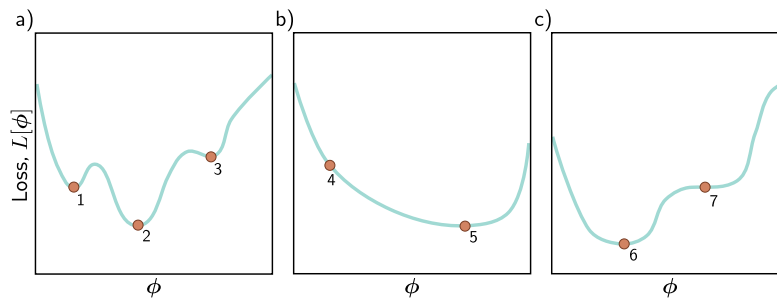


Local Search Study Guide

Artificial Intelligence

1 Performance Surfaces



1. Which of the functions above is/are convex?

Solution: Convex: b

2. Which of the points above is/are a local minimum?

Solution: Local minima: 1, 3, 7

3. Which of the points above is/are a global minimum?

Solution: Global minima: 2, 5, 6

2 Local Search Algorithms

1. Write the basic hill-climbing algorithm.

```
function HILL-CLIMBING(problem) returns a state that is a local maximum
  current ← problem.INITIAL
  while true do
    neighbor ← a highest-valued successor state of current
    if VALUE(neighbor) ≤ VALUE(current) then return current
  current ← neighbor
```

Solution:

2. What is the main weakness of hill-climbing algorithms?

Solution: They get stuck in local minima.

3. How can the basic hill-climbing algorithm be modified to overcome its weaknesses?

- Allow "sideways" moves
- Stochastic hill climbing chooses randomly from uphill moves. Stochastic beam search does this with k states in parallel.
- Random restart hill climbing restarts from multiple initial states.

4. How does simulated annealing avoid getting stuck in local minima?

Solution: Instead of picking the best move, simulated annealing picks a random move. If the move is better than current state, it is always accepted. Otherwise, the algorithm accepts the move with some probability less than 1. The probability decreases exponentially with the "badness" of the move – the amount ΔE by which the evaluation is worsened. The probability also decreases as the "temperature" T goes down: "bad" moves are more likely to be allowed at the start when T is high, and they become more unlikely as T decreases. If the schedule lowers T to 0 slowly enough, then a property of the Boltzmann distribution, $e^{\Delta E/T}$, is that all the probability is concentrated on the global maxima, which the algorithm will find with probability approaching 1.

5. What is (stochastic) beam search?

Solution: Local beam search keeps track of k states rather than just one. It begins with k randomly generated states. At each step, all the successors of all k states are generated. If any one is a goal, the algorithm halts. Otherwise, it selects the k best successors from the complete list and repeats. In this way, information is passed among the parallel search threads.

Local beam search can suffer from lack of diversity – all k states clustered in a small region of the state space. Stochastic beam search fixes this problem by not simply choosing the top k successors, but successors with probability proportional to the successor's value, thus increasing diversity.

6. What does the mixing number parameter, ρ , in the basic algorithm control?

Solution: The number of "parents" used to generate new candidate solutions.

7. What do you have when you set $\rho = 1$ in the basic genetic algorithm?

Solution: When $\rho = 1$, the basic genetic algorithm is equivalent to stochastic beam search.

8. In gradient descent algorithms, what happens if you set the step size/learning rate parameter too high?

Solution: The algorithm might "skip over" the minimum.