CS540 Introduction to Artificial Intelligence

Lecture 20

Young Wu

Based on lecture slides by Jerry Zhu and Yingyu Liang

August 3, 2022

Local Search

Motivation

- Local search is about searching through a state space by iteratively improving the cost to find an optimal or near-optimal state.
- The successor states are called the neighbors (sometimes move set).
- The assumption is that similar (nearby) solutions have similar costs.

Local Search Application

Motivation

- Optimization problems (gradient descent methods are all local search methods)
- Traveling salesman
- Boolean satisfiability (SAT)
- Scheduling

Hill Climbing (Valley Finding)

Description

- Start at a random state.
- Move to the best neighbor state (one of the successors).
- Stop when all neighbors are worse than the current state.
- The idea is similar to gradient descent.

Simulated Annealing

Description

- Each time, a random neighbor is generated.
- If the neighbor has a lower cost, move to the neighbor.
- If the neighbor has a higher cost, move to the neighbor with a small probability.
- Stop until bored.
- It is a version of Metropolis-Hastings Algorithm.

Genetic Algorithm

Description

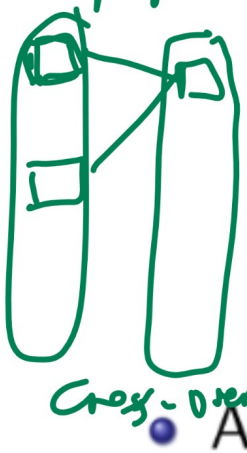
- Start with a fixed population of initial states.
- Find the successors by:

1. Cross over.
2. Mutation.

Reproduction Probability

Definition

- Each state in the population has probability of reproduction proportional to the fitness. Fitness is the opposite of the cost: higher cost means lower fitness. Use F to denote the fitness function, for example, $F(s) = \frac{1}{f(s)}$ is a valid fitness function.



$$p_i = \frac{F(s_i)}{\sum_{j=1}^{N} F(s_j)}, i = 1, 2, ..., N$$

opposite of cost
score
prob of selected as a parent

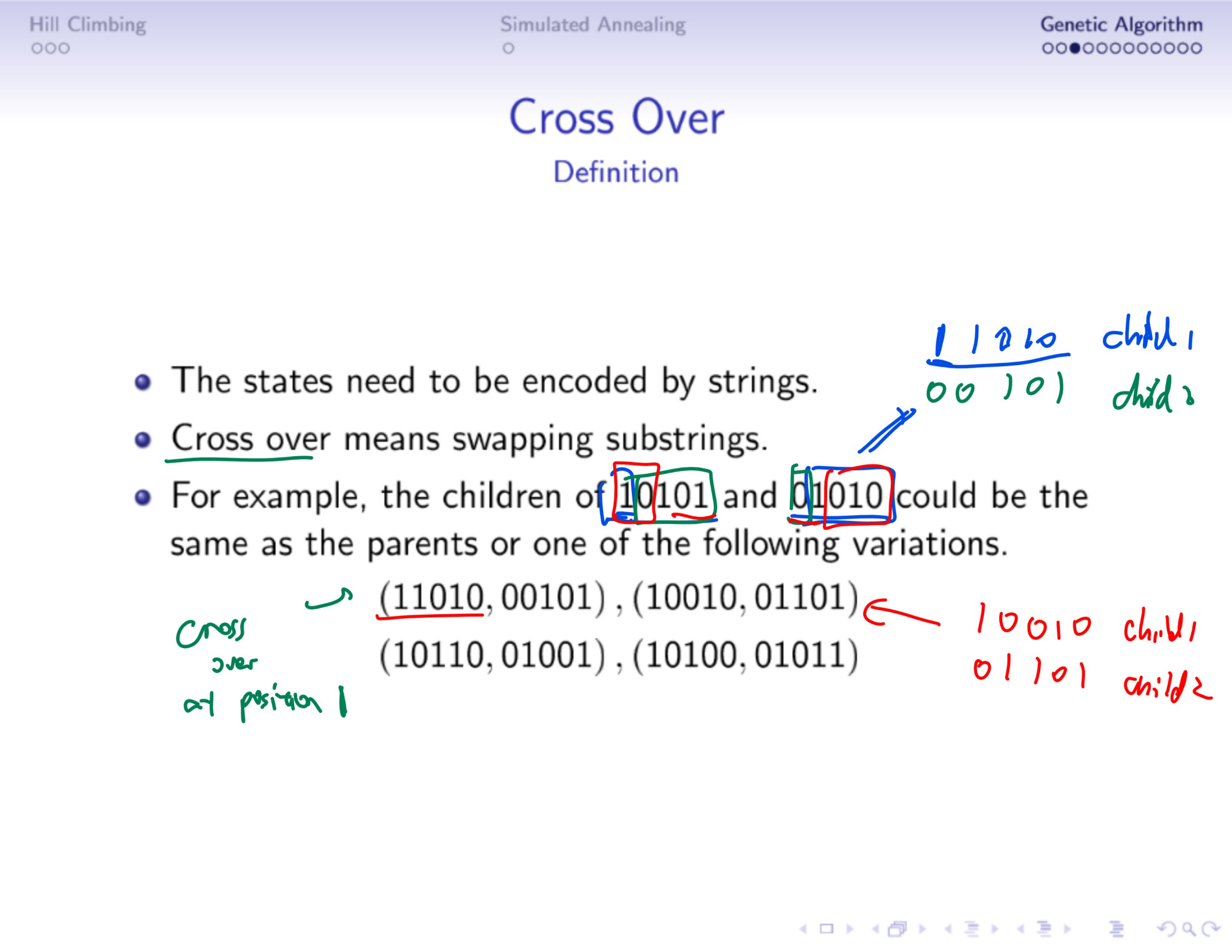
- A pair of states are selected according to the reproduction probabilities (using CDF inversion).

Cross Over

Definition

- The states need to be encoded by strings.
- Cross over means swapping substrings.
- For example, the children of 10101 and 01010 could be the same as the parents or one of the following variations.

$$(11010, 00101), (10010, 01101)$$
$$(10110, 01001), (10100, 01011)$$



Mutation

Definition

- The states need to be encoded by strings.
- Mutation means randomly updating substrings. Each character is changed with small probability q, called the mutation rate.
- For example, the mutated state from 000 could stay the same or be one of the following.
  - one of $001, 010, 100$, with probability $q(1-q)^2$
  - one of $011, 101, 110$, with probability $q^2(1-q)$
  - and 111, with probability q^3

Cross Over, Modifications

Definition

- The previous cross over method is called 1 point cross over.
- It is also possible to divide the string into N parts. The method is called N point cross over.
- It is also possible to choose each character from one of the parents randomly. The method is called uniform cross over.

Mutation, Modifications

Definition

- For specific problems, there are ways other than flipping bits to mutate a state.

1. Two-swap: ABCDE to EBCDA
2. Two-interchange: ABCDE to EDCBA

Travelling salesperson.

Genetic Algorithm TSP Example

Discussion

Fitness Example 1

Quiz

- Which one of the following states have the highest reproduction probability? The fitness function is $f(x) = \min\{t \in \{1,2,3,4,5,6\} : x_t = 1\}$ with $x_6 = 1$.
- $A : (0,0,1,0,0)$ → fitness 3 → $\frac{3}{18}$
- $B : (0,1,0,0,1)$ → 2 → $\frac{2}{18}$
- $C : (0,0,1,1,0)$ → 3 → $\frac{3}{18}$
- $D : (0,0,0,1,0)$ → 4 → $\frac{4}{18}$
- $E : (0,0,0,0,0)$ → 6 → $\frac{6}{18}$

18

cost

score

max

prob picked as parent.

Fitness Example 2

Quiz

- Which one of the following states have the highest reproduction probability? The fitness function is $f(x) = \max\{t \in \{0, 1, 2, 3, 4, 5\} : x_t = 1\}$ with $x_0 = 1$.
- $A : (0, 0, 1, 0, 0)$
- $B : (0, 1, 0, 0, 1)$
- $C : (0, 0, 1, 1, 0)$
- $D : (0, 0, 0, 1, 0)$
- $E : (0, 0, 0, 0, 0)$

Q4

3 → 3/16

5 → 5/16

4 → 4/16

4 → 4/16

0 → 0/16

3
1
2
2
4

ranking

Variations

Discussion

- Parents can survive.
- Use ranking instead of $F(s)$ to compute reproduction probabilities.
- Cross over random bits instead of chunks.

Genetic Algorithm Performance

Discussion

- Use hill-climbing first.
- State design is the most important.
- In theory, cross over is much more efficient than mutation.

Summary

Discussion

- Search:
- Uninformed.
- Informed.
- Local Search: Hill Climbing (Valley Finding): Start at a random state → Move to the best successor → Repeat.
- Local Search: Simulated Annealing: Start at a random state → Generate a random successor → Move if better, Move with small probability if worse → Repeat.
- Local Search: Genetic Algorithm: Start with many random states → Cross-over according to fitness → Mutation → Repeat.
- Adversarial (next time).