TECHNISCHE UNIVERSITÄT DRESDEN

Computational Logic ∴ Group

Hannes Strass

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

Repeated Play

Lecture 8, 19th Jun 2023 // Algorithmic Game Theory, SS 2023

Previously ...

- A **behaviour strategy** assigns move probabilities to information sets.
- A **belief system** assigns probabilities to histories in information sets.
- An **assessment** is a pair (behaviour strategy profile, belief system).
- A **sequentially rational** assessment plays best responses “everywhere”.
- An assessment satisfies **consistency of beliefs** whenever the belief system’s probabilities match what is expected from everyone playing according to the behaviour strategy profile.
- An assessment is a **weak sequential equilibrium** iff it is both sequentially rational and satisfies consistency of beliefs.
- Mixed Nash equilibria for normal-form games and subgame perfect equilibria for sequential perfect-information games are special cases of weak sequential equilibria for extensive-form games.

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:

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

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 = Cooperate, D = Defect:

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

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

Stage k	1	2	3	4	5	6	Mean
Eli's action $s_{\mathtt{Eli}}^{(k)}$	C	C	C	D	D	C	
Fyn's action $s_{\mathtt{Fyn}}^{(k)}$	C	C	D	D	C	C	
Eli's payoff $u_{\mathtt{Eli}}(\mathbf{s}^{(k)})$	3	3	0	1	5	3	$2\frac{1}{2}$
Fyn's payoff $u_{\mathtt{Fyn}}(\mathbf{s}^{(k)})$	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 $i \in \{\mathtt{Eli}, \mathtt{Fyn}\}$ 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 $\mathbf{s}_{-i}^{(k-1)} = s_i^{(k-1)}$ for $k > 1$.

`GrimTrigger`: Play `Cooperate` for $k = 1$ and for $k > 1$ as long as $\mathbf{s}_{-i}^{(k-1)} = \mathtt{Cooperate}$.
If $\mathbf{s}_{-i}^{(k')} = \mathtt{Defect}$ 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
$s_{\texttt{Eli}}^{(k)}$	C	C	...	C	C	C	
$s_{\texttt{Fyn}}^{(k)}$	C	C	...	C	C	C	
$u_{\texttt{Eli}}(\mathbf{s}^{(k)})$	3	3	...	3	3	3	3
$u_{\texttt{Fyn}}(\mathbf{s}^{(k)})$	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, $(\texttt{AllDefect}, \texttt{AllDefect})$ is the strict Nash equilibrium.

Proof.

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

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

  Now $\sigma'_{\texttt{Eli}}$ yields a higher (than $\sigma_{\texttt{Eli}}$) payoff against $\sigma_{\texttt{Fyn}}$, contradiction. □

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.
- δ is the **continuation probability**:
  At stage k, the game passes into stage $k + 1$ with probability δ.
- Overall utilities are again the arithmetic mean over stages.

Observation

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

Random Repeated Prisoner's Dilemma (1)

Theorem

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

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 δ can be suitably chosen. □

Random Repeated Prisoner's Dilemma (2)

- (AllDefect, AllDefect) continues to be an equilibrium for all δ.
- There are various other non-punitive outcomes, e.g. (Pavlov, Pavlov) or (TitForTat, TitForTat) (for sufficiently large δ).
- 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 δ, 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 η.
- With probability $(1 - \xi)$, player i perceives the actual outcome, and
- with probability ξ, 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 fight over 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 fight (and win) 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 π;
- the remaining fraction ε of individuals (**mutants**) play strategy ρ.

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

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

Definition

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

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

Evolutionarily Stable Strategies: Variant

Observation (1)

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

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

Thus if π is an ESS, then:

- $U(\pi \mid \pi) \geq U(\rho \mid \pi)$ for all $\rho \neq \pi$, whence
- π is a best response to itself, and
- (π, π) 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 $U(\cdot \mid \cdot)$.
A mixed strategy π is an evolutionarily stable strategy for G if and only if:

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

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:
- If $p = \rho'(\text{Hawk}) > \frac{V}{C}$, then responding Dove would improve payoff.
- If $q = \rho''(\text{Hawk}) < \frac{V}{C}$, then responding Hawk would improve payoff.
- If $\rho = \{\text{Hawk} \mapsto r, \text{Dove} \mapsto 1 - r\}$ with $r \neq \frac{V}{C}$ is a best response to π, we get $U(\rho \mid \pi) = U(\pi \mid \pi)$. This yields r and we can verify $U(\pi \mid \rho) > U(\rho \mid \rho)$.

Computational Complexity of ESS

Observation

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

Example: Rock-Paper-Scissors

- Rock-Paper-Scissors has the unique Nash equilibrium $\pi^* = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.
- Every pure strategy $s \in \{\texttt{Rock}, \texttt{Paper}, \texttt{Scissors}\}$ is a best response to π^*.
- Moreover, for every pure s, $U(\pi^* \mid s) = U(s \mid s)$ and π^* 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 $V = \{1, 2, \ldots, n\}$.
- To define $G_{(V,E)}$, set $S_1 = S_2 = V \cup \{0\}$ and define $U(\cdot \mid \cdot)$ as follows:

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

- 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 π is the uniform distribution on C, then π 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. □

$\rightsquigarrow$ 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 π on slide 25 is an ESS by showing $U(\pi \mid \rho) > U(\rho \mid \rho)$.