Artificial Intelligence Practicum 2
Bayesian Networks

Garben Tanghe
Olivier Van den Nest
Group 16

November 10, 2019

Question 1

- Query variables (i.e. diseases)
  - Bronchitis 9 %
  - Common cold 35 %
  - Lung cancer 4 %
  - Pneumonia 2 %
- Evidence variables (i.e. symptoms and risk factors)
  - Chest pain (symptom) 14 %
  - Fever (symptom) 9 %
  - Loose cough (symptom) 19 %
  - Runny nose (symptom)
  - Shortness of breath (symptom) 18 %
  - Smoking (risk factor) 29 %
  - Weakened immune system (risk factor) 5 %

Question 2

The network structure (topology) can be found in Figure 1. The risk factors are displayed in red, the diseases in blue and symptoms in green.

It is hereby important to note that some of the logical causalities were not explicitly stated in the assignment. There also needed to be dealt with incomplete data, which did not allow to specify the Conditional Probability Tables (CPTs) completely. There was made up for the missing information by using common sense reasoning.

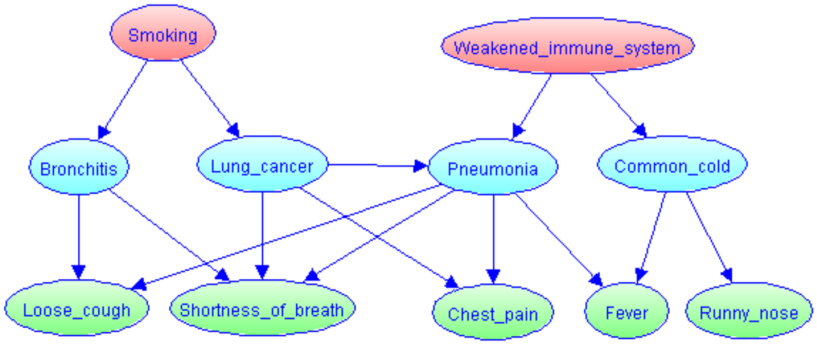

Figure 1: The network structure (topology)

The first added assumption is that a Weakened immune system is a risk factor for the Common cold besides Pneumonia. The values
$P(Common\_cold|Weakened\_immune\_system) = 0.90$ and
$P(Common\_cold|\neg Weakened\_immune\_system) = 0.32$ are added. The total occurrence of the Common cold in function of conditional probabilities on Weakened immune system is
$0.05 * 0.90 + 0.95 * 0.32 = 0.349$, which is approximately $35\,\%$.

Second, no information about the occurrence of a Runny nose is given. Since it is quite common, a value of $34\,\%$ will be given. There is thus assumed that this symptom occurs more often than the other symptoms. This number is again calculated from conditional probabilities, but this time on Common cold:
$0.35 * 0.95 + 0.65 * 0.01 = 0.339$, which is approximately $34\,\%$.

A third remark is that not all values in the assignment are exact. For example there is stated that Lung cancer has an occurrence rate of $4\,\%$. However, when calculating this number from other given values, there can be seen that this percentage has been rounded off. All people either smoke ($29\,\%$) or do not smoke ($71\,\%$). The probability of Lung cancer can now be written in function of the conditional probabilities for having Lung cancer while Smoking ($10\,\%$) and having Lung cancer while not Smoking ($1\,\%$). This results in a total probability of
$0.29 * 0.10 + 0.71 * 0.01 = 0.0361$ which is not equal to the $4\,\%$ from before. Further, this does not mean that the occurrence rate of Lung cancer should be around $3.6\,\%$ as all other used values to get this number could also have been rounded off.

The last and also biggest assumption made, is that if there is no causal relation drawn in the topology, the nodes are independent. (For example, a Weakened immune system is no risk factor for Lung cancer and Lung cancer does not increase the chance of the symptom Fever.)

Question 3

The number of parameters follows from the following expression:

$$\#parameters = \sum_{n} 2^{\#parents(n)} \tag{1}$$

The number of parents for each node (from left to right and from top to bottom) is: 0, 0, 1, 1, 2, 1, 2, 3, 2, 2, 1. This results in:
$1 + 1 + 2 + 2 + 4 + 2 + 4 + 8 + 4 + 4 + 2 = 34$.

Question 4

The correlations and risk factors are listed below.

- Correlations
  - $P(Shortness\_of\_breath|Bronchitis) = 0.80$
  - $P(Shortness\_of\_breath|\neg Bronchitis \wedge (Pneumonia \vee Lung\_cancer)) = 0.50$
  - $P(Shortness\_of\_breath|Other) = 0.10$
  - $P(Chest\_pain|Pneumonia \vee Lung\_cancer) = 0.90$
  - $P(Chest\_pain|Other) = 0.10$
  - $P(Loose\_cough|Pneumonia \vee Bronchitis) = 0.90$
  - $P(Loose\_cough|Other) = 0.10$
  - $P(Fever|Common\_cold) = 0.20$
  - $P(Runny\_nose|Common\_cold) = 0.95$
  - $P(Runny\_nose|\neg Common\_cold) = 0.01$
  - $P(Fever|Other) = 0.001$
  - $P(Fever|Pneumonia) = 0.90$
- Risk factors
  - $P(Lung\_cancer|Smoking) = 0.10$
  - $P(Lung\_cancer|\neg Smoking) = 0.01$
  - $P(Bronchitis|Smoking) = 0.30$
  - $P(Bronchitis|\neg Smoking) = 0.01$
  - $P(Pneumonia|Weakened\_immune\_system) = 0.30$
  - $P(Pneumonia|Lung\_cancer) = 0.05$
  - $P(Pneumonia|Other) = 0.001$

Note: $Other$ *is equivalent to* $\neg Common\_cold \wedge \neg Pneumonia \wedge \neg Lung\_cancer \wedge \neg Bronchitis$.

Question 5

The CPT of Pneumonia is given in Figure 2. The conditional probabilities for Pneumonia, given Lung cancer, a Weakened immune system and Other are all given. Remember that there was assumed that Smoking, Bronchitis and the Common cold cannot cause Pneumonia, since no causality between them is stated in the assignment. The conditional probability for Pneumonia, given both Lung cancer and a Weakened immune system is added by making an educated guess. There is reasoned that this probability will slightly increase when the patient also has a Weakened immune system, compared to when they only have Lung cancer.

Lung_canc...	Yes		No	
Weakened...	Yes	No	Yes	No
Yes	0.32	0.05	0.3	0.001
No	0.68	0.95	0.7	0.999

Figure 2: The CPT for Pneumonia

The one for Fever can be found in Figure 3. The conditional probabilities for Fever, given Pneumonia, the Common cold and Other are all given. Again, keep in mind that there was assumed that Bronchitis, Lung cancer or any of the other symptoms cannot cause Fever, since no causality between them is stated in the assignment. Also, the diseases Pneumonia and the Common cold are completely independent of each other for that same reason. The conditional probability for Fever, given both Pneumonia and the Common cold is once more added by making an educated guess. Similarly as before, there is reasoned that this probability will slightly increase when the patient also has the Common cold, compared to when they only have Pneumonia.

Pneumonia	Yes		No	
Common_...	Yes	No	Yes	No
Yes	0.94	0.9	0.2	0.001
No	0.06	0.1	0.8	0.999

Figure 3: The CPT for Fever

Question 6

Note: The initial probabilities for each disease can be found in Figure 4.

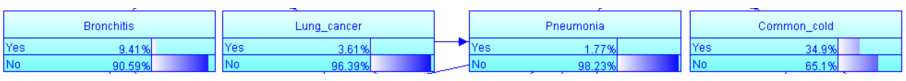

Figure 4: The initial probabilities for each disease

When there is Fever without a Runny nose, there is a $47.51\,\%$ chance of Pneumonia. The Common cold must also be looked at, since this has an even higher probability

(52.94 %). All probabilities for each disease when there is evidence of Fever and no Runny nose can be found in Figure 5.

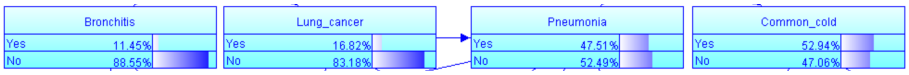

Figure 5: probabilities for each disease when there is evidence of Fever and no Runny nose

Question 7

In case there is evidence of Chest pain, Shortness of breath, no Fever and no Loose cough, there is a 20.82 % chance on Lung cancer. However, Bronchitis is more likely with 37.84 %. All probabilities are added in Figure 6

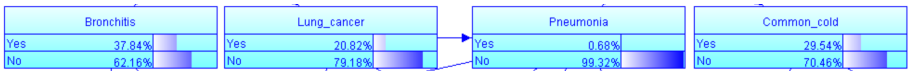

Figure 6: probabilities for each disease when there is evidence of Chest pain, Shortness of breath, no Fever and no Loose cough

Question 8

Bronchitis and Lung cancer can both be caused by Smoking (and developed independently).

Question 9

Assuming the patient speaks nasally because of a Runny nose, the Common cold would be a good diagnosis. It is even 99.95 % certain. The probabilities are found in Figure 7.

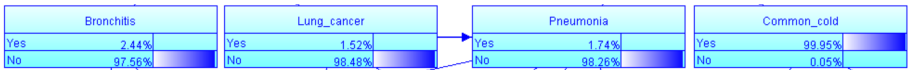

Figure 7: probabilities for each disease when there is evidence of Loose cough, no Smoking, no Weakened immune system, Fever and a Runny nose

Question 10

For this case, there is assumed that the homeless person has a Weakened immune system. Combining this with the symptoms of Fever and Loose cough, diagnosing him for Pneumonia or the Common cold are the best options. The probabilities are found in Figure 8.

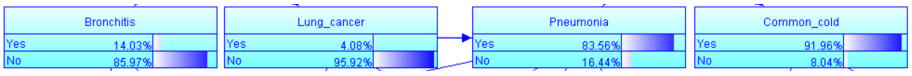

Figure 8: probabilities for each disease when there is evidence of Loose cough, Fever and a Weakened immune system

If additional evidence of no Runny nose is given, the probability of the Common cold drops to $36.6\,\%$, while the probability of Pneumonia rises a little further to $93.39\,\%$. For completeness, all probabilities are listen in Figure 9.

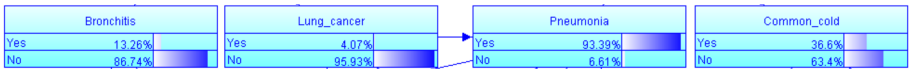

Figure 9: probabilities for each disease when there is evidence of Loose cough, Fever, no Runny nose and a Weakened immune system

Question 11

When a patient is Smoking, has a Weakened immune system and Shortness of breath, the Common cold, Bronchitis and Pneumonia are all quite likely. The probabilities are listen in Figure 10.

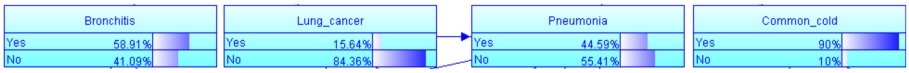

Figure 10: probabilities for each disease when there is evidence of Smoking, a Weakened immune system and Shortness of breadth

Depending on weather or not they have a Runny nose, the Common cold will be a certainty or almost impossible. Whatever evidence is further provided, the probabilities for Bronchitis and Pneumonia vary around $50\,\%$. Here, it becomes clear that there is more going on. It is possible that the patient has both of these diseases. This could be explained by another risk factor that causes these two, but is not provided in the topology. It is also possible that the patient has only one of these diseases, but since they cause similar symptoms, more information is needed. Information on a new symptom, for example Slime (present for Pneumonia, but not for Bronchitis) is necessary to make a final decision.