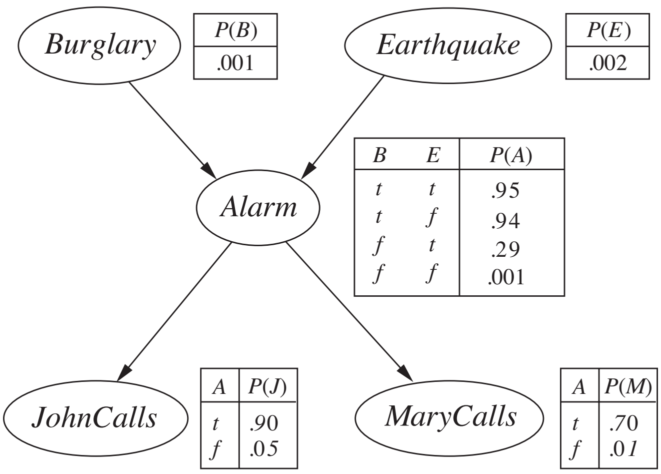
Worksheet 12: PGMs

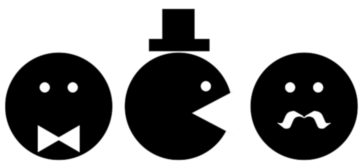
COMP90051 Statistical Machine Learning

Semester 1, 2023

Exercise 1 (Based on RN 14.15)**.** Leo is a botanist who lives in the Bay Area. His neighbourhood is a hotspot for burglars, so his house is fitted with an alarm system. Unfortunately, the alarm is not perfectly reliable: it doesn't always trigger during a home invasion, and it may be erroneously triggered during minor earthquakes, which occur occasionally. Leo has asked his neighbours John and Mary (who don't know each other) to call him if they hear the alarm. Leo would like to determine the likelihood that his home is being burgled if he receives a phone call from both John and Mary. Use Bayes' theorem and marginalisation and the following PGM to help him answer this question.



Exercise 2. Mr. and Ms. Pacman have been searching for each other in the Pacman world [1]. Ms. Pacman has been pregnant with a baby, and this morning she has given birth to Pacbaby (congratulations, Pacmans!). To train Pacbaby to avoid encountering ghosts in the maze,[2] the Pacmans are trying to teach Pacbaby to distinguish Pacmen (pl.) from ghosts using discriminative visual features such as the presence of a bowtie, hat, mustache, etc.

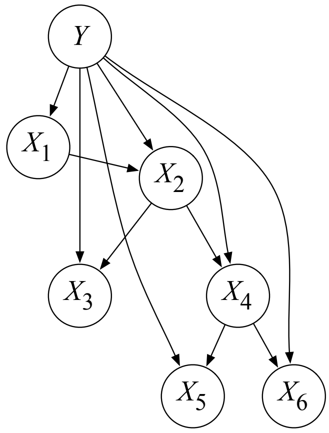


Pacbaby has noticed that the features are not independent—nearly everyone who has a hat has a mustache, while those with bowties are always clean shaven. She decides to use a tree-augmented Naive Bayes model (TANB) to account for conditional dependencies. A TANB

[1] Based on Berkeley CS188 section, see http://ai.berkeley.edu/project_overview.html

[2] Ghosts are nice enough not to eat Pacbaby, but they will take all her money.

is an extension of a Naive Bayes model, where features are no longer assumed conditionally independent given the binary class $Y \in \{1, -1\}$ (Pacman or not-Pacman, respectively). Let $X_1, X_2, \ldots, X_6$ be the random variables corresponding to the features that Pacbaby observes. The TANB model arranges vertices in a tree-structured Bayes net with Y at the root:



(a) Assume all features $X_1, \ldots, X_6$ are observed in the TANB model. What is the classification rule? Your answer should be in terms of the prior and conditional probabilities.

When we perform a marginalisation operation—i.e. removing a variable from a joint distribution, we perform a sum over the product of all factors that include that random variable. For example, marginalising over X_4 in the joint distribution above involves a factor containing four random variables.

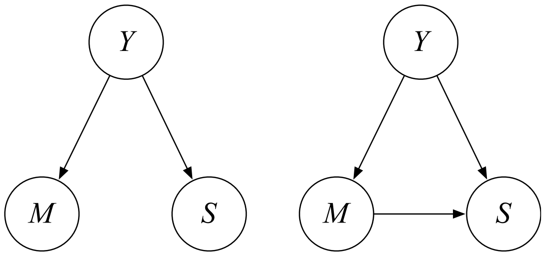
$$\sum_{X_4} \underbrace{p(X_4|X_2, Y)p(X_5|X_4, Y)p(X_6|X_4, Y)}_{\phi(X_2, X_4, X_5, X_6)}$$

This induces a dependency between all the random variables in the factor except the variable being marginalised—all subsequent operations will have to treat X_2, X_5, X_6 together (X_4 is summed out). Assuming there is no special algebraic structure in the summand that can be exploited, the complexity is exponential in the number of different random variables in the summand. Thus the overall complexity of the variable elimination algorithm is dominated by the number of variables in the largest elimination factor, $\phi(\ldots)$. Determining the optimal (lowest-complexity) elimination ordering is intractable, but a useful heuristic is to find an ordering that minimises the size of the largest factor generated.

(b) Specify an elimination order that is efficient for the query $p(Y|X_5 = x_5)$ in the TANB model above. How many variables are in the biggest factor induced by variable elimination with your ordering? Which variables are they?

(c) Specify an elimination order that is efficient for the query $p(X_3|X_5 = x_5)$ in the TANB model above. How many variables are in the biggest factor induced by variable elimination with your ordering? Which variables are they?

Exercise 3. Consider the Bayes nets below over the nodes Y (Pacbaby sees Pacman or not), M (Pacbaby sees a moustache), and S (Pacbaby sees sunglasses).



(NB) (TANB)

Empirically:

- Pacbaby observes $Y = 1$ or $Y = -1$ (Pacman or not) 50% of the time.
- Given $Y = 1$, Pacbaby observes $M = 1$ (moustache) 50% of the time and $S = 1$ (sunglasses) 50% of the time.
- When Pacbaby observes $Y = -1$, the frequency of observations are identical (equal probabilities of $M = 1, -1$, $S = 1, -1$).
- When Pacbaby observes $Y = 1$, anyone with a moustache wears sunglasses and anyone without a moustache does not wear sunglasses.
- If $Y = -1$ the presence/absence of a moustache has no influence on sunglasses.

(a) Based on the above information, fill in Pacbaby's conditional probability tables.

(b) Pacbaby sees someone with a moustache and wearing a pair of sunglasses. What prediction does the NB model make? What probability does it assign to its prediction? What prediction does Pacbaby's TANB model make? What probability does it assign to its prediction?