Summarization of CS486/686

Wenhu Chen

Lecture 23
Outline

Search Algorithm

Uncertainty Estimation

Decision Theory

Machine Learning
Search Algorithm

Uncertainty Estimation

Decision Theory

Machine Learning
Overview

- How to formulate a search problem?
- What is a search tree?
- What is generic Search algorithm?
- What is DFS and what is BFS?
- What is the space/time complexity of DFS and BFS?
- What is the iterative deepening space complexity?
How to use heuristic search?

- What is LCFS (lowest-cost first)?
- What is GBFS (lowest-heuristic first)?
- What is A* search (combination of two)?
A* Search Algorithm

- Space and Time Complexities.
- Completeness and Optimality.
- Admissible Heuristics $\rightarrow$ Optimality
- Consistent Heuristics $\rightarrow$ Multi-Path Pruning
## Summary of Search Strategies

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Frontier Selection</th>
<th>Halts?</th>
<th>Space</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth-first</td>
<td>Last node added</td>
<td>No</td>
<td>Linear</td>
<td>Exp</td>
</tr>
<tr>
<td>Breadth-first</td>
<td>First node added</td>
<td>Yes</td>
<td>Exp</td>
<td>Exp</td>
</tr>
<tr>
<td>Lowest-cost-first</td>
<td>min $cost(n)$</td>
<td>Yes</td>
<td>Exp</td>
<td>Exp</td>
</tr>
<tr>
<td>Greedy Best-first</td>
<td>min $h(n)$</td>
<td>No</td>
<td>Exp</td>
<td>Exp</td>
</tr>
<tr>
<td>A*</td>
<td>min $cost(n) + h(n)$</td>
<td>Yes</td>
<td>Exp</td>
<td>Exp</td>
</tr>
</tbody>
</table>
Constraint Satisfaction Problem

- Generate-and-Test is way too slow
- Why do We need to model the internal structure of the state?
- What is Backtracking Algorithm?
- What is Arc consistency Algorithm?
- AC-3 Algorithm, using Arc consistency to eliminate Arc
- AC-3 Algorithm complexity
Local Search

- Why do we need local search?
- How do we perform greedy descent?
- How can we avoid local minima?
- What is Simulated Annealing?
Search Algorithm

Uncertainty Estimation

Decision Theory

Machine Learning
Independence

- What is unconditional independence?
- What is conditional independence?
- What is chain rule/product rule/sum rule/bayes rule?
- Universal approach to calculate a probability.
Independence

- Given joint probability distribution, derive the independence step by step.
- Why do we need to use Bayesian Networks?
- How can we compute joint probability over a Bayesian Network?
D-Separation

- What is D-Separation Rule 1?
- What is D-Separation Rule 2?
- What is D-Separation Rule 3?
- How do you apply these D-Separation rules to understand independence between different nodes?
Constructing Bayesian Network

- Pick an order
- Add nodes to the graph
- Pick the minimum subset as parents
- Form a Bayesian Network
Variable Elimination Algorithm

- Define Factors
- Restrict Factors to reflect Evidence
- Multiply factors with shared variables
- Sum out hidden variables
- Normalize to obtain probability
Hidden Markov Model (Smoothing)

- Derive the forward recursion \( P(S_k \mid o_{0:k}) \)
- Compute the forward recursion \( P(S_k \mid o_{0:k}) \)
- Derive the backward recursion \( P(o_{k+1:t-1} \mid S_k) \)
- Compute the backward recursion \( P(o_{k+1:t-1} \mid S_k) \)
- Combine forward and backward recursion to calculate smoothing probability
Viterbi Algorithm

- Dynamic Programming to reuse intermediate variables
- Backtracking to derive the most likely hidden variables
Search Algorithm

Uncertainty Estimation

Decision Theory

Machine Learning
Decision Network

- Understand what is utility.
- Compute expected utility
- Apply Variable Elimination Algorithm in Decision Networks
Markov Decision Process

- What is discounted reward?
- What is policy and value function?
- Bellman equation to update value function.
- Policy iteration to update policy function.
Reinforcement Learning

- How to implement passive ADP?
- How to implement Active ADP?
- How to do Q-Learning?
- What is the difference between ADP and Q-Learning?
- How fast are ADP and Q-Learning?
Supervised Learning

- Classification vs. Regression
- Cross-Validation
- How to avoid Over-fitting?
- Trade-offs between bias and variance.
Unsupervised Learning

- How to do K-means clustering?
- What is Principled component Analysis?
- What is auto-encoder?
Decision Tree

- How to compute entropy?
- How to grow a full tree?
- How to determine the order of testing features?
Neural Networks

▶ What is activation function and what qualifies as activation functions?
▶ How to perform back-propagation?
▶ How to perform gradient descent?
▶ What are the existing optimizers?
Exam

- Non-programmable calculator
- Two-sided A4 Study Note
- A total of 8 problems
- Tentative: 84 marks, 76+ marks count as 100%
- Tentative: Pass Threshold: 42 marks