Workshop 10

COMP90051 Machine Learning

Semester 2, 2020

Learning Outcomes

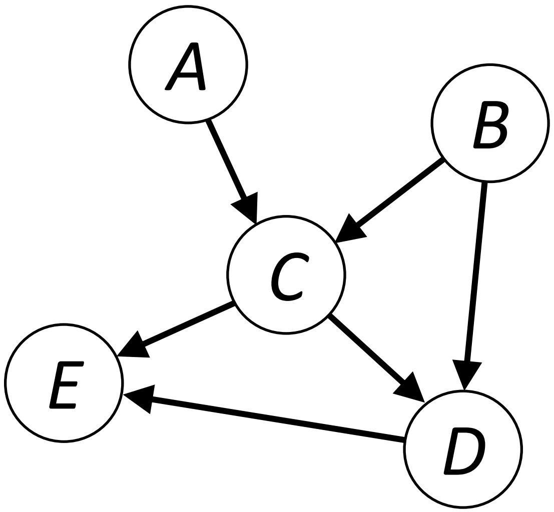
By the end of this workshop you should be able to:

1. write down the factorised joint distribution represented by a directed PGM
2. answer probabilistic queries on simple directed PGMs by variable elimination
3. reason about independence between variables

Q1: Bayes net

a) Find the factorised joint distribution

b) Count the number of parameters in the CPTs

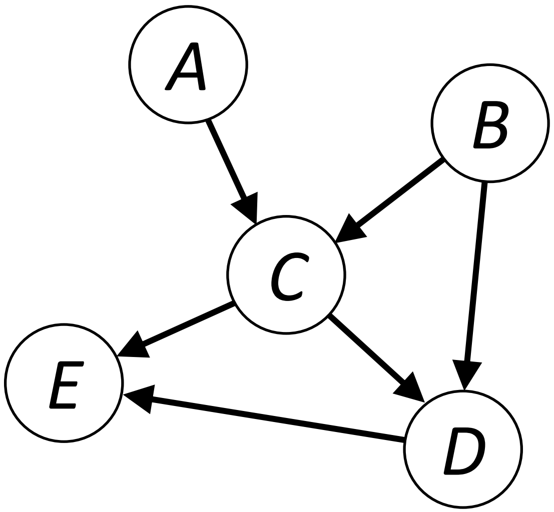


Q1: Bayes net

a) Find the factorised joint distribution
$p(A,B,C,D,E) = p(E|C,D)p(D|C,B)p(C|A,B)p(A)p(B)$

b) Count the number of parameters in the CPTs

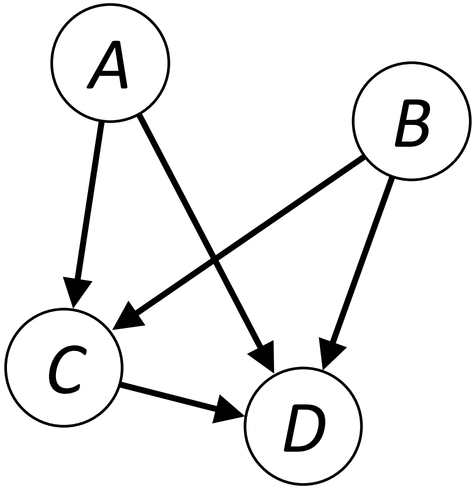
Node	# free params
A	$2^0 = 1$
B	$2^0 = 1$
C	$2^2 = 4$
D	$2^2 = 4$
E	$2^2 = 4$
Total	14



Q1: Bayes net

a) Find the factorised joint distribution

b) Count the number of parameters in the CPTs



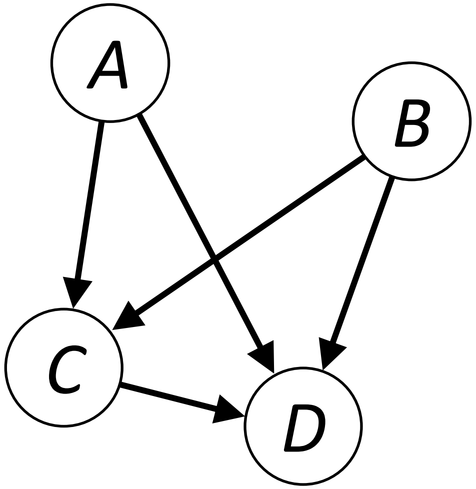
Q1: Bayes net

a) Find the factorised joint distribution

$$p(A, B, C, D) = p(D|A, B, C)p(C|A, B)p(A)p(B)$$

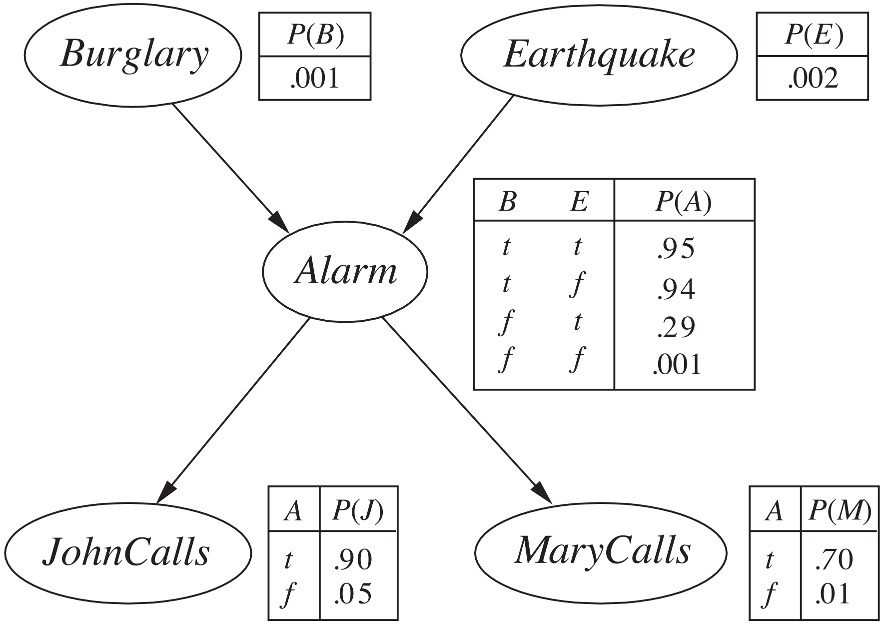
b) Count the number of parameters in the CPTs

Node	# free params
A	$2^0 = 1$
B	$2^0 = 1$
C	$2^2 = 4$
D	$2^3 = 8$
Total	14



Q2: Variable elimination

- Leo's house has an alarm to detect burglars
- The alarm is occasionally set off by an earthquake
- Leo's neighbours John and Mary (who don't know each other) sometimes call if they hear the alarm
- If Leo receives a call from John and Mary, what's the likelihood his house has been burgled?



Q2: Variable elimination

- Query analysis:
  - Query nodes: *Burglary*
  - Evidence (observed) nodes: *JohnCalls*, *MaryCalls*
  - Latent (unobserved) nodes: *Earthquake*, *Alarm*
- Need to compute $p(B|j,m)$

Here lowercase means the observed value

- Bayes' rule gives

$$p(B|j,m) = \frac{p(B,j,m)}{p(j,m)}$$

- Use the full joint distribution + marginalisation to compute the numerator and denominator

Q2: Variable elimination

Numerator:

$$p(B,j,m) = \sum_E \sum_A p(A,B,E,j,m) = p(B) \overbrace{\sum_E p(E) \underbrace{\sum_A p(A|B,E)\, \underbrace{p(j|A)p(m|A)}_{f_{j,m}(A)}}_{f_{j,m}(B,E)}}^{f_{j,m}(B)}$$

$f_{j,m}(A) =$

A	$f_{j,m}(A)$
0	0.0005
1	0.63

$=$

A	$p(j\|A)$
0	0.05
1	0.90

$*$

A	$p(m\|A)$
0	0.01
1	0.70

$f_{j,m}(B,E) =$

B	E	$f_{j,m}(B,E)$
0	0	0.0011295
0	1	0.183055
1	0	0.59223
1	1	0.598525

$=$

		$p(A\|B,E)$	
B	E	A=0	A=1
0	0	0.999	0.001
0	1	0.71	0.29
1	0	0.06	0.94
1	1	0.05	0.95

$\times$

Marginalising over A

A	$f_{j,m}(A)$
0	0.0005
1	0.63

Q2: Variable elimination

Marginalising over E

$f_{j,m}(B) =$

B	$f_{j,m}(B)$
0	0.00149335
1	0.59224259

$=$

	$f_{j,m}(B,E)$	
B	E=0	E=1
0	0.0011295	0.183055
1	0.59223	0.598525

$\times$

E	$p(E)$
0	0.998
1	0.002

$p(B,j,m) =$

B	$p(B)$
0	0.00149186
1	0.00059224

$=$

B	$p(B)$
0	0.999
1	0.001

$*$

B	$f_{j,m}(B)$
0	0.00149335
1	0.59224259

Denominator:

$$p(j,m) = \sum_B p(B,j,m) = 0.00149186 + 0.00059224 = 0.002084100239$$

Putting the results together:

$$p(B = 1|j,m) = \frac{p(B = 1,j,m)}{p(j,m)} = \frac{0.00059224}{0.002084100239} = 0.28417184$$

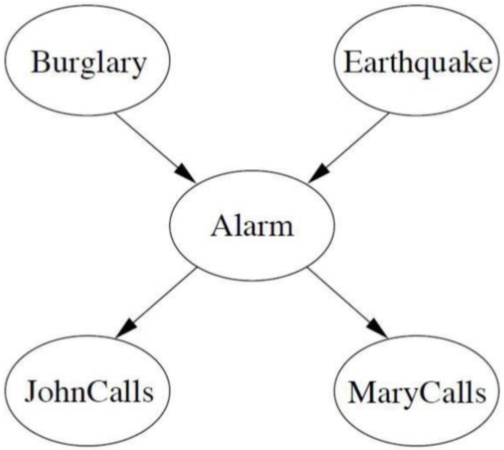
Q3: Independence

Returning to the previous Bayes net:

a) If no evidence is observed, are the *Burglary* and *Earthquake* nodes independent?

b) What if we observe *Alarm* = t?

c) Explain your answer to (b) intuitively.



Q3: Independence

a) Check for (marginal) independence:

$$p(B,E) = p(B)p(E)\underbrace{\sum_A p(A|B,E)\underbrace{\sum_J p(J|A)}_{=1}\underbrace{\sum_M p(M|A)}_{=1}}_{=1}$$

$$= p(B)p(E)$$

b) Check for (conditional) independence when we observe $a = t$:

$$p(B,E|a) \propto p(B)p(E)p(a|B,E)\underbrace{\sum_J p(J|a)}_{=1}\underbrace{\sum_M p(M|a)}_{=1}$$

$$\neq p(B|a)p(E|a)$$

Hint: see supplemental slides 'Independence in PGMs' for graphical rules

Q3: Independence

c) Why are B and E dependent given $a = t$?

* The sub-graph consisting of B, E, A is known as a *common effect* structure (a.k.a. v-structure)
* Observing the effect (*Alarm*) puts the *Burglar* and *Earthquake* in competition as an explanation for the effect
* In other words, we expect B and E to be anti-correlated given $a = t$
* This is indeed the case—e.g. $p(B = t|e = t, a = t) = 0.003$ and $p(E = t|b = t, a = t) = 0.002$