Optimal Strategies of the Iterated Prisoner's Dilemma for Multiple Conflicting Objectives

Shashi Mittal and Kalyanmoy Deb, IIT Kanpur

Email: deb@iitk.ac.in, http://www.iitk.ac.in/kangal

The game

		Player 2	
	Decision	Cooperate	Defect
Player 1	Cooperate	R=3 R=3	S=0 T=5
	Defect	T=5 S=0	P=1 P=1

Encoding strategies

Axelrod's method of encoding strategies: Assign a value to each possible pair of Moves: e.g. R = 0, T = 1, S = 2 and P = 3. Specify a move for a history of three moves. Using this scheme, a particular strategy can be defined as a 70-bit binary string (64 for history 3 moves, and 6 for pre-game behavior)

```
Say, previous three moves are:
        Player 1  Player2   Code
Move 1     C         C       R
Move 2     D         C       T
Move 3     C         C       R
                  RTR=(010)  = 4
                           4
Player 1 chooses 5-th position
```

```
An example EA Solution:
CDDCC..........CDC  CCDDCC
  64 positions          6 pos.
                  (for initial move)
Outcome: (C) or Cooperate
```

Axelrod's work

- He used single-objective genetic algorithm (GA) to evolve optimal strategies against a set of opponents
- The strategies so obtained performed quite well in a round robin tournament, and defeated other optimal strategies (e.g. *Tit for Tat*)

Our approach

Use a **multi-objective** genetic algorithm:

- *Maximize* self score
- *Minimize* opponent score

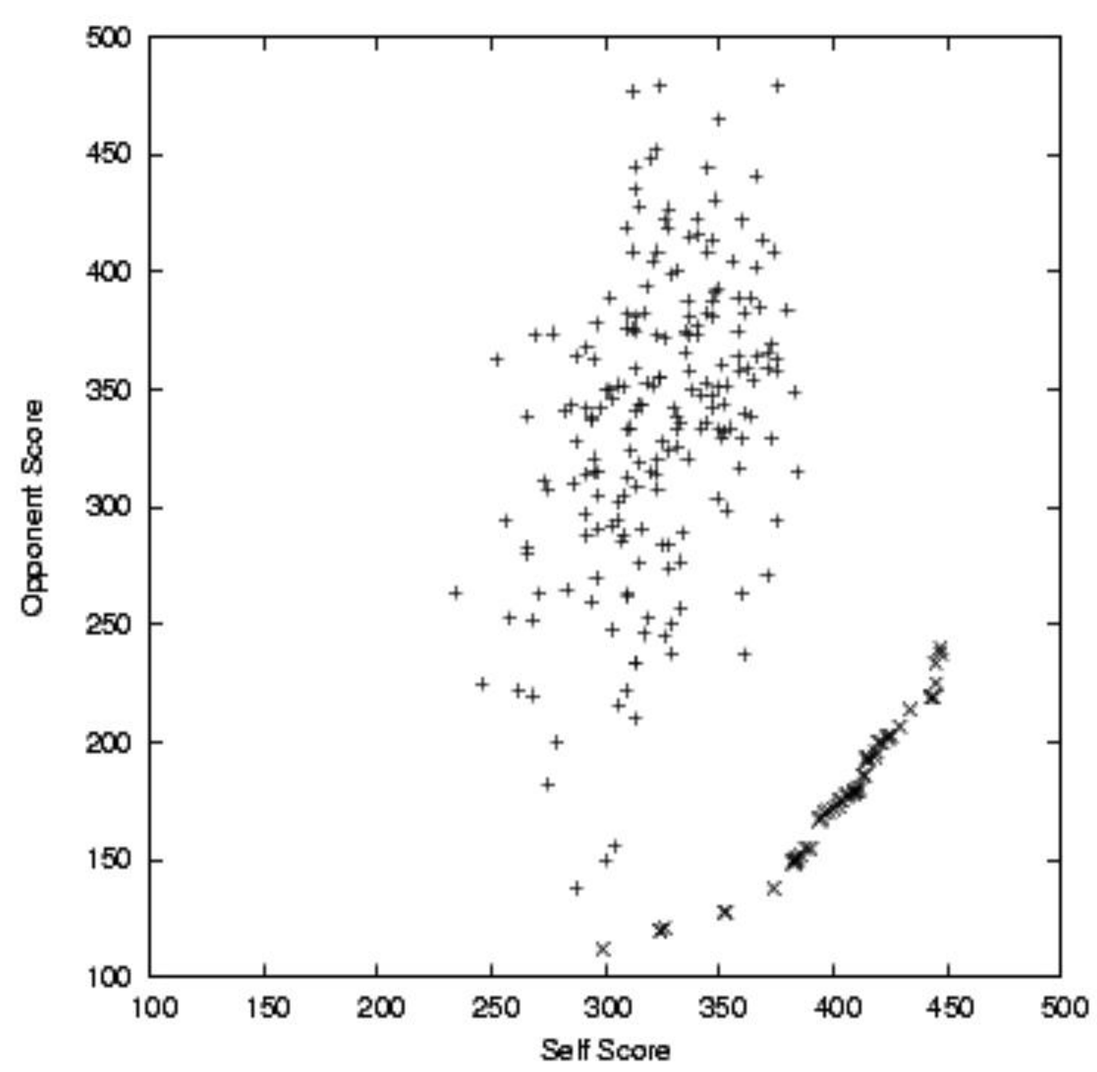
16 other players were included in the round robin tournament.

NSGA-II algorithm was used to obtain the set of Pareto-optimal strategies.

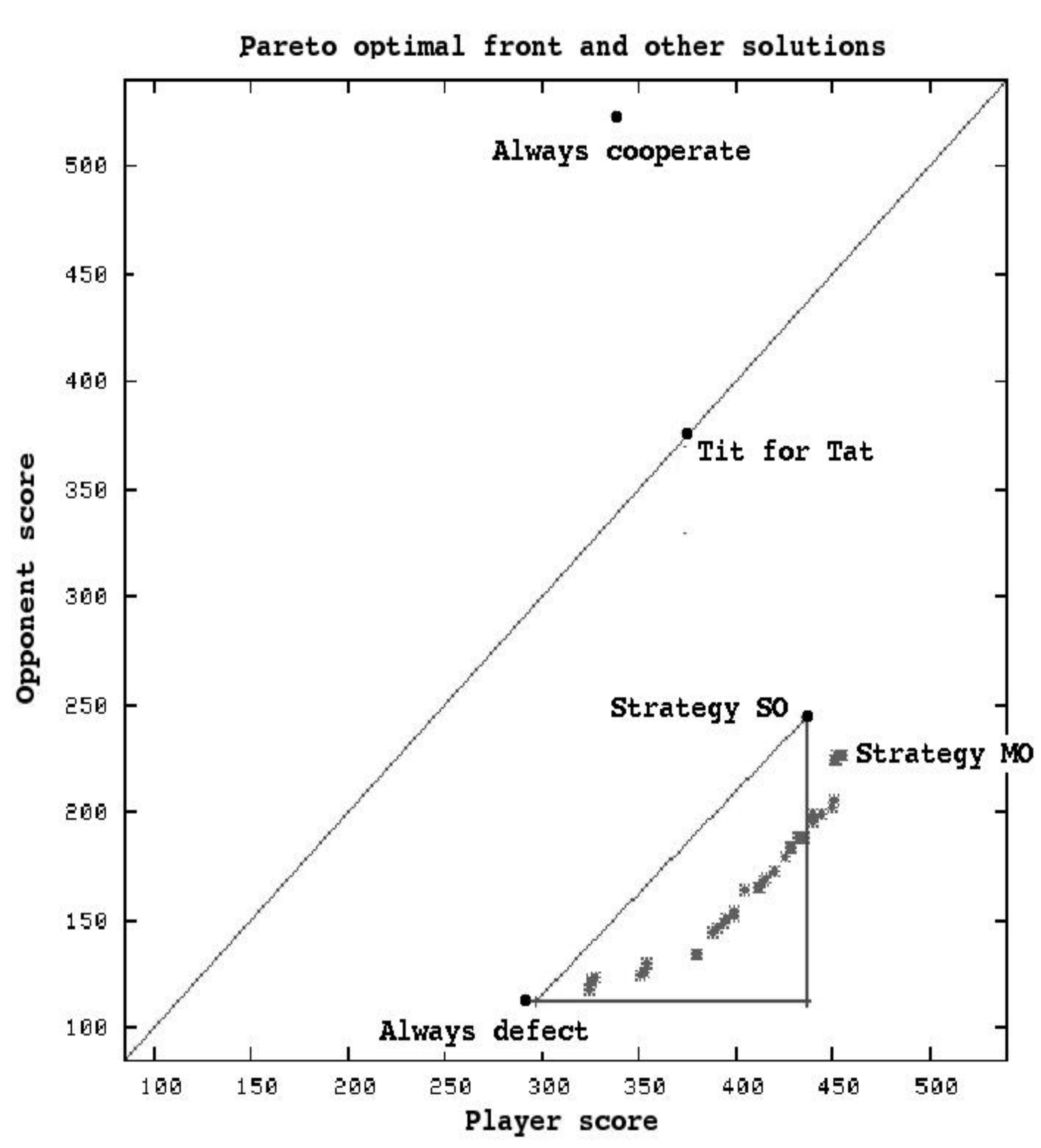
Why multi-objective GA?

- The two objectives, maximizing self-score and minimizing opponent score are conflicting
- A game can also be won by minimizing the opponent's score
- Multi-objective will give a set of Pareto-optimal strategies, which can give useful insight about optimal strategies

Results obtained by using NSGA-II



The above figure shows the initial random solutions (shown with '+') and the Pareto-optimal front (shown in 'x') obtained using NSGA-II



This figure shows the Pareto optimal front together with a few other strategies. It shows that the strategy obtained using single objective GA is dominated by the Pareto-optimal front.

Tournament scores

Player	Average score
Strategy MO	431
Strategy SO	421
Tit for Tat	394
Hard Tit for Tat	379
Soft Majority	375
Tit for two tats	374
Spiteful	368
Naïve prober	362
Remorseful prober	351
Always cooperate	343
Pavlov	341
Suspicious tit for tat	327
Periodic player CD	320
Periodic player CCD	320
Hard majority	307
Random player	296
Always defect	288
Periodic player DDC	286

Strategy SO: Strategy obtained using single-objective GA

Strategy MO: Strategy obtained using NSGA-II

Results

- There is indeed a trade-off between the two objectives
- NSGA-II gives strategies which outperform the other strategies, as well as the one obtained using single-objective GA
- Strategies lying on the Pareto-optimal front share some interesting common features

Other strategies used

Always cooperate: Cooperates on every move

Always defect: Defects on every move

Tit for tat: Cooperates on first move, then copies opponent's last move

Suspicious tit for tat: Same as tit for tat but defects on first move

Pavlov: Cooperates on first move, and defects only when both players disagreed on prev. move

Spiteful: Cooperates until the opponent defects

Random player: Makes a random move

Periodic player CD: Plays C,D periodically

Tit for two tats: Cooperates on first move, and defects only when the opponent defects two times

Soft majority: Begins by cooperating, and cooperates if the majority of opponent's moves are cooperate

Hard majority: Begins by defecting and defects if the majority of opponent's moves are defect