

Lab 9 - Co-evolutionary Learning of Strategies for 2-Player Iterated Prisoner's Dilemma Games

CSE, SUSTech

Outline of This Lab

- Teaching Assistant
- What is 2-Player Iterated Prisoner's Dilemma Games?
- Representation of Strategy
- Experimental Setup
- Illustrate the Results!

Teaching assistant

- Ms. Honglin Jin (12531321@mail.sustech.edu.cn)

Iterated Prisoner's Dilemma Games

- Non-zero sum, non-cooperative games

	Cooperate	Defect
Cooperate	R	T
Defect	R	S

	C	D
C	3	5
D	0	1

Player B

Player A

	C	D
C	3	5
D	0	1

$$T > R > P > S \text{ and } R > (S+T)/2$$

- How can an EA learn to play the game optimally starting from a population of random strategies?

Representation of 2IPD Strategy - Part I

To be more specific, consider the set of strategies that are deterministic and use the outcomes of the three (two, or one) previous moves to make a choice in the current move. Since there are four possible outcomes for each move, there are $4 \times 4 \times 4$ (4×4 or $4 = 64$ (16 or 4) (or 2^6 , 2^4 , and 2^2) different histories of the three (two or one) moves. Therefore, to determine its choice of cooperation or defection, a strategy would only need to determine what to do in each of the situation that could arise. This could be specified by a list of sixty-four (sixteen or 4) C's and D's.

For example a strategy with one previous move will be $a_1 a_2 a_3 a_4$.

the length of individual = $3 \times 2 + 2^{3 \times 2} = 70$

pre-game history to start the game

number of previous moves

number of players

cc (00)	a_1
CD (01)	a_2
DC (10)	a_3
DD (11)	a_4

Representation of 2IPD Strategy -Part II

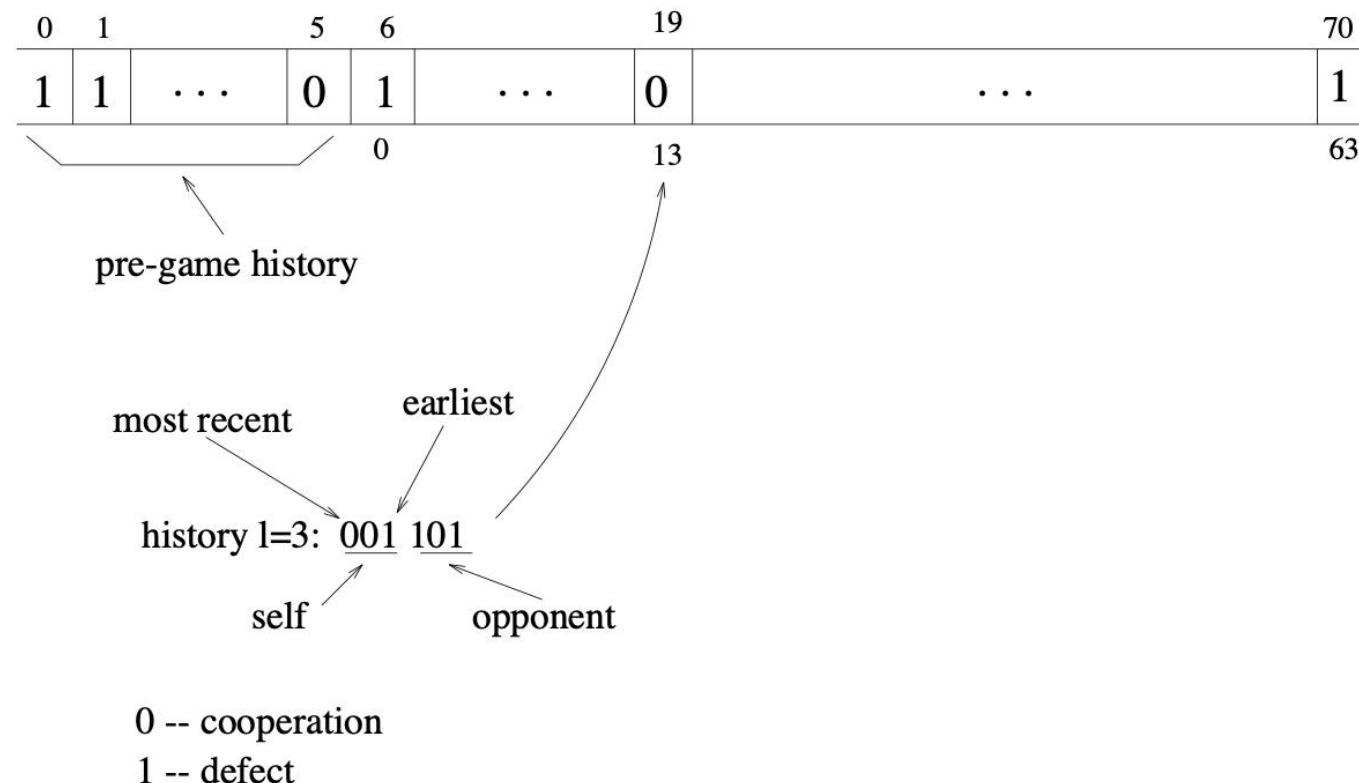


Figure 3: Encoding of strategies, assuming history (memory) length 3.

Experimental Setup

Implement a genetic algorithm to evolve strategies for playing the 2IPD with three, two, and one previous move, respectively.

- ✓ The payoff matrix →
- ✓ each game plays for 50 moves
- ✓ crossover rate: 0.6, mutation rate: 0.001
- ✓ population size: 100
- ✓ Rank-based selection
- ✓ maximum generation numbers: 200
- ✓ run 20 times

		Player B	
		C	D
Player A	C	3	5
	D	3	0
	C	0	1
	D	5	1

Illustrate The Results!

Plot 3 figures for 3, 2, and 1 previous move, respectively

- ✓ x-axis: generation number.
- ✓ y-axis: average payoff over 20 runs.