CS540 Introduction to Artificial Intelligence

Lecture 19

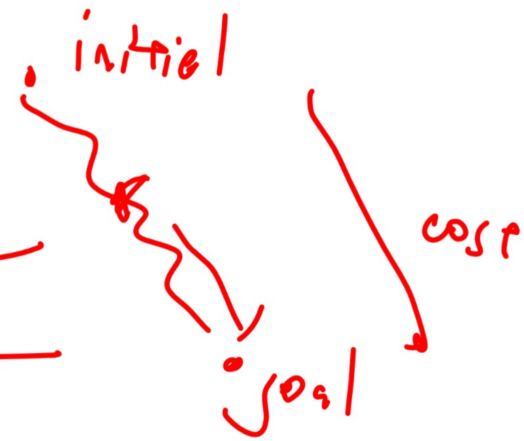
Young Wu

Based on lecture slides by Jerry Zhu, Yingyu Liang, and Charles Dyer

August 3, 2022

Summary

Discussion



- Search:
- Uninformed.
- Informed.
- Local Search: Hill Climbing (Valley Finding).
- Local Search: Simulated Annealing.
- Local Search: Genetic Algorithm.
- Adversarial (next time).

Coordination Game

Admin

- There will be around 10 new questions on the final exam. I will post n of them before the exam (next week):
- $A : n = 0$.
- $B : n = 1$ if more than 50 percent of you choose B.
- $C : n = 2$ if more than 75 percent of you choose C.
- $D : n = 3$ if more than 95 percent of you choose D.
- $E : n = 0$.
- I will repeat this question a second time. If you fail to coordinate both times, I will not post any of the new questions.

Q1

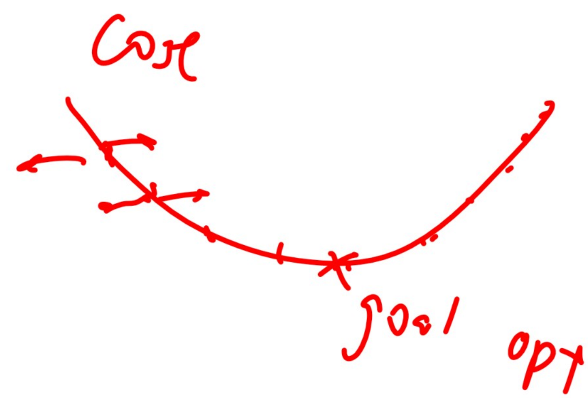
Traveling Salesperson Example

Motivation

Local Search

Motivation

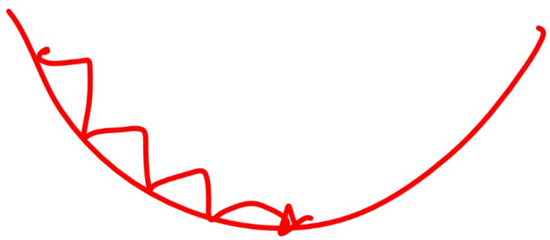
- Local search is about searching through a state space by iteratively improving the cost to find an optimal or near-optimal state.
- The successor states are called the neighbors (sometimes move set).
- The assumption is that similar (nearby) solutions have similar costs.



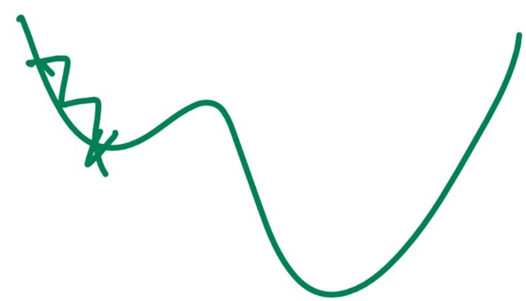
Hill Climbing (Valley Finding)

max min

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.



Boolean Satisfiability Example 1

Quiz

- Assume all variables A, B, C, D, E are set to True. How many of the following clauses are satisfied and which one of the variables should be changed to False to maximize the number of clauses satisfied?
  - $A \vee \neg B \vee C$ → T
  - $\neg A \vee C \vee D$ → T
  - $B \vee D \vee \neg E$ → T
  - $\neg C \vee \neg D \vee \neg E$ → F
  - $\neg A \vee \neg C \vee E$ → T

and ($\wedge$), or ($\vee$) → 4

define successors (neighbors) as flip one variable: T → F, F → T; flip A, B, C, D, E

flip	A	B	C	D	E
score	4	4	5	5	4

max → flip C or D

Boolean Satisfiability Example 2

Quiz

Q2

- Assume all variables A, B, C, D, E are set to True. Which one of the variables should be changed to False to maximize the number of clauses satisfied?
  - $\neg A \vee \neg B \vee \neg E$ F
  - $\neg A \vee \neg B \vee \neg D$ F
  - $\neg A \vee \neg C \vee \neg D$ F
  - $\neg B \vee \neg C \vee \neg D$ F
  - $\neg C \vee \neg D \vee \neg E$ F
  - E : I don't understand. Score = 0

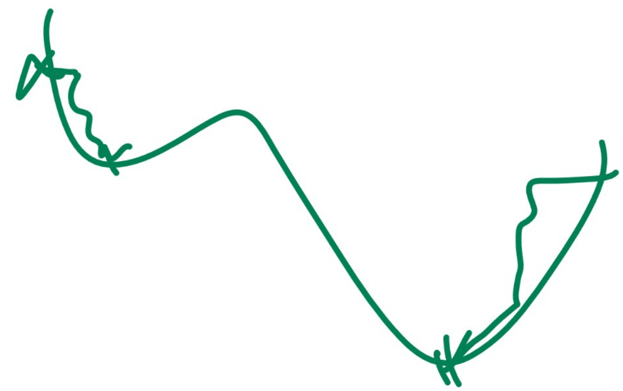
A, B, C score = 3

D score = 4

D

Random Restarts

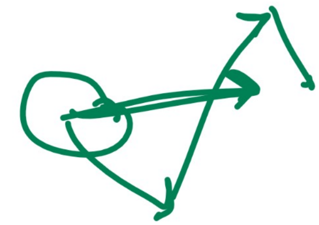
Discussion



- A simple modification is picking random initial states multiple times and finding the best among the local minima.

First Choice Hill Climbing

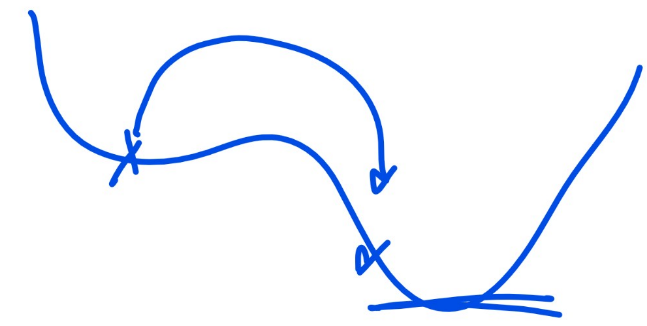
Discussion



- If there are too many neighbors, randomly generate neighbors until a better neighbor is found.
- This method is called first choice hill climbing.

Walk SAT Example

Discussion



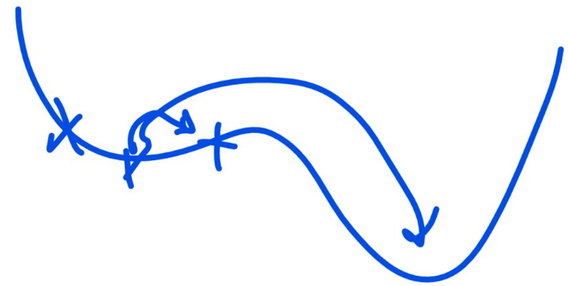
- Pick a random unsatisfied clause.
- Select and flip a variable from that clause:

1. With probability p, pick a random variable.
2. With probability $1 - p$, pick the variable that maximizes the number of satisfied clauses.

- Repeat until the solution is found.
- Walk SAT uses the idea of stochastic hill climbing.

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.

Annealing

Definition

- The annealing process of heated solids.
- Anneal: to subject (glass or metal) to a process of heating and slow cooling to toughen and reduce brittleness.
- Alloys manage to find a near global minimum energy state when heated and then slowly cooled.

Acceptance Probability

Definition

- The probability of moving to a state with a higher cost should be small.

1. Constant: $p = 0.1$
2. Decreases with time: $p = \frac{1}{t}$
3. Decreases with time and as the energy difference increases:

$$p = \exp\left(-\frac{|f(s') - f(s)|}{T(t)}\right)$$

cost of s' - cost of s

- The algorithm corresponding to the third idea is called simulated annealing. The Temperature function $T(t)$ should be a decreasing in time t (iteration number).

Temperature

Definition

- T represents temperature which decreases over time. For example, the temperature can change arithmetically or geometrically.

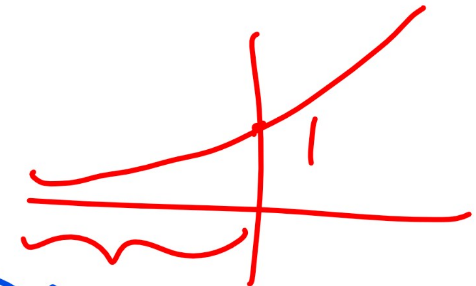
$$T(t+1) = \max\{T(t) - 1, 1\}, T(0) = \text{ large}$$
$$T(t+1) = 0.9T(t), T(0) = \text{ large}$$

- High temperature: almost always accept any s'.
- Low temperature: first choice hill climbing.

Simulated Annealing Example 1

Quiz

SA



- Suppose we are minimizing and $f(s) = 6, f(s') = 5, T = 4.$
  What is the probability we move from s to s' in the next step?
  What is the probability we move from s' to s in the next step?

$6 \to 5$ improving

prob of 1

$$p = e^{-\frac{|f(s') - f(s)|}{T}}$$

move to worse state

$5 \to 6$

$$p = e^{-\frac{1}{4}} = \ldots$$

Simulated Annealing Example 2

Quiz

- Suppose we are minimizing and $f(s) = 0, f(s') = \log(5), T = 1$. What is the probability we move from s to s'.
- A : 0
- B : $\frac{1}{5}$
- C : $\frac{4}{5}$
- D : 1
- E : I don't understand.

Q3

$0 \to \log(s') > 0$

$\log_e$

$$P = e^{-\frac{|f(s) - f(s')|}{T(t)}}$$

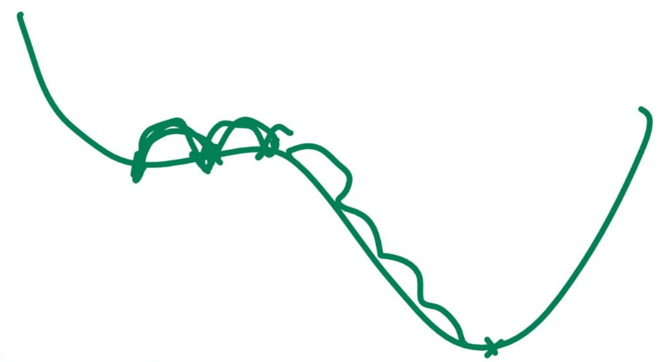
$$= e^{-\frac{\log 5}{1}}$$

$$= e^{-\log 5}$$

$$= e^{\log \frac{1}{5}} = \frac{1}{5}$$

Simulated Annealing Performance

Discussion



random restart

GD

- Use hill-climbing first. SGD → First choice HC.
- Neighborhood design is the most important.
- In theory, with infinitely slow cooling rate, Simulated Annealing finds global minimum with probability 1.