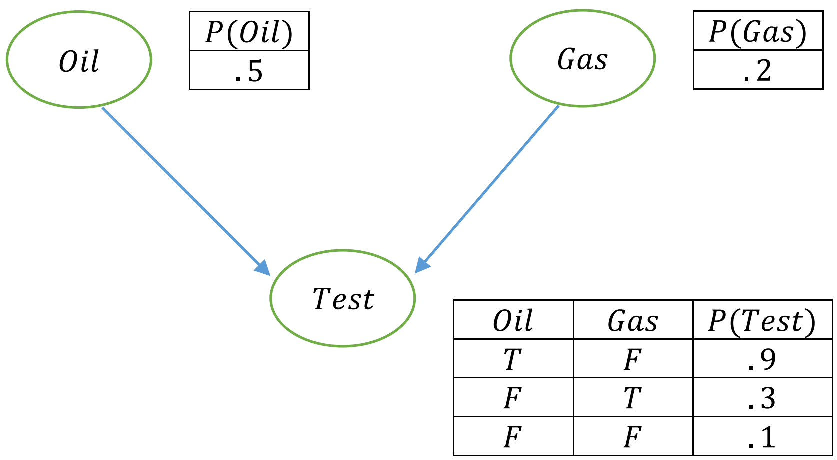
CS 161
HW6
Xiaopei Zhang
004309991

1.

(a). The problem can be modeled as a Bayesian network in the following manner:



$P(Oil)$
$.5$

$P(Gas)$
$.2$

Oil	Gas	$P(Test)$
T	F	$.9$
F	T	$.3$
F	F	$.1$

(b). The probability we need to calculate is actually $P(Oil|Test)$.
According to Bayes' Rule, we have

$$P(Oil|Test) = \frac{P(Test|Oil) * P(Oil)}{P(Test)}$$

In this expression,

$$P(Test) = \sum_{o,g} P(Test|o,g) * P(o,g) = .5 * .9 + .2 * .3 + .3 * .1 = .54$$

$$P(Test|Oil) = .9, P(Oil) = .5$$

Therefore, if the test comes back positive, the probability that oil is present is

$$P(Oil|Test) = \frac{.9 * .5}{.54} = .83$$

2.

(a).

$$\begin{aligned} &P(A,B,C,D,E,F,G,H) \\ &\quad = P(G|A,B,C,D,E,F,H) * P(H|A,B,C,D,E,F) * P(F|A,B,C,D,E) \\ &\quad * P(E|A,B,C,D) * P(D|A,B,C) * P(C|A,B) * P(B|A) * P(A) \end{aligned}$$

According to the condition independence indicated in the Bayesian network, we can simplify this expression as

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

(b). Let's define all factors as the following:

$$f_1(F,G) \equiv P(G|F), f_2(E,F,H) \equiv P(H|E,F)$$
$$f_3(C,D,F) \equiv P(F|C,D), f_4(B,E) \equiv P(E|B)$$
$$f_5(A,B,D) \equiv P(D|A,B), f_6(A,C) \equiv P(C|A)$$
$$f_7(B) \equiv P(B), f_8(A) \equiv P(A)$$

Therefore,

$$P(A,B,C,D,E,F,G,H) \equiv f_1(F,G) * f_2(E,F,H) * f_3(C,D,F) * f_4(B,E) * f_5(A,B,D) * f_6(A,C) * f_7(B) * f_8(A)$$

(c).

$$P(E,F,G,H) = \sum_a \sum_b \sum_c \sum_d P(A,B,C,D,E,F,G,H)$$
$$\equiv \sum_a \sum_b \sum_c \sum_d \big(f_1(F,G) * f_2(E,F,H) * f_3(C,D,F) * f_4(B,E) * f_5(A,B,D) * f_6(A,C) * f_7(B) * f_8(A)\big) = \sum_a \sum_b \sum_c \sum_d f_9(A,B,C,E,F,G,H)$$
$$= \sum_a \sum_b \sum_c f_{10}(A,B,C,E,F,G,H) = \sum_a \sum_b f_{11}(A,B,E,F,G,H)$$
$$= \sum_a f_{12}(A,E,F,G,H) = f_{13}(E,F,G,H)$$

(d).

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

According to given CPTs, we know that

$$P(\neg e|\neg b) = .1$$
$$P(d|a,\neg b) = .6$$
$$P(\neg b) = .3$$
$$P(a) = .2$$

Hence,

$$P(a,\neg b,c,d,\neg e,f,\neg g,h) = .1 * .6 * .3 * .2 * \ P(\neg g|f) * P(h|\neg e,f) * P(f|c,d) * P(c|a)$$
$$= .0036 * P(\neg g|f) * P(h|\neg e,f) * P(f|c,d) * P(c|a)$$

(e). From the Bayesian network, we observe that A, B are conditional, so

$$P(\neg a, b) = P(\neg a) * P(b) = .8 * .7 = .56$$

From the Bayesian network, we also observe that A, E are conditional, so

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

To calculate $P(\neg e)$, we need to make use of CPT of $P(E|B)$

$$P(\neg e) = P(\neg e|b) * P(b) + P(\neg e|\neg b) * P(\neg b) = .9 * .7 + .1 * .3 = .66,$$
which means $P(\neg e|a) = .66$.

(f). The Markovian assumptions encoded in the Bayesian network structure is that a node is conditionally independent of all other nodes in the network, given its parents, children, and children's parents.

(g). The Markov blanket for variable D contains: A, B, C, F.

(h). The resulting factor of multiplying the factors corresponding to $P(D|A,B)$ and $P(E|B)$ is

$$f_{14}(A,B,D,E) = f_5(A,B,D) * f_4(B,E)$$

A	B	D	$f_5(A,B,D)$	B	E	$f_4(B,E)$	A	B	D	E	$f_{14}(A,B,D,E)$
T	T	T	$.5$	T	T	$.1$	T	T	T	T	$.5 * .1 = .05$
T	T	F	$.5$	T	F	$.9$	T	T	T	F	$.5 * .9 = .45$
T	F	T	$.6$	F	T	$.9$	T	T	F	T	$.5 * .1 = .05$
T	F	F	$.4$	F	F	$.1$	T	F	T	T	$.6 * .9 = .54$
F	T	T	$.1$				T	T	F	F	$.5 * .9 = .45$
F	T	F	$.9$				T	F	T	F	$.6 * .1 = .06$
F	F	T	$.8$				T	F	F	T	$.4 * .9 = .36$
F	F	F	$.2$				T	F	F	F	$.4 * .1 = .04$
							F	T	T	T	$.1 * .1 = .01$
							F	T	T	F	$.1 * .9 = .09$
							F	T	F	T	$.9 * .1 = .09$
							F	F	T	T	$.8 * .9 = .72$
							F	T	F	F	$.9 * .9 = .81$
							F	F	T	F	$.8 * .1 = .08$
							F	F	F	T	$.2 * .9 = .18$
							F	F	F	F	$.2 * .1 = .02$

(i).

$$\sum_d f_{14}(A,B,D,E) = f_{14}(A,B,d,E) + f_{14}(A,B,\neg d,E) = f_{15}(A,B,E)$$

A	B	E	$f_{15}(A,B,E)$
T	T	T	$.05 + .05 = .1$
T	T	F	$.45 + .45 = .9$
T	F	T	$.54 + .36 = .9$
T	F	F	$.06 + .04 = .1$
F	T	T	$.01 + .09 = .1$
F	T	F	$.09 + .81 = .9$
F	F	T	$.72 + .18 = .9$
F	F	F	$.08 + .02 = .1$