Heuristic Machine Decision Making

By Eric Bright

Overview

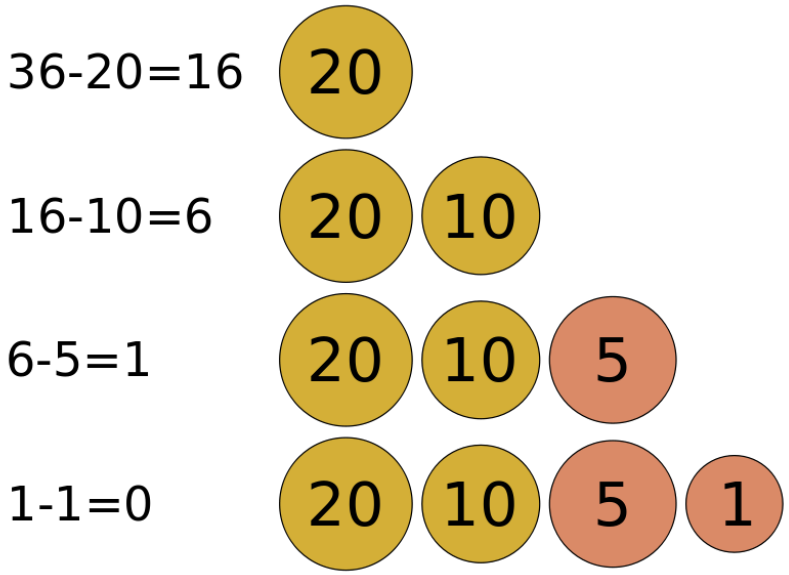
- Definitions
- Examples
  - Greedy Algorithm
  - Travelling Salesman Problem
- Real-world application: Heuristic analysis

Definition

- **Heuristic** (/hjʊˈrɪstɪk/; Greek: "Εὑρίσκω", "**find**" or "**discover**") refers to experience-based techniques for problem solving, learning, and discovery that give a solution which is not guaranteed to be optimal – Wikipedia
- involving or serving as an aid to learning, discovery, or problem-solving by experimental and especially trial-and-error methods – Merriam-Webster
- Specifically, in Computer Science: a technique designed for solving a problem more quickly when classic methods are too slow, or for finding an approximate solution when classic methods fail to find any exact solution
- Trades optimality, completeness, accuracy, or precision for speed
  - Optimality: Many solutions may exist for a problem, some better than others. Is the "best" solution necessary?
  - Completeness: Several solutions may exist for a problem, but do we need to find them all?
  - Accuracy and precision: How large of a margin of error is acceptable in our results?

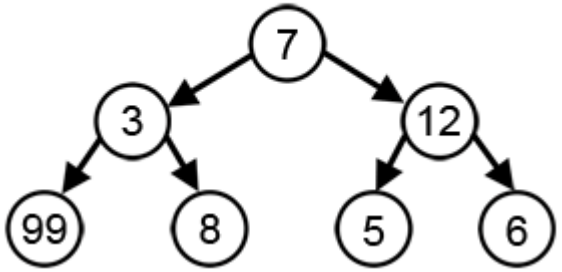
Example: Greedy algorithm

- Makes the locally optimal choice at each stage in hope of finding a global optimum.
- Making change – determine the minimum number of coins to give while making change using 20, 10, 5, and 1 cent coins.
  - The coin of the highest value, less than the remaining change owed, is the local optimum.

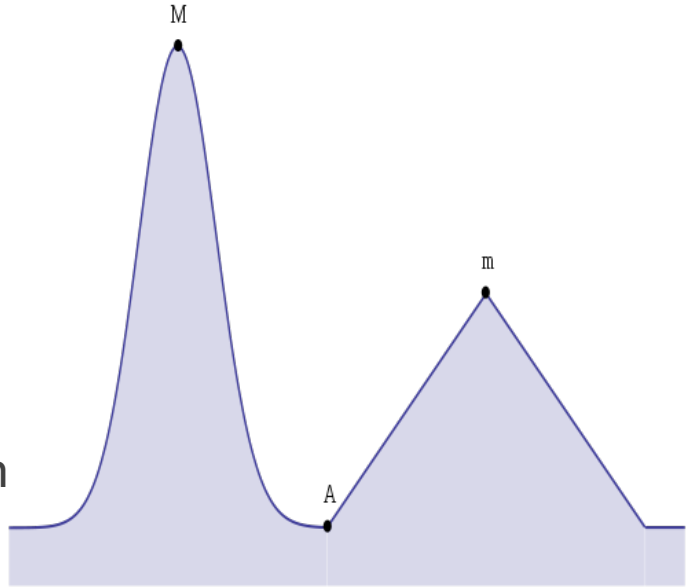


Example: Greedy algorithm

- Goal of reaching the largest sum – does not produce an optimal solution in this cas



- Dependent on starting point. Starting at “A”, a greedy algorithm will find the local maximum “m”, not the global maximum “M”.



Example: Travelling Salesman Problem

- Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?
- Problem was first formulated in 1930 and has been intensively studied
- It is a computationally difficult problem that is used as a benchmark for many optimization methods.
- Even with just 20 cities, the time to solve the problem becomes unreasonable
- Heuristics allow approximate results in a much more reasonable amount of time
  - Nearest Neighbor: the salesman starts at a random city and repeatedly visits the nearest city until all have been visited – within 25% of an optimal solution
  - Greedy – within 15-20% of an optimal solution

Application: Heuristic analysis

- Many anti-virus applications use heuristic analysis to detect new computer viruses (called Zero-day viruses)
- Also detects new variants of viruses already in the wild
- Questionable program or script is executed inside a specialized virtual machine, allowing the anti-virus program to simulate what would happen
- Commands performed by the program are monitored for common viral activities
- Advantages: Easy implementation and high performance
- Disadvantages
  - Detection rate of viruses that use completely never-before-seen exploits is low
  - False positive rate is high

Review

- Heuristics are techniques that trade optimality, completeness, accuracy, or precision for speed
- Greedy algorithm and Nearest neighbor are two well known Heuristic algorithms
- Travelling Salesman Problem is a computationally complex problem that can be approximated using heuristic algorithms
- Heuristic algorithms can be applied to real-world situations, such as anti-virus programs

References

- "Algorithms and Heuristics." *Algorithms and Heuristics*. Indiana University, n.d. Web. 24 Oct. 2013.
- "Heuristic Analysis in Kaspersky Anti-Virus 2013." *Kaspersky Anti-Virus 2013*. Kaspersky Lab, 01 Mar. 2013. Web. 22 Oct. 2013.
- Nilsson, Christian. *Heuristics for the Traveling Salesman Problem*. Linköping University, n.d. Web.