

Hannes Strass

Faculty of Computer Science, Institute of Artificial Intelligence, Computational Logic Group

# Repeated Play

Lecture 9, 17th Jun 2024 // Algorithmic Game Theory, SS 2024

# Previously ...

- The **regret** is the difference between a player's best possible strategy and their actual strategy.
- A **correlated equilibrium** can be seen as providing players with private signals they can use to best-respond to each other's strategies.
- The **regret matching** algorithm uses self-play to steer play towards the set of correlated equilibria.
- In the case of two-player zero-sum games, regret matching tends towards the set of (mixed) Nash equilibria.
- The **counterfactual regret minimisation** algorithm applies regret matching to every information set of an (imperfect-information) extensive-form game (with perfect recall).

# Motivation

- Previous models have focussed on games that are played **once**.
- We also want to model **repeated** interactions (of the same individuals).
- A particular focus will be the **emergence of cooperation**:
  - In the (one-shot) **Prisoner's Dilemma**, **Confess** dominates **Silent**.
  - Yet people (and businesses, animals, ...) cooperate with each other. Why?
    - (A) Incorrect modelling of preferences (e.g. not including integrity and fidelity).
    - (B) Incorrect modelling of outcomes (e.g. not including in-game retaliation).
    - (C) **Not modelling anticipation of future (and memory of past) interactions.**
- For simplicity, we assume (non-cooperative) **games in normal form**.

We will explore option (C) in this lecture.

# Overview

Fixed Repetition

Open-Ended Repetition

Modelling Noise

Evolutionary Game Theory

# Fixed Repetition

# Finite Repeated Game: Definition

## Definition

A **finite repeated game** consists of a game  $G$  and a natural number  $m \geq 2$ .

- The value of  $m$  indicates the number of repetitions (stages) of  $G$ .
- In each stage  $k$  with  $1 \leq k \leq m$ ,  $G$  is played with all players knowing all **actions** (strategies) chosen by all players in the previous stage (if any).
- The overall utilities of the players are obtained as the arithmetic mean over stages:

$$\bar{u}_i := \frac{u_i(\mathbf{s}^{(1)}) + \dots + u_i(\mathbf{s}^{(m)})}{m}$$

where  $\mathbf{s}^{(k)}$  denotes the **joint action** (strategy profile) chosen in stage  $k$ .

# Recall: Prisoner's Dilemma

## Prisoner's Dilemma

Two robbers, **Eli** and **Fyn**, have been caught by police. Each can stay silent, i.e. **Cooperate** with the other player, or confess to police, i.e. **Defect**. If both cooperate, they get the reward  $R$ ; if both defect, they get the punishment  $P$ . If one cooperates and the other defects, one gets "sucker's payoff"  $S$  and the other gets "temptation payoff"  $T$ .

(Eli, Fyn)	Cooperate	Defect
Cooperate	$(R, R)$	$(S, T)$
Defect	$(T, S)$	$(P, P)$

- To make this a **dilemma**, it must hold that  $S < P < R < T$ .
- For **repeated** play, we additionally require that  $S + T < 2R$ .
- For simplicity, we use our earlier values:  $S = 0$ ,  $P = 1$ ,  $R = 3$ , and  $T = 5$ .

# Repeated Prisoner's Dilemma: Examples (1)

Consider the repeated games with  $m = 6$  and  $C = \text{Cooperate}$ ,  $D = \text{Defect}$ :

Stage $k$	1	2	3	4	5	6	Mean
Eli's action $s_{\text{Eli}}^{(k)}$	C	C	C	C	C	C	
• Fyn's action $s_{\text{Fyn}}^{(k)}$	C	C	C	C	C	C	
Eli's payoff $u_{\text{Eli}}(\mathbf{s}^{(k)})$	3	3	3	3	3	3	3
Fyn's payoff $u_{\text{Fyn}}(\mathbf{s}^{(k)})$	3	3	3	3	3	3	3
Stage $k$	1	2	3	4	5	6	Mean
Eli's action $s_{\text{Eli}}^{(k)}$	C	C	C	C	C	C	
• Fyn's action $s_{\text{Fyn}}^{(k)}$	C	C	C	C	D	D	
Eli's payoff $u_{\text{Eli}}(\mathbf{s}^{(k)})$	3	3	3	3	0	0	2
Fyn's payoff $u_{\text{Fyn}}(\mathbf{s}^{(k)})$	3	3	3	3	5	5	$3\frac{2}{3}$



## Repeated Prisoner's Dilemma: Examples (2)

Stage $k$	1	2	3	4	5	6	Mean
Eli's action $s_{\text{Eli}}^{(k)}$	C	C	C	D	D	D	
• Fyn's action $s_{\text{Fyn}}^{(k)}$	C	C	D	D	D	D	
Eli's payoff $u_{\text{Eli}}(\mathbf{s}^{(k)})$	3	3	0	1	1	1	$1\frac{1}{2}$
Fyn's payoff $u_{\text{Fyn}}(\mathbf{s}^{(k)})$	3	3	5	1	1	1	$2\frac{1}{3}$
Stage $k$	1	2	3	4	5	6	Mean
Eli's action $s_{\text{Eli}}^{(k)}$	C	C	C	D	D	C	
• Fyn's action $s_{\text{Fyn}}^{(k)}$	C	C	D	D	C	C	
Eli's payoff $u_{\text{Eli}}(\mathbf{s}^{(k)})$	3	3	0	1	5	3	$2\frac{1}{2}$
Fyn's payoff $u_{\text{Fyn}}(\mathbf{s}^{(k)})$	3	3	5	1	0	3	$2\frac{1}{2}$

What is a good (meta) strategy for the repeated prisoner's dilemma?

# Repeated Prisoner's Dilemma: Strategies

## Definition

A **meta strategy** for a repeated game states, for each possible history of previous joint actions, a probability distribution on possible actions.

Some possible (meta) strategies for player  $i \in \{E1i, Fyn\}$  are:

**AllDefect**: Play **Defect** in all stages.

**TitForTat**: Play **Cooperate** in stage  $k = 1$ . For stages  $k > 1$ , play  $\mathbf{s}_{-i}^{(k-1)}$ .

**Pavlov**: Play **Cooperate** for  $k = 1$  and whenever  $\mathbf{s}_{-i}^{(k-1)} = \mathbf{s}_i^{(k-1)}$  for  $k > 1$ .

**GrimTrigger**: Play **Cooperate** for  $k = 1$  and for  $k > 1$  as long as  $\mathbf{s}_{-i}^{(k-1)} = \text{Cooperate}$ .  
If  $\mathbf{s}_{-i}^{(k')} = \text{Defect}$  for some  $k' < m$ , play **Defect** for all  $k' < k'' \leq m$ .

What is the “best” (meta) strategy for the finite repeated prisoner's dilemma?

# Repeated Finite Prisoner's Dilemma: Analysis

Stage $k$	1	2	...	$m-2$	$m-1$	$m$	Mean
$s_{Eli}^{(k)}$	C	C	...	C	C	C	
$s_{Fyn}^{(k)}$	C	C	...	C	C	C	
$u_{Eli}(\mathbf{s}^{(k)})$	3	3	...	3	3	3	3
$u_{Fyn}(\mathbf{s}^{(k)})$	3	3	...	3	3	3	3

Repetition does not lead to cooperation if the number of stages is known!

# Repeated Finite Prisoner's Dilemma: Equilibria

## Theorem

In a finite repeated prisoner's dilemma,  $(\text{AllDefect}, \text{AllDefect})$  is the strict Nash equilibrium.

## Proof.

- $(\text{AllDefect}, \text{AllDefect})$  is a Nash equilibrium: Let  $E_{li}$  play  $\text{AllDefect}$ . If  $F_{yn}$  plays anything other than  $\text{AllDefect}$ , then there is some  $k \leq m$  with  $s_{F_{yn}}^{(k)} = C$ . But then  $s_{F_{yn}}^{(k)} = D$  yields a higher payoff, so  $F_{yn}$  does not play a best response. So only  $\text{AllDefect}$  is a best response to  $\text{AllDefect}$ .
- Let  $(\sigma_{E_{li}}, \sigma_{F_{yn}})$  be a (mixed) Nash equilibrium and assume  $\sigma_{E_{li}}^{(k)}(C) > 0$ . Define  $\sigma'_{E_{li}}$  by

$$\sigma'_{E_{li}}^{(\ell)} := \begin{cases} \sigma_{E_{li}}^{(\ell)} & \text{if } \ell < k, \\ D & \text{otherwise.} \end{cases}$$

Now  $\sigma'_{E_{li}}$  yields a higher (than  $\sigma_{E_{li}}$ ) payoff against  $\sigma_{F_{yn}}$ , contradiction.  $\square$

# Open-Ended Repetition

# Random Repeated Game: Definition

## Definition

A **random repeated game** consists of a game  $G$  and a number  $\delta \in [0, 1)$ .

- In each stage,  $G$  is played with full knowledge about previous stages.
- $\delta$  is the **continuation probability**:  
At stage  $k$ , the game passes into stage  $k + 1$  with probability  $\delta$ .
- Overall utilities are again the arithmetic mean over stages.

## Observation

The expected number of stages is  $1 + \delta + \delta^2 + \dots = \frac{1}{1-\delta}$ .

# Random Repeated Prisoner's Dilemma (1)

## Theorem

In the random repeated prisoner's dilemma, (`GrimTrigger`, `GrimTrigger`) is a Nash equilibrium for sufficiently large  $\delta$ .

## Proof.

- If both play `GrimTrigger`, they get overall payoff  $R$ .
- To obtain a higher payoff, `Eli` must `Defect` at some stage  $k$  (and later).
- If it pays to `Defect` in some stage, it pays to `Defect` from the first stage on.
- Thus `Eli`'s payoff against `GrimTrigger` is  $\frac{1}{m} \cdot (T + (m - 1) \cdot P)$  for  $m = \frac{1}{1-\delta}$ .
- Therefore `GrimTrigger` is a best response to itself whenever

$$(1 - \delta) \cdot T + \delta \cdot P \leq R, \quad \text{or equivalently} \quad \frac{T - R}{T - P} \leq \delta.$$

- Since  $P < R$  implies  $T - R < T - P$ , it follows that  $\delta$  can be suitably chosen.  $\square$

## Random Repeated Prisoner's Dilemma (2)

- $(AllDefect, AllDefect)$  continues to be an equilibrium for all  $\delta$ .
- There are various other non-punitive outcomes, e.g.  $(Pavlov, Pavlov)$  or  $(TitForTat, TitForTat)$  (for sufficiently large  $\delta$ ).
- In fact, the equilibrium concept makes very unspecific predictions:

### Theorem ("Folk Theorem")

Let  $(a, b)$  be an overall payoff pair with  $a, b > P$ .

For every  $\varepsilon > 0$  and sufficiently large continuation probability  $\delta$ , there exists a Nash equilibrium of the random repeated prisoner's dilemma with mean expected payoff pair  $(c, d)$  such that  $|a - c| < \varepsilon$  and  $|b - d| < \varepsilon$ .

### Example

To achieve  $a = \frac{2}{3} \cdot R + \frac{1}{3} \cdot T$  and  $b = \frac{2}{3} \cdot R + \frac{1}{3} \cdot S$ , **Eli** and **Fyn** play as follows:

- **Eli** repeats **C, C, D** as long as **Fyn** cooperates; **Eli** only plays **D** otherwise.
- **Fyn** plays **C** as long as **Eli** plays the pattern **CCD**; **Fyn** repeats **D** otherwise.



# Modelling Noise

# Noisy Repeated Game: Definition

## Definition

A **noisy repeated game** consists of a game  $G$  and the following:

- a continuation probability  $\delta \in [0, 1)$ ,
- a **perception error** probability  $\xi \in [0, 1]$ , and
- an **implementation error** probability  $\eta \in [0, 1]$ .

In each stage  $k$ , the game  $G$  is played such that for each  $i \in P$ :

- Player  $i$ 's chosen action is being used with probability  $(1 - \eta)$ , and
- a different action is being used with probability  $\eta$ .
- With probability  $(1 - \xi)$ , player  $i$  perceives the actual outcome, and
- with probability  $\xi$ , player  $i$  perceives a different outcome.
- A (meta) strategy for player  $i$  in a noisy repeated game only considers:
  - intended and implemented actions of  $i$  (but not how those were perceived),
  - joint actions  $\mathbf{s}_{-i}$  as perceived by  $i$  (but not intentions or actual actions).

# Noisy Repeated Game: Example

Eli and Fyn are both trying to play **TitForTat** in a noisy repeated game:

Stage $k$	1	2	3	4	5	6	Mean
Eli's intended action	C	C	D	C	D	D	
Eli's implemented action	C	C	D	C	D	D	
Eli's action as perceived by Fyn	C	C	D	C	D	D	
Fyn's intended action	C	C	C	D	C	D	
Fyn's implemented action	C	D	C	D	C	D	
Fyn's action as perceived by Eli	C	D	C	D	D	D	
Eli's payoff	3	0	5	0	5	1	$2\frac{1}{3}$
Fyn's payoff	3	5	0	5	0	1	$2\frac{1}{3}$

- Nash equilibria make no specific predictions.
- A “noisy version” of the folk theorem exists.

# Evolutionary Game Theory

# Evolutionary Game Theory: Setting

## Evolutionary Game Theory: Basic Modelling Assumptions

- Organisms (individuals) in a population interact (compete for resources).
- Two individuals of the (infinite) population are chosen at random ...
- ...and play a symmetric ( $u_1(s_1, s_2) = u_2(s_2, s_1)$  for all  $s_1, s_2 \in S_1 = S_2$ ) game.
- The game's payoffs will increase/decrease the individuals' fitness ...
- ...and thereby the (individual's) strategy's frequency in the population via:
- **replication**, where payoffs increase fitness to reproduce (create identical copies of the individual, playing the same strategy); or
- **imitation**, where individuals observe payoffs of others and adopt strategies that yield higher utilities.

- Questions:
1. How do distributions of strategies evolve over time?
  2. Which strategies are resilient to mutant invasions?

We will briefly consider 2. in the remainder of this lecture.

# Example: Hawks v. Doves

## Hawks v. Doves

Two individuals compete for a resource. Obtaining the resource leads to a fitness increase of  $V$ ; losing a fight leads to a fitness decrease of  $C$ . A **Hawk** will win the resource against a **Dove**. If two **Doves** play each other, they will split the resource equally. If two **Hawks** play each other, they will both fight, with equal chances of winning.

(1, 2)	Hawk	Dove
Hawk	$\frac{V-C}{2}$	$V$
Dove	$0$	$\frac{V}{2}$

- (Dove, Dove) is not a Nash equilibrium whenever  $V > \frac{V}{2}$ .
- (Hawk, Hawk) is a Nash equilibrium whenever  $V \geq C$ .

Which (mixed) strategies are resilient to mutant invasions?

# Evolutionarily Stable Strategies: Definition

- Suppose a fraction  $1 - \varepsilon$  of individuals (**incumbents**) that play strategy  $\pi$ ;
- the remaining fraction  $\varepsilon$  of individuals (**mutants**) play strategy  $\rho$ .

**Idea:** Strategy  $\pi$  is **stable** if the expected payoff of  $\pi$  is higher than that of all  $\rho \neq \pi$ , for sufficiently small  $\varepsilon$ .

- Denote the expected (one-shot) payoff of playing  $\pi$  against  $\rho$  by  $U(\pi | \rho)$ .
- The (overall) expected payoff of an incumbent:  $(1 - \varepsilon) \cdot U(\pi | \pi) + \varepsilon \cdot U(\pi | \rho)$
- The (overall) expected payoff of a mutant:  $(1 - \varepsilon) \cdot U(\rho | \pi) + \varepsilon \cdot U(\rho | \rho)$

## Definition

A (mixed) strategy  $\pi$  is an **evolutionarily stable strategy** (ESS) iff for every (mixed) strategy  $\rho \neq \pi$ , there exists an  $\varepsilon_\rho$  such that for all  $0 < \varepsilon < \varepsilon_\rho$ :

$$(1 - \varepsilon) \cdot U(\pi | \pi) + \varepsilon \cdot U(\pi | \rho) > (1 - \varepsilon) \cdot U(\rho | \pi) + \varepsilon \cdot U(\rho | \rho)$$

# Evolutionarily Stable Strategies: Variant

## Observation (1)

A strategy  $\pi$  is evolutionarily stable if and only if, for all  $\rho \neq \pi$ , either:

- $U(\pi | \pi) > U(\rho | \pi)$ , or
- $U(\pi | \pi) = U(\rho | \pi)$  and  $U(\pi | \rho) > U(\rho | \rho)$ .

Thus if  $\pi$  is an ESS, then:

- $U(\pi | \pi) \geq U(\rho | \pi)$  for all  $\rho \neq \pi$ , whence
- $\pi$  is a best response to itself, and
- $(\pi, \pi)$  is a Nash equilibrium.

## Example: Hawks v. Doves

- If  $V > C$ , then  $\frac{V-C}{2} > 0$  and **Hawk** is an ESS.
- **(Dove, Dove)** is not a Nash equilibrium, so **Dove** is not an ESS.



# ESSs and Nash Equilibria

## Theorem

Let  $G$  be a two-player normal-form game with symmetric payoffs  $U(\cdot | \cdot)$ . A mixed strategy  $\pi$  is an evolutionarily stable strategy for  $G$  if and only if:

1.  $(\pi, \pi)$  is a Nash equilibrium of  $G$ , and
2. for every best response  $\rho$  to  $\pi$  with  $\rho \neq \pi$ , we have  $U(\pi | \rho) > U(\rho | \rho)$ .

## Example: Hawks v. Doves

- Assume  $V \leq C$  and  $C > 0$  then  $\pi = \left\{ \text{Hawk} \mapsto \frac{V}{C}, \text{Dove} \mapsto 1 - \frac{V}{C} \right\}$  is an ESS:
- $U(\text{Hawk} | \pi) = \frac{V}{C} \cdot \frac{V-C}{2} + (1 - \frac{V}{C}) \cdot V = \frac{V^2 - VC + 2CV - 2V^2}{2C} = \frac{CV - V^2}{2C} = U(\pi | \pi)$
- $U(\text{Dove} | \pi) = (1 - \frac{V}{C}) \cdot \frac{V}{2} = \frac{CV - V^2}{2C} = U(\pi | \pi)$
- Any  $\rho = \{ \text{Hawk} \mapsto r, \text{Dove} \mapsto 1 - r \}$  is a best response to  $\pi$ ;  $r \neq \frac{V}{C}$  allows to verify  $U(\pi | \rho) > U(\rho | \rho)$ .

# Computational Complexity of ESS

## Observation

Not every normal-form game has an evolutionarily stable strategy.

## Example: Rock-Paper-Scissors

- The unique Nash equilibrium is  $\pi^* = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$  with  $U(\pi^* | \pi^*) = 0$ .
- Every  $\rho = \{\text{Rock} \mapsto x, \text{Paper} \mapsto y, \text{Scissors} \mapsto z\}$  is a best response to  $\pi^*$ .
- $U(\pi^* | \rho) = U(\rho | \pi^*) = 0$  by symmetry and since the game is zero-sum.
- Finally,  $U(\rho | \rho) = 0$  as well, thus  $\pi^*$  is no ESS.

Thus the following decision problem is relevant:

### Exists-ESS

**Given:** A two-player symmetric game  $G$  in normal form.

**Question:** Does  $G$  have an ESS?

# Computational Complexity of ESS (1)

## Theorem

**Exists-ESS** is both NP-hard and coNP-hard.

Proof (Sketch, I).

- We reduce from the following problem **Max-Clique**:
  - Given:** An undirected Graph  $(V, E)$  and a number  $k \in \mathbb{N}$ .
  - Question:** Does  $(V, E)$  have a maximal clique of size exactly  $k$ ?
- The nodes  $V$  will become pure strategies.
- The edges  $E$  will determine pairs of pure strategies with positive payoff.
- An ESS must play well against itself, so must prefer edges of the graph.
- Playing strategies that mix the nodes of a clique serves to achieve this.

# Computational Complexity of ESS (2)

Proof (Sketch, II).

- Assume w.l.o.g. that  $V = \{1, 2, \dots, n\}$ .
- To define  $G_{(V,E)}$ , set  $S_1 = S_2 = V \cup \{0\}$  and define  $U(\cdot | \cdot)$  as follows:

$$U(s_i | s_j) := \begin{cases} 1 & \text{if } i, j \neq 0 \text{ and } i \neq j \text{ and } (i, j) \in E, \\ 0 & \text{if } i, j \neq 0 \text{ and } i \neq j \text{ and } (i, j) \notin E, \\ \frac{1}{2} & \text{if } i, j \neq 0 \text{ and } i = j, \\ 1 - \frac{1}{2k} & \text{otherwise, that is, } i = 0 \text{ or } j = 0. \end{cases}$$

- It remains to show that  $G_{(V,E)}$  has an ESS if and only if the largest clique in  $(V, E)$  has a size other than  $k$ .

# Computational Complexity of ESS (3)

Proof (Sketch, III).

- If  $C$  is a maximal clique in  $(V, E)$  of size  $k' > k$  and  $\pi$  is the uniform distribution on  $C$ , then  $\pi$  is an ESS.
- If the maximal clique of  $(V, E)$  has size  $k' < k$ , then pure strategy 0 is an ESS.
- If the maximal clique of  $(V, E)$  has size at least  $k$ , then 0 is not an ESS.
- If the maximal clique of  $(V, E)$  has size at most  $k$ , then any strategy other than 0 is not an ESS. □

↪ Even when an ESS exists, it is unlikely that a finite population will quickly converge to it.

# Conclusion

## Summary

- In a **finite repeated game**, a two-player normal-form game is repeated for a fixed number of times; cooperation cannot be expected in this case.
- In a **random repeated game**, the end of interaction can not be predicted for sure; cooperation can emerge for large enough continuation probabilities, but equilibria make no specific predictions.
- A **noisy repeated game** may have implementation/perception errors.
- An **evolutionarily stable strategy** is a Nash equilibrium that performs better against “mutants” than the “mutants” against themselves.
- Deciding whether a game has an ESS is NP-hard and coNP-hard.

## Action Points

Prove that  $\pi$  on slide 25 is an ESS by showing  $U(\pi | \rho) > U(\rho | \rho)$ .