Lunds Tekniska Högskola
Institutionen för datavetenskap

EDAF70 (tidigare EDA132)
VT 2018

Tillämpad Artificiell Intelligens
Applied Artificial Intelligence
Tentamen 2018–03–15, 14.00–19.00, Gasque

You can give your answers in English or Swedish.
You are welcome to use a combination of figures and text in your answers.
100 points, 50% needed for pass.

1 Search (JM): 12 points

Pelle lives in a small village Sylthult somewhere in Syldavia. He is approaching the age when he will choose the university he will study computer science at. There are two choices: University of Syldavia denoted by G2 on the map below (see Fig. 1), and a newcomer, University of Maldivia, denoted by G1. Pelle reasons as follows:

1. Our province's bus connections are so well matched, that switching between buses takes no time;
2. Every minute saved on trips to the university counts, therefore 1 minute saving one way gives 2 minutes/day (two ways), gives 600 min/year (students use to have some vacations, don't they), gives 3000 min/5 years of studies. More than two full days of life! (Pelle has a calculator and found this out.) Who would like to waste two full days of life for unproductively sitting in a bus?
3. Pelle's older syster, Lotta, has just taken an AI course and she is good at finding best solutions to such complicated problems.

So Pelle asks her to find him the fastest way to a university, given the map of bus connections below. As Lotta has no information except the travel times provided in the map, she uses the breadth-first exploration of the graph.

Question 1 (2p) Which university will Lotta recommend to Pelle? What will be the path taken to the university?

After getting the answer Pelle realizes that he has read some article in the local Syldavian Nyheter, where two experts have compared the quality of the bus network. In particular, those experts provided their estimates for

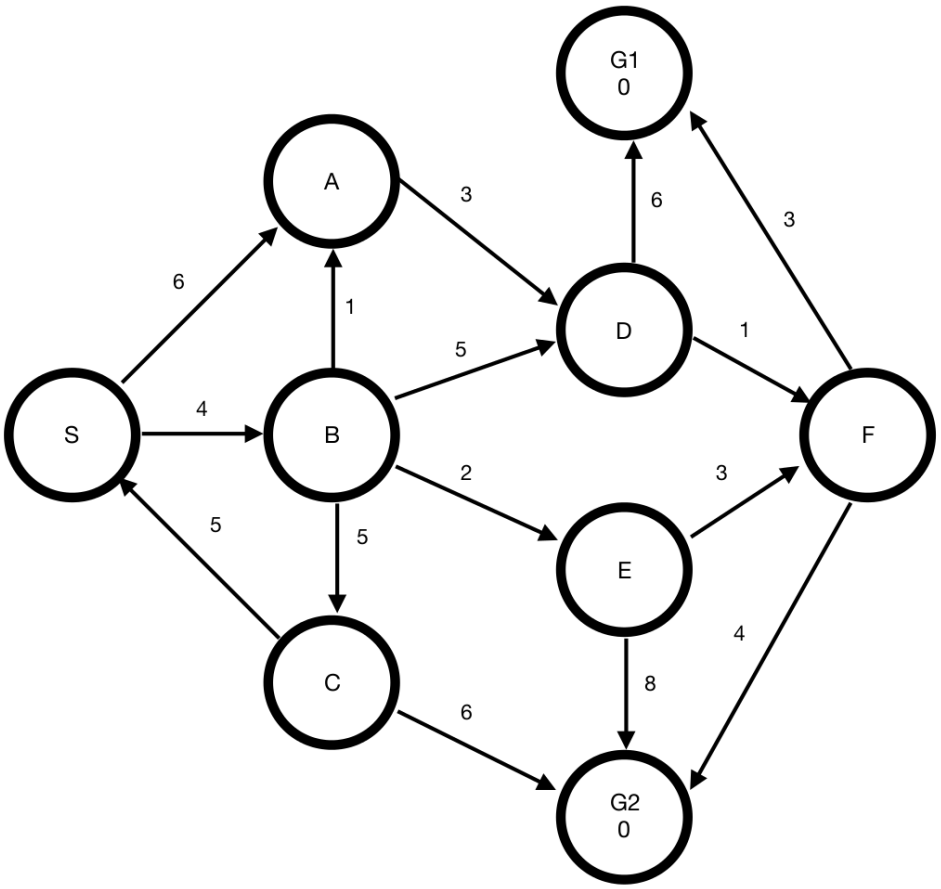

Figure 1: The Syldavian bus network.

arrival to a university from any of the villages included in the bus network (they wanted to stress that no citizen has farther to a university than 10 minutes of travel). Their estimates look as follows:

Given this information Lotta decides to use an informed search algorithm, A*, letting the solution to be found optimally (sometimes).

Question 2 (4p) Assuming Lotta uses estimates of expert #1, which university will Lotta recommend to Pelle? What will be the path taken to the university? What will be the EXACT order of node expansion during search (in case of ties assume alphabetical or lexicographic order)?

Question 3 (1p) Is expert #1 really an expert? Are his estimates correct (i.e., is this an admissible heuristic)? If not, why?

	expert 1	expert 2	you	admissible but not-optimal
S	10	10		
A	9	8		
B	8	6		
C	5	5		
D	3	4		
E	6	4		
F	2	3		
G1	0	0		
G2	0	0		

Table 1: How long would it take to a university? Excerpt from *Syldavian Nyheter*, 15/3/2019

Question 4 (1p) How about expert #2: is he really an expert? Are his estimates correct (i.e., is this an admissible heuristic)? If not, why?

Question 5 (3p) In the column entitled "YOU" write the **optimal** heuristic function for this problem. What would be the order of node expansion then and what would be the answer?

Question 6 (1p) Fill in the last column with a non-optimal, but admissible heuristic (you may of course copy one of the first two columns, if the function provided there is admissible).

You may, of course, use a separate sheet of paper for your solution.

2 Planning (JM): 15 points

In figure 2 there is a two handed robot ABB YuMi capable of manipulation and use of simple tools. Your task will be to describe the domain of YuMi's action in such manner that you can use the automatic planning algorithms for assembly planning. Assume that your YuMi has the following capabilities:

1. The right hand can drill.
2. Drilling makes holes in plates.
3. The left hand can screw.
4. Screwing connects two plates together.
5. Each hand can pick a plate from a feeder or a fixture.

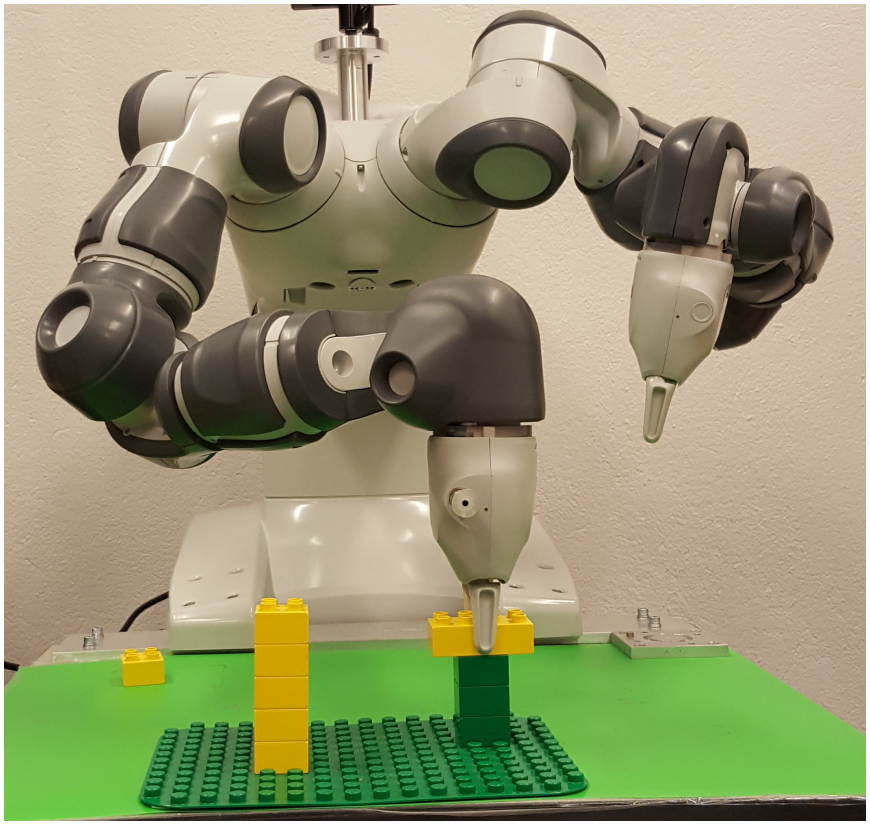

Figure 2: The ABB YuMi robot.

6. Each hand can place (held) plate on a fixture, or in the output tray. There may be more fixtures than one.
7. In order to let YuMi drill, the other hand needs to hold the plate to be drilled through.
8. In order to screw, both hands need to hold one plate each (both plates must have holes made earlier). This is enough to get the plates screwed, getting a screw in place is outside the scope of this text:-)

Do the following:

1. Define the predicates needed to describe every state in this assembly domain. (2p)
2. Describe the initial state, where plates do not have any holes and are all in the feeder, and YuMi arms, fixtures and the output tray are empty. (3p)
3. Define the necessary operators (i.e. name them, list their parameters and provide preconditions and effects for each of them. (7p)

4. Formulate a goal describing two screwed plates lying in the output tray. (2p)

5. Is this a STRIPS domain?

3 Reasoning (JM): 3+5+5=13 points

University of Syldavia has lots of employed teachers. Some of them are professors. A teacher must be a nominated docent. To be nominated as docent one must have defended a PhD or, if one is an artist, painted a big picture. Painters do not defend PhDs as they devote all times to painting (and teaching).

Assume Lotta meets a teacher in the corridor. Can she with certainty say that the teacher is a professor? That the teacher is a docent? That the teacher has a PhD?

If your answer to all three questions is no then motivate it. If there is a positive answer, then use

1. backward chaining

2. resolution

to prove it, if possible. Motivate if the answer was yes, but you could not prove.

4 Machine learning (PN): 30 points

In this question, you will implement a simple course recommender system based on the marks given in each course. The objective of the recommender will be to help a students obtain the best possible marks. It does not evaluate the quality of the courses.

In this question, you will have to write programs. You can only use Python or Java. If you do not know these languages, use an algorithmic notation.

4.1 The Dataset

Table 2 shows an except of KODAL, the student records system of HTL, an avant-garde university of Syldavia[1]. Each row in KODAL corresponds to a student and each column to a course.

[1] **Disclaimer**: The story, all names, characters, and incidents portrayed in this question are fictitious. No identification with actual persons (living or deceased), places, buildings, and products is intended or should be inferred. No person or entity associated with this question received payment or anything of value, or entered into any agreement. No animals were harmed in the writing of this question.

HTL has 20,000 students and offers 500 different courses. Each cell of the table contains the mark obtained by the student, either U, 3, 4, or 5, where U means failed. For instance, student tad12aaa got a mark of 3 in course ABC01 and failed in course ABC02.

If a student never registered to a course or if her/his mark is unknown, the cell is filled with a question mark.

Courses / **Students**	ABC01	ABC02	ABC03	ABC04	CBA01	...
tad12aaa	3	U	?	4	5	...
yft14xxx	4	?	3	?	5	...
tad10bbb	5	?	5	3	?	...
tad11ccc	3	?	?	4	5	...
...	...	...				

Table 2: An except from the KODAL system

4.2 Analyzing the Dataset

In this section, we will analyze the dataset and understand its structure.

1. What courses did student tad11ccc take and what are her/his marks?
2. Intuitively, given the marks shown in Table 2, which student is the most similar to tad11ccc?
3. Supposing that the students perform similarly across the courses, what result can we expect for tad11ccc in course ABC02?

4.3 Encoding the Dataset

Prof. Alembick, the general maintainer of KODAL, encoded the dataset using a dictionary (and a `Map` in Java). The keys are the students and the values, the rows containing the mark records. Each row is also a dictionary, where the key is the course name and the value, the mark.

In Python, Table 2 is encoded as:

```
kodal = {
 'tad12aaa':
  {'ABC01': 3, 'ABC02': 'U', 'ABC03': '?', 'ABC04': 4, 'CBA01': 5},
 'yft14xxx':
  {'ABC01': 4, 'ABC02': '?', 'ABC03': 3, 'ABC04': '?', 'CBA01': 5},
 'tad10bbb':
  {'ABC01': 5, 'ABC02': '?', 'ABC03': 5, 'ABC04': 3, 'CBA01': '?'},
 'tad11ccc':
```

```
  {'ABC01': 3, 'ABC02': '?', 'ABC03': '?', 'ABC04': 4, 'CBA01': 5}
}
```

We access the values with the keys as index:

```
print(kodal['tad12aaa'])
{'ABC04': 4, 'CBA01': 5, 'ABC01': 3, 'ABC03': '?', 'ABC02': 'U'}

print(kodal['tad12aaa']['ABC01'])
3
```

1. The teachers using KODAL replace the numerical mark with the U letter when a student fails a course. Describe how we can convert this value into a number? Justify your answer.

2. Write a function in Python or Java to replace the U values with a number. You will suppose you have access to the dictionary (or Java `Map`) structure with the `kodal` variable. (You do not need to read it from a file).

4.4 A Baseline Technique to Select a Course

A baseline technique is a starting point to solve a problem. Here we will select a course from a minimal analysis of the marks. Note that we do not to analyze the quality or usefulness of a course.

4.4.1 Manual Computations

1. Taking an easy course for the grades is a no-brainer strategy (although not always a good idea). Having a look at Table 2, which course would you select to pass with honors (effortlessly)? Which course would you avoid? Justify your answers.

2. Compute manually the mean of the known marks for each of the courses ABC01, ABC02, ABC03, ABC04, and CBA01 in Table 2 and rank their "easiness". You will have to process the U values.

4.4.2 Programming

1. Write a function in Python or Java to collect all the marks of a given course, if the mark is know. You will suppose you have access to the dictionary (or `Map`) structure with the `kodal` variable. You will return a dictionary like the one below:

   ```
   {'ABC04': [4, 4, 3], 'ABC01': [3, 4, 3, 5],
   'CBA01': [5, 5, 5], 'ABC03': [3, 5], 'ABC02': ['U']}
   ```

2. Write a function in Python or Java to rank the courses using the mean of the marks the students obtained in it. You will return the ranked list of courses, the first one being the easiest.

4.5 Comparing Profiles

Students at HTL certainly deserve better advice than "Take an easy course." In this exercise, we will compare the result profiles to make smarter recommendations.

During the lecture on natural language processing, we computed the similarity of two documents using a cosine function (Lecture of March 2, 2018, Part I, slide 10). We will use this cosine similarity to compare the mark profiles of two students.

Let us denote $\mathbf{q}$ and $\mathbf{d}$, two vectors representing the marks of two students, where q_i and d_i are the respective marks in course i. We have:

$$\begin{aligned} \cos(\mathbf{q}, \mathbf{d}) &= \frac{\mathbf{q} \cdot \mathbf{d}}{||\mathbf{q}|| \cdot ||\mathbf{d}||} \\ &= \frac{\sum_{i=1}^{n} q_i d_i}{\sqrt{\sum_{i=1}^{n} q_i^2} \sqrt{\sum_{i=1}^{n} d_i^2}}. \end{aligned}$$

where n is the total number of courses. Students with identical results will have a cosine of 1, while dissimilar ones will tend to 0.

4.5.1 Manual Computations

1. Compute manually the cosine between the marks of tad11ccc and tad12aaa; you will replace the question marks with 0. You do not need to reduce the results (if you do not have a calculator): You can write $\frac{2}{\sqrt{5}}$, for instance, instead of 0.894 (the value of this fraction);

2. Compute manually the cosine between tad11ccc and yft14xxx; you will replace the question marks with 0;

3. Compute manually the cosine between tad11ccc and tad10bbb; you will replace the question marks with 0;

4.5.2 Programming

1. Write a function that computes the dot product of two mark vectors from two students.

2. Write a function that computes the cosine similarity between two mark vectors from two students. You will use the previous function.

3. Write a function that computes the cosine similarity between all the pairs of students. You will use the previous function.

4.6 Recommending a Course

We will now recommend courses to a student using his/her mark profile. We will consider all the courses the student did not take, for instance ABC02 and ABC03 for student tad11ccc and we will predict the mark s/he would get. We will call these courses, the candidates. A realistic system would then propose a set of candidate courses, where the student would obtain his/her best marks. This last part is set aside in this examination.

There are many recommendation algorithms; we will now implement two elementary techniques:

1. Given a student and a candidate course, for instance tad11ccc and ABC02, a first technique is to determine the most similar student in KODAL, who has taken the course and predict that the outcome will be the same.

   (a) Using the results in Sect. 4.5.1, what would be the predicted result of tad11ccc for ABC02?

   (b) And for ABC03?

   (c) Program a function to implement the technique above.

2. We can derive a second technique from the first one: Given a student, take the weighted average of the marks of the other students, where the weights will be the cosine similarities between the students.

   (a) What could be the advantage and disadvantages of this technique compared with the first one?

   (b) Using the results in Sect. 4.5.1, what what would be the predicted result of tad11ccc for ABC02?

   (c) And for ABC03? You can just write the product or sum you obtain and not reduce it to a value. If you reduce it, take the floor of the result to get the mark (if the result is 4.3, the mark will be 4).

   (d) Program a function to implement the technique above.

5 Probabilistic Reasoning / Bayesian Networks (EAT): 8 + 4 + 6 + 10 + 2 = 30 points

The Tasty Eggs Company produces Easter Eggs with two fillings, “Toffee”, and “Chili”. 60% of the produced pieces have toffee filling, the rest (40%) have

chili (the kind of hot chili). 50% of each type of eggs have dark chocolate, the other half milk chocolate coating.

Normally, the eggs are produced using differently patterned moulds for the different fillings (dots for toffee, waves for chili), and also the wrappers (brown for the toffee eggs and red for the chili ones) should make a clear distinction possible. However, during one week each spring the company produces their "Funny Eggs special edition", which means that patterns and wrappers are more or less randomly chosen for the eggs. In the end of the special week, only 70% of the toffee flavoured eggs have little dots on them, and only 70% have a brown wrapper, while only 80% of the chili eggs have waves and only 80% have red wrappers. The chocolate coating distribution is, however, not changed, so still 50% of the toffee and 50% of the chili eggs have dark, the other milk chocolate coating.

You buy one bag of classic Tasty Eggs and one bag of the Funny Eggs edition. Each bag contains 20 eggs following the above distributions. You bring the bags to some family gathering, and you plan to have the classic bag's content put in a bowl on the table where the kids will be seated, as you know that some of them really appreciate the strong chili flavour, but others cannot really cope with it, so you want to give them the possibility to know up-front what they set their teeth in. For your adult family members, you want to have the fun part, as you know that everybody can deal with a hot surprise.

You know your family and assume, that things might not work out as planned, so you prepare for the worst (at least for four different scenarios) and have everything calculated just in case:

a) Your aunt does not listen to your instructions and pours both bag contents into one big bowl. You have prepared for this by drawing a Bayesian network to represent the mess. What does the optimal network for the mixed bowl look like, and what is in the CPTs? What is the probability for getting chili filling in a dark, dotted egg with a brown wrapper?

b) Your cousin thinks it looks much nicer with only red eggs in one bowl and brown ones in the other. What is the probability of grabbing a chili egg from the bowl with brown eggs if you do not look at the pattern or chocolate coating?

c) The bowls with the bag contents have been correctly filled and placed, but your uncle has recently discovered that he feels uncomfortable when he eats spicy food. He still thinks he wants the fun part, and grabs an egg with brown wrapper from the bowl with the funny eggs. It has dots on it and is coated in dark chocolate. What is the probability for it to actually have toffee filling?

d) One of your nephews wants to play a trick and exchanges the bowls - or not, you simply do not know which bowl is which, but you know that the

bag contents have not been mixed. Are two observations of dark, wavy eggs with red wrappers in a row that have chili filling from one bowl enough to know that this was the bowl filled from the classic bag if you use MAP reasoning? What would finding a milk chocolate coated, wavy egg in a brown wrapper tell you?

e) All these calculations made you think of different aspects of Bayesian learning and reasoning over time that you learned about a while ago. How are the *optimal Bayesian classifier* and the basic algorithm for *forward filtering* connected? Explain briefly!

General hint: Getting the right formula written down is more important than the result in numbers!

Good Luck!