Multi-agent learning

Repeated games

Gerard Vreeswijk, Intelligent Systems Group, Computer Science Department, Faculty of Sciences, Utrecht University, The Netherlands.

Friday 3rd May, 2019

Repeated games: motivation

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.
2. Much learning in MASs can therefore be modelled through *learning in games*.

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.
2. Much learning in MASs can therefore be modelled through *learning in games*.
3. Learning in games usually takes place through the (gradual) adaptation of strategies (hence, behaviour) in a repeated game.

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.
2. Much learning in MASs can therefore be modelled through *learning in games*.
3. Learning in games usually takes place through the (gradual) adaptation of strategies (hence, behaviour) in a repeated game.
4. In most repeated games, one game (a.k.a. stage game) is played repeatedly. Possibilities:

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.
2. Much learning in MASs can therefore be modelled through *learning in games*.
3. Learning in games usually takes place through the (gradual) adaptation of strategies (hence, behaviour) in a repeated game.
4. In most repeated games, one game (a.k.a. stage game) is played repeatedly. Possibilities:
   - A finite number of times.

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.
2. Much learning in MASs can therefore be modelled through *learning in games*.
3. Learning in games usually takes place through the (gradual) adaptation of strategies (hence, behaviour) in a repeated game.
4. In most repeated games, one game (a.k.a. stage game) is played repeatedly. Possibilities:
   - A finite number of times.
   - An indefinite (same: *indeterminate*) number of times.

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.
2. Much learning in MASs can therefore be modelled through *learning in games*.
3. Learning in games usually takes place through the (gradual) adaptation of strategies (hence, behaviour) in a repeated game.
4. In most repeated games, one game (a.k.a. stage game) is played repeatedly. Possibilities:
   - A finite number of times.
   - An indefinite (same: *indeterminate*) number of times.
   - An infinite number of times.

Repeated games: motivation

1. Much interaction in MASs can be modelled through *games*.
2. Much learning in MASs can therefore be modelled through *learning in games*.
3. Learning in games usually takes place through the (gradual) adaptation of strategies (hence, behaviour) in a repeated game.
4. In most repeated games, one game (a.k.a. stage game) is played repeatedly. Possibilities:
   - A finite number of times.
   - An indefinite (same: *indeterminate*) number of times.
   - An infinite number of times.
5. Therefore, familiarity with the basic concepts and results from the theory of repeated games is essential to understand MAL.

Plan for today

Plan for today

- NE in normal form games that are repeated a finite number of times.

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.
  - Discount factor. Models the probability of continuation.

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.
  - Discount factor. Models the probability of continuation.
  - Folk theorem. (Actually many FT's.) Repeated games generally do have infinitely many Nash equilibria.

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.
  - Discount factor. Models the probability of continuation.
  - Folk theorem. (Actually many FT's.) Repeated games generally do have infinitely many Nash equilibria.
  - Trigger strategy.

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.
  - Discount factor. Models the probability of continuation.
  - Folk theorem. (Actually many FT's.) Repeated games generally do have infinitely many Nash equilibria.
  - Trigger strategy. "on-path" vs. "off-path" play

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.
  - Discount factor. Models the probability of continuation.
  - Folk theorem. (Actually many FT's.) Repeated games generally do have infinitely many Nash equilibria.
  - Trigger strategy. "on-path" vs. "off-path" play, "minmax" as a threat.

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.
  - Discount factor. Models the probability of continuation.
  - Folk theorem. (Actually many FT's.) Repeated games generally do have infinitely many Nash equilibria.
  - Trigger strategy. "on-path" vs. "off-path" play, "minmax" as a threat.

This presentation draws heavily on (Peters, 2008).

Plan for today

- NE in normal form games that are repeated a finite number of times.
  - Principle of backward induction.
- NE in normal form games that are repeated an indefinite number of times.
  - Discount factor. Models the probability of continuation.
  - Folk theorem. (Actually many FT's.) Repeated games generally do have infinitely many Nash equilibria.
  - Trigger strategy. "on-path" vs. "off-path" play, "minmax" as a threat.

This presentation draws heavily on (Peters, 2008).

* H. Peters (2008): *Game Theory: A Multi-Leveled Approach*. Springer, ISBN: 978-3-540-69290-4. Ch. 8: Repeated games.

Part I:
Nash equilibria

Part I:
Nash equilibria
in normal form games

Part I:
Nash equilibria
in normal form games
that are repeated

Part I:
Nash equilibria
in normal form games
that are repeated
a finite number of times

Nash equilibria in playing the PD twice

Prisoners' Dilemma		*Other:* Cooperate	Defect
You:	Cooperate	$(3, 3)$	$(0, 5)$
	Defect	$(5, 0)$	$(1, 1)$

Nash equilibria in playing the PD twice

Prisoners' Dilemma		*Other:* Cooperate	Defect
You:	Cooperate	$(3, 3)$	$(0, 5)$
	Defect	$(5, 0)$	$(1, 1)$

- Even if mixed strategies are allowed, the PD possesses one Nash equilibrium, viz. (D, D) with payoffs $(1, 1)$.

Nash equilibria in playing the PD twice

Prisoners' Dilemma		*Other:* Cooperate	Defect
You:	Cooperate	$(3, 3)$	$(0, 5)$
	Defect	$(5, 0)$	$(1, 1)$

- Even if mixed strategies are allowed, the PD possesses one Nash equilibrium, viz. (D, D) with payoffs $(1, 1)$.
- This equilibrium is Pareto sub-optimal.

Nash equilibria in playing the PD twice

Prisoners' Dilemma		*Other:* Cooperate	Defect
You:	Cooperate	$(3, 3)$	$(0, 5)$
	Defect	$(5, 0)$	$(1, 1)$

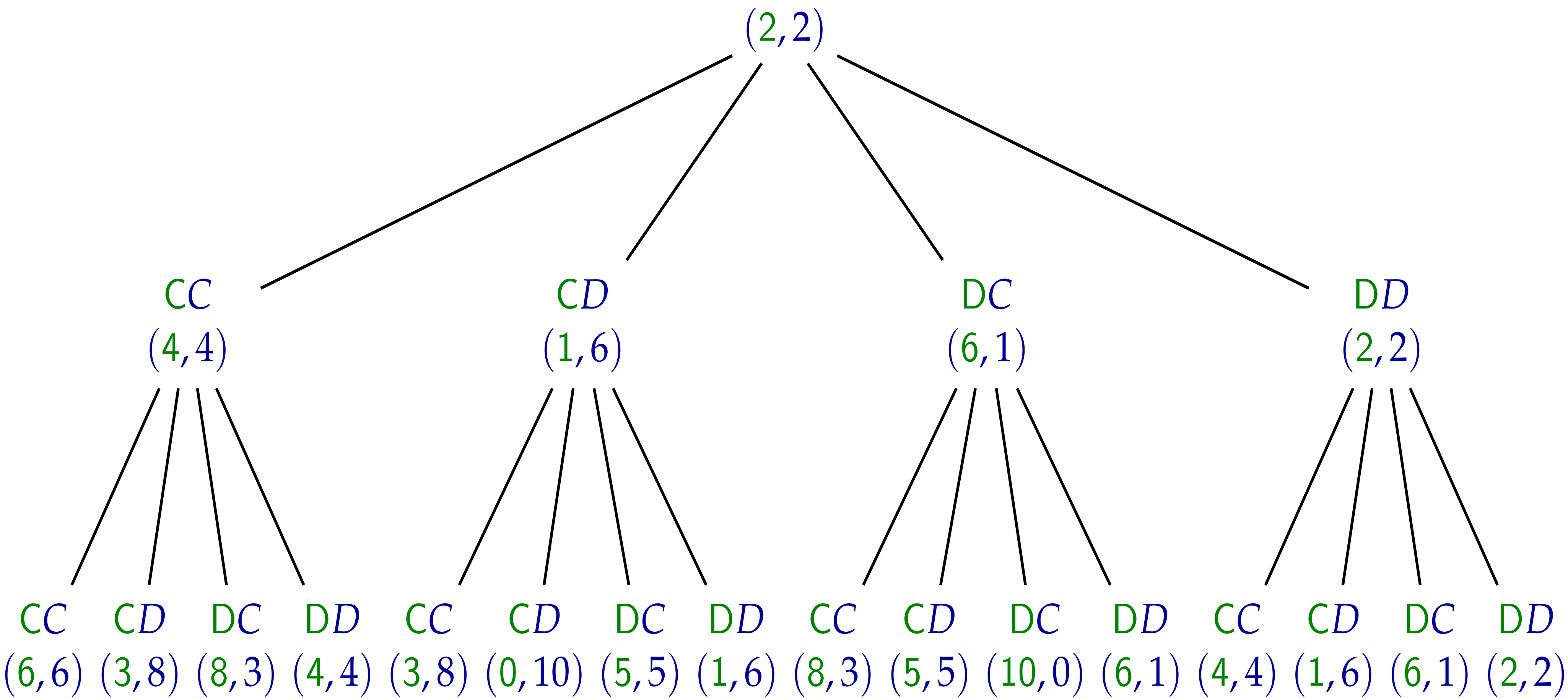
- Even if mixed strategies are allowed, the PD possesses one Nash equilibrium, viz. (D, D) with payoffs $(1, 1)$.
- This equilibrium is Pareto sub-optimal.
- Does the situation change if two parties get to play the Prisoners' Dilemma two times in succession?

Nash equilibria in playing the PD twice

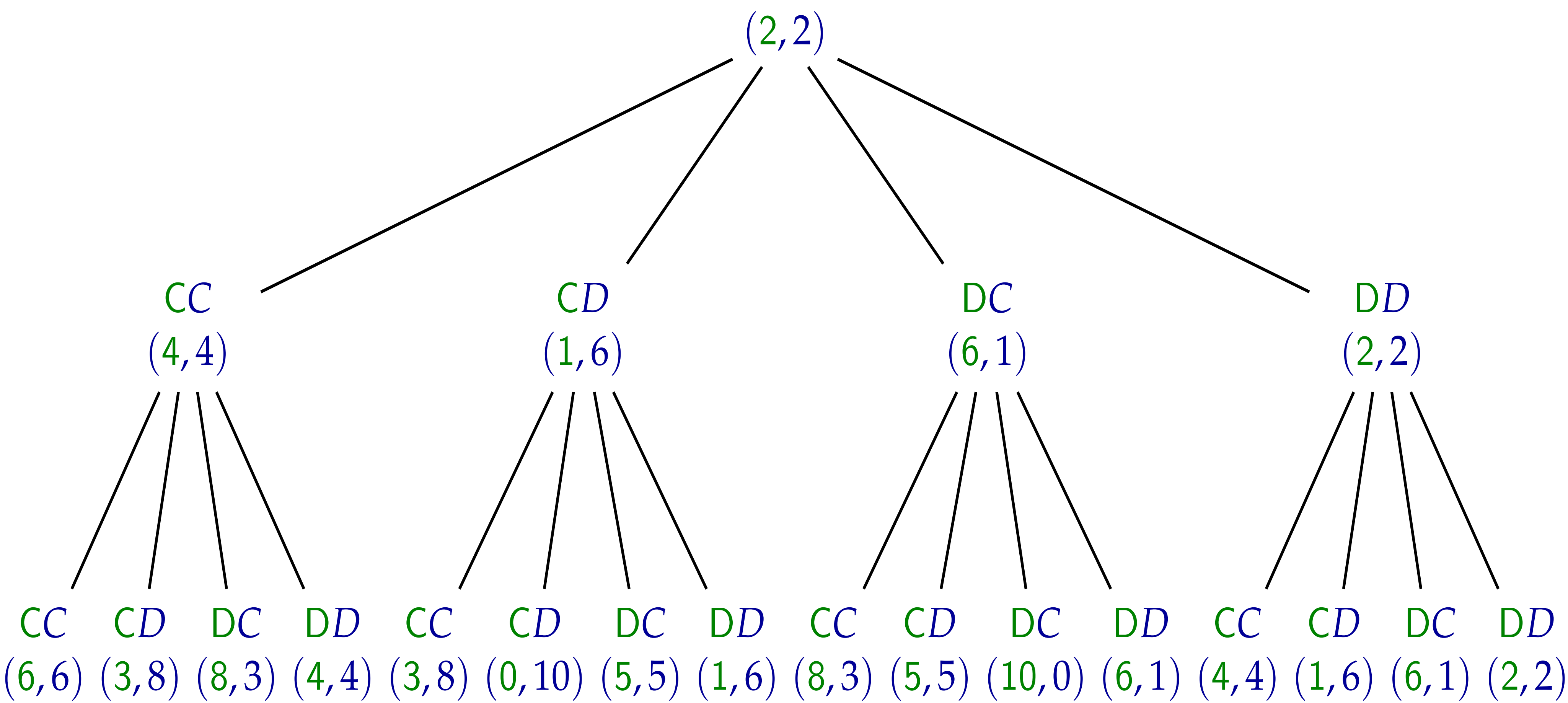
Prisoners' Dilemma		*Other:* Cooperate	Defect
You:	Cooperate	$(3,3)$	$(0,5)$
	Defect	$(5,0)$	$(1,1)$

- Even if mixed strategies are allowed, the PD possesses one Nash equilibrium, viz. (D,D) with payoffs $(1,1)$.
- This equilibrium is Pareto sub-optimal.
- Does the situation change if two parties get to play the Prisoners' Dilemma two times in succession?
- The following diagram (hopefully) shows that playing the PD two times in succession does not yield an essentially new NE.

Nash equilibria in playing the PD twice



Nash equilibria in playing the PD twice



P.S. This is just a payoff tree, not a game in extensive form!

Nash equilibria in playing the PD twice

In normal form:

		Other:			
		CC	CD	DC	DD
You:	CC	(6, 6)	(3, 8)	(3, 8)	(0, 10)
	CD	(8, 3)	(4, 4)	(5, 5)	(1, 6)
	DC	(8, 3)	(5, 5)	(4, 4)	(1, 6)
	DD	(10, 0)	(6, 1)	(6, 1)	(2, 2)

Nash equilibria in playing the PD twice

In normal form:

		Other: CC	CD	DC	DD
You:	CC	$(6, 6)$	$(3, 8)$	$(3, 8)$	$(0, 10)$
	CD	$(8, 3)$	$(4, 4)$	$(5, 5)$	$(1, 6)$
	DC	$(8, 3)$	$(5, 5)$	$(4, 4)$	$(1, 6)$
	DD	$(10, 0)$	$(6, 1)$	$(6, 1)$	$(2, 2)$

- The action profile (DD, DD) is the only Nash equilibrium.

Nash equilibria in playing the PD twice

In normal form:

		Other: CC	CD	DC	DD
You:	CC	$(6,6)$	$(3,8)$	$(3,8)$	$(0,10)$
	CD	$(8,3)$	$(4,4)$	$(5,5)$	$(1,6)$
	DC	$(8,3)$	$(5,5)$	$(4,4)$	$(1,6)$
	DD	$(10,0)$	$(6,1)$	$(6,1)$	$(2,2)$

- The action profile (DD, DD) is the only Nash equilibrium.
- With 3 successive games, we obtain a $2^3 \times 2^3$ matrix, where the action profile (DDD, DDD) still would be the only Nash equilibrium.

Nash equilibria in playing the PD twice

In normal form:

		Other:			
		CC	CD	DC	DD
You:	CC	$(6,6)$	$(3,8)$	$(3,8)$	$(0,10)$
	CD	$(8,3)$	$(4,4)$	$(5,5)$	$(1,6)$
	DC	$(8,3)$	$(5,5)$	$(4,4)$	$(1,6)$
	DD	$(10,0)$	$(6,1)$	$(6,1)$	$(2,2)$

- The action profile (DD, DD) is the only Nash equilibrium.
- With 3 successive games, we obtain a $2^3 \times 2^3$ matrix, where the action profile (DDD, DDD) still would be the only Nash equilibrium.
- Generalise to N repetitions: $(\mathrm{D}D^{N-1}, DD^{N-1})$ still is the only Nash equilibrium in a repeated game where the PD is played N times in succession.

Part II:
Nash equilibria

Part II:
Nash equilibria
in normal form games

Part II:
Nash equilibria
in normal form games
that are repeated

Part II:
Nash equilibria
in normal form games
that are repeated
an indefinite number of times

Terminology: finite, indefinite, infinite

To repeat an experiment . . .

Terminology: finite, indefinite, infinite

To repeat an experiment ...

- ... **ten** times.

Terminology: finite, indefinite, infinite

To repeat an experiment . . .

- . . . **ten** times. That's hopefully clear.

Terminology: finite, indefinite, infinite

To repeat an experiment ...

- ... **ten** times. That's hopefully clear.
- ... a **finite** number of times.

Terminology: finite, indefinite, infinite

To repeat an experiment . . .

- . . . **ten** times. That's hopefully clear.
- . . . a **finite** number of times. May mean: a fixed number of times, where the number of repetitions is determined beforehand.

Terminology: finite, indefinite, infinite

To repeat an experiment . . .

- . . . **ten** times. That's hopefully clear.
- . . . a **finite** number of times. May mean: a fixed number of times, where the number of repetitions is determined beforehand. Or it may mean: an indefinite number of times.

Terminology: finite, indefinite, infinite

To repeat an experiment ...

- ... **ten** times. That's hopefully clear.
- ... a **finite** number of times. May mean: a fixed number of times, where the number of repetitions is determined beforehand. Or it may mean: an indefinite number of times. Depends on the context!

Terminology: finite, indefinite, infinite

To repeat an experiment . . .

- . . . **ten** times. That's hopefully clear.
- . . . a **finite** number of times. May mean: a fixed number of times, where the number of repetitions is determined beforehand. Or it may mean: an indefinite number of times. Depends on the context!
- . . . an **indefinite** number of times.

Terminology: finite, indefinite, infinite

To repeat an experiment . . .

- . . . **ten** times. That's hopefully clear.
- . . . a **finite** number of times. May mean: a fixed number of times, where the number of repetitions is determined beforehand. Or it may mean: an indefinite number of times. Depends on the context!
- . . . an **indefinite** number of times. Means: a finite number of times, but nothing is known beforehand about the number of repetitions.

Terminology: finite, indefinite, infinite

To repeat an experiment ...

- ... **ten** times. That's hopefully clear.
- ... a **finite** number of times. May mean: a fixed number of times, where the number of repetitions is determined beforehand. Or it may mean: an indefinite number of times. Depends on the context!
- ... an **indefinite** number of times. Means: a finite number of times, but nothing is known beforehand about the number of repetitions.
- ... an **infinite** number of times.

Terminology: finite, indefinite, infinite

To repeat an experiment . . .

- . . . **ten** times. That's hopefully clear.
- . . . a **finite** number of times. May mean: a fixed number of times, where the number of repetitions is determined beforehand. Or it may mean: an indefinite number of times. Depends on the context!
- . . . an **indefinite** number of times. Means: a finite number of times, but nothing is known beforehand about the number of repetitions.
- . . . an **infinite** number of times. When throwing a dice this must mean a countably infinite number of times.

Indefinite number of repetitions

Indefinite number of repetitions

- A Pareto-suboptimal outcome can be avoided in case the following three conditions are met.

Indefinite number of repetitions

- A Pareto-suboptimal outcome can be avoided in case the following three conditions are met.

  1. The Prisoners' Dilemma is repeated an indefinite number of times (rounds).

Indefinite number of repetitions

- A Pareto-suboptimal outcome can be avoided in case the following three conditions are met.

  1. The Prisoners' Dilemma is repeated an indefinite number of times (rounds).
  2. A so-called discount factor $\delta \in [0, 1]$ determines the probability of continuing the game after each round.

Indefinite number of repetitions

- A Pareto-suboptimal outcome can be avoided in case the following three conditions are met.

  1. The Prisoners' Dilemma is repeated an indefinite number of times (rounds).
  2. A so-called discount factor $\delta \in [0, 1]$ determines the probability of continuing the game after each round.
  3. The probability to continue, δ, is large enough.

Indefinite number of repetitions

- A Pareto-suboptimal outcome can be avoided in case the following three conditions are met.

  1. The Prisoners' Dilemma is repeated an indefinite number of times (rounds).

  2. A so-called discount factor $\delta \in [0, 1]$ determines the probability of continuing the game after each round.

  3. The probability to continue, δ, is large enough.

- Under these conditions suddenly infinitely many Nash equilibria exist. This is sometimes called an embarrassment of richness (Peters, 2008).

Indefinite number of repetitions

- A Pareto-suboptimal outcome can be avoided in case the following three conditions are met.

  1. The Prisoners' Dilemma is repeated an indefinite number of times (rounds).

  2. A so-called discount factor $\delta \in [0, 1]$ determines the probability of continuing the game after each round.

  3. The probability to continue, δ, is large enough.

- Under these conditions suddenly infinitely many Nash equilibria exist. This is sometimes called an embarrassment of richness (Peters, 2008).

- Various Folk theorems state the existence of multiple equilibria in games that are repeated an indefinite number of times.

Indefinite number of repetitions

- A Pareto-suboptimal outcome can be avoided in case the following three conditions are met.

  1. The Prisoners' Dilemma is repeated an indefinite number of times (rounds).

  2. A so-called discount factor $\delta \in [0, 1]$ determines the probability of continuing the game after each round.

  3. The probability to continue, δ, is large enough.

- Under these conditions suddenly infinitely many Nash equilibria exist. This is sometimes called an embarrassment of richness (Peters, 2008).

- Various Folk theorems state the existence of multiple equilibria in games that are repeated an indefinite number of times.

- Here we discuss one version of "the" Folk Theorem.

Family of Folk Theorems

There actually exist many Folk Theorems.

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case)

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case)

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case), or certain parts of the history may be hidden.

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case), or certain parts of the history may be hidden.
- **Reward**. Players may collect their payoff through a discount factor (present case)

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case), or certain parts of the history may be hidden.
- **Reward**. Players may collect their payoff through a discount factor (present case) or through average rewards.

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case), or certain parts of the history may be hidden.
- **Reward**. Players may collect their payoff through a discount factor (present case) or through average rewards.
- **Subgame perfectness**. Subgame perfect equilibria (present case)

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case), or certain parts of the history may be hidden.
- **Reward**. Players may collect their payoff through a discount factor (present case) or through average rewards.
- **Subgame perfectness**. Subgame perfect equilibria (present case) or plain Nash equilibria.

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case), or certain parts of the history may be hidden.
- **Reward**. Players may collect their payoff through a discount factor (present case) or through average rewards.
- **Subgame perfectness**. Subgame perfect equilibria (present case) or plain Nash equilibria.
- **Equilibrium**. We may be interested in Nash equilibria (present case)

Family of Folk Theorems

There actually exist many Folk Theorems.

- **Horizon**. The game may be repeated indefinitely (present case) or there may be an upper bound to the number of plays.
- **Information**. Players may act on the basis of CKR (present case), or certain parts of the history may be hidden.
- **Reward**. Players may collect their payoff through a discount factor (present case) or through average rewards.
- **Subgame perfectness**. Subgame perfect equilibria (present case) or plain Nash equilibria.
- **Equilibrium**. We may be interested in Nash equilibria (present case), or other types of equilibria, such as so-called ϵ-*Nash equilibria* or so-called *correlated equilibria*.

The concept of a repeated game

The concept of a repeated game

- Let G be a game in normal form.

The concept of a repeated game

- Let G be a game in normal form.
- The repeated game $G^*(\delta)$, is G, played an indefinite number of times

The concept of a repeated game

- Let G be a game in normal form.
- The repeated game $G^*(\delta)$, is G, played an indefinite number of times, where δ represents the probability that the game will be played another time.

The concept of a repeated game

- Let G be a game in normal form.
- The repeated game $G^*(\delta)$, is G, played an indefinite number of times, where δ represents the probability that the game will be played another time.

  Exercise: give $P\{\ G^*(\delta) \text{ lasts at least } t \text{ rounds }\}$.

The concept of a repeated game

- Let G be a game in normal form.
- The repeated game $G^*(\delta)$, is G, played an indefinite number of times, where δ represents the probability that the game will be played another time.

  Exercise: give $P\{ G^*(\delta) \text{ lasts at least } t \text{ rounds } \}$. Answer: δ^t.

The concept of a repeated game

- Let G be a game in normal form.
- The repeated game $G^*(\delta)$, is G, played an indefinite number of times, where δ represents the probability that the game will be played another time.

  Exercise: give $P\{\ G^*(\delta)$ lasts at least t rounds $\}$. Answer: δ^t.
- G is called the stage game.

The concept of a repeated game

- Let G be a game in normal form.
- The repeated game $G^*(\delta)$, is G, played an indefinite number of times, where δ represents the probability that the game will be played another time.

  Exercise: give $P\{\ G^*(\delta)$ lasts at least t rounds $\}$. Answer: δ^t.
- G is called the stage game.
- A history h of length t of a repeated game is a sequence of action profiles of length t.

The concept of a repeated game

- Let G be a game in normal form.
- The repeated game $G^*(\delta)$, is G, played an indefinite number of times, where δ represents the probability that the game will be played another time.

  Exercise: give $P\{\ G^*(\delta) \text{ lasts at least } t \text{ rounds } \}$. Answer: δ^t.
- G is called the stage game.
- A history h of length t of a repeated game is a sequence of action profiles of length t.

  Example: (for the prisoner's dilemma):

$$\begin{array}{rcccccccccc} \text{Row player:} & C & D & D & D & C & C & D & D & D & D \\ \text{Column player:} & C & D & D & D & D & D & D & C & D & D \\ & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array}$$

The concept of a repeated game (II)

The concept of a repeated game (II)

- The set of all possible histories (of any length) is denoted by H.

The concept of a repeated game (II)

- The set of all possible histories (of any length) is denoted by H.
- A strategy for Player i is a function $s_i : H \to \Delta\{C, D\}$ such that

$$\Pr(\text{ Player } i \text{ plays C in round } |h| + 1 \mid h\,) = s_i(h)\,(\text{C}).$$

The concept of a repeated game (II)

- The set of all possible histories (of any length) is denoted by H.
- A strategy for Player i is a function $s_i : H \rightarrow \Delta\{C, D\}$ such that

$$\Pr(\text{ Player } i \text{ plays C in round } |h| + 1 \mid h\ ) = s_i(h)\,(\mathrm{C}).$$

- A strategy profile s is a combination of strategies, one for each player.

The concept of a repeated game (II)

- The set of all possible histories (of any length) is denoted by H.
- A strategy for Player i is a function $s_i : H \to \Delta\{C, D\}$ such that

$$\Pr(\text{ Player } i \text{ plays C in round } |h| + 1 \mid h\ ) = s_i(h)\,(\mathrm{C}).$$

- A strategy profile s is a combination of strategies, one for each player.
- The expected payoff for player i given s can be computed. It is

$$\text{Expected payoff}_i(s) = \sum_{t=0}^{\infty} \delta^t\, \text{Expected payoff}_{i,t}(s).$$

Example on next page.

The concept of a repeated game (II)

Repeated from the previous slide:

The expected payoff for player i given s can be computed. It is

$$\text{Expected payoff}_i(s) = \sum_{t=0}^{\infty} \delta^t \,\text{Expected payoff}_{i,t}(s).$$

The concept of a repeated game (II)

Repeated from the previous slide:

The expected payoff for player i given s can be computed. It is

$$\text{Expected payoff}_i(s) = \sum_{t=0}^{\infty} \delta^t \, \text{Expected payoff}_{i,t}(s).$$

Example: prisoner's dilemma, strategy Player 1 is $s_1 =$ "always cooperate 80%";

The concept of a repeated game (II)

Repeated from the previous slide:

The expected payoff for player i given s can be computed. It is

$$\text{Expected payoff}_i(s) = \sum_{t=0}^{\infty} \delta^t \, \text{Expected payoff}_{i,t}(s).$$

Example: prisoner's dilemma, strategy Player 1 is $s_1 =$ "always cooperate 80%"; strategy Player 2 is $s_2 =$ "always cooperate 70%";

The concept of a repeated game (II)

Repeated from the previous slide:

The expected payoff for player i given s can be computed. It is

$$\text{Expected payoff}_i(s) = \sum_{t=0}^{\infty} \delta^t \, \text{Expected payoff}_{i,t}(s).$$

Example: prisoner's dilemma, strategy Player 1 is $s_1 =$ "always cooperate 80%"; strategy Player 2 is $s_2 =$ "always cooperate 70%"; $\delta = 1/2$.

The concept of a repeated game (II)

Repeated from the previous slide:

The expected payoff for player i given s can be computed. It is

$$\text{Expected payoff}_i(s) = \sum_{t=0}^{\infty} \delta^t \, \text{Expected payoff}_{i,t}(s).$$

Example: prisoner's dilemma, strategy Player 1 is $s_1 =$ "always cooperate 80%"; strategy Player 2 is $s_2 =$ "always cooperate 70%"; $\delta = 1/2$.

$$\begin{aligned}\text{Expected payoff}_1(s) &= \sum_{t=0}^{\infty} \left[\left(\frac{1}{2}\right)^t [0.8(0.7 \cdot 3 + 0.3 \cdot 0) + 0.2(0.7 \cdot 5 + 0.3 \cdot 1)] \right] \\ &= \frac{1}{1-1/2} [\ldots] \approx \frac{1}{1-1/2} 2.44 = 2 \times 2.44 = 4.88.\end{aligned}$$

Subgame perfect equilibria of $G^*(\delta)$: D^*

Definition. A strategy profile s for $G^*(\delta)$ is a subgame-perfect Nash equilibrium if (1) it is a Nash equilibrium of this repeated game, and (2) for every subgame (i.e., tail game) of this repeated game, s restricted to that subgame is a Nash equilibrium as well.

Subgame perfect equilibria of $G^*(\delta)$: D^*

Definition. A strategy profile s for $G^*(\delta)$ is a subgame-perfect Nash equilibrium if (1) it is a Nash equilibrium of this repeated game, and (2) for every subgame (i.e., tail game) of this repeated game, s restricted to that subgame is a Nash equilibrium as well.

Consider the strategy of iterated defection D^*: "always defect, no matter what".[1]

Subgame perfect equilibria of $G^*(\delta)$: D^*

Definition. A strategy profile s for $G^*(\delta)$ is a subgame-perfect Nash equilibrium if (1) it is a Nash equilibrium of this repeated game, and (2) for every subgame (i.e., tail game) of this repeated game, s restricted to that subgame is a Nash equilibrium as well.

Consider the strategy of iterated defection D^*: "always defect, no matter what".[1]

Claim. *The strategy profile* (D^*, D^*) *is a subgame perfect equilibrium in* $G^*(\delta)$.

Subgame perfect equilibria of $G^*(\delta)$: D^*

Definition. A strategy profile s for $G^*(\delta)$ is a subgame-perfect Nash equilibrium if (1) it is a Nash equilibrium of this repeated game, and (2) for every subgame (i.e., tail game) of this repeated game, s restricted to that subgame is a Nash equilibrium as well.

Consider the strategy of iterated defection D^*: "always defect, no matter what".[1]

Claim. *The strategy profile* (D^*, D^*) *is a subgame perfect equilibrium in* $G^*(\delta)$.

Proof. Consider any tailgame starting at round $t \geq 0$.

Subgame perfect equilibria of $G^*(\delta)$: D^*

Definition. A strategy profile s for $G^*(\delta)$ is a subgame-perfect Nash equilibrium if (1) it is a Nash equilibrium of this repeated game, and (2) for every subgame (i.e., tail game) of this repeated game, s restricted to that subgame is a Nash equilibrium as well.

Consider the strategy of iterated defection D^*: "always defect, no matter what".[1]

Claim. *The strategy profile (D^*, D^*) is a subgame perfect equilibrium in $G^*(\delta)$.*

Proof. Consider any tailgame starting at round $t \geq 0$. We are done if we can show that (D^*, D^*) is a NE for this subgame.

Subgame perfect equilibria of $G^*(\delta)$: D^*

Definition. A strategy profile s for $G^*(\delta)$ is a subgame-perfect Nash equilibrium if (1) it is a Nash equilibrium of this repeated game, and (2) for every subgame (i.e., tail game) of this repeated game, s restricted to that subgame is a Nash equilibrium as well.

Consider the strategy of iterated defection D^*: "always defect, no matter what".[1]

Claim. *The strategy profile* (D^*, D^*) *is a subgame perfect equilibrium in* $G^*(\delta)$.

Proof. Consider any tailgame starting at round $t \geq 0$. We are done if we can show that (D^*, D^*) is a NE for this subgame. This is true: given that one player always defects, it never pays off for the other player to play C at any time.

Subgame perfect equilibria of $G^*(\delta)$: D^*

Definition. A strategy profile s for $G^*(\delta)$ is a subgame-perfect Nash equilibrium if (1) it is a Nash equilibrium of this repeated game, and (2) for every subgame (i.e., tail game) of this repeated game, s restricted to that subgame is a Nash equilibrium as well.

Consider the strategy of iterated defection D^*: "always defect, no matter what".[1]

Claim. *The strategy profile (D^*, D^*) is a subgame perfect equilibrium in $G^*(\delta)$.*

Proof. Consider any tailgame starting at round $t \geq 0$. We are done if we can show that (D^*, D^*) is a NE for this subgame. This is true: given that one player always defects, it never pays off for the other player to play C at any time. Therefore, everyone sticks to D^*.

[1] A notation like D^* or (worse) D^∞ is suggestive. Mathematically it makes no sense, but intuitively it does.

Part III:
Trigger strategies

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$*, provided the probability of continuation,* δ*, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N.

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$*, provided the probability of continuation,* δ*, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1}$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$, *provided the probability of continuation,* δ, *is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$*, provided the probability of continuation,* δ*, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$, *provided the probability of continuation,* δ, *is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty}$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$*, provided the probability of continuation,* δ*, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

$$\left[\sum_{t=0}^{\infty} \delta^t \cdot 3 \right]$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$, *provided the probability of continuation,* δ, *is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

$$\left[\sum_{t=0}^{\infty} \delta^t \cdot 3 \right] - \left[\sum_{t=0}^{N-1} \right.$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

$$\left[\sum_{t=0}^{\infty} \delta^t \cdot 3\right] - \left[\sum_{t=0}^{N-1} \delta^t \cdot 3\right.$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile* (T, T) *is a subgame perfect equilibrium in* $G^*(\delta)$, *provided the probability of continuation,* δ, *is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

$$\left[\sum_{t=0}^{\infty} \delta^t \cdot 3\right] - \left[\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5\right.$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

$$\left[\sum_{t=0}^{\infty} \delta^t \cdot 3\right] - \left[\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1\right]$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

$$\left[\sum_{t=0}^{\infty} \delta^t \cdot 3\right] - \left[\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1\right] = -2\delta^N$$

Cost of deviating in Round N

Consider the so-called trigger strategy T: "always play C unless D has been played at least once. In that case play D forever".

Claim. *The strategy profile (T, T) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

Proof. Suppose one player starts to defect at Round N. By doing so he expects a payoff of

$$\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1$$

By playing C throughout, he *could* have earned $\sum_{t=0}^{\infty} \delta^t \cdot 3$ which means he forfeited

$$\left[\sum_{t=0}^{\infty} \delta^t \cdot 3\right] - \left[\sum_{t=0}^{N-1} \delta^t \cdot 3 + \delta^N \cdot 5 + \sum_{t=N+1}^{\infty} \delta^t \cdot 1\right] = -2\delta^N + 2 \sum_{t=N+1}^{\infty} \delta^t.$$

Cost of deviating in Round N

So if

$$-2\delta^N + 2\sum_{t=N+1}^{\infty} \delta^t > 0$$

there is forfeit from payoff.

Cost of deviating in Round N

So if

$$-2\delta^N + 2\sum_{t=N+1}^{\infty}\delta^t > 0$$

there is forfeit from payoff. This is when

$$\sum_{t=N+1}^{\infty}\delta^t > \delta^N$$

Cost of deviating in Round N

So if

$$-2\delta^N + 2 \sum_{t=N+1}^{\infty} \delta^t > 0$$

there is forfeit from payoff. This is when

$$\sum_{t=N+1}^{\infty} \delta^t > \delta^N$$

$$\Leftrightarrow \quad \delta^{N+1} \sum_{t=0}^{\infty} \delta^t > \delta^N$$

Cost of deviating in Round N

So if

$$-2\delta^N + 2\sum_{t=N+1}^{\infty} \delta^t > 0$$

there is forfeit from payoff. This is when

$$\sum_{t=N+1}^{\infty} \delta^t > \delta^N$$

$$\Leftrightarrow \quad \delta^{N+1}\sum_{t=0}^{\infty} \delta^t > \delta^N \quad \Leftrightarrow \quad \delta = 1 \text{ xor } \delta^{N+1}\frac{1}{1-\delta} > \delta^N$$

Cost of deviating in Round N

So if

$$-2\delta^N + 2\sum_{t=N+1}^{\infty} \delta^t > 0$$

there is forfeit from payoff. This is when

$$\sum_{t=N+1}^{\infty} \delta^t > \delta^N$$

$$\Leftrightarrow \quad \delta^{N+1}\sum_{t=0}^{\infty} \delta^t > \delta^N \quad \Leftrightarrow \quad \delta = 1 \text{ xor } \delta^{N+1}\frac{1}{1-\delta} > \delta^N$$

$$\Leftrightarrow \quad \delta = 1 \text{ xor } \delta\frac{1}{1-\delta} > 1,\ \delta \neq 0$$

Cost of deviating in Round N

So if

$$-2\delta^N + 2\sum_{t=N+1}^{\infty} \delta^t > 0$$

there is forfeit from payoff. This is when

$$\sum_{t=N+1}^{\infty} \delta^t > \delta^N$$

$$\Leftrightarrow \quad \delta^{N+1}\sum_{t=0}^{\infty}\delta^t > \delta^N \quad \Leftrightarrow \quad \delta = 1 \ \text{xor} \ \delta^{N+1}\frac{1}{1-\delta} > \delta^N$$

$$\Leftrightarrow \quad \delta = 1 \ \text{xor} \ \delta\frac{1}{1-\delta} > 1,\ \delta \neq 0 \quad \Leftrightarrow \quad \delta > 1-\delta$$

Cost of deviating in Round N

So if

$$-2\delta^N + 2\sum_{t=N+1}^{\infty} \delta^t > 0$$

there is forfeit from payoff. This is when

$$\sum_{t=N+1}^{\infty} \delta^t > \delta^N$$

$$\Leftrightarrow \quad \delta^{N+1}\sum_{t=0}^{\infty} \delta^t > \delta^N \quad \Leftrightarrow \quad \delta = 1 \text{ xor } \delta^{N+1}\frac{1}{1-\delta} > \delta^N$$

$$\Leftrightarrow \quad \delta = 1 \text{ xor } \delta\frac{1}{1-\delta} > 1,\ \delta \neq 0 \quad \Leftrightarrow \quad \delta > 1-\delta \quad \Leftrightarrow \quad \delta > \frac{1}{2}.$$

Cost of deviating in Round N

So if

$$-2\delta^N + 2\sum_{t=N+1}^{\infty} \delta^t > 0$$

there is forfeit from payoff. This is when

$$\sum_{t=N+1}^{\infty} \delta^t > \delta^N$$

$$\Leftrightarrow \quad \delta^{N+1}\sum_{t=0}^{\infty}\delta^t > \delta^N \quad \Leftrightarrow \quad \delta = 1 \text{ xor } \delta^{N+1}\frac{1}{1-\delta} > \delta^N$$

$$\Leftrightarrow \quad \delta = 1 \text{ xor } \delta\frac{1}{1-\delta} > 1,\ \delta \neq 0 \quad \Leftrightarrow \quad \delta > 1-\delta \quad \Leftrightarrow \quad \delta > \frac{1}{2}.$$

Therefore, if $\delta > 1/2$ every player forfeits payoff by deviating from T. $\square$

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Informal definition of strategies. A and B tacitly agree to alternate actions, i.e.,

$$(C, D), (D, C), (C, D), (D, C) \ldots.$$

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Informal definition of strategies. A and B tacitly agree to alternate actions, i.e.,

$$(C, D), (D, C), (C, D), (D, C) \ldots.$$

If one of them deviates, the other party plays D forever.

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Informal definition of strategies. A and B tacitly agree to alternate actions, i.e.,

$$(C, D), (D, C), (C, D), (D, C) \ldots.$$

If one of them deviates, the other party plays D forever. (The party who originally deviated plays D forever thereafter as well.)

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Informal definition of strategies. A and B tacitly agree to alternate actions, i.e.,

$$(C, D), (D, C), (C, D), (D, C) \ldots.$$

If one of them deviates, the other party plays D forever. (The party who originally deviated plays D forever thereafter as well.) CKR is at work here!

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Informal definition of strategies. A and B tacitly agree to alternate actions, i.e.,

$$(C, D), (D, C), (C, D), (D, C) \dots.$$

If one of them deviates, the other party plays D forever. (The party who originally deviated plays D forever thereafter as well.) CKR is at work here!

Let A be the strategy that plays C in Round 1. Let B be the other strategy.

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Informal definition of strategies. A and B tacitly agree to alternate actions, i.e.,

$$(C, D), (D, C), (C, D), (D, C) \ldots.$$

If one of them deviates, the other party plays D forever. (The party who originally deviated plays D forever thereafter as well.) CKR is at work here!

Let A be the strategy that plays C in Round 1. Let B be the other strategy.

Claim. *The strategy profile* (A, B) *is a subgame perfect equilibrium in* $G^*(\delta)$*, provided the probability of continuation,* δ*, is sufficiently large.*

Example: an alternating trigger strategy

Yet another subgame perfect equilibrium:

Informal definition of strategies. A and B tacitly agree to alternate actions, i.e.,

$$(C, D), (D, C), (C, D), (D, C) \ldots.$$

If one of them deviates, the other party plays D forever. (The party who originally deviated plays D forever thereafter as well.) CKR is at work here!

Let A be the strategy that plays C in Round 1. Let B be the other strategy.

Claim. *The strategy profile (A, B) is a subgame perfect equilibrium in $G^*(\delta)$, provided the probability of continuation, δ, is sufficiently large.*

An analysis of this situation and a proof of this claim can be found in (Peters, 2008), pp. 104-105.*

*H. Peters (2008): *Game Theory: A Multi-Leveled Approach*. Springer, ISBN: 978-3-540-69290-4.

Generalisation of trigger strategies

Generalisation of trigger strategies

- Every convex combination[2] of payoffs

$$\alpha_1(3,3) + \alpha_2(0,5) + \alpha_3(5,0) + \alpha_4(1,1)$$

can be established by smartly picking appropriate strategy patterns.

Generalisation of trigger strategies

- Every convex combination[2] of payoffs

$$\alpha_1(3,3) + \alpha_2(0,5) + \alpha_3(5,0) + \alpha_4(1,1)$$

can be established by smartly picking appropriate strategy patterns. E.g.: “We play 4 times (C, C). Then we play 7 times $(C, D), (D, C)$, then 4 times (C, C) and so on”.

Generalisation of trigger strategies

- Every convex combination[2] of payoffs

$$\alpha_1(3,3) + \alpha_2(0,5) + \alpha_3(5,0) + \alpha_4(1,1)$$

  can be established by smartly picking appropriate strategy patterns. E.g.: "We play 4 times (C,C). Then we play 7 times $(C,D),(D,C)$, then 4 times (C,C) and so on".

- Ensure that (C,C) occurs (in the long run) in α_1, (C,D) in α_2, (D,C) in α_3, and (D,D) in α_4 percent of the stages.

Generalisation of trigger strategies

- Every convex combination[2] of payoffs

$$\alpha_1(3,3) + \alpha_2(0,5) + \alpha_3(5,0) + \alpha_4(1,1)$$

  can be established by smartly picking appropriate strategy patterns. E.g.: "We play 4 times (C, C). Then we play 7 times $(C, D), (D, C)$, then 4 times (C, C) and so on".

- Ensure that (C, C) occurs (in the long run) in α_1, (C, D) in α_2, (D, C) in α_3, and (D, D) in α_4 percent of the stages.

- As long as these limiting average payoffs exceed $\text{payoff}(\{D, D\})$ for each player (which is 1), associated trigger strategies can be formulated that lead to these payoffs and trigger eternal play of (D, D) after a deviation.

Generalisation of trigger strategies

- Every convex combination[2] of payoffs

$$\alpha_1(3,3) + \alpha_2(0,5) + \alpha_3(5,0) + \alpha_4(1,1)$$

  can be established by smartly picking appropriate strategy patterns. E.g.: "We play 4 times (C,C). Then we play 7 times $(C,D),(D,C)$, then 4 times (C,C) and so on".

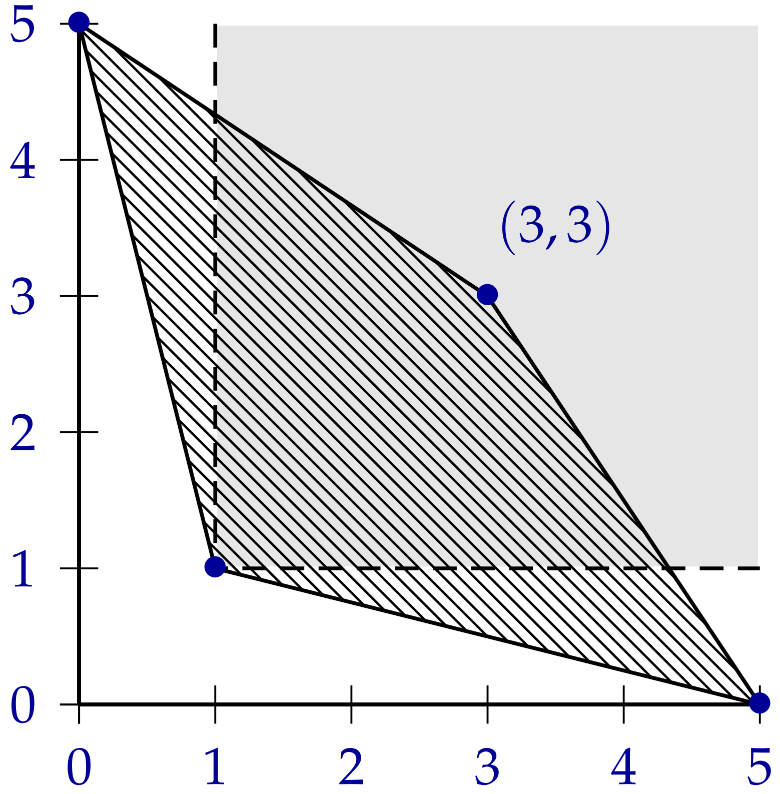
- Ensure that (C,C) occurs (in the long run) in α_1, (C,D) in α_2, (D,C) in α_3, and (D,D) in α_4 percent of the stages.

- As long as these limiting average payoffs exceed $\text{payoff}(\{D,D\})$ for each player (which is 1), associated trigger strategies can be formulated that lead to these payoffs and trigger eternal play of (D,D) after a deviation.

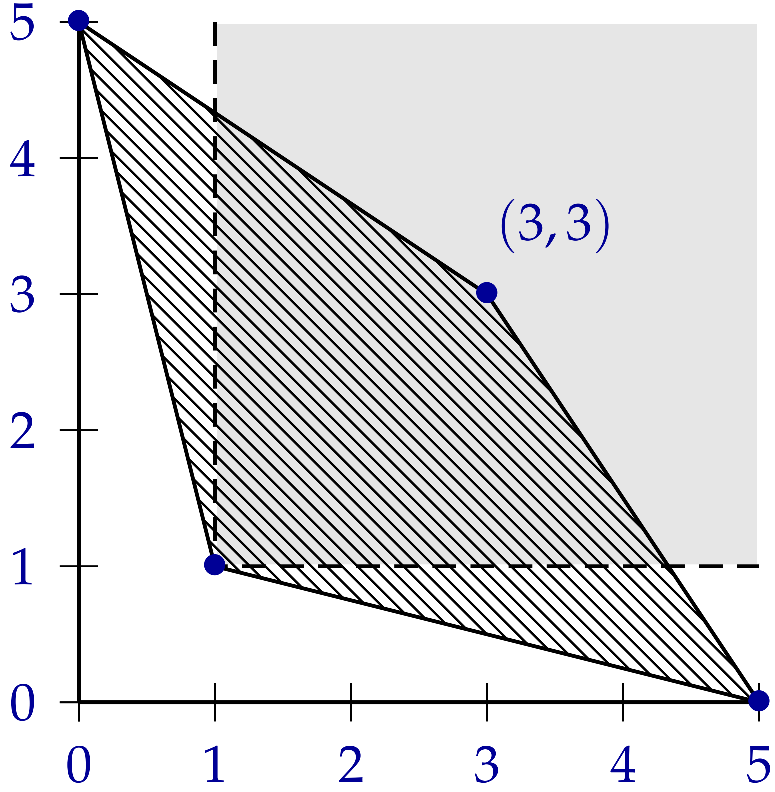
- For δ high enough, these strategies again form a SGP NE.

[2]Meaning $\alpha_i \geq 0$ and $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$.

Folk theorem for SGP NE in a repeated PD

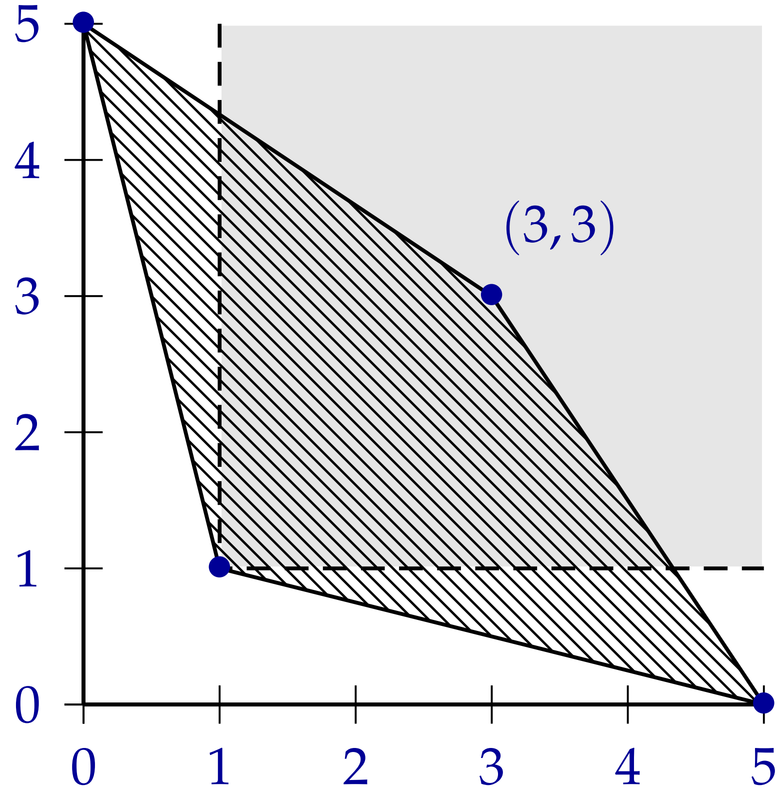


Folk theorem for SGP NE in a repeated PD



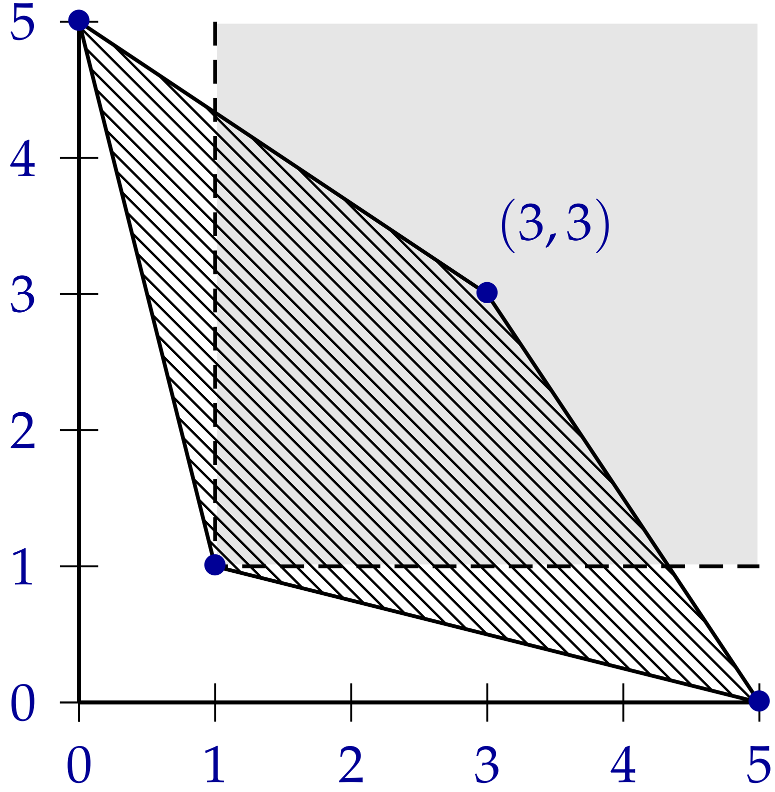
1. **Feasible payoffs** (striped): payoff combos that can be obtained by jointly repeating patterns of actions (more accurately: patterns of action profiles).

Folk theorem for SGP NE in a repeated PD



1. **Feasible payoffs** (striped): payoff combos that can be obtained by jointly repeating patterns of actions (more accurately: patterns of action profiles).
2. **Enforceable payoffs** (shaded): everyone resides above punishment minmax.

Folk theorem for SGP NE in a repeated PD

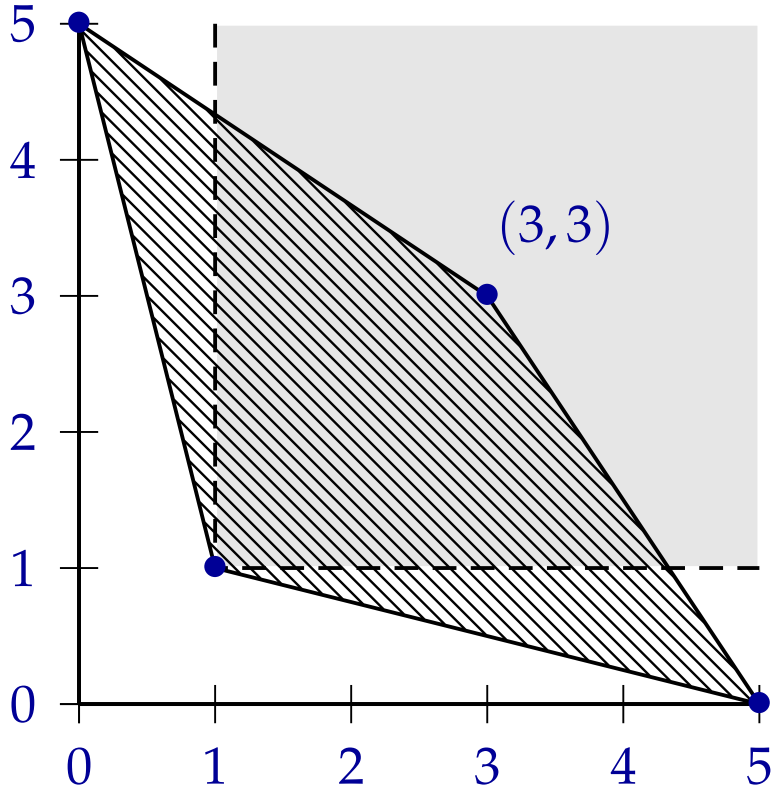


1. **Feasible payoffs** (striped): payoff combos that can be obtained by jointly repeating patterns of actions (more accurately: patterns of action profiles).

2. **Enforceable payoffs** (shaded): everyone resides above punishment minmax.

For every payoff pair (x, y) in (1) $\cap$ (2),

Folk theorem for SGP NE in a repeated PD

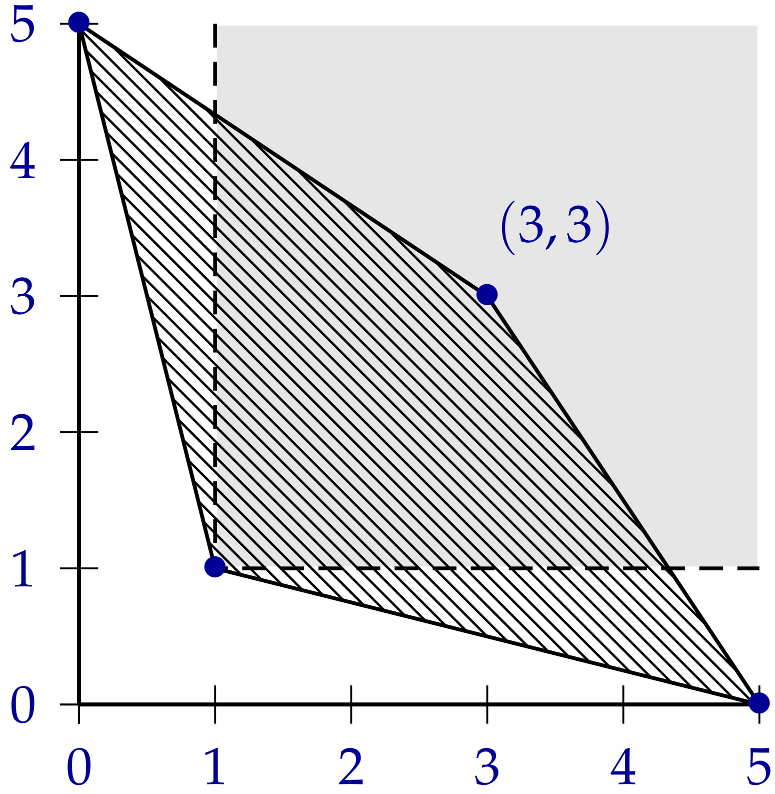


1. **Feasible payoffs** (striped): payoff combos that can be obtained by jointly repeating patterns of actions (more accurately: patterns of action profiles).

2. **Enforceable payoffs** (shaded): everyone resides above punishment minmax.

For every payoff pair (x, y) in (1) $\cap$ (2), there is a $\delta(x, y) \in (0, 1)$,

Folk theorem for SGP NE in a repeated PD

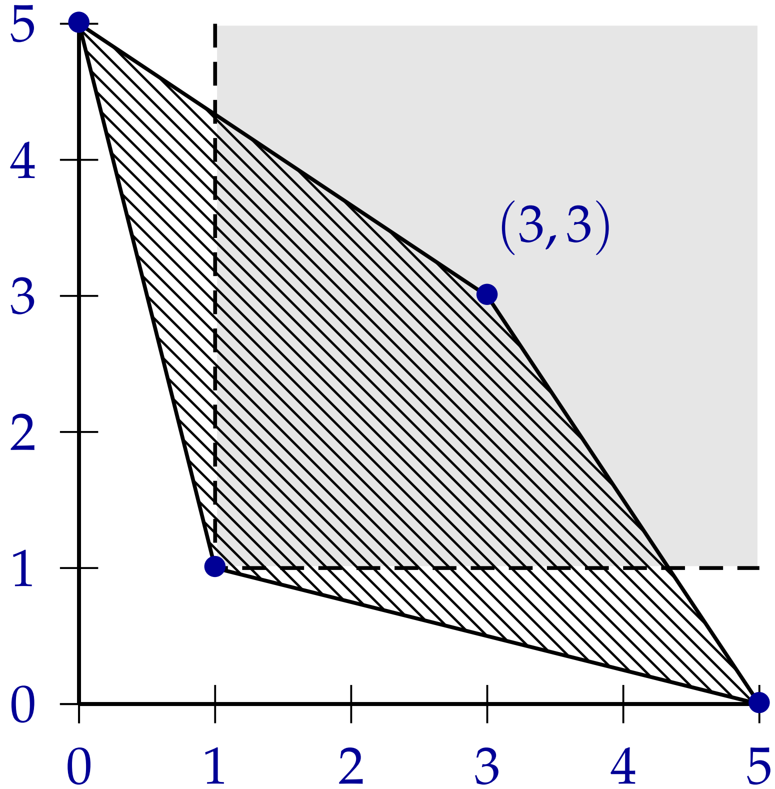


1. **Feasible payoffs** (striped): payoff combos that can be obtained by jointly repeating patterns of actions (more accurately: patterns of action profiles).

2. **Enforceable payoffs** (shaded): everyone resides above punishment minmax.

For every payoff pair (x, y) in (1) $\cap$ (2), there is a $\delta(x, y) \in (0, 1)$, such that for all $\delta \geq \delta(x, y)$

Folk theorem for SGP NE in a repeated PD



1. **Feasible payoffs** (striped): payoff combos that can be obtained by jointly repeating patterns of actions (more accurately: patterns of action profiles).

2. **Enforceable payoffs** (shaded): everyone resides above punishment minmax.

For every payoff pair (x, y) in (1) $\cap$ (2), there is a $\delta(x, y) \in (0, 1)$, such that for all $\delta \geq \delta(x, y)$ the payoff (x, y) can be obtained as the limiting average in a subgame perfect equilibrium of $G^*(\delta)$.

Part IV:
non-SGP Nash equilibria

Existence of non-SGP Nash equilibria

Existence of non-SGP Nash equilibria

- We have seen that many subgame perfect equilibria exist for repeated games.

Existence of non-SGP Nash equilibria

- We have seen that many subgame perfect equilibria exist for repeated games.
- What about the existence of non-SGP Nash equilibria in repeated games, i.e., equilibria that are not necessarily subgame perfect?

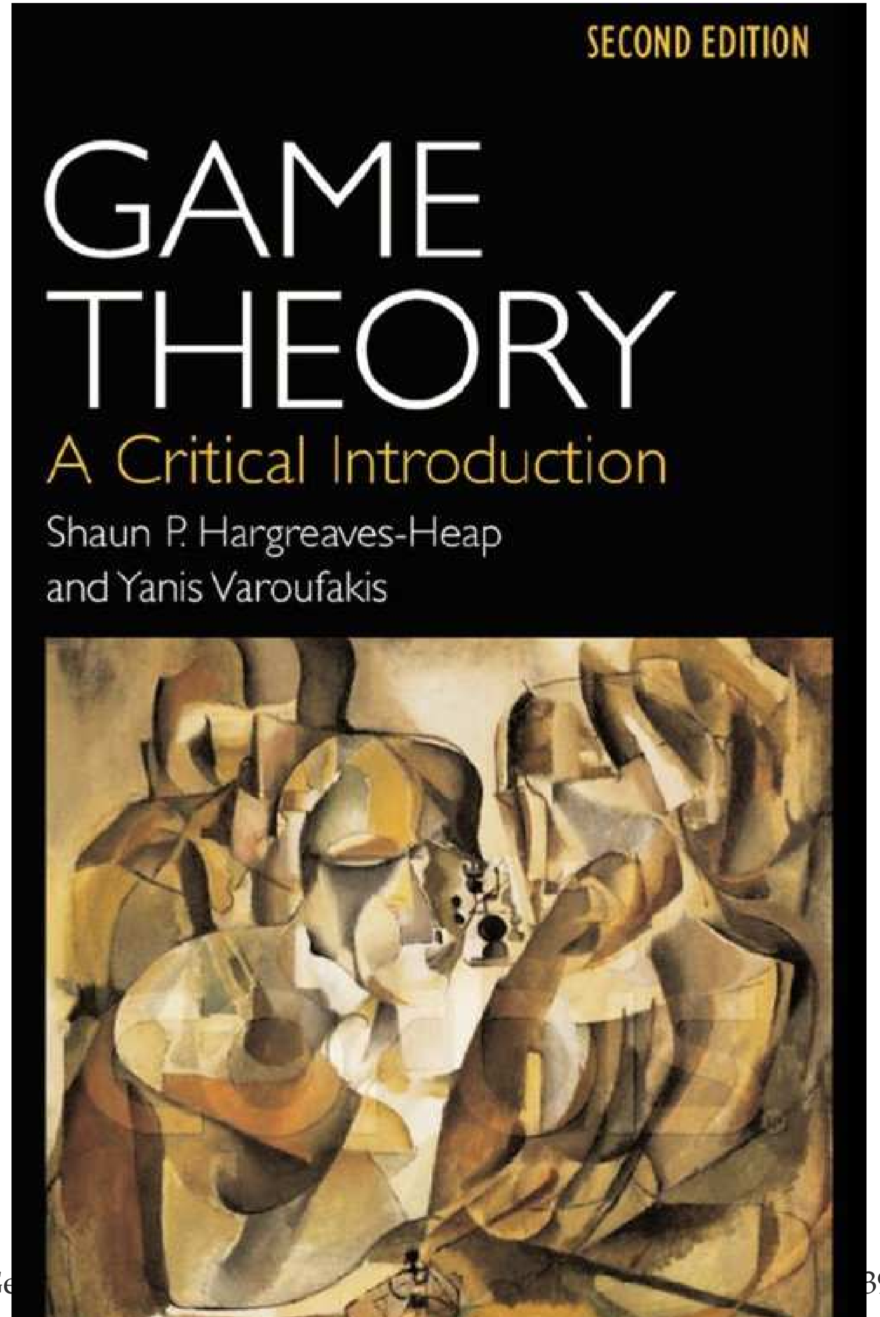
Existence of non-SGP Nash equilibria

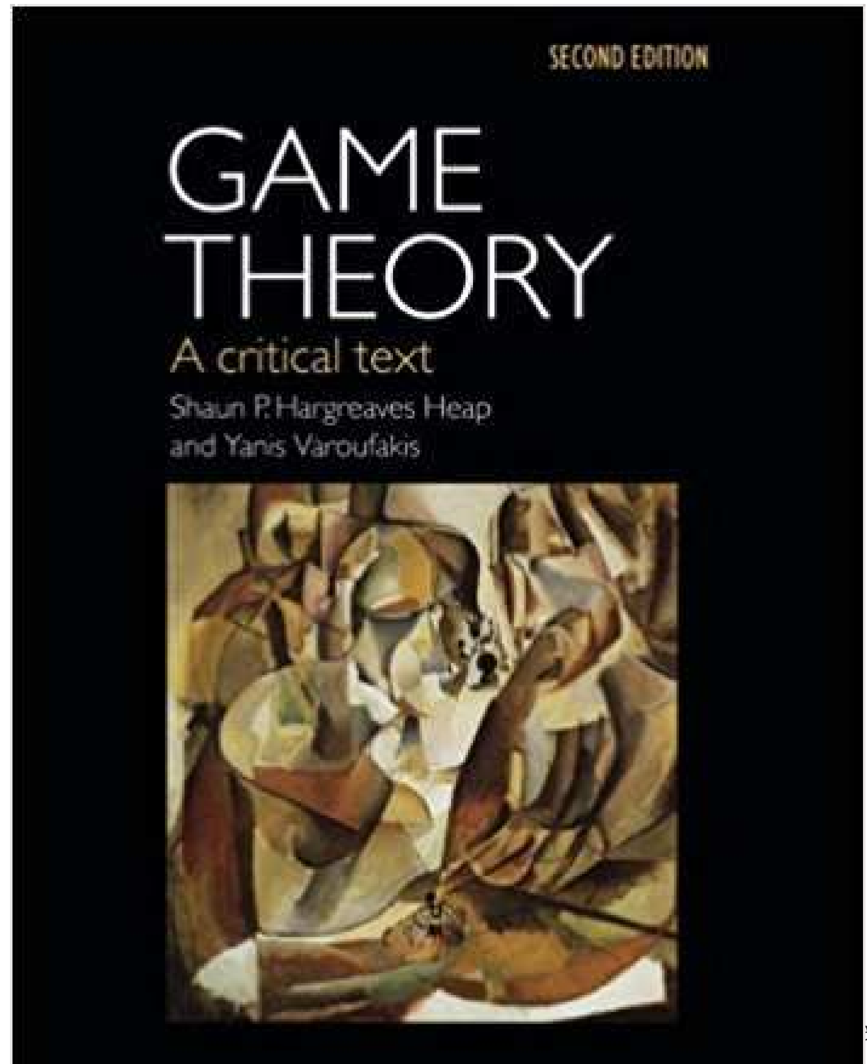
- We have seen that many subgame perfect equilibria exist for repeated games.
- What about the existence of non-SGP Nash equilibria in repeated games, i.e., equilibria that are not necessarily subgame perfect?
- Without the requirement of subgame perfection, deviations can be punished more severely: the equilibrium does not have to induce SGPs.

Existence of non-SGP Nash equilibria

- We have seen that many subgame perfect equilibria exist for repeated games.
- What about the existence of non-SGP Nash equilibria in repeated games, i.e., equilibria that are not necessarily subgame perfect?
- Without the requirement of subgame perfection, deviations can be punished more severely: the equilibrium does not have to induce SGPs.
- However, non-SGPs implies threats that are not credible.

Game Theory: A Critical [what?]





Example: A repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

Example: A repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

1. For row, U is a dominating strategy.

Example: A repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

1. For row, U is a dominating strategy.
2. The pure profile (U, L) is the only mixed strategy profile that is a NE.

Example: A repeated game with a non-SGP NE

Some game		Col: Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

1. For row, U is a dominating strategy.
2. The pure profile (U, L) is the only mixed strategy profile that is a NE.
3. Define trigger-strategies $(T1, T2)$ such that the pattern $[(D, R), (U, L)^3]^*$ is played indefinitely.

Example: A repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

1. For row, U is a dominating strategy.
2. The pure profile (U, L) is the only mixed strategy profile that is a NE.
3. Define trigger-strategies $(T1, T2)$ such that the pattern $[(D, R), (U, L)^3]^*$ is played indefinitely.

   If this pattern is violated, both parties fall back to punishment strategies:

Example: A repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

1. For row, U is a dominating strategy.
2. The pure profile (U, L) is the only mixed strategy profile that is a NE.
3. Define trigger-strategies $(T1, T2)$ such that the pattern $[(D,R),(U,L)^3]^*$ is played indefinitely.

   If this pattern is violated, both parties fall back to punishment strategies:

   - The punishment strategy of row is mixed $(0.8, 0.2)^*$.

Example: A repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

1. For row, U is a dominating strategy.
2. The pure profile (U, L) is the only mixed strategy profile that is a NE.
3. Define trigger-strategies $(T1, T2)$ such that the pattern $[(D, R), (U, L)^3]^*$ is played indefinitely.

   If this pattern is violated, both parties fall back to punishment strategies:

   - The punishment strategy of row is mixed $(0.8, 0.2)^*$.
   - The punishment strategy of col is pure R^*.

Example: A repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

1. For row, U is a dominating strategy.
2. The pure profile (U, L) is the only mixed strategy profile that is a NE.
3. Define trigger-strategies $(T1, T2)$ such that the pattern $[(D,R),(U,L)^3]^*$ is played indefinitely.

   If this pattern is violated, both parties fall back to punishment strategies:

   - The punishment strategy of row is mixed $(0.8, 0.2)^*$.
   - The punishment strategy of col is pure R^*.

   This combination of strategies is not a NE.

Example: A repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

1. For row, U is a dominating strategy.
2. The pure profile (U, L) is the only mixed strategy profile that is a NE.
3. Define trigger-strategies $(T1, T2)$ such that the pattern $[(D,R),(U,L)^3]^*$ is played indefinitely.

   If this pattern is violated, both parties fall back to punishment strategies:

   - The punishment strategy of row is mixed $(0.8, 0.2)^*$.
   - The punishment strategy of col is pure R^*.

   This combination of strategies is not a NE. (For R^* induces U^*.)

Example: a repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

Example: a repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

Claim. *The combination of trigger strategies* $(T1, T2)$ *is a Nash-equilibrium for large enough* $\delta \in [0, 1]$.

Example: a repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

Claim. *The combination of trigger strategies* $(T1, T2)$ *is a Nash-equilibrium for large enough* $\delta \in [0, 1]$.

- $T1 \Rightarrow T2$. If row plays (the non-degenerated part of) $T1$, then col must play $T2$, for $T2$ is a best response to $T1$.

Example: a repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

Claim. *The combination of trigger strategies $(T1, T2)$ is a Nash-equilibrium for large enough $\delta \in [0, 1]$.*

- $T1 \Rightarrow T2$. If row plays (the non-degenerated part of) $T1$, then col must play $T2$, for $T2$ is a best response to $T1$.
- $T2 \Rightarrow T1$. If at all, the best moment for row to deviate is at D, for that would give row an incidental advantage of 1. After that row's opponent falls back to R^*.

Example: a repeated game with a non-SGP NE

	Some game	Col: Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

Claim. *The combination of trigger strategies $(T1, T2)$ is a Nash-equilibrium for large enough $\delta \in [0,1]$.*

- $T1 \Rightarrow T2$. If row plays (the non-degenerated part of) $T1$, then col must play $T2$, for $T2$ is a best response to $T1$.
- $T2 \Rightarrow T1$. If at all, the best moment for row to deviate is at D, for that would give row an incidental advantage of 1. After that row's opponent falls back to R^*.

  Total payoff for row: 0 (for cheating)

Example: a repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

Claim. *The combination of trigger strategies* $(T1, T2)$ *is a Nash-equilibrium for large enough* $\delta \in [0,1]$.

- $T1 \Rightarrow T2$. If row plays (the non-degenerated part of) $T1$, then col must play $T2$, for $T2$ is a best response to $T1$.
- $T2 \Rightarrow T1$. If at all, the best moment for row to deviate is at D, for that would give row an incidental advantage of 1. After that row's opponent falls back to R^*.

  Total payoff for row: 0 (for cheating) $+ 0 + \cdots + 0$ (for being punished by col).

Example: a repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

Example: a repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player).

Example: a repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1 \cdot \delta + 1 \cdot \delta^2 + 1 \cdot \delta^3)$$

Example: a repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1 \cdot \delta + 1 \cdot \delta^2 + 1 \cdot \delta^3) + (-1 \cdot \delta^4 + 1 \cdot \delta^5 + 1 \cdot \delta^6 + 1 \cdot \delta^7)$$

Example: a repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1\cdot \delta + 1\cdot \delta^2 + 1\cdot \delta^3) + (-1\cdot \delta^4 + 1\cdot \delta^5 + 1\cdot \delta^6 + 1\cdot \delta^7) + \dots$$

Example: a repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1 \cdot \delta + 1 \cdot \delta^2 + 1 \cdot \delta^3) + (-1 \cdot \delta^4 + 1 \cdot \delta^5 + 1 \cdot \delta^6 + 1 \cdot \delta^7) + \dots$$

$$= \sum_{k=0}^{\infty} \delta^k$$

Example: a repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1 \cdot \delta + 1 \cdot \delta^2 + 1 \cdot \delta^3) + (-1 \cdot \delta^4 + 1 \cdot \delta^5 + 1 \cdot \delta^6 + 1 \cdot \delta^7) + \ldots$$

$$= \sum_{k=0}^{\infty} \delta^k - 2 \sum_{k=0}^{\infty} \delta^{4k}$$

Example: a repeated game with a non-SGP NE

Some game		Col: Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1\cdot\delta + 1\cdot\delta^2 + 1\cdot\delta^3) + (-1\cdot\delta^4 + 1\cdot\delta^5 + 1\cdot\delta^6 + 1\cdot\delta^7) + \ldots$$

$$= \sum_{k=0}^{\infty} \delta^k - 2\sum_{k=0}^{\infty} \delta^{4k} = \frac{1}{1-\delta}$$

Example: a repeated game with a non-SGP NE

		Col:	
	Some game	Left (L)	Right (R)
Row:	Up (U)	$(1, 1)$	$(0, 0)$
	Down (D)	$(0, 0)$	$(-1, 4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1 \cdot \delta + 1 \cdot \delta^2 + 1 \cdot \delta^3) + (-1 \cdot \delta^4 + 1 \cdot \delta^5 + 1 \cdot \delta^6 + 1 \cdot \delta^7) + \dots$$

$$= \sum_{k=0}^{\infty} \delta^k - 2 \sum_{k=0}^{\infty} \delta^{4k} = \frac{1}{1-\delta} - 2\frac{1}{1-\delta^4}.$$

Example: a repeated game with a non-SGP NE

	Some game	*Col:* Left (L)	Right (R)
Row:	Up (U)	$(1,1)$	$(0,0)$
	Down (D)	$(0,0)$	$(-1,4)$

- $T2 \Rightarrow T1$ **(continued)**. Total payoff for row player: 0 (for cheating) + $0 + \cdots + 0$ (for being punished by the column player). Payoff for row player if he was loyal:

$$(-1 + 1\cdot \delta + 1\cdot \delta^2 + 1\cdot \delta^3) + (-1\cdot \delta^4 + 1\cdot \delta^5 + 1\cdot \delta^6 + 1\cdot \delta^7) + \dots$$

$$= \sum_{k=0}^{\infty} \delta^k - 2\sum_{k=0}^{\infty} \delta^{4k} = \frac{1}{1-\delta} - 2\frac{1}{1-\delta^4}.$$

This expression is positive only if $\delta \geq 0.54$. (Solve 3rd-degree equation.)

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1, 1)$	$(0, 0)$
D	$(0, 0)$	$(-1, 4)$

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1,1)$	$(0,0)$
D	$(0,0)$	$(-1,4)$

- Col can punish row maximally by playing R^*. How row can punish col is less obvious.

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1, 1)$	$(0, 0)$
D	$(0, 0)$	$(-1, 4)$

- Col can punish row maximally by playing R^*. How row can punish col is less obvious.
- If row plays D^* then col will play R^*. If row plays U^*, then col will play L^*.

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1, 1)$	$(0, 0)$
D	$(0, 0)$	$(-1, 4)$

- Col can punish row maximally by playing R^*. How row can punish col is less obvious.
- If row plays D^* then col will play R^*. If row plays U^*, then col will play L^*.
- Row can punish col even more by playing a minmax strategy.

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1,1)$	$(0,0)$
D	$(0,0)$	$(-1,4)$

- Col can punish row maximally by playing R^*. How row can punish col is less obvious.
- If row plays D^* then col will play R^*. If row plays U^*, then col will play L^*.
- Row can punish col even more by playing a minmax strategy.

  Given row's mixed strategy (u,d), col maximises his expected payoff by choosing the right mix (l,r):

$$\max_{l,r} ul \cdot 1 + dr \cdot 4$$

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1,1)$	$(0,0)$
D	$(0,0)$	$(-1,4)$

- Col can punish row maximally by playing R^*. How row can punish col is less obvious.
- If row plays D^* then col will play R^*. If row plays U^*, then col will play L^*.
- Row can punish col even more by playing a minmax strategy.

  Given row's mixed strategy (u,d), col maximises his expected payoff by choosing the right mix (l,r):

$$\max_{l,r} ul \cdot 1 + dr \cdot 4$$

$$= \max_{l} ul + 4(1-u)(1-l)$$

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1,1)$	$(0,0)$
D	$(0,0)$	$(-1,4)$

- Col can punish row maximally by playing R^*. How row can punish col is less obvious.
- If row plays D^* then col will play R^*. If row plays U^*, then col will play L^*.
- Row can punish col even more by playing a minmax strategy.

  Given row's mixed strategy (u, d), col maximises his expected payoff by choosing the right mix (l, r):

$$\max_{l,r} ul \cdot 1 + dr \cdot 4$$

$$= \max_{l} ul + 4(1-u)(1-l)$$

$$= \max_{l} (5u-4)l + 4 - 4u.$$

Example: A repeated game with a non-SGP NE

Col:	L	R
Row: U	$(1,1)$	$(0,0)$
D	$(0,0)$	$(-1,4)$

- Col can punish row maximally by playing R^*. How row can punish col is less obvious.
- If row plays D^* then col will play R^*. If row plays U^*, then col will play L^*.
- Row can punish col even more by playing a minmax strategy.

  Given row's mixed strategy (u, d), col maximises his expected payoff by choosing the right mix (l, r):

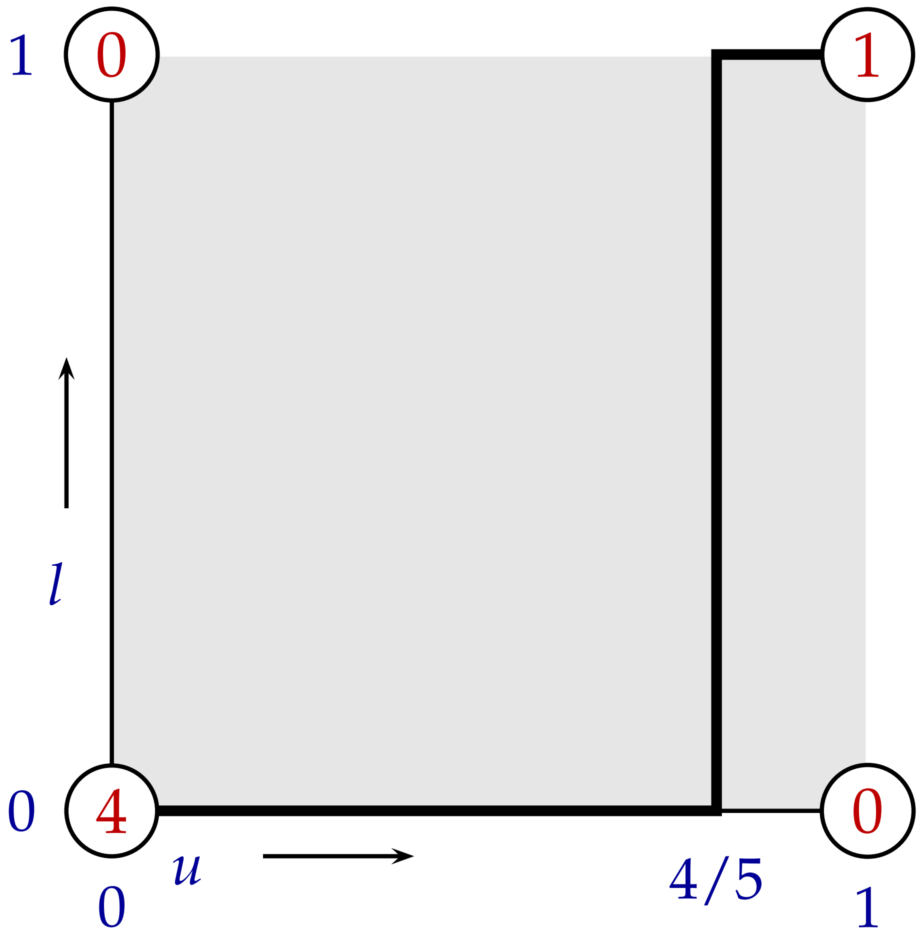
$$\max_{l,r} ul \cdot 1 + dr \cdot 4$$

$$= \max_{l} ul + 4(1-u)(1-l)$$

$$= \max_{l} (5u-4)l + 4 - 4u.$$

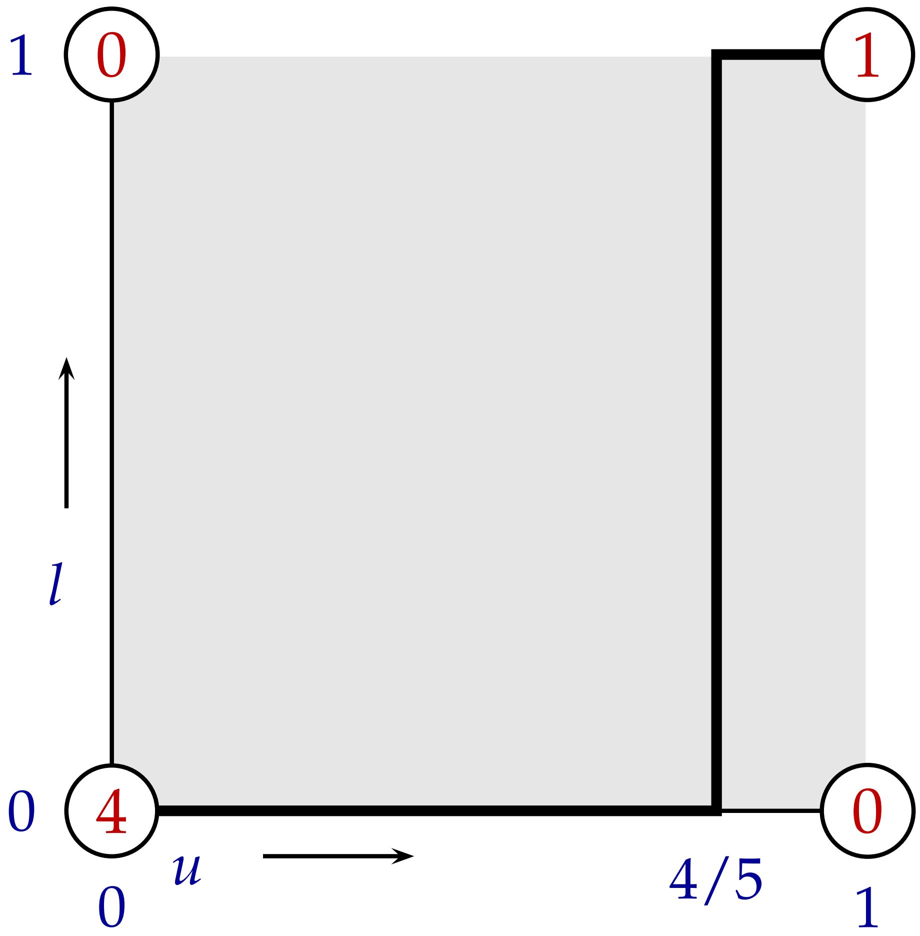
- If $5u - 4 = 0$, it does not matter what col chooses for l—his expected payoff is always $4 - 4(4/5) = 4/5$.

Example: A repeated game with a non-SGP NE



Draw a picture of the payoff surface of col.

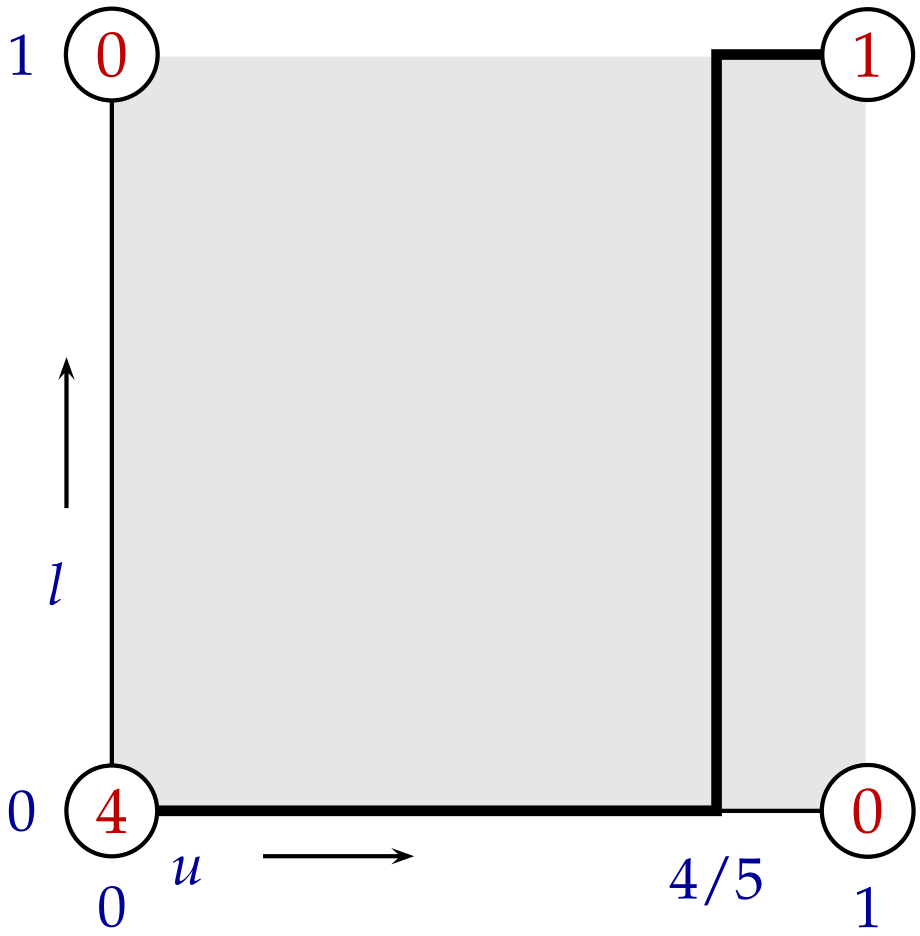
Example: A repeated game with a non-SGP NE



Draw a picture of the payoff surface of col.

- If $5u - 4 = 0$, it does not matter what col chooses for l. He expects $4/5$.

Example: A repeated game with a non-SGP NE

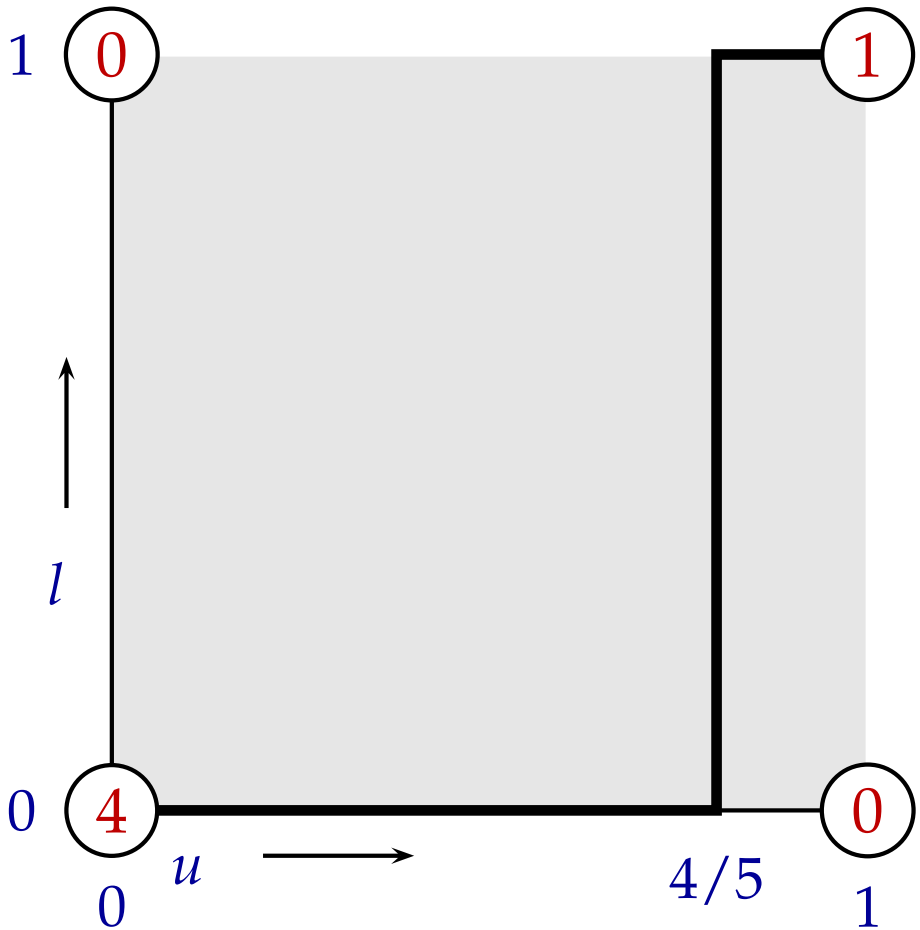


Draw a picture of the payoff surface of col.

- If $5u - 4 = 0$, it does not matter what col chooses for l. He expects $4/5$.

- If $5u - 4 > 0$ col will play $l = 1$, and expects to earn u and $u > 4/5$.

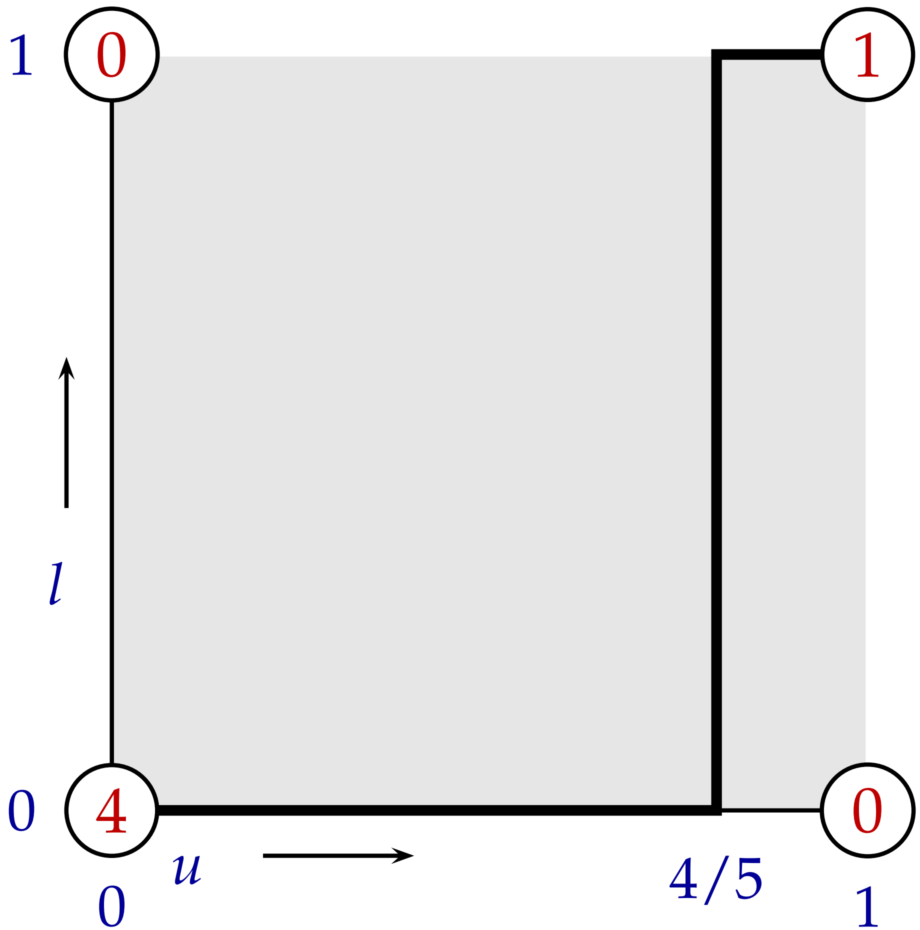
Example: A repeated game with a non-SGP NE



Draw a picture of the payoff surface of col.

- If $5u - 4 = 0$, it does not matter what col chooses for l. He expects $4/5$.
- If $5u - 4 > 0$ col will play $l = 1$, and expects to earn u and $u > 4/5$.
- If $5u - 4 < 0$ col will play $l = 0$, and expects to earn $4 - 4u > 4 - 4(4/5) = 4/5$.

Example: A repeated game with a non-SGP NE



Draw a picture of the payoff surface of col.

- If $5u - 4 = 0$, it does not matter what col chooses for l. He expects $4/5$.
- If $5u - 4 > 0$ col will play $l = 1$, and expects to earn u and $u > 4/5$.
- If $5u - 4 < 0$ col will play $l = 0$, and expects to earn $4 - 4u > 4 - 4(4/5) = 4/5$.

These calculations are done by hand, and do not easily generalise to higher dimensions.

What next?

Now that we know that infinitely many equilibria exist in repeated games (an embarrassment of richness), there are a number of ways in which we may proceed.

What next?

Now that we know that infinitely many equilibria exist in repeated games (an embarrassment of richness), there are a number of ways in which we may proceed.

- **Reinforcement Learning**. Agents simply execute the action(s) with maximal rewards in the past.

What next?

Now that we know that infinitely many equilibria exist in repeated games (an embarrassment of richness), there are a number of ways in which we may proceed.

- **Reinforcement Learning**. Agents simply execute the action(s) with maximal rewards in the past.
- **No-regret learning**. Agents execute the action(s) with maximal virtual rewards in the past.

What next?

Now that we know that infinitely many equilibria exist in repeated games (an embarrassment of richness), there are a number of ways in which we may proceed.

- **Reinforcement Learning**. Agents simply execute the action(s) with maximal rewards in the past.
- **No-regret learning**. Agents execute the action(s) with maximal virtual rewards in the past.
- **Fictitious Play**. Sample the actions of opponent(s) and play a best response.

What next?

Now that we know that infinitely many equilibria exist in repeated games (an embarrassment of richness), there are a number of ways in which we may proceed.

- **Reinforcement Learning**. Agents simply execute the action(s) with maximal rewards in the past.
- **No-regret learning**. Agents execute the action(s) with maximal virtual rewards in the past.
- **Fictitious Play**. Sample the actions of opponent(s) and play a best response.
- **Gradient Dynamics**. This is to approximate NE of single-shot games (stage games) through gradient ascent (hill-climbing).