Artificial Intelligence

Evolutionary Algorithms

Lesson 10: Behavioral Simulation

Vincenzo Piuri

Università degli Studi di Milano

Evolutionary Algorithms for Behavioral Simulation

- Theoretical basis of behavioral simulation: game theory
  - Analyzes social and economic situations
  - Models agents and their actions as game moves in a formally specified framework
- Evolutionary algorithms:
  - Encode the moves of an agent in a chromosome
  - Agents interact with each other and evaluate their success
  - Agents reproduce or die due to their achieved success

The Prisoner's Dilemma (1)

- Most thoroughly studied problem of the game theory
  - 2 persons robbed a bank and were arrested
  - The available circumstantial evidence is not sufficient for a conviction because of the bank robbery
  - There is, however, sufficient evidence for a conviction because of a lesser criminal offense (say, illegal possession of firearms)
  - Prosecutor offers both prisoners to become a key witness

The Prisoner's Dilemma (2)

- If one of them confesses to the bank robbery, he/she is exempted from punishment
- The other prisoner will be punished with the full force of the law (10 years imprisonment)
- Problem:
  - Both prisoners are offered this possibility and thus both may be tempted to confess
  - Since they both pleaded guilty, they receive a mitigated sentence
  - Both of them have to spend 5 years in prison

The Prisoner’s Dilemma (3)

- Payoff matrix

B / A	keeps silent	confesses
keeps silent	A: -1, B: -1	A: -10, B: 0
confesses	A: 0, B: -10	A: -5, B: -5

- Keeping silent is favorable for both
  - But: a double confession is the so-called Nash equilibrium
  - No agent can improve its payoff by changing its action (each payoff matrix has at least one Nash equilibrium

The Prisoner's Dilemma (4)

- Payoff matrix

A \ B	cooperate	defect
cooperate	-1 -1	-10 0
defect	0 -10	-5 -5

A \ B	cooperate	defect
cooperate	R R	S T
defect	T S	P P

R: Reward for mutual cooperation **P**: Punishment for mutual defection
T: Temptation to defect **S**: Sucker's payoff

  - Exact values for R, P, T and S are not important
  - It must hold
    - $T > R > P > S$

      there is actually a temptation; cooperation must be better than defection; incentive to avoid getting exploited
    - $2R > T + S$

      ongoing cooperation must be better than alternate exploitation

The Prisoner’s Dilemma (5)

- Iterated prisoner’s dilemma
  - Dilemma is performed several times consecutively where all previous actions of the agents are known
  - If the dilemma is performed once, it is favorable to choose the Nash equilibrium
  - If the dilemma is performed several times, one agent can react on the uncoorperative behavior (possibility of retaliation)

The Prisoner's Dilemma (6)

- Questions:
  1. Is cooperation created in the iterated prisoner's dilemma?
  2. What is the best strategy in the iterated prisoner's dilemma?
- Payoff matrix

A \ B	cooperate	defect
cooperate	3 / 3	0 / 5
defect	5 / 0	1 / 1

  - Smallest non-negative integer numbers that satisfy the two conditions

The Prisoner's Dilemma (7)

- Equivalent Retaliation (Tit-for-Tat)
  - Cooperate in the first game
  - React in all following games with the move of the opponent's previous played game
  - May react inadequately to mistakes
    - If two instances of Tit-for-Tat play against each other and one instance "accidentally" plays defect, this results in mutual retaliations
- Tit-for-Two-Tat
  - Strategy starts retaliating only after having been exploited twice

Genetic Approach (1)

- Encoding of the strategies
  - Consider all possible sequences of three consecutive games (6 moves: $2^6 = 64$ possible sequences)
  - Store what move should be played in the next game (C – cooperate, D – defect)

		1. game	2. game	3. game	
1. Bit:	response to	(C,C),	(C,C),	(C,C):	C
2. Bit:	response to	(C,C),	(C,C),	(C,D):	D
3. Bit:	response to	(C,C),	(C,C),	(D,C):	C
⋮		⋮		⋮	⋮
64. Bit:	response to	(D,D),	(D,D),	(D,D):	D

  - 6 Bit that encode the course of the game “before” the first move
  - Each chromosome has 70 binary genes (either C or D)

Genetic Approach (2)

- Initial population
  - Randomly sampling bit sequences of length 70
- Current population
  - Choose pairs of individuals randomly
  - Pairs play the prisoner's dilemma 200 times
  - On the first 3 games: use the stored history of the beginning of the games to determine the moves

over-average result	$(x \geq \mu + \sigma)$:	2 children
average result	$(\mu - \sigma < x < \mu + \sigma)$:	1 child
below-average result	$(\mu - \sigma \geq x)$:	no offspring

genetic operators: Bit-Mutation, one-point crossover

Genetic Approach (3)

- Deduced behavioral general patterns
  - Don’t rock the boat: Cooperate after three times (C,C), (C,C), (C,C) → C
  - Be provokable: Play defect after a sudden defect of the opponent (C,C), (C,C), (C,D) → D
  - Accept an apology: Cooperate after mutual exploitation (C,C), (C,D), (D,C) → C
  - Forget: (Do not be resentful) Cooperate after cooperation has been restored after one defect (also without retaliation) (C,C), (C,D), (C,C) → C
  - Accept a rut: Play defect after three times defect of the opponent (D,D), (D,D), (D,D) → D