DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

TU Delft

Exam CSE3300: Algorithms for NP-hard Problems

7 April 2022 | 13:30–16:30

- Start each question on a new page.
- This is a closed-book individual examination with 7 questions worth 33 points in total.
- If your score is n points, then your grade for this exam will be $1 + \frac{9}{33}n$.
- Use of materials including but not limited to course books, readers, notes, slides, etc is not permitted.
- Use of mobile computing devices (including mobile phones) is not permitted.
- By submitting your answers this exam, you agree to the following Code of Honour pledge: "I promise that I have not used unauthorized help from people or other sources for completing my exam. I created the submitted answers all by myself during the time slot that was allocated."

  Name: ____________________________ Signature: ____________________________

- Write clearly, use correct English, and avoid verbose explanations. Giving irrelevant information may lead to a reduction in your score. *Almost all question parts can be answered in a few lines!*
- This exam covers all information on the slides of the course, and everything discussed in lectures.
- This exam assumes a familiarity with the stated background of the course.
- The total number of pages of this exam is 6 (excluding this front page).
- Exam prepared by E. Demirović, A. Lukina, N. Yorke-Smith.

1. (a) (1 point) How many strings of length n can you construct using only three different characters, under the constraint that the first and last character in the string must be the same? Assume that $n \geq 3$.

Solution: $3 \cdot 3^{n-2} = 3^{n-1}$

2. (a) (2 points) Describe the propagation algorithm for the constraint $\sum w_i \cdot x_i \leq K$, where the weights $w_i \in \mathbf{N}$ are integer constants and the variables $x_i \in \{0, 1\}$ are binary.

Solution: Several equivalent answers possible, e.g., let the $s = K - \sum_i w_i \cdot LB(x_i)$ [1], then we have that $s + w_i > K \rightarrow x_i = 0$ [1].

(b) (2 points) Discuss two different ways of representing the domain of a discrete variable. Make sure to indicate the advantage and disadvantage of the two approaches relative to each other.

Solution: Many answers possible. One such answer is to state that we can represent the domain as a set of integers, and the other as two integers denotes the lower and upper bounds of the variable [1]. The latter is more compact but with the former more inference is typically possible [1]. wtf

3. (a) (2 points) State one strength and one weakness of MCTS.

Solution: [any 2 points]

We create a program to play a card game called *Texas Hold 'Em poker*. Our program involves search and a heuristic for state evaluation, and other components.

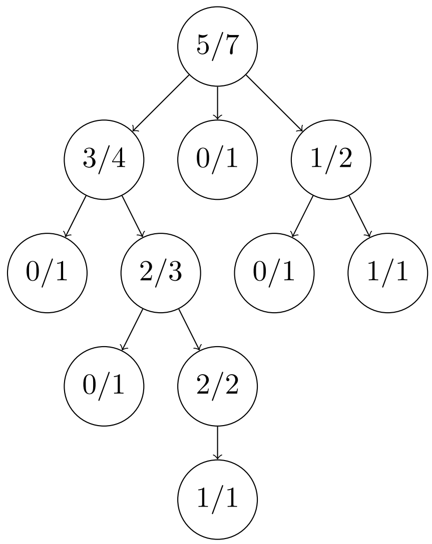
(b) (2 points) What are some pros and cons of using a neural network for the state evaluation function?

Solution: [any 2 points]

(c) (2 points) Suppose we use plain MCTS for the search (i.e., no improvements like UCT). Consider the below search tree in Fig. 1. In each node, the numbers a/b indicate a successful paths out of b. Why does MCTS selection choose the highlighted leaf node?

Solution: Consider the moves with winning possibility 2/3, 0/1 & 1/2 after the first move 4/6: the node 2/3 has the highest possibility of winning. From node 2/3, its child with highest probability is the node 1/2. From node 1/2, its child with highest probability is the node 1/1.

(d) (4 points) Suppose the highlighted node is selected and expanded, and, after roll-outs, evaluation finds a score of $1/1$. Back-propagation then occurs. Draw the final tree that results.



Solution:
[-1 for each mistake]

(e) (2 points) Briefly discuss one ethical issue with respect to a poker-playing program.

Solution: [any 1 point with discussion, or any 2 points]

4. (6 points) Let us approximately compute the integral

$$I = \int_0^{\pi/2} \sin x \, dx.$$

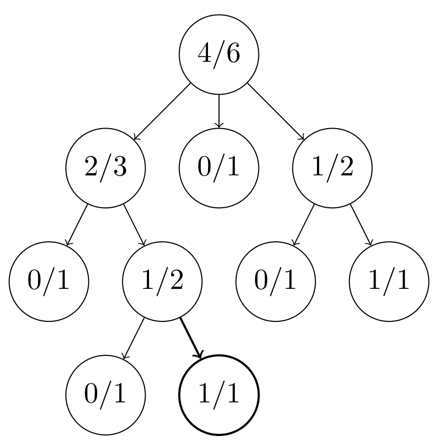


Figure 1: MCTS tree

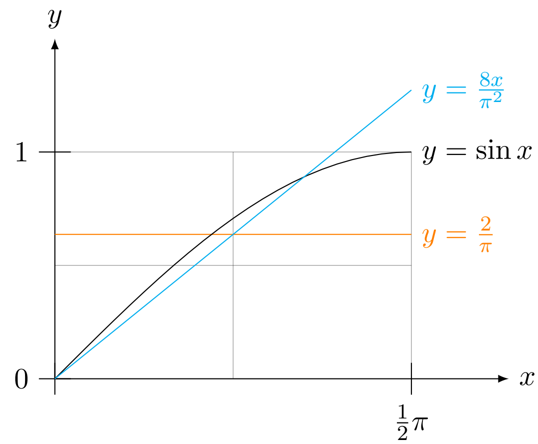


Figure 2: The integrand $y = \sin x$ and two densities.

The exact value of this integral is known:

$$I = \int_0^{\pi/2} \sin x \, dx = [-\cos x]_0^{\pi/2} = 1.$$

don't know how to do this at all

Consider two different random variables ξ for calculating this integral using importance sampling (see Fig. 2): one with constant density $p(x) \equiv 2/\pi$ (i.e., uniform distribution in the interval $0 < x < \pi/2$), and the other with linear density $p(x) = 8x/\pi^2$. Describe modelling of ξ in each case and discuss their suitability for importance sampling.

Solution: It is evident that the linear density agrees better with the recommendation of importance sampling, that it is desirable for $p(x)$ to be proportional to $\sin x$. Thus one may expect that the second approach will yield the better result.

Let $p(x) = 2/\pi$ for all x in $(0, \pi/2)$. The formula for modelling ξ can be obtained for $a = 0$ and $b = n/2$.

The random variable ξ is said to be uniformly distributed over the interval (a, b) if its density is constant over this interval:

$$p(x) = 1/(b - a) \ \forall a < x < b.$$

To model ξ, we need to compute integral

$$\int_a^{\xi} p(x)dx = \int_a^{\xi} dx/(b - a) = \gamma$$

This is easily computed as $(\xi - a)/(b - a) = \gamma$. Hence, we arrive at $\xi = a + \gamma(b - a)$. Then $\xi = \pi/2\gamma$.

$$I \approx \frac{\pi}{2N} \sum_{j=1}^{N} \sin \xi_i.$$

On the other hand, let $p(x) = 8x/\pi^2$. For the modelling of ξ we use the same approach:

$$\int_0^{\xi} \frac{8x}{\pi^2} dx = \gamma$$

hence, after some simple calculations, we obtain

$$\xi = \pi/2\sqrt{\gamma}.$$

Then

$$I \approx \frac{\pi^2}{8N} \sum_{j=1}^{N} \frac{\sin \xi_i}{\xi_i}.$$

As we anticipated, the second approach gives the more accurate result. One can approximate the variances for both cases.

5. (4 points) Consider the problem of placing wifi transmitting devices on a graph so that each node in the graph has access to wifi.

   There are n wifi devices. You are given a 'transmission reach' function $R(i, n_1, n_2)$, which returns one if wifi device i may grant wifi access to node n_2 when the wifi device is placed at node n_1, and zero otherwise. A wifi device may be placed at any node in the graph, but no two wifi devices may be placed on the same node. The question is whether it is possible to grant wifi access to each node in the graph by using at most k wifi devices.

   - Model the problem, i.e., specify variables and their corresponding domains, and constraints.
   - Does your model have symmetries? If no, provide an explanation. If yes, propose a symmetry breaking constraint along with a brief motivation.
   - Propose a possible look-ahead strategy.

Solution: The problem described is a set covering problem. It could be modelled by having a binary variable $x_{d,n}$ stating whether device d is assigned to node n, or having a variable for each node x_n indicating which device is allocated to node n (if any). There are other ways to model. Constraints then follow based on the chosen variable definition.

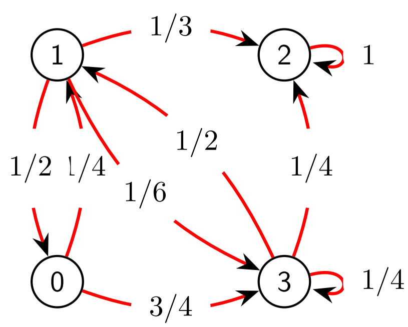
- 0.5 marks for variable and domain definition.
- 0.5 marks to state that no two devices may be at the same node.
- 0.5 marks to state that each node needs to be covered by at least one device.
- 0.5 marks to state that at most k devices may be used.
- 0.5 marks for identifying a symmetry.
- 0.5 marks for providing a correct symmetry breaking constraint.
- ? 1 mark for look-ahead.

6. (2 points) Let $X = \{0, 1, 2, 3\}$ and consider the probability transition matrix

$$||p_{ij}|| = \begin{vmatrix} 0 & 1/4 & 0 & 3/4 \\ 1/2 & 0 & 1/3 & 1/6 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 1/4 & 1/4 \end{vmatrix}.$$

Draw a graph of the corresponding Markov chain and explain it.

Solution:



Here state 2 is said to be absorbing: if the particle gets into this state it remains there, since $p_{22} = 1$. From state 0 the particle goes to the adjacent states 3 or 1; state 3 has the property that the particle remains there with probability $1/4$ and goes to state 2 with the same probability or to state 1 with twice more likely. From state 1 the particle can reach any other state.

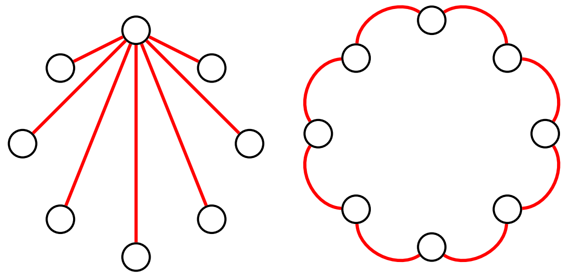


Figure 3: Social networks: Star (left) and Ring (right).

7. (4 points) Consider the two social network communication structures in Fig. 3. Discuss which implementation of the particle swarm optimization would choose each of these networks and why.

Solution: The star social structure, where all particles are interconnected. Each particle can therefore communicate with every other particle. In this case each particle is attracted towards the best solution found by the entire swarm. Each particle therefore imitates the overall best solution. The first implementation of the PSO used a star network structure, with the resulting algorithm generally being referred to as the gbest PSO. The gbest PSO has been shown to converge faster than other network structures, but with a susceptibility to be trapped in local minima. The gbest PSO performs best for unimodal problems.

The ring social structure, where each particle communicates with its nN immediate neighbours. In the case of $n_N = 2$, a particle communicates with its immediately adjacent neighbours. Each particle attempts to imitate its best neighbour by moving closer to the best solution found within the neighbourhood. It is important to note that neighbourhoods overlap, which facilitates the exchange of information between neighbourhoods and, in the end, convergence to a single solution. Since information flows at a slower rate through the social network, convergence is slower, but larger parts of the search space are covered compared to the star structure. This behaviour allows the ring structure to provide better performance in terms of the quality of solutions found for multi-modal problems than the star structure. The resulting PSO algorithm is generally referred to as the lbest PSO.

