Search, Data, and Knowledge Reasoning

Exercise Sheet #2

In addition to the PDF version of this exercise sheet, we also offer a matching online version. We hope to integrate the aggregated results or exemplary (and anonymous) excerpts to make our online exercise more fruitful for you. Therefore, we kindly ask you to indicate your answers in this survey (as well). Of course, all answers remain anonymously at all time.

To access the online version, please scan the QR code or use the following URL:

https://forms.gle/7RwvorumjjGAsByB7

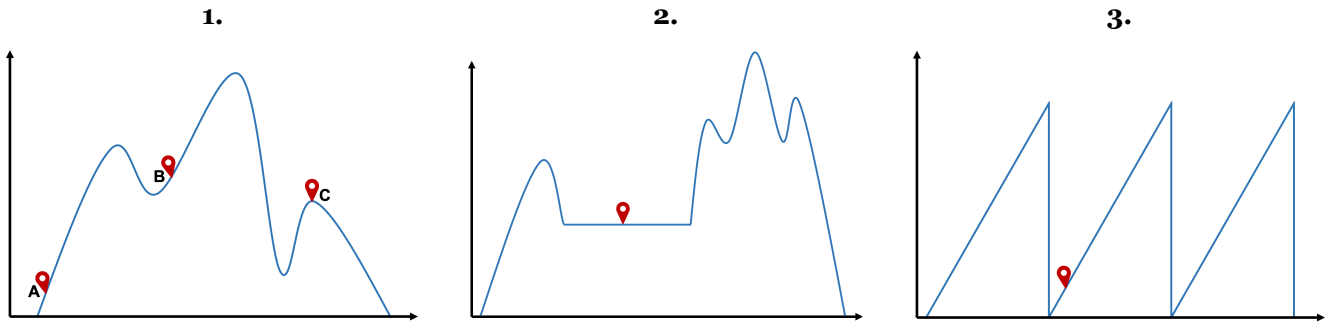
Exercise 2.1: *Local Search Algorithms*

After your exciting walk a few weeks ago – during which you and your best mate had a deep talk about the fascinating world of intelligent agents – you two meet again. After explaining how you faced your problem with "Okea" and what a new sofa has to do with search, your friend wonders: "Hmm, your solutions make sense in these contexts. However, when I was procrastinating during the last exam period, I watched a video about algorithms called 'Local Search Algorithms' that are frequently used to optimize machine learning techniques. Can you tell me how these algorithms differ from the ones that you have mentioned so far?". Instantly, your friend raises below follow-up questions that you try to address:

a. When does it make sense to apply *local search algorithms* and what are their advantages?

b. What is the main problem of these algorithms and how are they trying to cope with it?

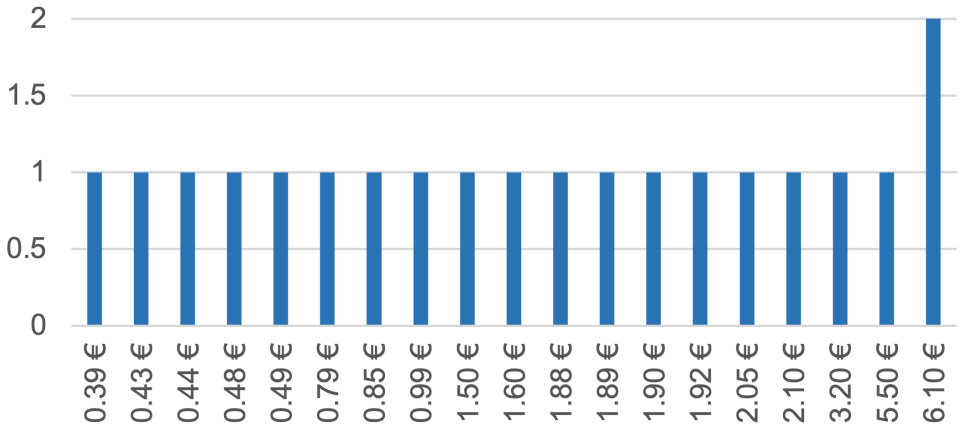
c. Suddenly, your friend pulls out a pen and paper and starts to draw three charts and says: "Assume your agent faces the following unknown solution spaces and that your agent applies *Local Hill Climbing*. The red marker indicates the starting position (in 1. there are three alternative starting positions that should be thought through). Regarding each case, will the agent reliably find the global maximum?"



d. Regarding the three charts above, would *Simulated Annealing Search* work better or worse?

Exercise 2.2: Binning

Since you read about all the great health benefits of avocados, you decided that you want to regularly eat avocados (ignoring all the consequences for the environment). As an AI expert who do not like to make her/his own decisions, you decide to develop an AI solution called "AIocado" that recommends you where to buy your avocado on a weekly basis. You therefore decide to collect data that characterizes the offered avocados with features, such as their price, size, or weight. After weeks of sampling data, you notice that especially the avocado prices vary a lot. In order to prepare the data for AIocado, you experiment with different binning strategies. The below chart visualizes your sampled data describing how many supermarkets offer avocados for a certain price.



a. First, you want to compare *equal width* and *equal frequency binning*.
   - For *equal width binning*, you first try width=1€ and then width=2€ alternatively.
   - For equal frequency binning, you first try frequency=5 and then frequency=10.

   How does the distribution change when applying each strategy?

b. Which (dis-)advantages of each binning approach do you notice?

Exercise 2.3: Induction vs. Deduction vs. Abduction

After learning about the three different kinds of reasoning, you reflect on reasonings that you conducted in the past to understand which kind of reasoning was performed. You think about the following past experiences, trying to classify them either as *inductive*, *deductive,* or *abductive* reasoning:

a. A few weeks ago, when you read a lot about avocados and their health benefits, you read everywhere that hard avocados generally do not taste good. One day later, your significant other bought you an avocado as (s)he knew about your newly developed avocado love. Your excitement quickly stops as soon as you realized how hard the avocado felt like, thinking "it's very hard, so it must taste bad". Later on, your suspicion was confirmed when you tried to eat the avocado for dinner.

b. Back when you were a kid, you learned that all dinosaurs died because they were hit by a meteor. Shortly after, you heard in the news that an astronaut died as he got hit by a meteor. Instantly, you concluded: "The astronaut must have been a dinosaur!".

c. When you dog was still a puppy, you taught it several tricks, such as "sit" or "play dead". You achieved this by following a simple strategy: Every time when your dog performed the respective trick correctly, it was rewarded with a treat. If it performed the trick wrong, it did not receive a treat. After a few weeks of training, your dog understood that if you say "sit" and it sits that it can expect to receive a treat.

Exercise 2.4: Forward vs. Backward Chaining

a. Briefly compare the objective of *forward* and *backward chaining*.

b. When you were young, you enjoyed studying the characteristics of different animals. After a few weeks, you have developed the following knowledge base:

1. Has a beak ⟹ lays eggs
2. Has fur ⟹ gives milk to its newborns
3. Has feathers ⟹ can fly
4. Has fur ⟹ can sweat
5. gives milk to its newborns AND can sweat ⟹ is a mammal
6. Lays eggs AND can fly ⟹ is a bird
7. Lays eggs AND is a mammal ⟹ is a platypus
8. Lays eggs AND has no legs ⟹ is a snake
9. Is a Platypus ⟹ searches for food at the muddy bottom of rivers
10. Search for food at the muddy bottom of rivers AND is a bird ⟹ is a duck
11. Search for food at the muddy bottom of AND is a mammal AND lays eggs ⟹ has a venomous sting
12. Is a snake ⟹ has a venomous tooth
13. Has a venomous sting ⟹ is dangerous
14. Has a venomous tooth ⟹ is dangerous

One day, you stumble across a special animal named "Karl". Using your developed knowledge base, you want to understand which kind of animal it is and whether Karl is dangerous. Initially, you only know that Karl has a beak and fur. Of course, you apply *forward chaining* to "classify" Karl based on your developed knowledge base. Can you derive the desired answers?

c. To buy your new and luxurious sofa, you decide to get a credit from your bank. After exploring all the information about the process of granting a credit you developed the following knowledge base:

1. Is younger than 30 ⟹ is student
2. Is older than 30 ⟹ is employee
3. Is employee ⟹ has high income
4. Is student ⟹ has low income
5. Has rich parents ⟹ has low risk
6. Has low income AND has low risk ⟹ gets credit granted
7. Has high income ⟹ gets credit granted

Using *backward chaining*, you try to derive the conditions that you must fulfill to get a credit. Is it true that, in order to receive a credit, you need rich parents if you are younger than 30?