CS 367: Tutorial Week 11
Probabilistic Logic & NLP

Q1: Bayes Network

A patient goes to the doctor for a medical condition, the doctor suspects three diseases as the cause of the condition. The three diseases are D_1, D_2, D_3, which are marginally independent from each other. There are four symptoms S_1, S_2, S_3, S_4 which the doctor wants to check for presence in order to find the most probable cause of the condition. The symptoms are conditionally dependent to the three diseases as follows: S_1 depends only on D_1, S_2 depends on D_1 and D_2. S_3 is depends on D_1 and D_3, whereas S_4 depends only on D_3. Assume all random variables are Boolean, they are either 'true' or 'false'.

1. Draw the Bayesian network for this problem.

The Bayesian network is shown in Figure 2.

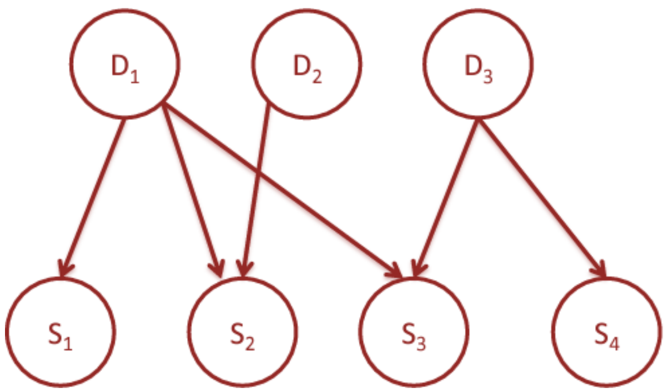

Figure 2: The Bayesian network for disease symptom problem.

2. If we observe the fourth symptom, (S4=true), for which diseases do we gain information?

   By observing S4=true we gain information about only disease D3. Notice D1 and D2 are independent of S4.

Q2: Bayesian Net in Problog

Living in Auckland in winter, you always take your umbrella to go out when its rain, when it is cloudy and when it is sunny too (in case it rains later). The probability of rain is 70%, the probability of cloudy is 85%, and the probability of sunny is 15% (that is why we are all longing for summer). But you are sometime in a rush and forget to take your umbrella. You forget your umbrella more often when it is sunny and cloudy than when it is rainy:

- When it is rainy, you take your umbrella 99% of the time.
- When it is cloudy, you take your umbrella 75% of the time.
- When it is cloudy, you take your umbrella 50% of the time.

1. Model this scenario in a Problog program.
2. Write down a query to find the probability of you taking your umbrella.

```
0.75::rainy.
0.85::cloudy.
0.15::sunny.

0.99::umbrella :- rainy.
0.75::umbrella :- cloudy.
0.5::umbrella :- sunny.


query(umbrella).
```

3. Implement your program in the Problog online editor (https://dtai.cs.kuleuven.be/problog/editor.html) and evaluate your query. What is the value of the probability?

Query ▼	Location	Probability
umbrella	12:7	0.91365703

4. You noticed that you are much more likely to take your umbrella when you had something for breakfast (your mind might be in a better state to remember to take it).
   The probability that you had breakfast given that you took your umbrella is 0.8, and the probability that you had breakfast given that you did not take your umbrella is 0.3.
   Add this to your previous model and implement it.

```
food(muesli).
food(juice).

0.75::rainy.
0.85::cloudy.
0.15::sunny.

0.99::umbrella :- rainy.
0.75::umbrella :- cloudy.
0.5::umbrella :- sunny.

0.8:: breakfast(X):- umbrella, food(X).
0.3:: breakfast(X):- \+umbrella, food(X).
```

5. Write statements and a query to find the probability that you take your umbrella, given the evidence that you have muesli and juice for breakfast. Implement them. What is the probability now?

```
evidence(breakfast(muesli), true).
evidence(breakfast(juice), true).

query(umbrella).
```

Query ▼	Location	Probability
umbrella	19:7	0.98688486

Q3: Naïve Bayes classification

Consider a corpus of text composed of 7 sentences, 4 in English and 3 in French:
“I go to school with a sandwich”, “My sandwich is good”, “I buy a sandwich”, “School is great”,
“Mon école est super”, “Je mange dans mon école”, “Mon école est loin”.

1. Training. Following the same process than in the lecture, build a table with the word counts and the probabilities for each word to appear in each language, based on the corpus. Your probabilities should include add-on smoothing.

Language	**I**	**go**	**to**	**school**	**with**	a	**sandwich**	**my**	is	**good**	**buy**	**great**
English	2	1	1	2	1	2	3	1	2	1	1	1
English+1	3	2	2	3	2	3	4	2	3	2	2	2
$\hat{p}(v_i\|en)$	0.081	0.054	0.054	0.081	0.054	0.081	0.11	0.054	0.081	0.054	0.054	0.054
French	0	0	0	0	0	0	0	0	0	0	0	0
French+1	1	1	1	1	1	1	1	1	1	1	1	1
$\hat{p}(v_i\|fr)$	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032

Language	mon	**école**	est	**je**	mange	dans	loin	Σ
English	0	0	0	0	0	0	0	18
English+1	1	1	1	1	1	1	1	37
$\hat{p}(v_i\|en)$	0.027	0.027	0.027	0.027	0.027	0.027	0.027	1
French	3	3	2	1	1	1	1	12
French+1	4	4	3	2	2	2	2	31
$\hat{p}(v_i\|fr)$	0.13	0.13	0.097	0.064	0.064	0.064	0.064	1

2. Classification. Based on the table you built previously, classify the following sentence as English or French.
“Je mange mon sandwich dans mon école”

$$P(en) \times P(\text{"je"}|en) \times P(\text{"mange"}|en) \times P(\text{"un"}|en) \times P(\text{"sandwich"}|en) \times P(\text{"dans"}|en) \times P(\text{"mon"}|en) \times P(\text{"école"}|en) = \frac{4}{7} \times \left(\frac{1}{37} \times \frac{1}{37} \times \frac{1}{37} \times \frac{4}{37} \times \frac{1}{37} \times \frac{1}{37} \times \frac{1}{37}\right) = 2.41e-11$$

$$P(fr) \times P(\text{"je"}|fr) \times P(\text{"mange"}|fr) \times P(\text{"un"}|fr) \times P(\text{"sandwich"}|fr) \times P(\text{"dans"}|fr) \times P(\text{"mon"}|fr) \times P(\text{"école"}|fr) = \frac{3}{7} \times \left(\frac{2}{31} \times \frac{2}{31} \times \frac{4}{31} \times \frac{1}{31} \times \frac{2}{31} \times \frac{4}{31} \times \frac{4}{31}\right) = 7.98e-9$$

So the language is French, under the assumptions of Naïve Bayes with Laplace (add one) smoothing (probability is higher).

3. What would have happened without Laplacian smoothing?

$$P(en) \times P(\text{"je"}|en) \times P(\text{"mange"}|en) \times P(\text{"un"}|en) \times P(\text{"sandwich"}|en) \times P(\text{"dans"}|en) \times P(\text{"mon"}|en) \times P(\text{"école"}|en) = \frac{4}{7} \times \left(0 \times 0 \times 0 \times \frac{4}{37} \times 0 \times 0 \times 0\right) = 0$$

$$P(fr) \times P(\text{"je"}|fr) \times P(\text{"mange"}\,|\,fr) \times P(\text{"un"}|fr) \times P(\text{"sandwich"}|fr) \times P(\text{"dans"}|fr) \times P(\text{"mon"}|fr) \times P(\text{"école"}|fr) = \frac{3}{7} \times \left(\frac{2}{37} \times \frac{2}{37} \times \frac{4}{37} \times 0 \times \frac{2}{37} \times \frac{4}{37} \times \frac{4}{37}\right) = 0$$

As soon as 1 conditional probability is 0, both probabilities are equal to 0 → not possible to classify.