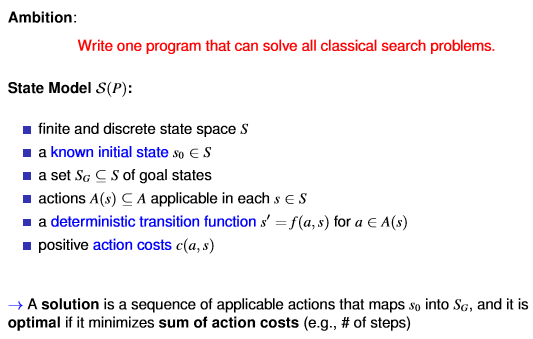
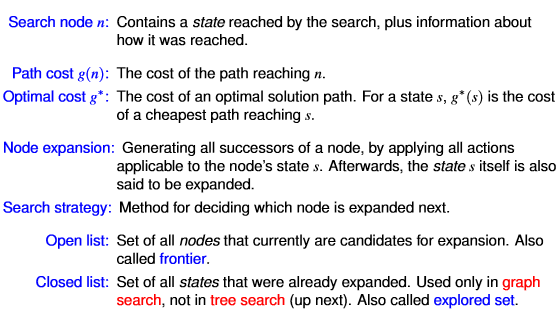
1 Search Algorithms

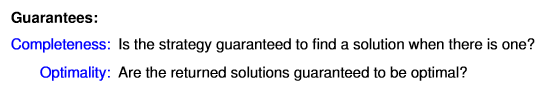
Basic State Model: Classical Planning

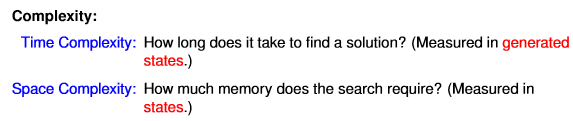


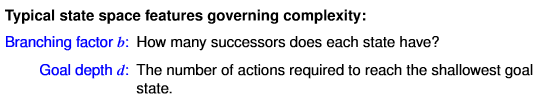
Search Terminology



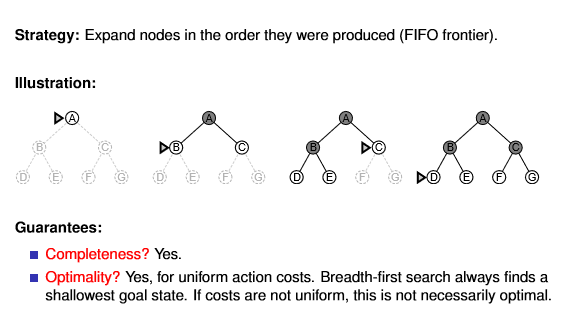
Criteria for Evaluating Search Strategies

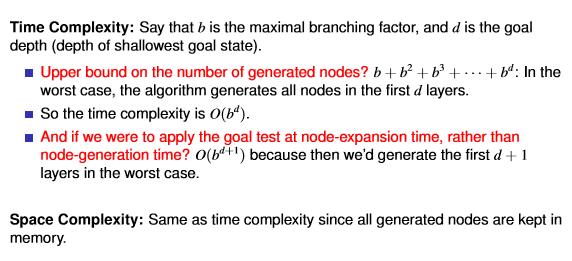




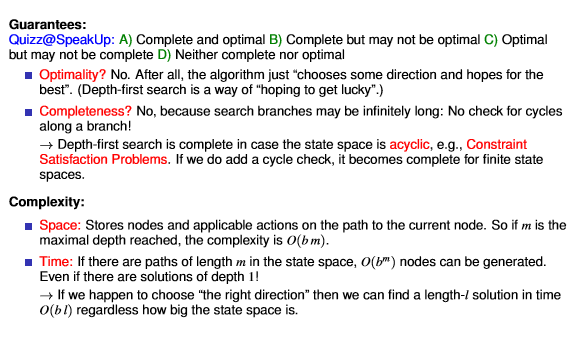
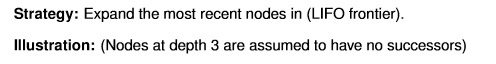


Breadth-First Search

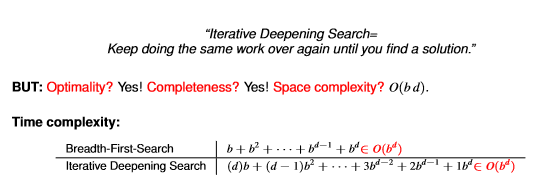




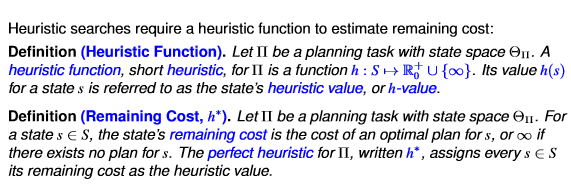
Depth-First Search



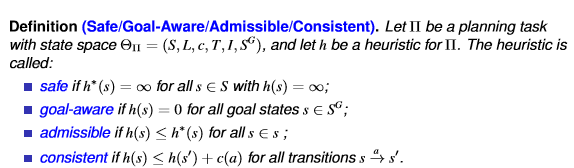
Iterative Deepening Search



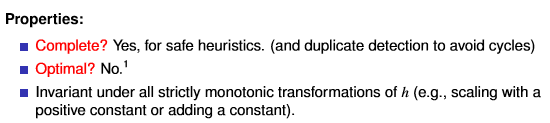
Heuristic Functions



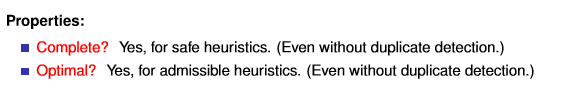
Properties of Heuristic Functions



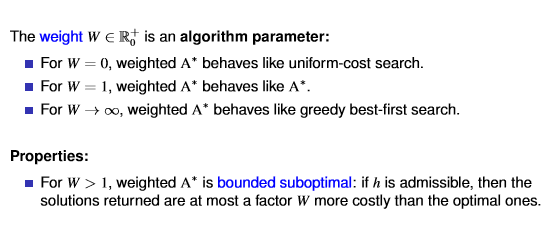
Greedy Best-First Search



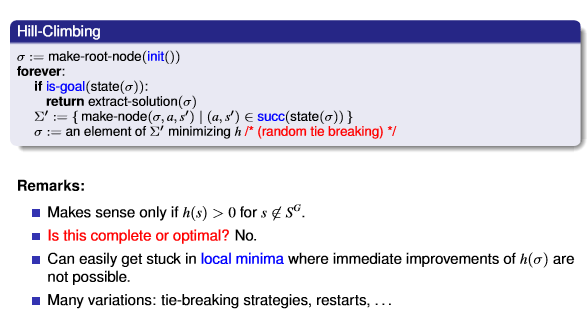
A∗



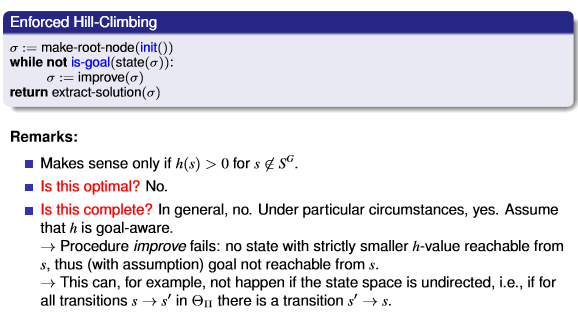
Weighted A∗



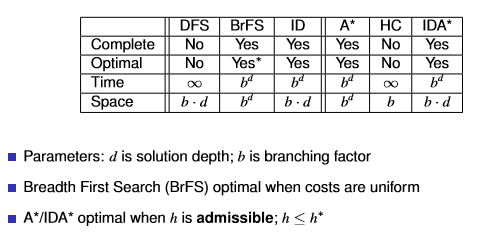
Hill-Climbing



Enforced Hill-Climbing



Properties of Search Algorithms



2 Introduction to Planning

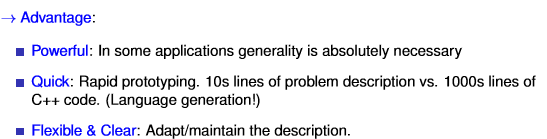
Programming-Based Approach



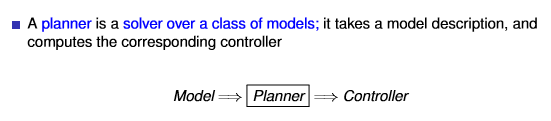
Learning-Based Approach



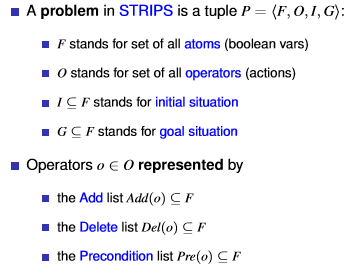
Model-Based Approach



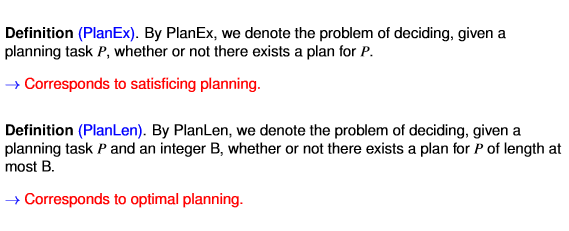
Models, Languages, and Solvers



Strips

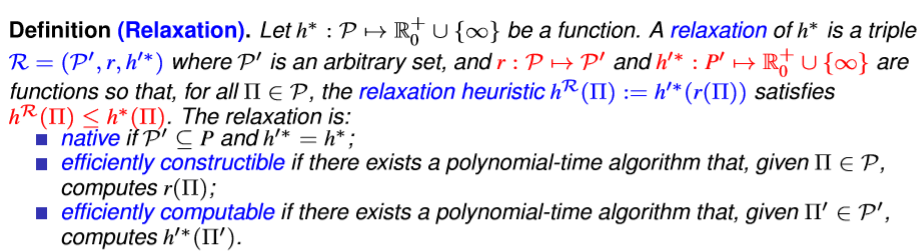


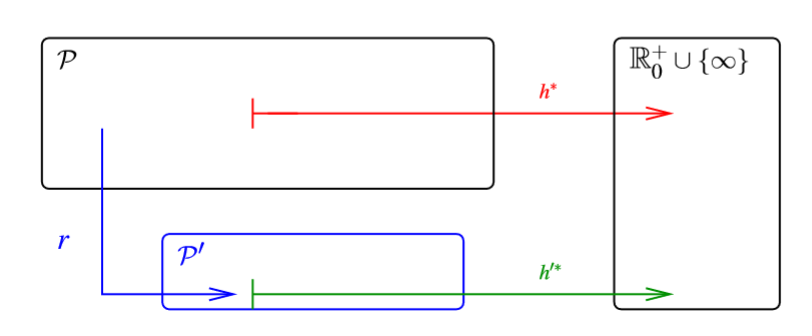
Decision Problems in Planning



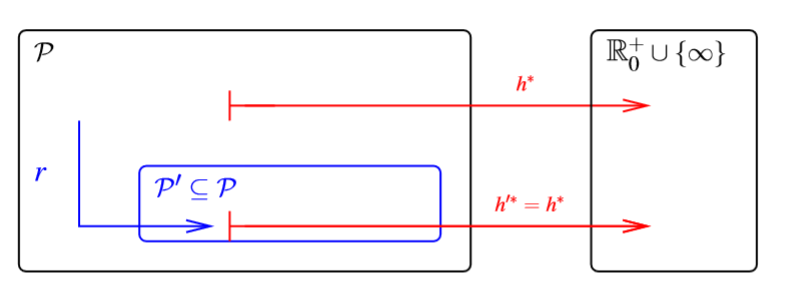
3 Generating Heuristic Functions

Relaxations

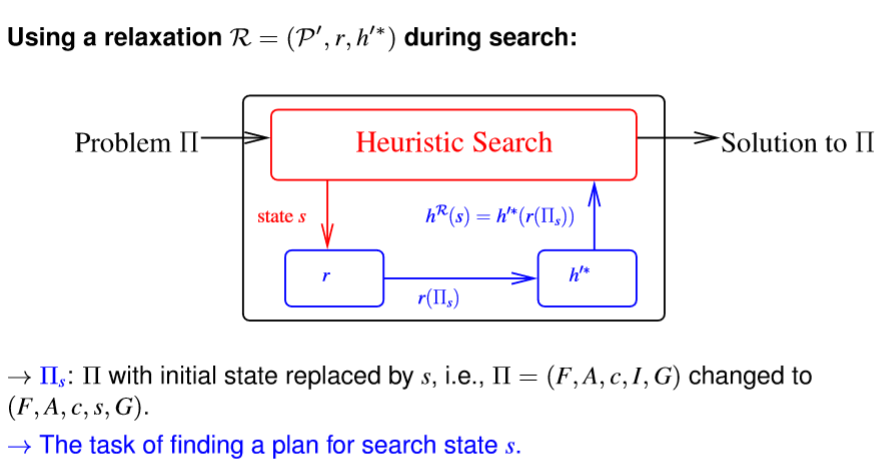




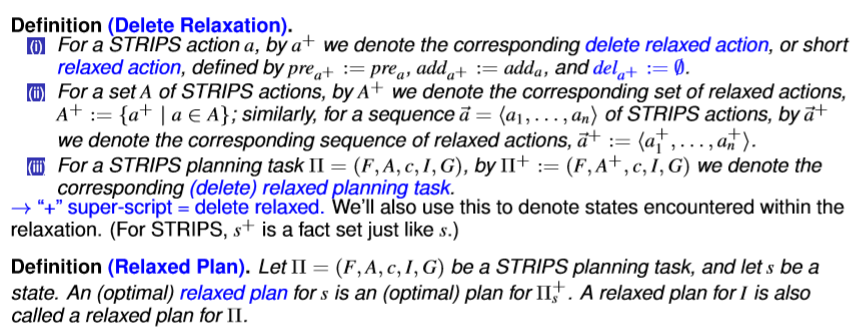
Native Relaxations



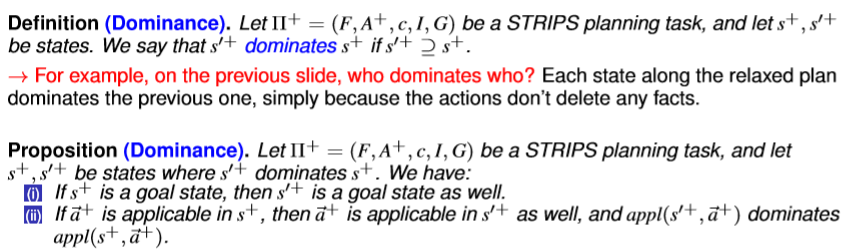
How to Relax During Search: Diagram



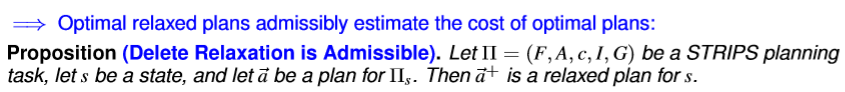
The Delete Relaxation



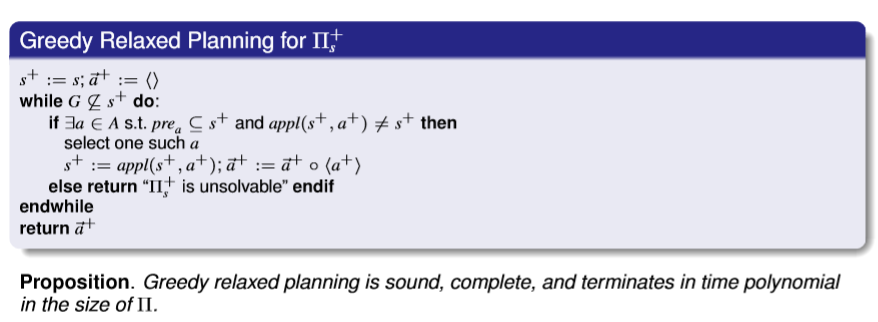
State Dominance

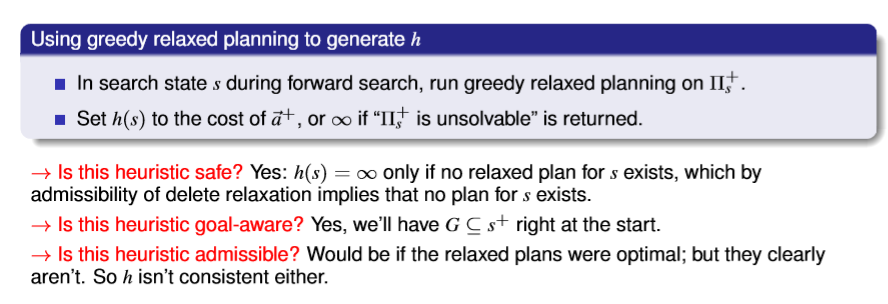




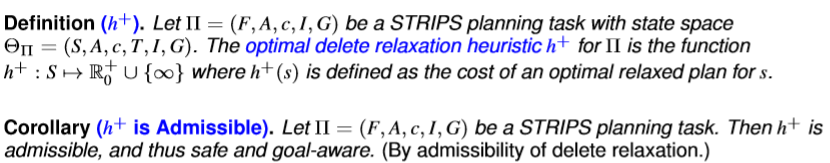


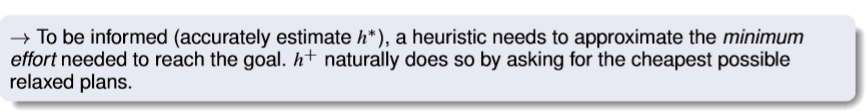
Greedy Relaxed Planning

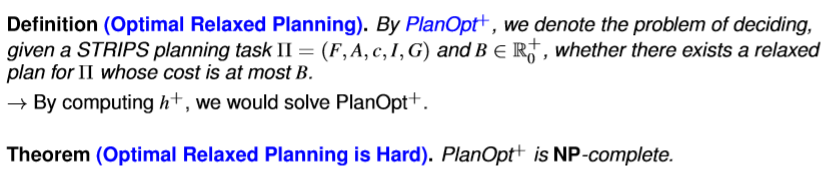




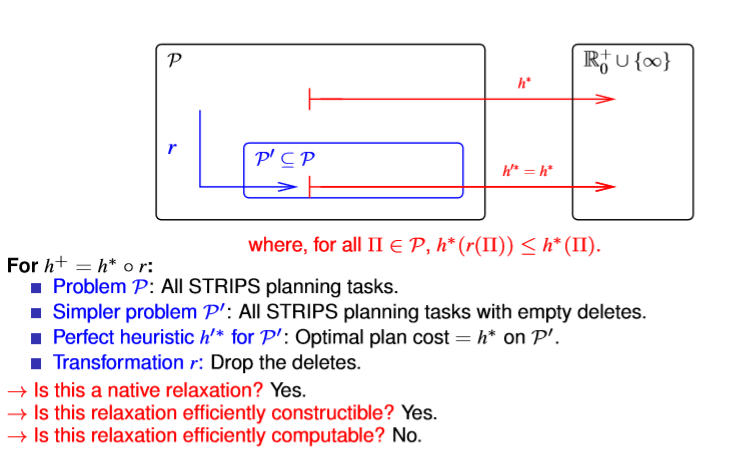
h+: The Optimal Delete Relaxation Heuristic



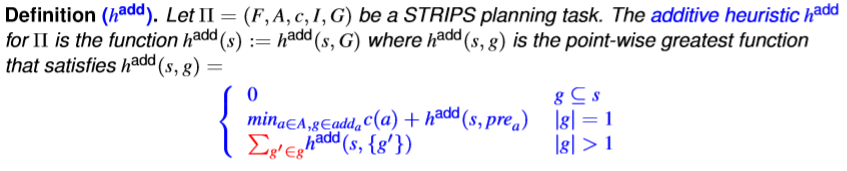


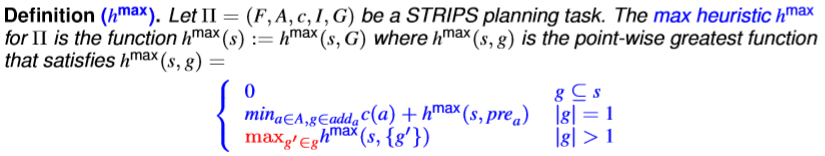


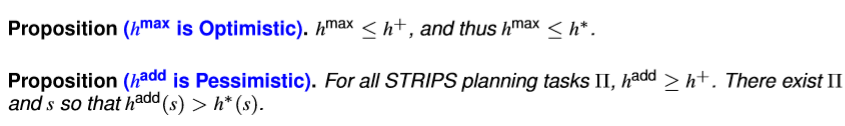
h+ as a Relaxation Heuristic



The Additive and Max Heuristics

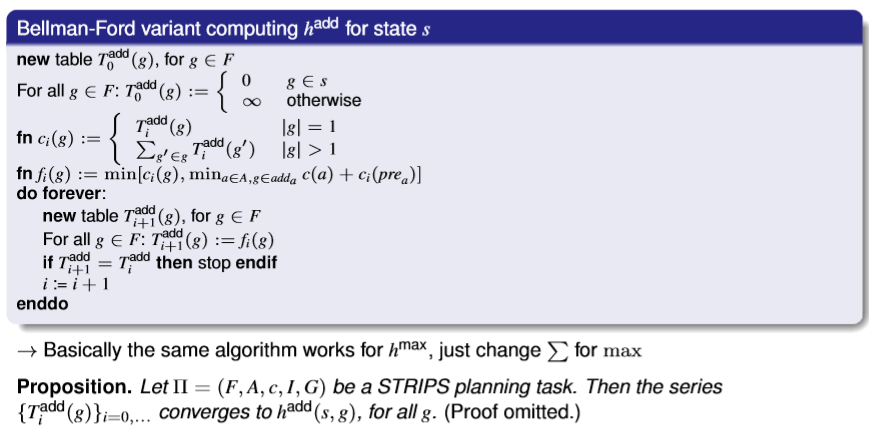


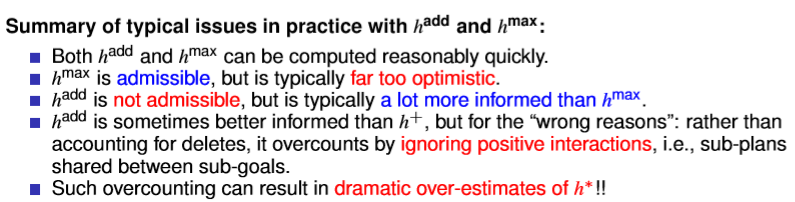




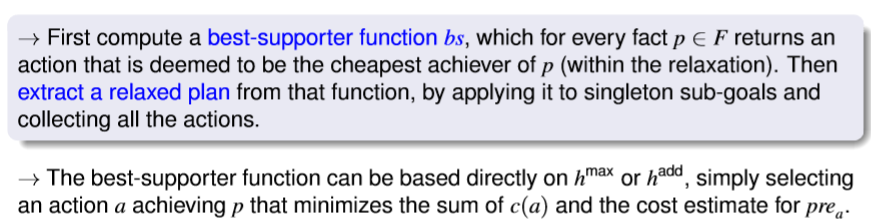


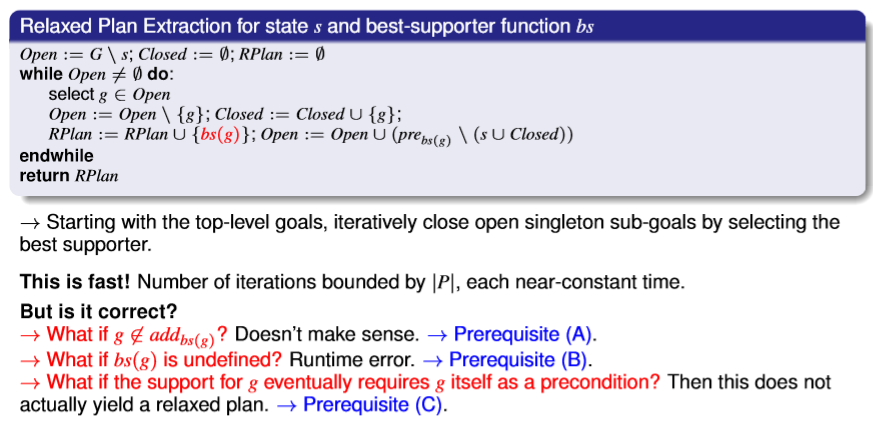
Bellman-Ford for hmax and hadd

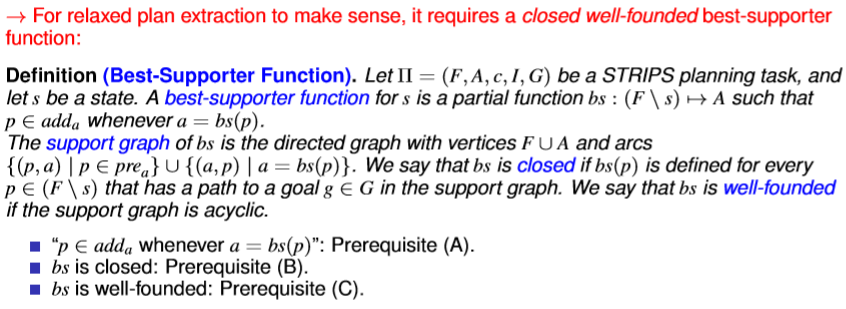


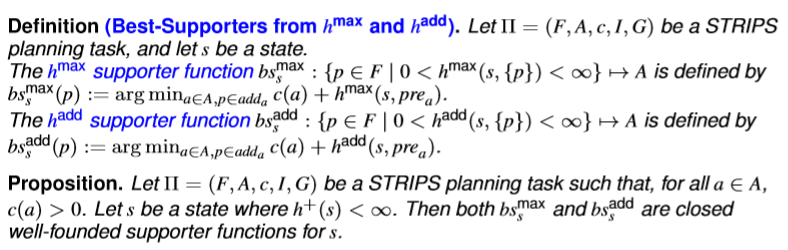


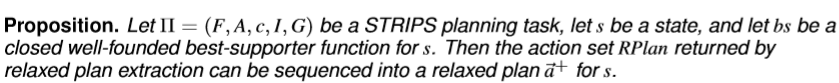
Relaxed Plan Extraction

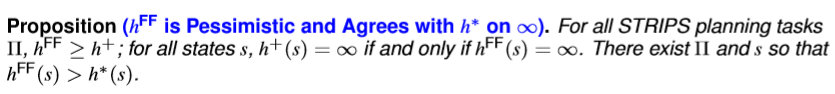




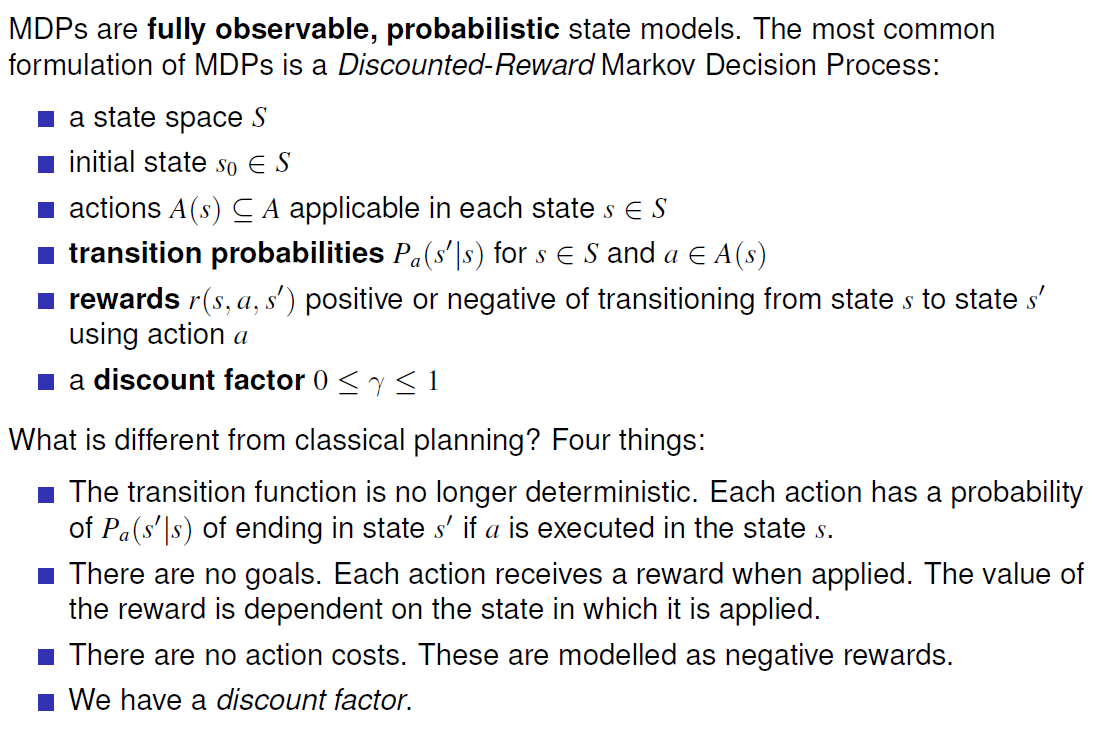




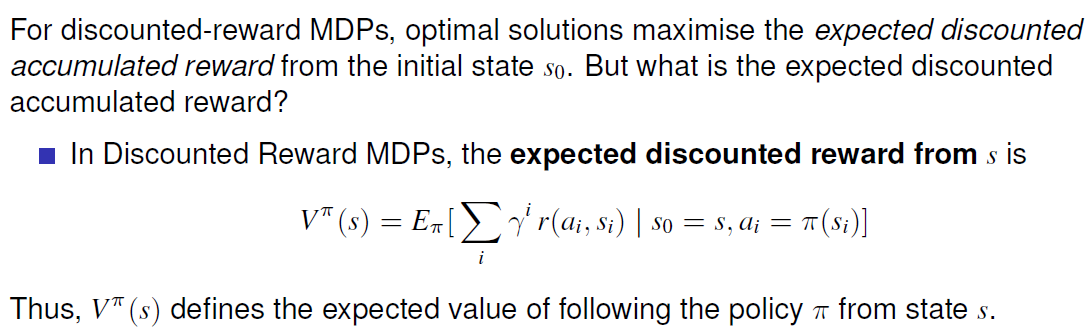




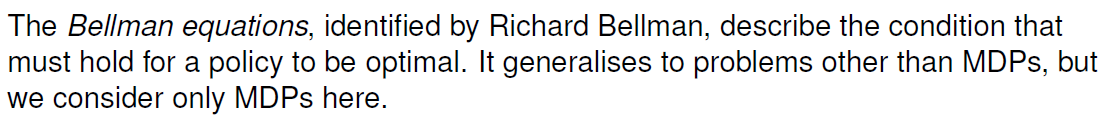
4 Markov Decision Processes (MDPs)

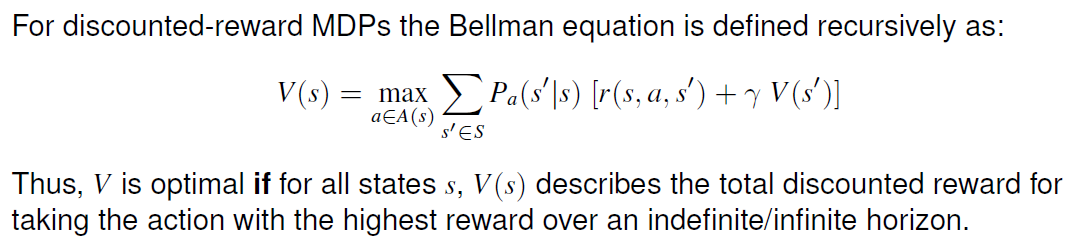


Expected Value of Policy for Discounted Reward MDPs

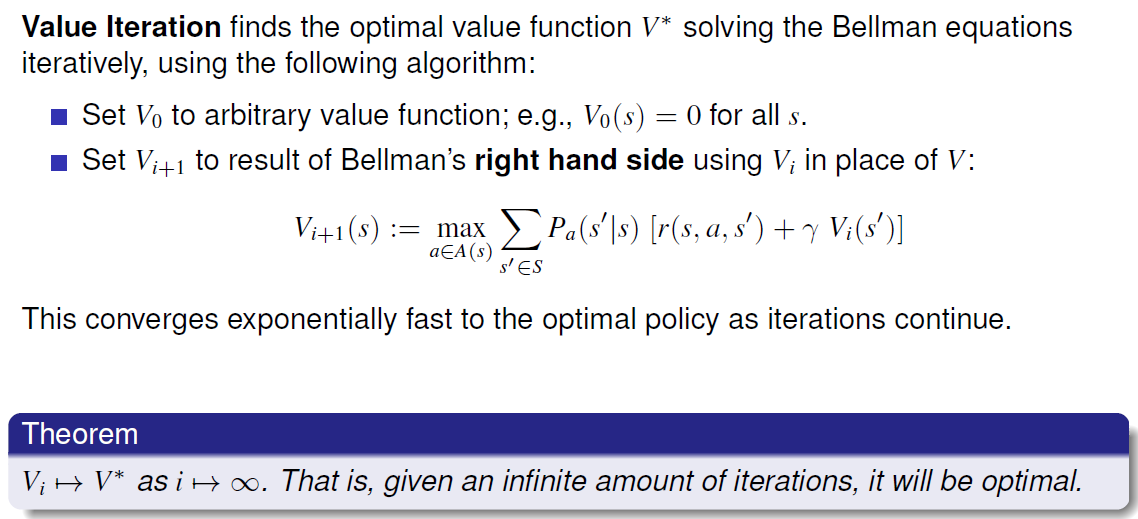


Bellman equations (Discounted-Reward MDPs)

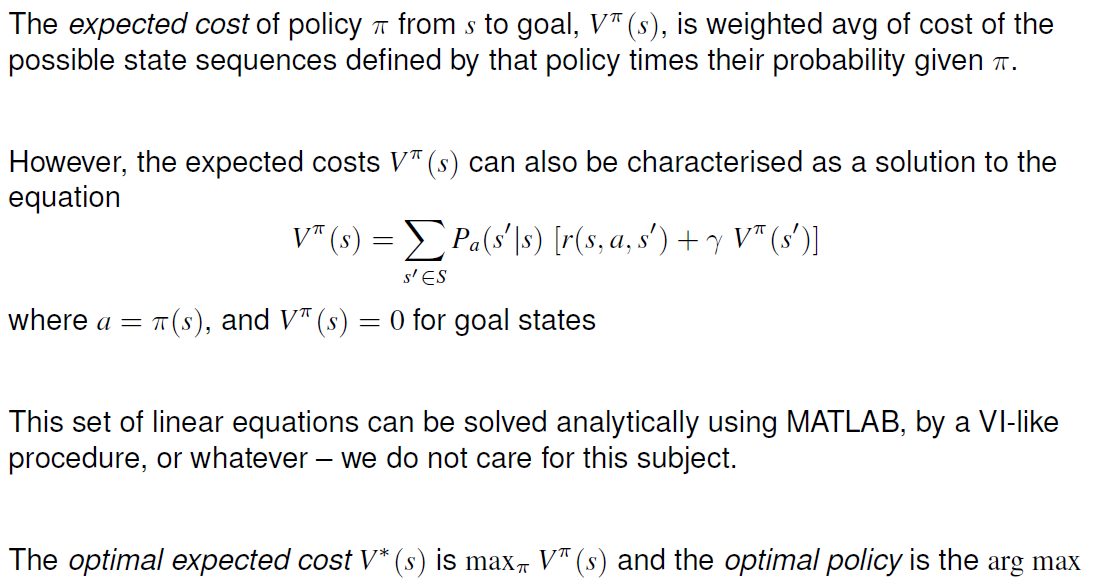




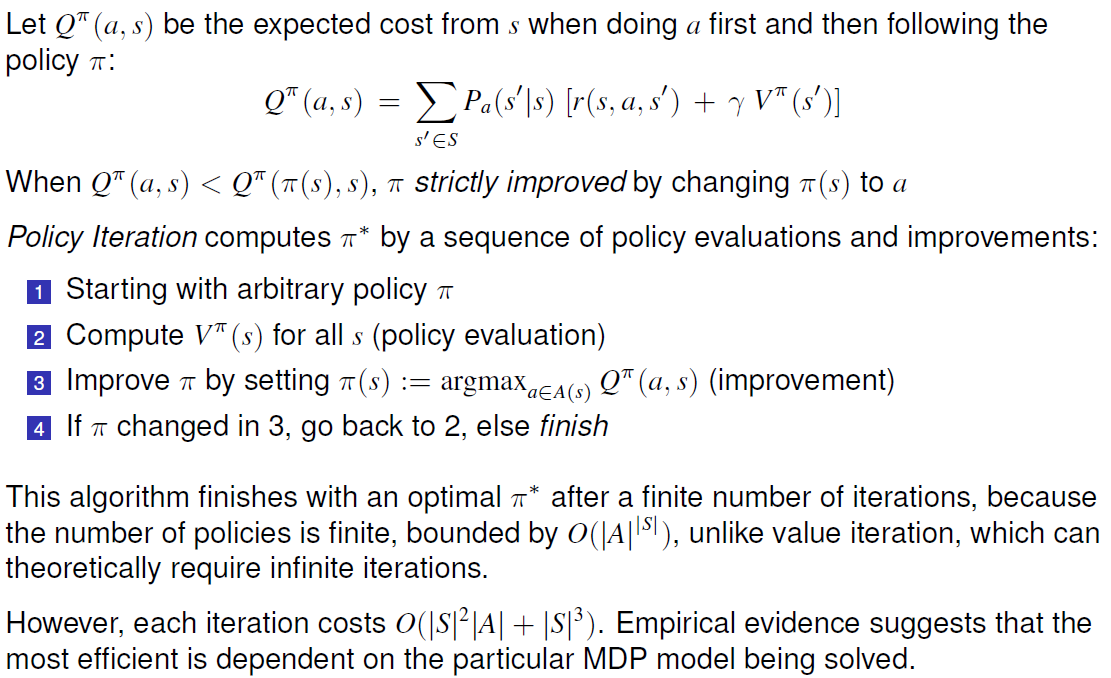
Value Iteration



Policy evaluation

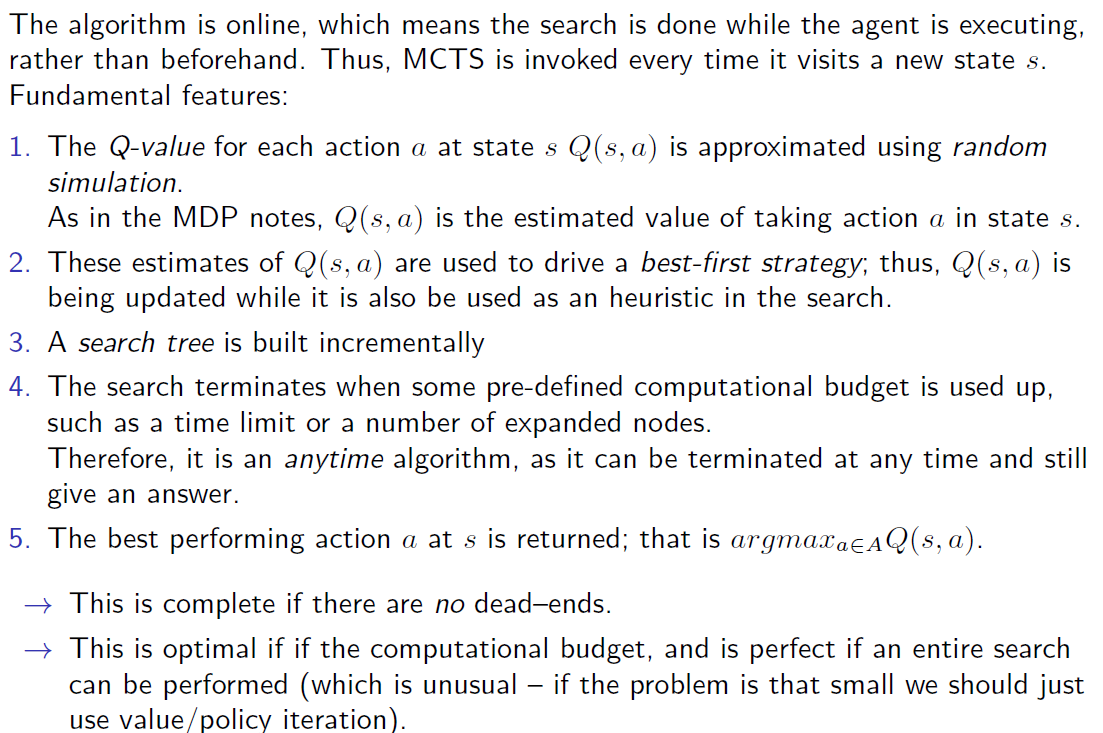


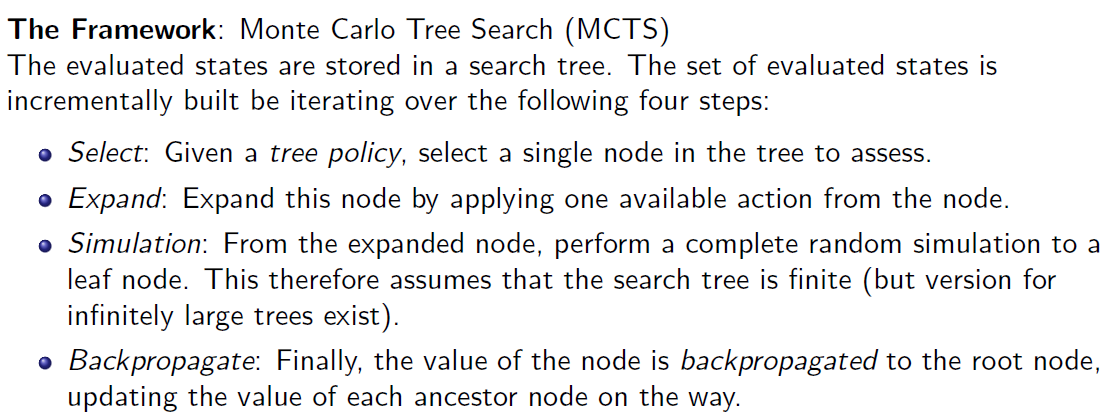
Policy Iteration



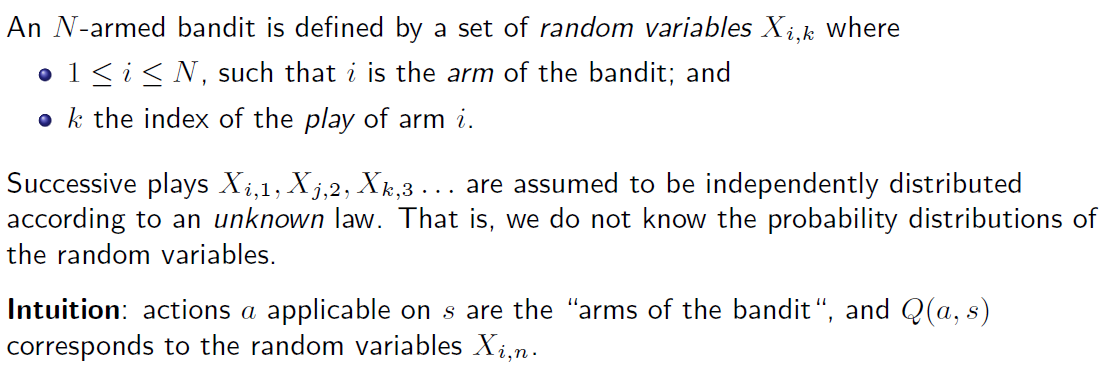
5 Monte Carlo Tree Search

Monte Carlo Tree Search

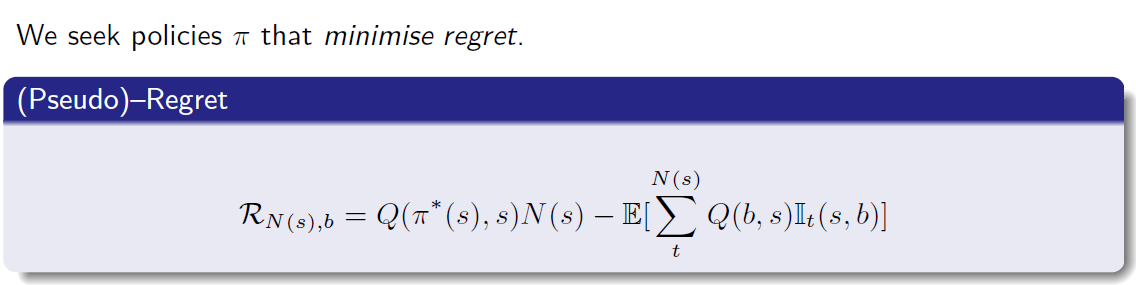


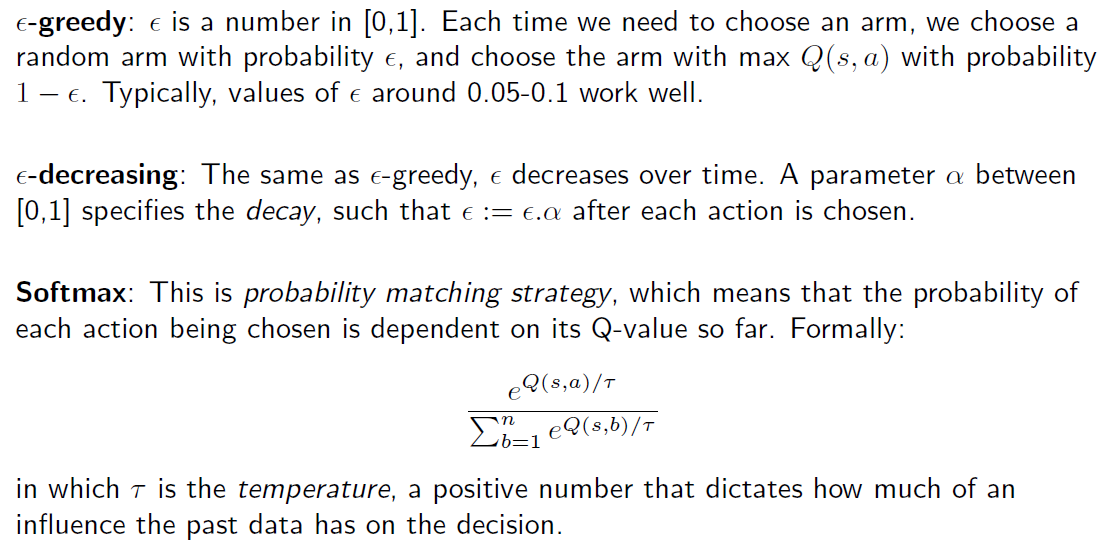


Multi-Armed Bandit

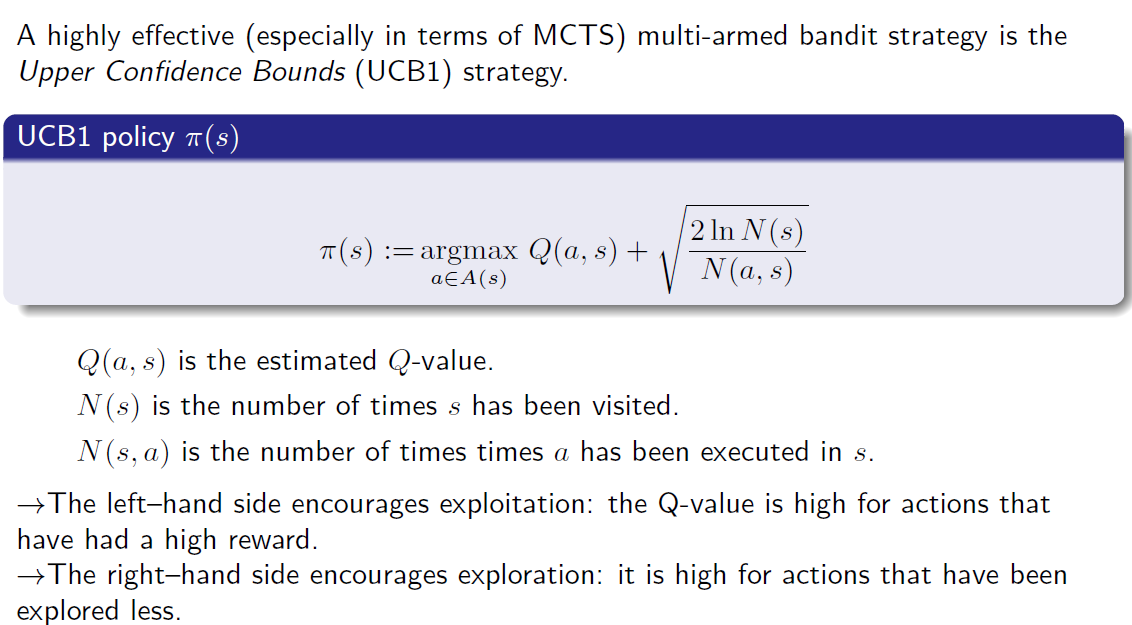


Exploration vs. Exploitation

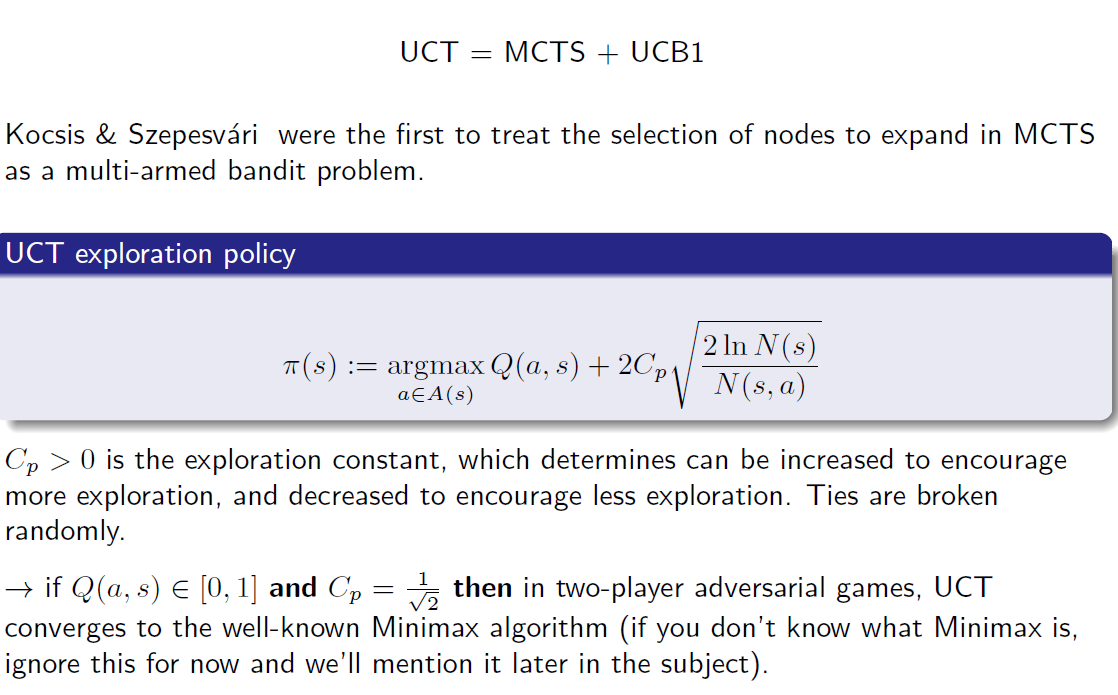




Upper Condence Bounds (UCB1)

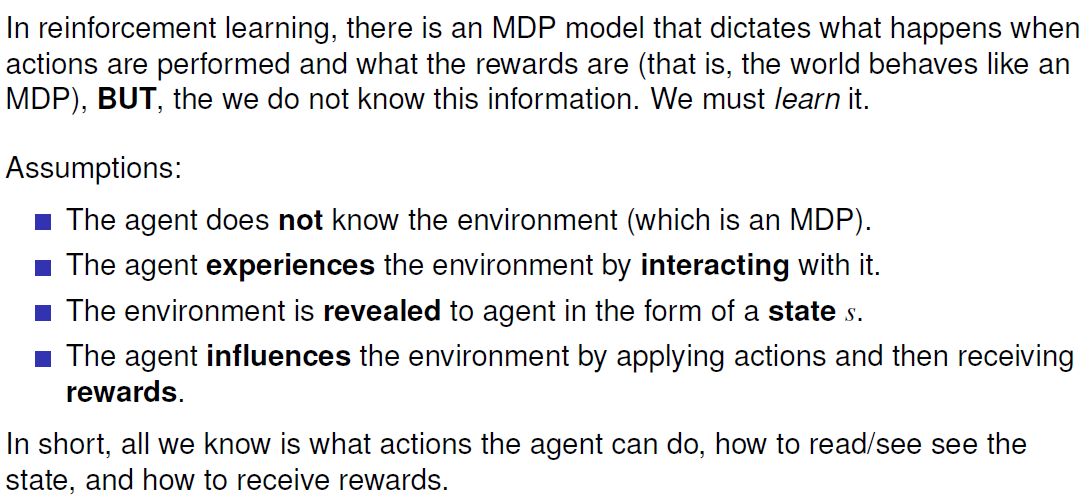


Upper Condence Trees (UCT)

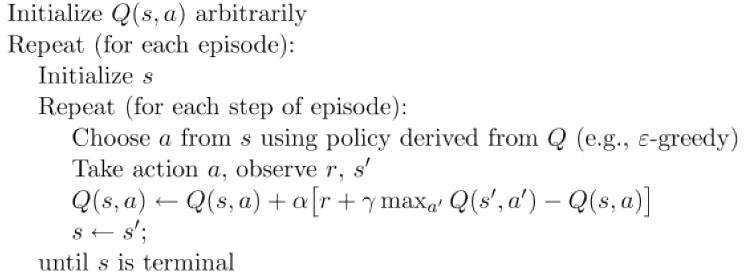


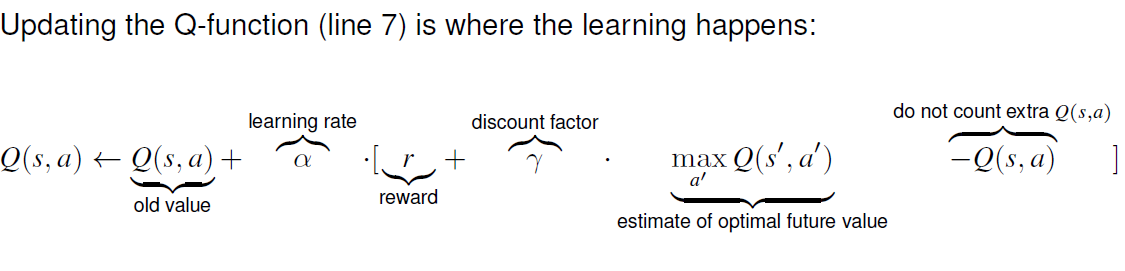
6 Reinforcement Learning

Reinforcement Learning

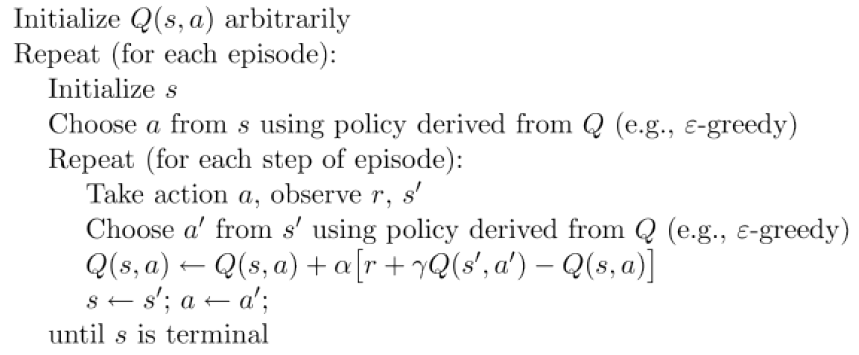


Q-Learning





SARSA: On-Policy Reinforcement Learning



Q-learning vs. SARSA



