Probabilistic Modelling and Reasoning, Tutorial Answer Sheet 2 (for Week 4)

School of Informatics, University of Edinburgh

Instructor: Amos Storkey

Prob. Sheet. ID: 22704.11

Thanks also to Kian Ming Adam Chai, Chris Williams for some of the answers.

1. Conditioned on s, we marginalise out (b, g, t), and query f.

$$p(f = \text{empty}|s = \text{no}) = \frac{p(f = \text{empty}, s = \text{no})}{p(s = \text{no})} = \frac{p(f = \text{empty}, s = \text{no})}{\sum_f p(f, s = \text{no})} \tag{1}$$

Below we use the decomposition of the joint probability according to the DAG: $p(b, f, g, t, s) = p(b)p(f)p(t|b)p(g|b, f)p(s|t, f)$.

$$\begin{aligned}
p(f, s{=}\text{ no}) &= \sum_{g,t,b} p(b, f, g, t, s{=}\text{ no}) && \text{(marginalisation)} \\
&= \sum_{g,t,b} p(b)p(f)p(t|b)p(s = \text{no}|t, f)p(g|b, f) && \text{(factorisation)} \\
&= p(f) \sum_t p(s = \text{no}|t, f) \sum_b p(b)p(t|b) \sum_g p(g|b, f) && (2) \\
&= p(f) \sum_t p(s = \text{no}|t, f) \sum_b p(b)p(t|b) && (\textstyle\sum_g p(g|b, f) = 1) \\
&= p(f) \sum_t p(s = \text{no}|t, f)p(t) && (\text{as } \textstyle\sum_b p(b)p(t|b) = p(t))
\end{aligned} \tag{3}$$

Using the CPTs in the figure we have

$$\begin{aligned}
p(t = \text{no}) &= p(b = \text{bad})p(t = \text{no}|b = \text{bad}) \\
&\quad + p(b = \text{good})p(t = \text{no}|b = \text{good}) = 0.98 \times 0.02 + 0.03 \times (1 - 0.02) = 0.049, \\
p(t = \text{yes}) &= 1 - 0.049 = 0.951.
\end{aligned}$$

Setting $f = \text{empty}$ we have

$$\begin{aligned}
p(f = \text{empty}, s = \text{no}) &= p(f = \text{empty}) \sum_t p(s = \text{no}|t, f = \text{empty})p(t) \\
&= 0.05[1.0 \times 0.049 + 0.92 \times 0.951] = 0.0462
\end{aligned}$$

and for $f =$ not empty

$$p(f = \text{not empty}, s = \text{no}) = p(f = \text{not empty}) \sum_t p(s = \text{no}|t, f = \text{not empty})p(t)$$
$$= 0.95[1.0 \times 0.049 + 0.01 \times 0.951] = 0.0556.$$

Thus

$$p(f = \text{empty}|s = \text{no}) = \frac{p(f = \text{empty}, s = \text{no})}{p(s = \text{no})} = \frac{0.0462}{0.0462 + 0.0556} = 0.4538.$$

2. This is a practical example - you should talk through the implementation and queries involved.

3. "Given a set of conditioning nodes C, if *every* path from any node in set A to any node in set B is blocked, then A is said to be d-seperated from B by C. This implies $I(A, B|C)$."

   "A path is said to be *blocked* if it includes a node such that either

   (a) the arrows on the path do not meet head-to-head at the node, and the node is in the conditioning set, or

   (b) the arrows do meet head-to-head, and neither the node, nor any of its descendents, is in the conditioning set."

   (a) $T - E - L - S$ is unblocked at E (head-to-head) because D is a descendent of E; unblocked at L (head-to-tail) because L is not in the evidence set. There exists a path that is unblocked, hence $\neg I(T, S|D)$.

   (b) i. $L - E - D - B$ is blocked at D (head-to-head) because $D \notin \{S\}$ and S is not a descendent of D.

   ii. $L - S - B$ is blocked at S (tail-to-tail) because $S \in \{S\}$.

   All paths are blocked, hence $I(L, B|S)$.

   (c) i. $A - T - E - L - S$ is blocked at L (head-to-tail) because $L \in \{L\}$.

   ii. $A - T - E - D - B - S$ is blocked at D (head-to-head) because $D \notin \{L\}$.

   All paths are blocked, hence $I(A, S|L)$.

   (d) $A - T - E - D - B - S$ is unblocked at

   i. D (head-to-head) because $D \in \{L, D\}$;

   ii. T, E, B are all head-to-tail and are not in the evidence set.

   $A - T - E - D - B - S$ unblocked, hence $\neg I(A, S|L, D)$. Note that $A - T - E - L - S$ remains blocked.

   Alternatively you could convert this into a directed factor graph and use factor graph separation rules. The results are the same.

4. This is about conversion between graphs. To convert from the directed graph to a Markov Network, we need to marry the parents. Try to insure to understand why marrying parents is necessary. To know it is minimal given the directed graph (but not necessarily minimal), it suffices to establish that marrying parents is necessary, as all other edges are clearly necessary, because a link $A \rightarrow B$ in the directed graph implies A, B are dependent given the rest and so those links are all necessary in the undirected case. Marrying parents is necessary as $I(A, B|\text{child node}, \text{all the rest})$ does not hold in the directed graph, but would hold in the undirected graph and so these links must be added.

   To get the factor graph from the directed graph, just make each $P(X|Pa(X))$ in the original directed graph a factor. Remember factor nodes are squares, variable nodes are circles and an edge F-V means variable V is part of the factor F.

   To get the factor from the undirected graph, find all cliques of the graph (Battery, Fuel, Gauge), (Battery, Fuel, Turn Over), (Turn Over, Fuel, Start). We need to allocate the conditional probabilities to potentials e.g. $P(B)P(F)P(G|B, F)$ to the first. $P(T|B)$ to the second, and $P(S|T, F)$ to the third. This is not unique. Factors are only defined up to a multiple, but also $P(B)$ and $P(F)$ could be allocated to the B-T-F factor instead.