!

Exercise

- Could we sustain cooperation in this infinitely-repeated game?

	C	***D***
C	$3, 3$	$0, 10$
D	$10, 0$	$1, 1$

Defection is more attractive!

Repeated Games

- Agents can condition future decisions based on past actions

- This generates many equilibria: Folk Theorems

- Check “The Evolution of Trust”, Nicky Case
  - https://ncase.me/trust/

Outline

- Introduction to Repeated games
- Finitely-repeated games
- Infinitely-repeated games
- Folk Theorem
- **Replicator dynamics**

Evolutionary Learning and other Large-population models

- Consider large populations
- Model the evolution of behavior of the whole population
- Standard model technique in biology, social networks, human population behavior.

MULTIAGENT SYSTEMS
Algorithmic, Game-Theoretic, and Logical Foundations
Yoav Shoham, Kevin Leyton-Brown, 2008

Evolutionary Learning and other Large-population models

The *replicator dynamic* models a population undergoing frequent replicator interactions.

Replicator dynamic

Basic interaction between any two agents

	A	**B**
A	x, x	u, v
B	v, u	y,y

agents have no distinct roles (symmetric game)

Evolutionary Learning and other Large-population models

Replicator dynamic

	A	B
A	x, x	u, v
B	v, u	y,y

- $\varphi_t(A)$ – number of agents playing A at time t
- $\theta_t(A) = \frac{\varphi_t(A)}{\sum_a \varphi_t(a)}$ - ratio of agents playing A at time t
- $u_t(a) = \sum_b \theta_t(b) u(a, b)$

Evolutionary Learning and other Large-population models

Replicator dynamic

- $u_t(a) = \sum_b \theta_t(b) u(a,b)$
- $u_t^* = \sum_a \theta_t(a) u_t(a)$

The number of agents change proportionally to their utility.

- $\dot{\varphi}_t(a) = \varphi_t(a) u_t(a)$
- $\theta_t(A) = \frac{\varphi_t(A)}{\sum_a \varphi_t(a)}$

- $\dot{\theta}_t(A) = \frac{\dot{\varphi}_t(A) \sum_a \varphi_t(a) - \varphi_t(A) \sum_a \dot{\varphi}_t(a)}{(\sum_a \varphi_t(a))^2} = \theta_t(a)[u_t(a) - u_t^*]$

Replicator dynamics

```
def dynreplicator(g,th):
    va = g.payoff_matrices[0]@th
    v = th@va
    dth = th * (va-v)

    return dth

ddx = np.zeros((100,100))
ddy = np.zeros((100,100))

for kk in Games.keys():
    g = Games[kk]
    for ix in np.linspace(0, .99, 10,):
        for iy in np.linspace(0, .99, 10):
            dth = dynreplicator(g,[ix,iy])
            ddx[int(ix*10),int(iy*10)],ddy[int(ix*10),int(iy*10)] = dth
    plt.figure()
    plt.subplot(1,2,1)
    plt.quiver(xx,yy,ddx,ddy)
    plt.title(kk)

    th = [0.7,0.3]
    TH = []
    for ii in np.arange(0,20,dt):

        dth = dynreplicator(g,th)
        th = np.clip(th + dth * dt, 0, 1)
        th /= np.sum(th)
        TH += [th]
    plt.subplot(1,2,2)
    plt.plot(TH)
    plt.title(kk)
```

Replicator dynamics

Definition 7.7.2 (Steady state) *A* steady state *of a population using the replicator dynamic is a population state* θ *such that for all* $a \in A$, $\dot{\theta}(a) = 0$.

Definition 7.7.3 (Stable steady state) *A steady state* θ *of a replicator dynamic is* stable *if for every neighborhood* U *of* θ *there is another neighborhood* U' *of* θ *such* state *that if* $\theta_0 \in U'$ *then* $\theta_t \in U$ *for all* $t > 0$.

Definition 7.7.4 (Asymptotically stable state) *A steady state* θ *of a replicator dynamic is* asymptotically stable *if it is stable, and in addition if for every neighborhood* U *of* θ *it is the case that if* $\theta_0 \in U$ *then* $\lim_{t\to\infty} \theta_t = \theta$.

What is the equilibria? Is it stable?

	A	B
A	$0, 0$	$1, 1$
B	$1, 1$	$0, 0$

Figure 7.9: The Anti-Coordination game.

Anti-Coordination Game

- $u(a)=0\,\theta + 1\,(1-\theta) = 1-\theta$
- $u(b)=1\,\theta + 0\,(1-\theta) = \theta$

- $u^*=(1-\theta)\,\theta+\theta\,(1-\theta) = 2\,(1-\theta)\,\theta$

- $d\theta= \theta(1-\theta-2\,\theta(1-\theta)) = \theta(1-3\,\theta+2\,\theta\text{^}2)$

- $d\theta=0$ for $\theta = 0.5$
- $d\theta>0$ for $\theta<0.5$
- $d\theta>0$ for $\theta>0.5$

Stag-Hare

- What are the Nash Eq?
- Is there an assymptotic equilibria?

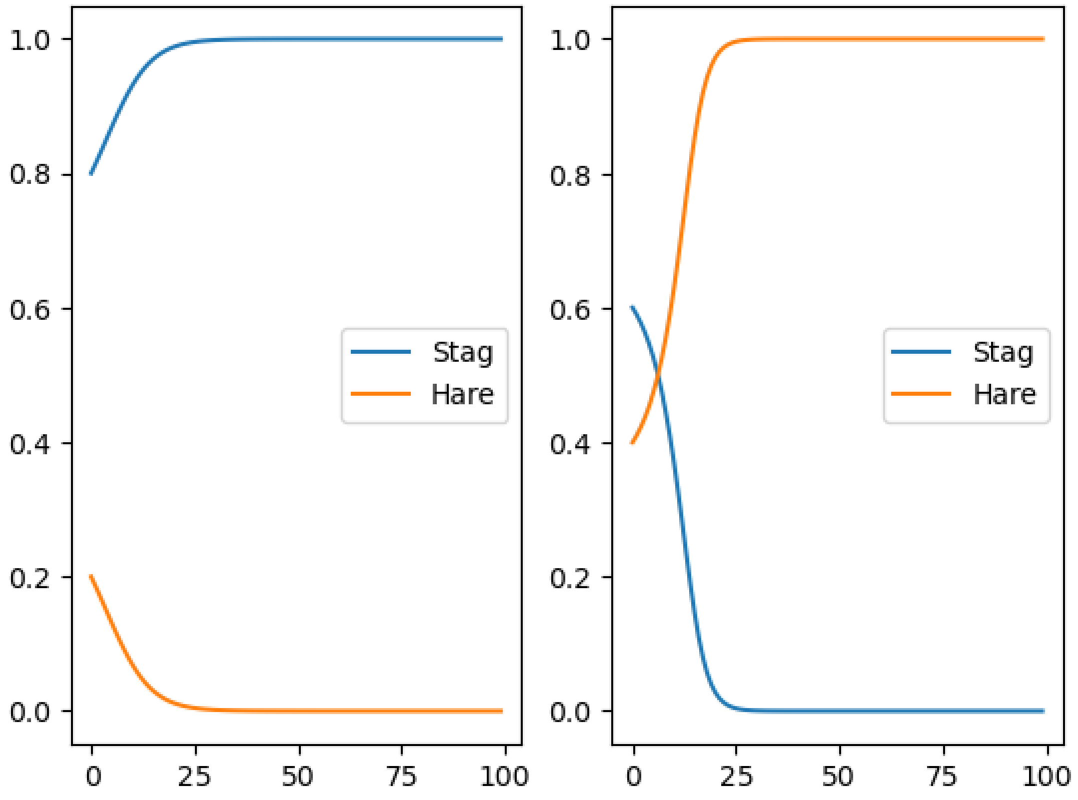
	Stag	Hare
Stag	10, 10	1, 8
Hare	8, 1	5, 5

Stag-Hare

	Stag	Hare
Stag	10, 10	1, 8
Hare	8, 1	5, 5

- u(s)=10 θ+ 1 (1- θ) = 9 θ+1
- u(h)=8 θ+5(1- θ) = 3 θ+5
- u* = (9 θ+1) θ+(3 θ+5)(1- θ)=6 θ^2- θ+5
- u(s)-u*=(9 θ+1)-(6 θ^2- θ+5)=-6θ^2+10θ-4
- u(s)-u*=0 => θ=1 V θ = 2/3

Stag-Hare



- u(s)=10 θ+ 1 (1- θ) = 9 θ+1
- u(h)=8 θ+5(1- θ) = 3 θ+3
- u* = (9 θ+1) θ+(3 θ+5)(1- θ)=6 θ^2- θ+5
- u(s)-u*=(9 θ+1)-(6 θ^2- θ+5)=-6θ^2+10θ-4
- u(s)-u*=0 => θ=1 V θ = 2/3

Replicator dynamics

- In a given game the populations will converge to the Nash equilibria?
- What if there is more than one?
- And if there is none?

Replicator dynamics

Theorem 7.7.5 *Given a normal-form game* $G = (\{1, 2\}, A = \{a_1, \ldots, a_k\}, u)$, *if the strategy profile* (S, S) *is a (symmetric) mixed strategy Nash equilibrium of* G *then the population share vector* $\theta = (S(a_1), \ldots, S(a_k))$ *is a steady state of the replicator dynamic of* G.

Theorem 7.7.6 *Given a normal-form game* $G = (\{1, 2\}, A\{a_1, \ldots, a_k\}, u)$ *and a mixed strategy* S, *if the population share vector* $\theta = (S(a_1), \ldots, S(a_k))$ *is a stable steady state of the replicator dynamic of* G, *then the strategy profile* (S, S) *is a mixed strategy Nash equilibrium of* G.

Replicator dynamics

Definition 7.7.7 (Trembling-hand perfect equilibrium) *A mixed strategy* S *is a* (trembling-hand) perfect equilibrium *of a normal-form game* G *if there exists a sequence* $S_0, S_1, \ldots$ *of fully mixed-strategy profiles such that* $\lim_{n\to\infty} S_n = S$, *and such that for each* S_k *in the sequence and each player* i, *the strategy* s_i *is a best response to the strategies* s_{k-i}.

Theorem 7.7.8 *Given a normal-form game* $G = (\{1, 2\}, A, u)$ *and a mixed strategy* S, *if the population share vector* $\theta = (S(a_1), \ldots, S(a_k))$ *is an asymptotically stable steady state of the replicator dynamic of* G, *then the strategy profile* (S, S) *is a Nash equilibrium of* G *that is trembling-hand perfect and isolated.*

Evolutionarily stable strategies

Unlike the steady states discussed earlier, it does not require the replicator dynamic; rather it is a static solution concept.
Roughly speaking, an evolutionarily stable strategy is a mixed strategy that is “resistant to invasion” by new strategies.

Definition 7.7.10 (Weak ESS) S *is a* weak evolutionarily weak stable strategy *if and only if for some* $\varrho > 0$ *and for all* S' *it is the case that either* $u(S, S) > u(S', S)$ *holds, or else both* $u(S, S) = u(S', S)$ *and* $u(S, S') \geq u(S', S')$ *hold.*

Evolutionarily stable strategies

Unlike the steady states discussed earlier, it does not require the replicator dynamic; rather it is a static solution concept.
Roughly speaking, an evolutionarily stable strategy is a mixed strategy that is “resistant to invasion” by new strategies.

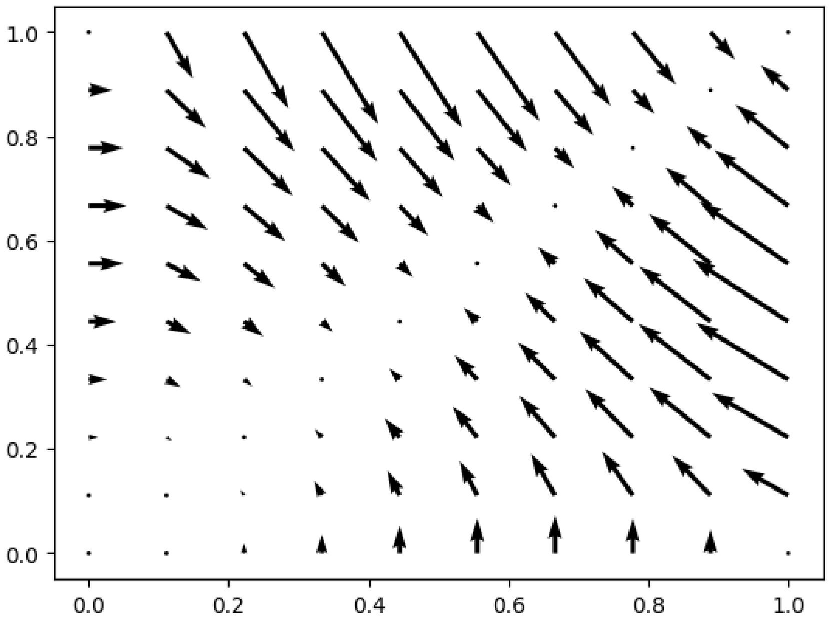
Definition 7.7.10 (Weak ESS) S *is a* weak evolutionarily weak stable strategy *if and only if for some* ϱ > 0 *and for all* S’ *it is the case that either* u(S, S) > u(S’, S) *holds, or else both* u(S, S) = u(S’, S) *and* u(S, S’) ≥ u(S’, S’) *hold.*

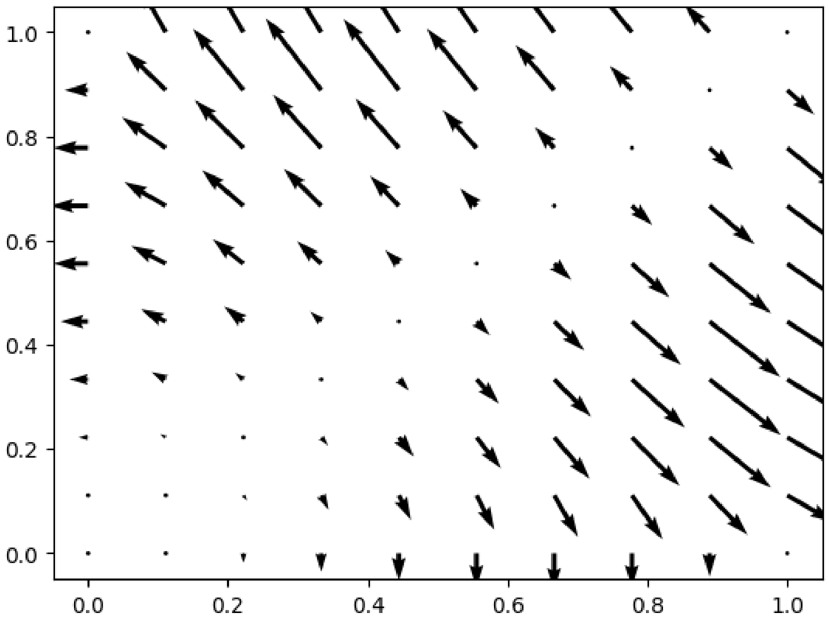
Theorem 7.7.11 *Given a symmetric two-player normal-form game* $G = (\{1, 2\}, A, u)$ *and a mixed strategy* S, *if* S *is an evolutionarily stable strategy then* (S, S) *is a Nash equilibrium of* G.

Theorem 7.7.12 *Given a symmetric two-player normal-form game* $G = (\{1, 2\}, A, u)$ *and a mixed strategy* S, *if* (S, S) *is a strict (symmetric) Nash equilibrium of* G, *then* S *is an evolutionarily stable strategy.*

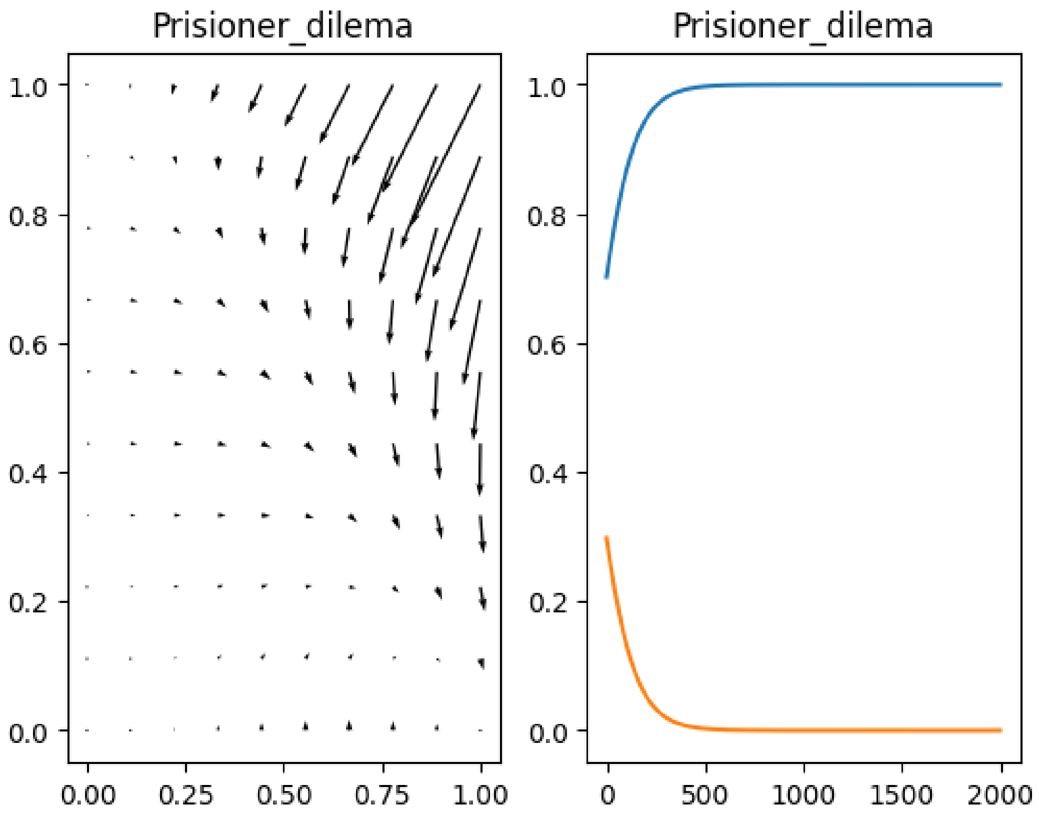
Theorem 7.7.13 *Given a symmetric two-player normal-form game* $G = (\{1, 2\}, A, u)$ *and a mixed strategy* S, *if* S *is an evolutionarily stable strategy then it is an asymptotically stable steady state of the replicator dynamic of* G.

Matching Pennies

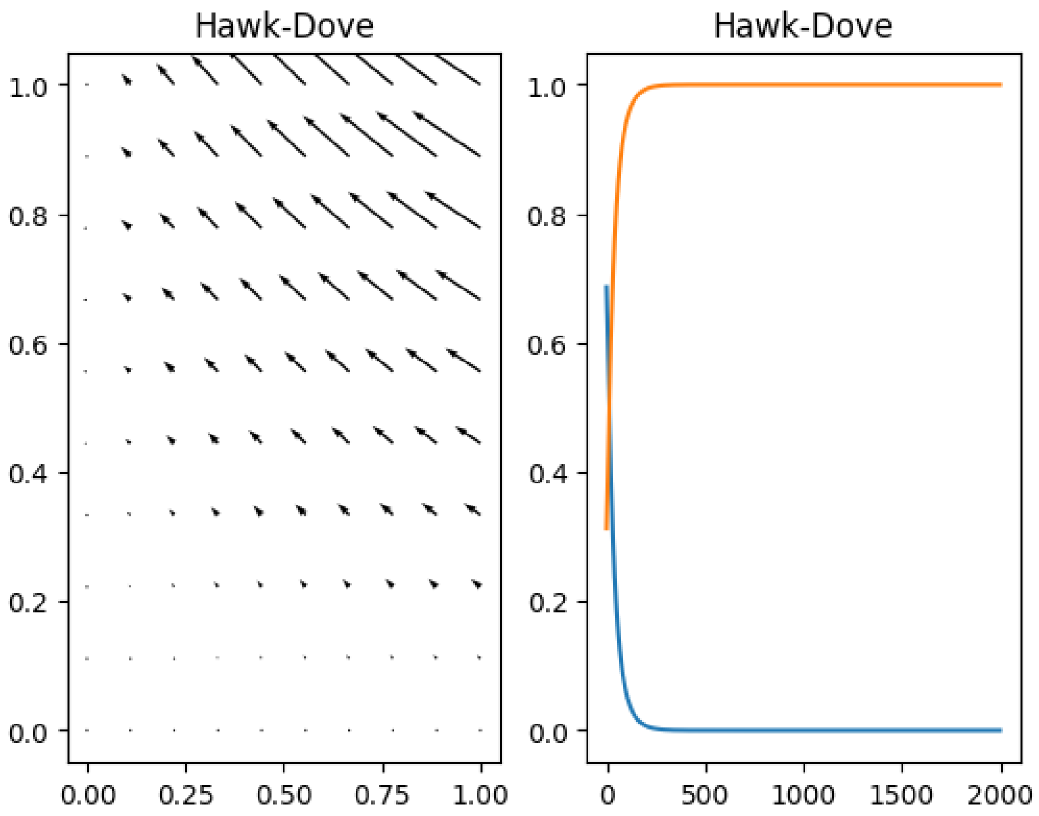




Prisioner dilema



Hawk-Dove



Hawk-Dove
Hawk-Dove
1.0
0.8
0.6
0.4
0.2
0.0
0.00
0.25
0.50
0.75
1.00
1.0
0.8
0.6
0.4
0.2
0.0
0
500
1000
1500
2000

Thank You

rui.prada@tecnico.ulisboa.pt
manuel.lopes@tecnico.ulisboa.pt
alberto.sardinha