CS 161 Homework 6

Charles Zhang

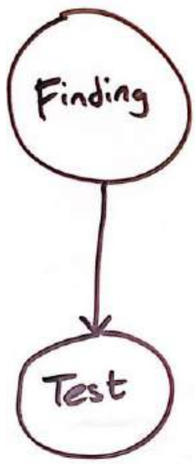
March 14, 2022

Problem 1

Consider the following:

> **An oil well may be drilled on Mr. Y's farm in Texas. Based on what has happened to similar farms, we judge the probability of only oil being present to be .5, the probability of only natural gas being present to be .2, and the probability of neither being present to be .3. Oil and gas never occur together. If oil is present, a geological test will give a positive result with probability .9; if natural gas is present, it will give a positive result with probability .3; and if neither are present, the test will be positive with probability .1.**

a) Model this problem as a Bayesian network over two variables: $Find$**, and** $Test$**.**



Finding	P(Finding)
Oil	0.5
Gas	0.2
None	0.3

Finding	Test	P(Test | Finding)
Oil	T	0.9
Gas	T	0.3
None	T	0.1

b) Suppose the test comes back positive. What's the probability that oil is present?

$$P(oil|test) = \frac{P(test|oil)P(oil)}{P(test)}$$

$$P(oil|test) = \frac{0.9 \times 0.5}{P(test)}$$

$$P(test) = P(test, oil) + P(test, \neg oil)$$

$$P(test) = (P(test|oil) \times P(oil)) + (P(test|\neg oil) \times P(\neg oil))$$

$$P(test) = (0.9 \times 0.5) + (0.4 \times 0.5)$$

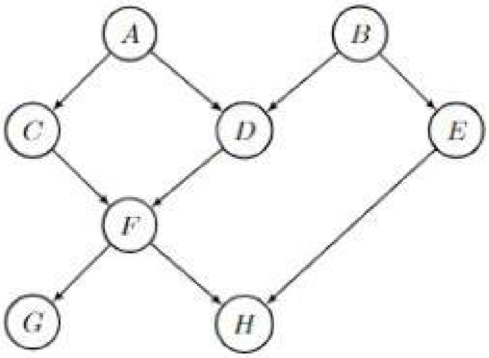
$$P(test) = 0.45 + 0.2 = 0.65$$

$$P(oil|test) = \frac{0.9 \times 0.5}{0.65}$$

$$\boxed{P(oil|test) = 0.69}$$

Problem 2

Consider the Bayesian network below:



A	$\Pr(A)$
1	.2
0	.8

B	$\Pr(B)$
1	.7
0	.3

B	E	$\Pr(E\|B)$
1	1	.1
1	0	.9
0	1	.9
0	0	.1

A	B	D	$\Pr(D\|AB)$
1	1	1	.5
1	1	0	.5
1	0	1	.6
1	0	0	.4
0	1	1	.1
0	1	0	.9
0	0	1	.8
0	0	0	.2

a) Express $\mathtt{Pr}(A, B, C, D, E, F, G, H)$ as a multiplication of conditional and marginal probabilities, according to the factorization encoded in the network structure.

$$P(A, B, C, D, E, F, G, H) = P(A)\times P(B)\times P(C|A)\times P(D|A, B)\times P(E|B)\times P(F|C, D)\times P(G|F)\times P(H|E, F)$$

c) Express $\mathtt{Pr}(a, \neg b, c, d, \neg e, f, \neg g, h)$ in terms of the parameters in the CPTs (a denotes $A = 1$ and $\neg a$ denotes $A = 0$). Use placeholder symbols for the parameters that are not shown in the CPTs.

$$P(a, \neg b, c, d, \neg e, f, \neg g, h) = P(a)\times P(\neg b)\times P(c|a)\times P(d|a, \neg b)\times P(\neg e, \neg b)\times P(f|c, d)\times P(\neg g|f)\times P(h|\neg e, f)$$

$$P(a, \neg b, c, d, \neg e, f, \neg g, h) = 0.2 \times 0.3 \times P(c|a) \times 0.6 \times 0.1 \times P(f|c, d) \times P(\neg g|f) \times P(h|\neg e, f)$$

$$\boxed{P(a, \neg b, c, d, \neg e, f, \neg g, h) = 0.0036 \times P(c|a) \times P(f|c, d) \times P(\neg g|f) \times P(h|\neg e, f)}$$

d) Compute $\mathtt{Pr}(\neg a, b)$ and $\mathtt{Pr}(\neg e|a)$. Justify your answers. Hint: leaf nodes that are not part of the probability query can be removed from the network without affecting the computed probability.

a and b are independent nodes, which means $P(\neg a, b)$ can be calculated as follows:

$$P(\neg a, b) = P(\neg a) \times P(b)$$

$$P(\neg a, b) = 0.8 \times 0.7$$

$$\boxed{P(\neg a, b) = 0.56}$$

Note that e is conditionally independent of a, given b. Therefore:

$$P(\neg e|a) = P(\neg e)$$

$$P(\neg e) = P(\neg e, b) + P(\neg e, \neg b)$$

$$P(\neg e) = P(\neg e|b) \times P(b) + P(\neg e|\neg b) \times P(\neg b)$$

$$P(\neg e) = 0.9 \times 0.7 + 0.1 \times 0.3$$

$$\boxed{P(\neg e) = 0.66}$$

e) List the Markovian assumptions (also known as topological semantics) encoded in the Bayesian network structure.

- $I(\{A\}; \{B, E\} \mid \emptyset)$
- $I(\{B\}; \{A, C\} \mid \emptyset)$
- $I(\{C\}; \{B, D, E\} \mid \{A\})$
- $I(\{D\}; \{C, E\} \mid \{A, B\})$
- $I(\{E\}; \{A, C, D, F, G\} \mid \{B\})$
- $I(\{F\}; \{A, B, E\} \mid \{C, D\})$
- $I(\{G\}; \{A, B, C, D, E, H\} \mid \{F\})$
- $I(\{H\}; \{A, B, C, D, G\} \mid \{E, F\})$

f) Provide the Markov blanket for variable D.

$$\boxed{A, B, C, F}$$