



# Workshop 12

COMP90051 Machine Learning

Semester 2, 2018

# Agenda

## 1. Worksheet 12 (35 min)

- \* Exercises on PGMs
- \* Pen and paper

## 2. Stan demo (15 min)

- \* Inference for nuclear power plant example
- \* (Optional) install Stan and follow along

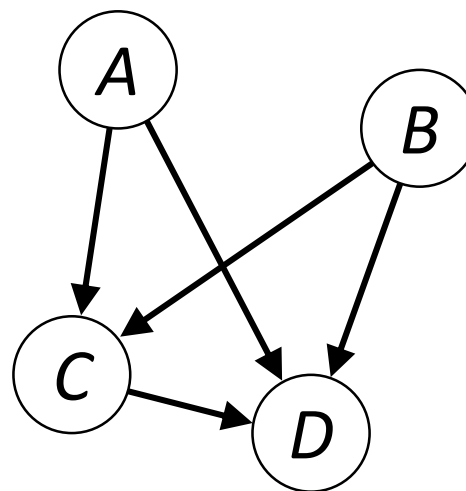
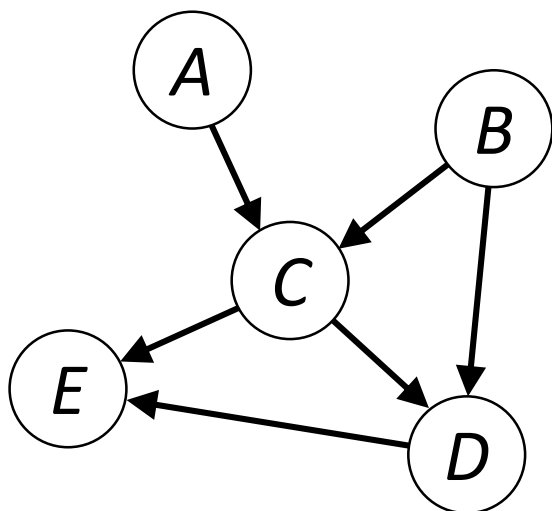
# Worksheet 12

## Probabilistic Graphical Models

# Q1: Bayes net

For the following Bayes nets:

- write down the **factorised joint distribution**
- count the **# of free parameters in the CPTs** (assuming each variable is Boolean).



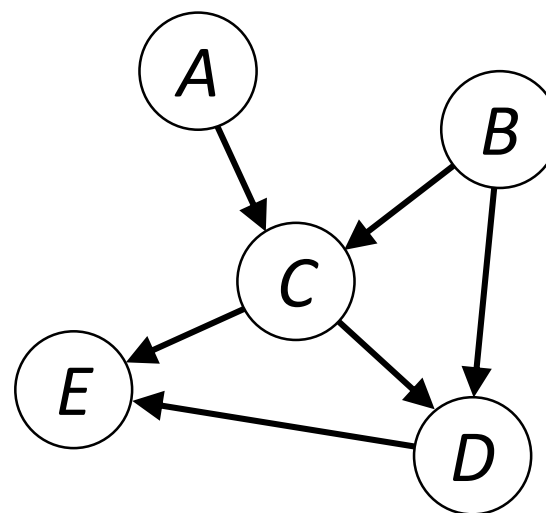
# Q1: Bayes net

Joint distribution:

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

CPTs:

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



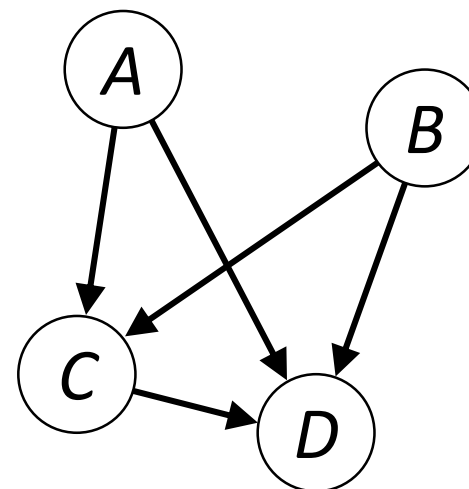
# Q1: Bayes net

Joint distribution:

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

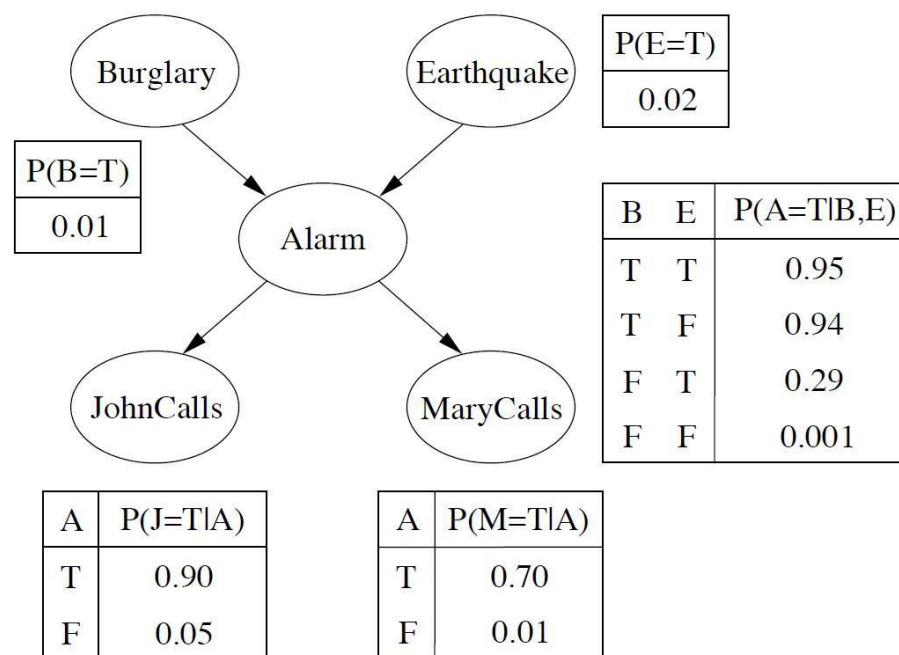
CPTs:

Node	# free params
<i>A</i>	$2^0 = 1$
<i>B</i>	$2^0 = 1$
<i>C</i>	$2^2 = 4$
<i>D</i>	$2^3 = 8$
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:**

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

$$f_{j,m}(A) = \begin{array}{c|c} A & f_{j,m}(A) \\ \hline 0 & 0.0005 \\ \hline 1 & 0.63 \end{array} = \begin{array}{c|c} A & p(j|A) \\ \hline 0 & 0.05 \\ \hline 1 & 0.90 \end{array} * \begin{array}{c|c} A & p(m|A) \\ \hline 0 & 0.01 \\ \hline 1 & 0.70 \end{array}$$

$$f_{j,m}(B, E) = \begin{array}{c|c|c} B & E & f_{j,m}(B, E) \\ \hline 0 & 0 & 0.0011295 \\ \hline 0 & 1 & 0.183055 \\ \hline 1 & 0 & 0.59223 \\ \hline 1 & 1 & 0.598525 \end{array} = \begin{array}{c|c|c|c} & & \overbrace{p(A|B, E)} & \\ B & E & A=0 & A=1 \\ \hline 0 & 0 & 0.999 & 0.001 \\ \hline 0 & 1 & 0.71 & 0.29 \\ \hline 1 & 0 & 0.06 & 0.94 \\ \hline 1 & 1 & 0.05 & 0.95 \end{array} \times \begin{array}{c|c} A & f_{j,m}(A) \\ \hline 0 & 0.0005 \\ \hline 1 & 0.63 \end{array}$$

Marginalising over  $A$

# Q2: Variable elimination

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

$B$	$f_{j,m}(B)$
0	0.00476801
1	0.5923559

$$=$$

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

$$\times$$

$E$	$p(E)$
0	0.98
1	0.02

Marginalising over  $E$

$$p(B, j, m) =$$

$B$	$p(B)$
0	0.0047203299
1	0.005923559

$$=$$

$B$	$p(B)$
0	0.99
1	0.01

$$*$$

$B$	$f_{j,m}(B)$
0	0.00476801
1	0.5923559

**Denominator:**

$$p(j, m) = \sum_B p(B, j, m) = 0.0047203299 + 0.005923559 = 0.0106438889$$

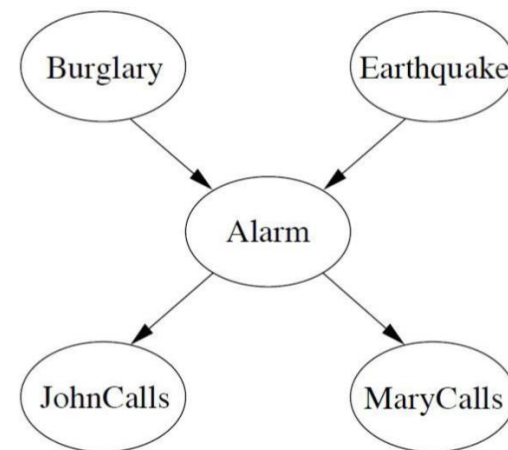
Putting the results together:

$$p(B = 1 | j, m) = \frac{p(B = 1, j, m)}{p(j, m)} = \frac{0.005923559}{0.0106438889} = 0.5565$$

## Q3: Independence

Returning to the previous Bayes net:

- Are the 'Burglary' and 'Earthquake' nodes independent?
- What if we observe 'MaryCalls' = T?
- What if we observe 'Alarm' = T?



# Q3: Independence

- (Marginal) independence:

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

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

- (Conditional) independence when 'MaryCalls' = T:

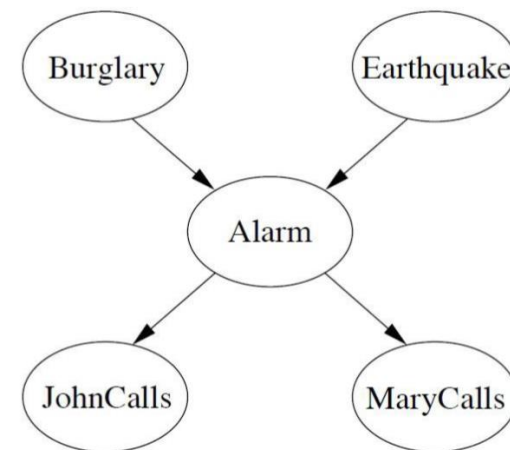
$$p(B, E|m) \propto p(B)p(E) \sum_A p(A|B, E) \underbrace{\sum_J p(J|A)}_{=1} p(m|A)$$

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

- (Conditional) independence when 'Alarm' = 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

# Stan Demo

# What is Stan?

- A probabilistic programming language
- Workflow:
  - \* declare data and parameters
  - \* declare log posterior
  - \* Stan automates the inference (MCMC, VB or MLE)
- Interfaces for R, Python, MATLAB, Julia, Stata
- Official website: <http://mc-stan.org>
- Learn more: [slide deck](#)

Some alternatives:

- PyMC3 (soon PyMC4)
- TensorFlow probability

# Nuclear power plant demo

See `nuclear.Rmd`