PROBLEMS UNIT 7

(2) We consider the following random variables:

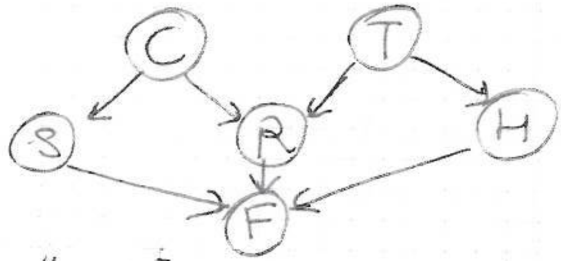
F: my house is flooded (boolean)
S: sprinklers were working (boolean)
R: rains on the weekend (boolean)
H: stay the weekend at home (boolean)
C: cloudy on Friday
T: temperature (three possible values: high, middle, low)

We are going to draw the variables in the following order:

C, T, S, R, H, F

This order tries to go from the fundamental causes, to the final effect.

The bayesian network could be the following:



The "arrows" in this network are based on the following conditional independencies assumptions:

$$P(T|C) = P(T)$$
$$P(S|C,T) = P(S|C)$$
$$P(R|S,C,T) = P(R|C,T)$$
$$P(H|R,S,C,T) = P(H|T)$$
$$P(F|H,R,S,C,T) = P(F|S,R,H)$$

The following could be "reasonable" probability tables:

$P(c) = 0.3$ $\quad P(T=low) = 0.2$ $\quad P(T=middle) = 0.7$ $\quad P(T=high) = 0.1$

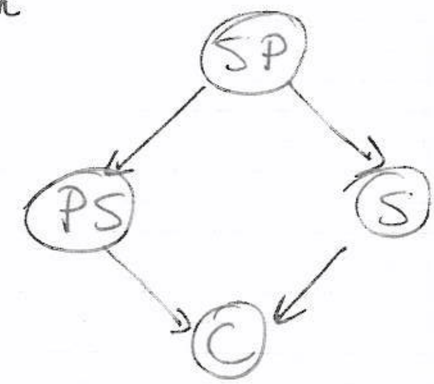
C	P(s|C)
c	0.05
¬c	0.8

C	T	P(r|C,T)
c	low	0.2
c	middle	0.6
c	high	0.7
¬c	low	0.05
¬c	middle	0.1
¬c	high	0.05

T	P(h|T)
low	0.1
middle	0.5
high	0.7

S	R	H	$P(b \mid S, R, H)$
s	r	h	0.8
s	r	¬h	0.6
s	¬r	h	0.6
s	¬r	¬h	0.6
¬s	r	h	0.4
¬s	r	¬h	0.4
¬s	¬r	h	0.4
¬s	¬r	¬h	0.01

5. The following is a possible model of the situation



To draw the network, we have assumed that "Smoker" is conditionally independent of "Passive smoker", given "Smoking parents". And also that "Cancer" is conditionally independent of "Smoking parents", given "Passive smoker" and "Smoker"

Now, we answer the following conditional independence questions using d-separation:

- Is PS c.i. of S given SP?

This is assumed when constructing the network, but we can also deduce it using d-separation:

Paths connecting PS and S

$PS \boxed{\leftarrow SP \rightarrow} S$ ✓ Yes ($SP \in \{SP\}$)

$PS \boxed{\rightarrow C \leftarrow} S$ ✓ Yes ($C \notin \{SP\}$ and has no descendants)

So the answer is yes.

- Is PS c.i of S given C?

  No, the path $PS \leftarrow SP \rightarrow S$ does not meet any criteria.

- Is SP c.i of C given PS?

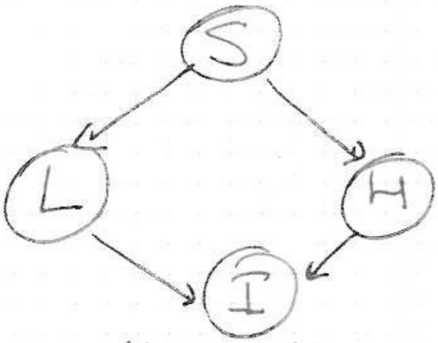
  No, the path $SP \rightarrow S \rightarrow C$ does not meet any criteria

- Is SP c.i. of C given S?

  No, the path $SP \rightarrow PS \rightarrow C$ does not meet any criteria.

7) We take the following order: S, L, H, I (from the cause to the effects).

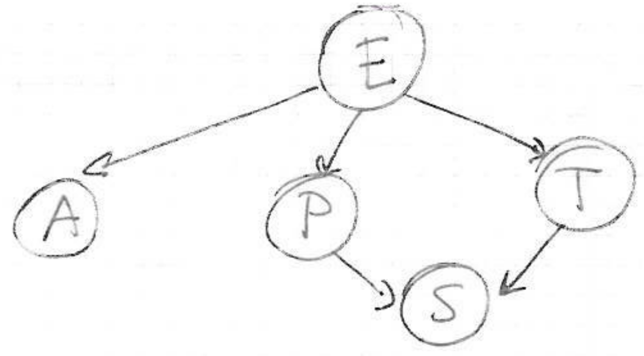
The bayesian network could be the following:



We have assumed that H is c.i. of L given S, and that I is c.i. of S, given L and H.

Formally
$$P(H \mid S, L) = P(H \mid S)$$
$$P(I \mid L, H, S) = P(I \mid L, H)$$

(11)



- Is it possible to calculate any entry in the FJD, from the probability tables of the network?

Yes, every bayesian network is a compact representation of the FJD. In this case

$$P(E, A, P, T, S) = P(E) \cdot P(A|E) \cdot P(P|E) \cdot P(T|E) \cdot P(S|P,T)$$

- a) Are P and A independent?

This is the same as asking if P and A are c.i. given $\emptyset$.
The answer is No. The following path connects A and P and does not meet the criteria:

$$A \leftarrow E \rightarrow P \quad (\text{since } E \notin \emptyset)$$

b) $P(S|T,E) = P(S|T)$?

This is the same as asking if S and E are c.i. given T
The answer si No. The following path connects S and E and does not meet the criteria

$$E \rightarrow P \rightarrow S$$

c) S and A c.i. given E ?

The answer is YES. We have two paths connecting S and A, and both meet the criteria:

$$A \boxed{\leftarrow E \rightarrow} P \rightarrow S \quad \checkmark \text{ Yes}$$

$$A \boxed{\leftarrow E \rightarrow} T \rightarrow S \quad \checkmark \text{ Yes}$$

d) P and T c.i. given S ?

The answer is No. The following path connects P and T and does not meet the criteria: $P \rightarrow S \leftarrow T$

- Compute $P(e, t, p, \neg a, \neg s)$.

$P(e, t, p, \neg a, \neg s) = P(e) \cdot P(\neg a|e) \cdot P(p|e) \cdot P(t|e) \cdot P(\neg s|p, t)$

$= 0{,}7 \cdot 0{,}8 \cdot 0{,}3 \cdot 0{,}8 \cdot 0{,}1 = 0.01344$

- Compute $P(E|s)$.

Let us apply the variable elimination algorithm

For this particular query, we may safely ignore variable A, since it is not an ancestor of query or evidence variables

The initial factors are:

$f_E(E) = P(E)$

E	P(E)
e	0.7
¬e	0.3

$f_P(E, P) = P(P|E)$

E	P	P(P|E)
e	p	0.3
e	¬p	0.7
¬e	p	0.8
¬e	¬p	0.2

$f_T(E, T) = P(T|E)$

E	T	P(T|E)
e	t	0.8
e	¬t	0.2
¬e	t	0.4
¬e	¬t	0.6

$f_S(P, T) = P(s|P, T)$

P	T	P(s|P,T)
p	t	0.9
p	¬t	0.6
¬p	t	0.5
¬p	¬t	0.1

* We eliminate variable P

$$f_{\times P}(E,T,P) = f_P(E,P) \times f_S(P,T)$$

$$f_{\bar{P}}(E,T) = \sum_P f_{\times P}(E,T,P)$$

E	T	p	¬p	$f_{\bar{P}}(E,T)$
e	t	0.3×0.9	0.7×0.5	0.62
e	¬t	0.3×0.6	0.7×0.1	0.25
¬e	t	0.8×0.9	0.2×0.5	0.82
¬e	¬t	0.8×0.6	0.2×0.1	0.5

Current factors = $\{ f_E(E), f_T(E,T), f_{\bar{P}}(E,T) \}$

* We eliminate variable T

$$f_{\times T}(E,T) = f_T(E,T) \times f_{\bar{P}}(E,T)$$

$$f_{\bar{T}}(E) = \sum_T f_{\times T}(E,T)$$

E	t	¬t	$f_{\bar{T}}(E)$
e	0.8×0.62	0.2×0.25	0.546
¬e	0.4×0.82	0.6×0.5	0.628

Current factors = $\{ f_E(E), f_{\bar{T}}(E) \}$

* Multiply and normalize

$$f_{\times E}(E) = f_E(E) \times f_{\bar{T}}(E)$$

E	$f_{\times E}(E)$
e	0.7 × 0.546 = 0.3822
¬e	0.3 × 0.628 = 0.1884

Normalizing $P(E|s) = \langle 0.66982, 0.330187$

(13) a) Compute $P(a, b, \neg c, d, \neg e, f, g, \neg h)$

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

$$= 0{,}3 \cdot 0{,}4 \cdot 0{,}6 \cdot 0{,}8 \cdot 0{,}1 \cdot 0{,}5 \cdot 0{,}1 \cdot 0{,}95 = 0{,}0002736$$

b) Compute $P(F \mid a, b, \neg d)$

For this particular query, we can ignore E, G and H, since they are not ancestors of any query or evidence variable

We apply the variable elimination algorithm

Initial factors:

$f_A() = P(a) = 0.3$

$f_B() = P(b) = 0.4$

$f_D() = P(\neg d \mid a, b) = 0.2$

→ These three factors can be also ignored, since they finally will be "absorbed" by the normalization constant

$f_C(C) = P(C \mid a, b)$

C	$P(C \mid a, b)$
c	0,4
¬c	0,6

$f_F(C, F) = P(F \mid C)$

C	F	$P(F \mid C)$
c	f	0.8
c	¬f	0.2
¬c	f	0.5
¬c	¬f	0.5

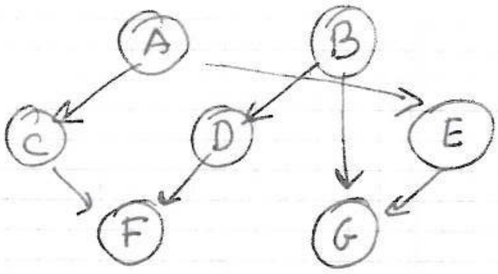
We eliminate C

F	c	¬c	$f_{\bar{C}}(F)$
f	0,4 × 0,8	0,6 × 0,5	0,62
¬f	0,4 × 0,2	0,6 × 0,5	0,38

We do not need to normalize (it is already normalized)

Then, $P(F \mid a, b, \neg d) = \langle 0{,}62, 0.38 \rangle$

14



1.

- For B, we have assumed that it is independent from A. That is, $P(B|A) = P(B)$
- For C, we have assumed that it is independent from B, given A: $P(C|A,B) = P(C|A)$
- For D, we have assumed that it is independent from A and C, given B: $P(D|A,B,C) = P(D|B)$
- For E, we have assumed that it is independent from B, C and D, given A: $P(E|A,B,C,D) = P(E|A)$
- For F, we have assumed that it is independent from A, B and E, given C and D: $P(F|A,B,C,D,E) = P(F|C,D)$
- For G, we have assumed that it is independent from A, C, D and F, given B and E: $P(G|A,B,C,D,E,F) = P(G|B,E)$

2. We will apply the d-separation criteria

- "If we know the value taken by A, our degree of belief in the value that C may take is not updated if in addition we knew the value taken by G"

This is exactly the definition of C and G being conditionally independent, given A, or equivalently $P(C|A,G) = P(C|A)$, or equivalently, that $\{A\}$ d-separates C and G.

There are two paths connecting C and G, and both meet the criteria, so the answer is YES:

① $C \rightarrow F \leftarrow D \leftarrow B \rightarrow G$ (since $F \notin \{A\}$)

② $C \leftarrow A \rightarrow E \rightarrow G$ (since $A \in \{A\}$)

- Are F and G conditionally independent, given A?

That is, does $\{A\}$ d-separates F and G.

There are two paths connecting F and G:

① $F \leftarrow C \leftarrow A \rightarrow E \rightarrow G$

② $F \leftarrow D \leftarrow B \rightarrow G$

But path ② does not meet any of the criteria, so the answer is <u>NO</u>

- $P(F|A,B) = P(F|A,B,G)$?

This is equivalent to say that $\{A,B\}$ d-separates F and G.

Both paths connecting F and G meet the criteria:

① $F \leftarrow \boxed{C \leftarrow A \rightarrow E} \rightarrow G$ $(A \in \{A,B\})$

② $F \leftarrow \boxed{D \leftarrow B \rightarrow G}$ $(B \in \{A,B\})$

Then, the answer is <u>YES</u>

3. We are asked to compute $P(\neg a | \neg b, g)$

For that, we will calculate $P(A|\neg b, g)$ using variable elimination and return the second component of that vector (the one corresponding to A = false

In this query, we do not have irrelevant variables.

<u>Initial factors:</u>

$f_A(A)$

A	P(A)
a	0.3
¬a	0.7

$f_B(B)$

B	P(B)
b	0.6
¬b	0.4

$f_C(A,C) = P(C|A)$

A	C	P(C|A)
a	c	0.7
a	¬c	0.3
¬a	c	0.1
¬a	¬c	0.9

$b_D(B,D) = P(D|B)$

B	D	P(D|B)
b	d	0.1
b	¬d	0.9
¬b	d	0.9
¬b	¬d	0.1

$b_E(A,E) = P(E|A)$

A	E	P(E|A)
a	e	0.3
a	¬e	0.7
¬a	e	0.8
¬a	¬e	0.2

$b_F(C,D) = P(\neg f|C,D)$

C	D	P(¬f|C,D)
c	d	0.1
c	¬d	0.3
¬c	d	0.5
¬c	¬d	0.8

$b_G(B,E) = P(g|B,E)$

B	E	P(g|B,E)
b	e	0.2
b	¬e	0.8
¬b	e	0.2
¬b	¬e	0.9

We eliminate B:

$b_{\times B}(B,D,E) = b_B(B) \times b_D(B,D) \times b_G(B,E)$

$b_{\overline{B}}(D,E) = \sum_B b_{\times B}(B,D,E)$

D	E	b	¬b	$b_{\overline{B}}(D,E)$
d	e	0.6·0.1·0.2	0.4·0.9·0.2	0.089
d	¬e	0.6·0.1·0.8	0.4·0.9·0.9	0.372
¬d	e	0.6·0.9·0.2	0.4·0.1·0.2	0.116
¬d	¬e	0.6·0.9·0.8	0.4·0.1·0.9	0.468

Current factors: $\{b_A(A), b_C(A,C), b_E(A,E), b_F(C,D), b_{\overline{B}}(D,E)\}$

We eliminate E:

$$f_{\times E}(A,D,E) = f_E(A,E) \cdot f_{\bar{B}}(D,E)$$

$$f_{\bar{E}}(A,D) = \sum_E f_{\times E}(A,D,E)$$

A	D	e	¬e	$f_{\bar{E}}(A,D)$
a	d	0.3·0.064	0.7·0.372	0.2856
a	¬d	0.3·0.116	0.7·0.468	0.3624
¬a	d	0.8·0.084	0.2·0.372	0.1416
¬a	¬d	0.8·0.116	0.2·0.468	0.1864

Current factors: $\{f_A(A), f_C(A,C), f_F(C,D), f_{\bar{E}}(A,D)\}$

We eliminate C.

$$f_{\times C}(A,C,D) = f_C(A,C) \cdot f_F(C,D) \quad f_{\bar{C}}(A,D) = \sum_C f_{\times C}(A,C,D)$$

A	D	c	¬c	$f_{\bar{C}}(A,D)$
a	d	0.7·0.1	0.3·0.5	0.22
a	¬d	0.7·0.3	0.3·0.8	0.45
¬a	d	0.1·0.1	0.9·0.5	0.46
¬a	¬d	0.1·0.3	0.9·0.8	0.75

Current factors: $\{f_A(A), f_{\bar{C}}(A,D), f_{\bar{E}}(A,D)\}$

We eliminate D:

$$f_{\times D}(A,D) = f_{\bar{C}}(A,D) \times f_{\bar{E}}(A,D) \quad f_{\bar{D}}(A) = \sum_D f_{\times D}(A,D)$$

A	d	¬d	$f_{\bar{D}}(A)$
a	0.2856·0.22	0.3624·0.45	0.22591
¬a	0.1416·0.46	0.1864·0.75	0.20494

Current factors: $\{f_A(A), f_{\bar{D}}(A)\}$

Multiply and normalize:

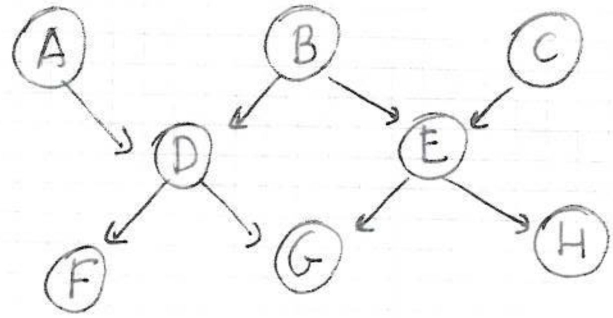
$$f_{\times A}(A) = f_A(A) \times f_{\bar{D}}(A)$$

A	$f_{\times A}(A)$
a	0.3·0.22591 = 0.06777
¬a	0.7·0.20494 = 0.14346

Normalizing: $P(A \mid \neg b, g) = \langle 0.32085, 0.679157 \rangle$

Then: $P(\neg a \mid \neg b, g) = 0.67915$

15



- Answer the following using d-separation

1. B and C independent?

Equivalently, does $\emptyset$ d-separate B and C?

Paths connecting B and C

$\boxed{B \rightarrow E \leftarrow C}$ ✓ $(E \notin \emptyset)$

$B \rightarrow \boxed{D \rightarrow G \leftarrow E} \leftarrow C$ ✓ $(G \notin \emptyset)$

Both paths meet the criteria, so the answer is <u>Yes</u>

2. Are B and C conditionally independent given E?

Equivalently, does $\{E\}$ d-separates B and C?

There is path: $B \rightarrow E \leftarrow C$ does not meet the criteria. So the answer is <u>NO</u>

3. Are F and H independent?

Equivalently, does $\emptyset$ d-separate F and H?

The path $F \leftarrow D \leftarrow B \rightarrow E \rightarrow H$ does not meet any d-separation criteria, so the answer is <u>NO</u>

4. Are F and H conditionally independent given B?

Equivalently, does $\{B\}$ d-separate F and H?

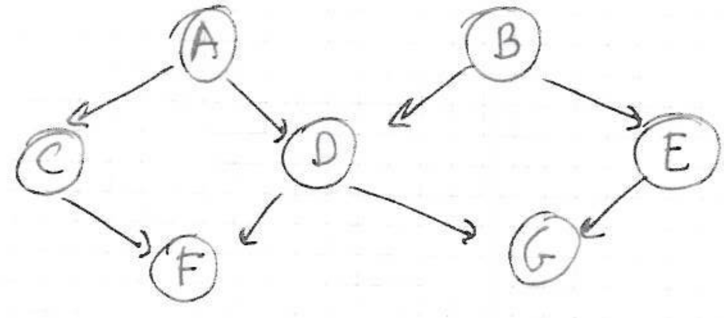
Two paths connecting F and H:

$F \leftarrow \boxed{D \leftarrow B \rightarrow E} \rightarrow H$ $(B \in \{B\})$

$F \leftarrow \boxed{D \rightarrow G \leftarrow E} \rightarrow H$ $(G \notin \{B\}$ and has no descendants$)$

Both paths meet the criteria, so the answer is <u>Yes</u>

16



- Compute $P(B|f)$, using variable elimination

For this query, we can ignore G and E, since they are not ancestors of B (query) or F (evidence).

Initial factors:

$f_A(A)$

A	P(A)
a	0.4
¬a	0.6

$f_B(B)$

B	P(B)
b	0.3
¬b	0.7

$f_C(A,C)$

A	C	P(C|A)
a	c	0.7
a	¬c	0.3
¬a	c	0.2
¬a	¬c	0.8

$f_D(A,B,D)$

A	B	D	P(D|A,B)
a	b	d	0.8
a	b	¬d	0.2
a	¬b	d	0.7
a	¬b	¬d	0.3
¬a	b	d	0.2
¬a	b	¬d	0.8
¬a	¬b	d	0.3
¬a	¬b	¬d	0.7

$f_F(C,D)$

C	D	P(f|C,D)
c	d	0.1
c	¬d	0.5
¬c	d	0.7
¬c	¬d	0.9

We eliminate C:

$$f_{\times C}(A,C,D) = f_C(A,C) \times f_F(D)$$
$$f_{\bar{C}}(A,D) = \sum_C f_{\times C}(A,C,D)$$

A	D	c	¬c	$f_{\bar{C}}(A,D)$
a	d	0.7 · 0.1	0.3 × 0.7	0.28
a	¬d	0.7 · 0.5	0.3 · 0.9	0.62
¬a	d	0.2 · 0.1	0.8 · 0.7	0.58
¬a	¬d	0.2 · 0.5	0.8 · 0.9	0.82

Current factors: { f (A) f (B) f (A,B,D) f (A,D) }

* <u>We eliminate D</u>:

$f_{\times D}(A,B,D) = f_D(A,B,D) \times f_{\bar{C}}(A,D)$

$f_{\bar{D}}(A,B) = \sum_{D} f_{\times D}(A,B,D)$

A	B	d	$\neg d$	$f_{\bar{D}}(A,B)$
a	b	$0.8 \cdot 0.28$	$0.2 \cdot 0.62$	0.348
a	$\neg b$	$0.2 \cdot 0.28$	$0.8 \cdot 0.62$	0.552
$\neg a$	b	$0.7 \cdot 0.58$	$0.3 \cdot 0.82$	0.652
$\neg a$	$\neg b$	$0.3 \cdot 0.58$	$0.7 \cdot 0.82$	0.748

Current factors: $\{ f_A(A), f_B(B), f_{\bar{D}}(A,B) \}$

* <u>We eliminate A</u>:

$f_{\times A}(A,B) = f_A(A) \times f_{\bar{D}}(A,B)$ $\quad f_{\bar{A}}(B) = \sum_{A} f_{\times A}(A,B)$

B	a	$\neg a$	$f_{\bar{A}}(B)$
b	0.4×0.348	$0.6 \cdot 0.652$	0.5304
$\neg b$	0.4×0.552	$0.6 \cdot 0.748$	0.6696

Current factors: $\{ f_B(B), f_{\bar{A}}(B) \}$

* Multiply and normalize:

$f_{\times B}(B) = f_B(B) \times f_{\bar{A}}(B)$

B	$f_{\times B}(B)$
b	$0.3 \times 0.5304 = 0.15912$
$\neg b$	$0.7 \times 0.6696 = 0.46872$

Normalizing: $P(B \mid b) = \langle 0.25344, 0.74656 \rangle$

- Apply likelihood weighting to compute an approximation of $P(B|d,\neg f)$

We can safely ignore E and G for the computation of this probability

Sample 1: A: Distribution $P(A)$: $<0.4, 0.6>$

Random number: 0.13

Value: a

B: Distr.: $P(B)$: $<0.3, 0.7>$

Random: 0.07

Value: b

C: Distr: $P(C|a)$: $<0.7, 0.3>$

Random: 0.57

Value: c

D: Observed value: d

$w \leftarrow w \cdot P(d|a,b) = 1 \cdot 0.8 = 0.8$

F: Observed value: $\neg f$

$w \leftarrow w \cdot P(\neg f|c,d) = 0.8 \cdot 0.9 = 0.72$

Sample generated: $<a, b, c, d, \neg f>$ Weight: 0.72

Sample 2: A: Distr. $P(A)$: $<0.4, 0.6>$

Random: 0.94

Value: $\neg a$

B: Distr: $P(B)$: $<0.3, 0.7>$

Random: 0.13

Value: b

C: Distr: $P(C|\neg a) = <0.2, 0.8>$

Aleat: 0.78

Value: $\neg c$

D: Observed value: d

$w \leftarrow w \cdot P(d|\neg a, b) = 1 \cdot 0.7 = 0.7$

F: Observed value: $\neg f$

$w \leftarrow w \cdot P(\neg f|\neg c, d) = 0.7 \cdot 0.3 = 0.21$

Sample generated: $<\neg a, b, \neg c, d, \neg f>$ Weight: 0.21

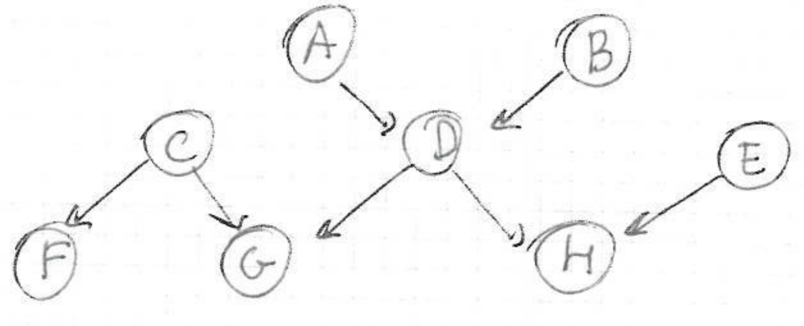
Sample 3 : A: Distr. $P(A) = \langle 0.4, 0.6 \rangle$
Random: 0.48
Value: $\neg a$
B: Distr. $P(B) = \langle 0.3, 0.7 \rangle$
Random: 0.38
Value: $\neg b$
C: Distr: $P(C|\neg a) = \langle 0.2, 0.8 \rangle$
Random: 0.75
Value: $\neg c$
D: Observed value: d
$\langle \neg a, \neg b, \neg c, d, \neg f \rangle$ $w \leftarrow w \times P(d|\neg a, \neg b) = 1 \times 0.3 = 0.3$
Weight: 0.09 F: Observed value: $\neg f$
$w \leftarrow w \times P(\neg f|\neg c, d) = 0.3 \times 0.3 = 0.09$

Sample 4
A: Distr. $P(A) = \langle 0.4, 0.6 \rangle$
Random: 0.93
Value: $\neg a$
B: Distr. $P(B) = \langle 0.3, 0.7 \rangle$
Random: 0.55
Value: $\neg b$
C: Distr: $P(C|\neg a) : \langle 0.2, 0.8 \rangle$
Random: 0.16
Value: c
D: Observed value: d
$\langle \neg a, \neg b, c, d, \neg f \rangle$ $w \leftarrow w \times P(d|\neg a, \neg b) = 1 \times 0.3 = 0.3$
Weight: 0.27. F: Observed value: $\neg f$
$w \leftarrow w \cdot P(\neg f|c, d) = 0.3 \times 0.9 = 0.27$

Sample 5
A: Distr. $P(A) = \langle 0.4, 0.6 \rangle$
Random: 0.91
Value: $\neg a$
B: Distr. $P(B) = \langle 0.3, 0.7 \rangle$
Random: 0.06
Value: b
C: Distr. $P(C|\neg a) : \langle 0.2, 0.8 \rangle$
Random: 0.74
Value: $\neg c$
D: Observed value: d
$\langle \neg a, b, \neg c, d, \neg f \rangle$ $w \leftarrow w \times P(d|\neg a, b) = 1 \times 0.7 = 0.7$
Weight: 0.21 F: Observed value: $\neg f$
$w \leftarrow w \times P(\neg f|\neg c, d) = 0.7 \times 0.3 = 0.21$

We estimate $P(B|d, \neg f)$ using these five samples
Sum of weights corresponding to B=true : $0.72 + 0.21 + 0.21 = 1.14$
" " " " to B=false: $0.09 + 0.27 = 0.36$

17



1. To compute $P(D \mid a, \neg g)$, we can ignore E, F and H, since they are not ancestors of D, A, or G, the query and evidence variables.

Let us now apply the variable elimination algorithm

<u>Initial factors:</u> $f_A() = P(a) = 0.3$ ← this can be ignored, since it will be integrated with the normalization constant

$f_B(B) = P(B)$

B	P(B)
b	0.4
¬b	0.6

$f_C(C) = P(C)$

C	P(C)
c	0.8
¬c	0.2

$f_D(B, D) = P(D \mid a, B)$

B	D	P(D | a, B)
b	d	0.8
b	¬d	0.2
¬b	d	0.7
¬b	¬d	0.3

$f_G(C, D) = P(\neg g \mid C, D)$

C	D	P(¬g | C, D)
c	d	0.9
c	¬d	0.3
¬c	d	0.5
¬c	¬d	0.6

We eliminate B:

$f_{\times B}(B, D) = f_B(B) \cdot f_D(B, D)$ $\quad f_{\overline{B}}(D) = \sum_B f_{\times D}(B, D)$

D	b	¬b	$f_{\overline{B}}(D)$
d	0.4 × 0.8	0.6 × 0.7	0.74
¬d	0.4 × 0.2	0.6 × 0.3	0.26

Current factors: $\{ f_{\overline{B}}(D), f_G(C, D), f_C(C) \}$

We eliminate C

$f_{\times C}(C,D) = f_C(C) \times f_G(C,D) \qquad f_{\bar{C}}(D) = \sum_c f_{\times C}(C,D)$

D	c	¬c	$f_{\bar{C}}(D)$
d	0.8×0.9	0.2×0.5	0.82
¬d	0.8×0.3	0.2×0.6	0.36

Current factors: $\{f_{\bar{B}}(D), f_{\bar{C}}(D)\}$

Multiply and normalize:

D	$f_{\times D}(D)$
d	0.74 × 0.82 = 0.6068
¬d	0.26 × 0.36 = 0.0936

$P(D \mid a, \neg g) = \langle 0.86636, 0.13364 \rangle$

Therefore, $P(d \mid a, \neg g) = 0.86636$

2. To compute $P(a \mid \neg g, \neg h)$ using likelihood weighting, we only generate samples with the observed values in the evidence variables G and H. Then, only samples (b) and (c) could be generated.

- The weight for (b) would be:

  $\langle \neg a, \neg b, \neg c, \neg d, \neg e, \neg f, \neg g, \neg h \rangle$

  $w = P(\neg g \mid \neg c, \neg d) \cdot P(\neg h \mid \neg d, \neg e) = 0.6 \cdot 0.7 = 0.42$

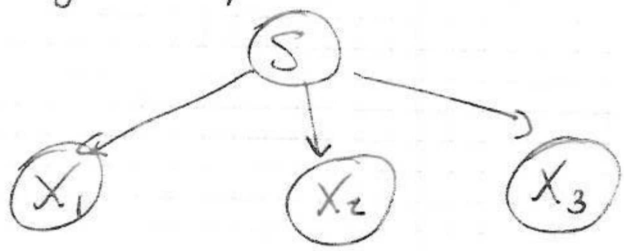
- The weight for (c) would be:

  $\langle a, b, c, d, \neg e, \neg f, \neg g, \neg h \rangle$

  $w = P(\neg g \mid c, d) \cdot P(\neg h \mid d, \neg e) = 0.9 \cdot 0.2 = 0.18$

(18)

- Let S be the random variable corresponding to the class (SPAM or NOT SPAM) of an email. And let X_1, X_2, X_3 random variables corresponding to the features. Then Naive Bayes assumes the following bayesian network:



- The estimates for the class probability are

$$P(S = \text{"SPAM"}) = \frac{250}{1000} = 0{,}25$$

$$P(S = \text{"NOT SPAM"}) = \frac{750}{1000} = 0{,}75$$

The conditional probabilities are (no smoothing):

$P(X_1 = true \mid S = \text{"SPAM"}) = 0.25$ $\quad$ $P(X_1 = false \mid S = \text{"SPAM"}) = 0.75$

$P(X_1 = true \mid S = \text{"NOT SPAM"}) = 0.5$ $\quad$ $P(X_1 = false \mid S = \text{"NOT SPAM"}) = 0.5$

$P(X_2 = true \mid S = \text{"SPAM"}) = 0{,}5$ $\quad$ $P(X_2 = false \mid S = \text{"SPAM"}) = 0.6$

$P(X_2 = true \mid S = \text{"NOT SPAM"}) = 0.25$ $\quad$ $P(X_2 = false \mid S = \text{"NOT SPAM"}) = 0.75$

$P(X_3 = true \mid S = \text{"SPAM"}) = \frac{100}{250} = 0.4$ $\quad$ $P(X_3 = false \mid S = \text{"SPAM"}) = 0.6$

$P(X_3 = false \mid S = \text{"NOT SPAM"}) = \frac{225}{750} = 0{,}3$ $\quad$ $P(X_3 = false \mid S = \text{"NOT SPAM"}) = 0{,}7$

- Classify new example: $X_1 = true$, $X_2 = X_3 = false$

$\boxed{S = \text{"SPAM"}}$

$$P(S = \text{"SPAM"}) \cdot P(X_1 = true \mid S = \text{"SPAM"}) \cdot P(X_2 = false \mid S = \text{"SPAM"}) \cdot P(X_3 = false \mid S = \text{"SPAM"}) =$$

$$= 0.25 \cdot 0.25 \cdot 0.6 \cdot 0.6 = 0{,}0225$$

$\boxed{S = \text{"NOT SPAM"}}$

$$P(S = \text{"NOT SPAM"}) \cdot P(X_1 = true \mid S = \text{"NOT SPAM"}) \cdot P(X_2 = false \mid S = \text{"NOT SPAM"}) \cdot P(X_3 = false \mid S = \text{"NOT SPAM"})$$

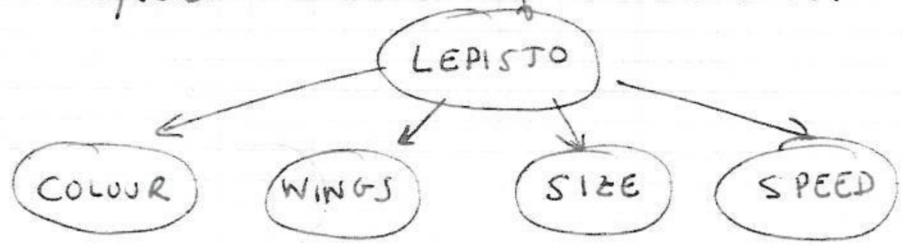
$$= 0.75 \cdot 0.5 \cdot 0.75 \cdot 0.7 = 0{,}196875$$ ← highest

The email is classified as "NOT SPAM"

(20) The random variables we consider are

LEPISTO: the class, with two possible values (yes, no)
COLOUR: with values black, white, yellow, red
WINGS: with values yes, no
SIZE: with values small, big, medium
SPEED: with values high, average, low

The bayesian network of the model is:



Probability estimates (using Laplace smoothing with k=1)

- P(COLOUR= black | LEPISTO= yes) = $\frac{1+1}{3+4} = \frac{2}{7}$

  [Explanation: we have 3 examples with class "yes", and 1 of them is "black". To smooth the proportion, we add 1 to the numerator and 4 to the denominator, since COLOUR has 4 possible values]

- P(COLOUR=yellow | LEPISTO=yes) = (0+1)/(3+4) = 1/7
- P(COLOUR=white | LEPISTO=yes) = (1+1)/(3+4) = 2/7
- P(COLOUR= red | LEPISTO=yes) = (1+1)/(3+4) = 2/7
- P(WINGS= yes | LEPISTO=yes) = (3+1)/(3+2) = 4/5
- P(WINGS= no | LEPISTO=yes) = (0+1)/(3+2) = 1/5
- P(SIZE =small | LEPISTO = yes) = (2+1)/(3+3) = 3/6
- P(SIZE = medium | LEPISTO = yes) = (1+1)/(3+3) = 2/6
- P(SIZE = big | LEPISTO = yes) = (0+1)/(3+3) = 1/6
- P(SPEED= high | LEPISTO = yes) = (3+1)/(3+3) = 4/6
- P(SPEED= average | LEPISTO = yes) = (0+1)/(3+3) = 1/6
- P(SPEED = low | LEPISTO= yes) = (0+1)/(3+3) = 1/6

- P(COLOR = black | LEPISTO = no) = (3+1)/(7+4) = 4/11
- P(COLOUR= yellow | LEPISTO=no) = (3+1)/(7+4) = 4/11
- P(COLOUR= white | LEPISTO = no) = (0+1)/(7+4) = 1/11
- P(COLOUR= red | LEPISTO = no) = (1+1)/(7+4) = 2/11
- P(WINGS= yes | LEPISTO = no) = (3+1)/(7+2) = 4/9
- P(WINGS= no | LEPISTO = no) = (4+1)/(7+2) = 5/9

- $P(SIZE = small \mid LEPISTO = no) =$ [illegible]
- $P(SIZE = medium \mid LEPISTO = no) = (2+1)/(7+3) = 3/10$
- $P(SIZE = big \mid LEPISTO = no) = (3+1)/(7+3) = 4/10$
- $P(SPEED = high \mid LEPISTO = no) = (1+1)/(7+3) = 2/10$
- $P(SPEED = average \mid LEPISTO = no) = (4+1)/(7+3) = 5/10$
- $P(SPEED = low \mid LEPISTO = no) = (2+1)/(7+3) = 3/10$
- $P(LEPISTO = yes) = \frac{3}{10}$
- $P(LEPISTO = no) = \frac{7}{10}$

• Classify the example

COLOUR	WINGS	SIZE	SPEED
yellow	no	small	high

$\boxed{LEPISTO = yes}$

$$P(LEPISTO=yes) \cdot P(COLOUR=yellow \mid LEPISTO=yes) \cdot P(WINGS=no \mid LEPISTO=yes) \cdot P(SIZE=small \mid LEPISTO=yes) \cdot P(SPEED=high \mid LEPISTO=yes) =$$

$$= \frac{3}{10} \cdot \frac{1}{7} \cdot \frac{1}{5} \cdot \frac{3}{6} \cdot \frac{4}{6} = 0.002857$$

$\boxed{LEPISTO = no}$

$$P(LEPISTO=no) \cdot P(COLOUR=yellow \mid LEPISTO=no) \cdot P(WINGS=no \mid LEPISTO=no) \cdot P(SIZE=small \mid LEPISTO=no) \cdot P(SPEED=high \mid LEPISTO=no) =$$

$$= \frac{7}{10} \cdot \frac{4}{11} \cdot \frac{5}{9} \cdot \frac{3}{10} \cdot \frac{2}{10} = \boxed{0.008484} \rightarrow \text{highest}$$

The example is predicted <u>NOT</u> to be a Lepisto.