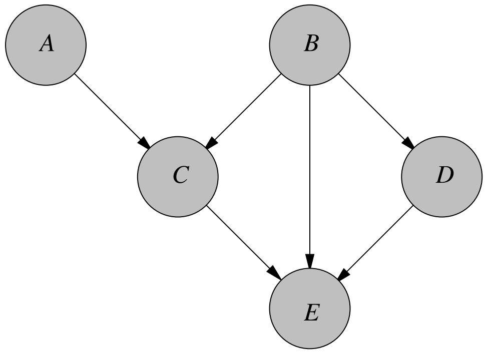
COMPUTER SCIENCE TRIPOS Part II – 2021 – Paper 9

10 Machine Learning and Bayesian Inference (sbh11)

Consider the following Bayesian Network:



All random variables (RVs) are Boolean. For an RV R we denote $R = T$ by r and $R = F$ by $\overline{r}$. We have $\Pr(a) = 0.1$, $\Pr(b) = 0.2$, $\Pr(d|b) = 0.7$ and $\Pr(d|\overline{b}) = 0.4$. For the remaining RVs we have

A	B	$\Pr(c\|A,B)$
F	F	0.2
F	T	0.2
T	F	0.5
T	T	0.6

C	B	D	$\Pr(e\|C,B,D)$
F	F	F	0.3
F	F	T	0.5
F	T	F	0.6
F	T	T	0.3
T	F	F	0.1
T	F	T	0.2
T	T	F	0.1
T	T	T	0.9

In this question you must use the *Variable Elimination* algorithm to compute $\Pr(A|\overline{e})$. You should begin with the factorisation

$$\Pr(A|\overline{e}) = \Pr(A) \sum_B \Pr(B) \sum_C \Pr(C|A,B) \sum_D \Pr(D|B) \Pr(\overline{e}|B,C,D).$$

You should express factors as tables of integers, leaving any necessary normalisation until the final step in Part (d).

(a) Define *conditional independence* of two RVs X and Y with respect to a third RV Z. [2 marks]

(b) Deduce the factor $F_{E,\overline{D}}(B,C)$ corresponding to the summation over D. [8 marks]

(c) Deduce the factor $F_{E,\overline{D},\overline{C}}(A,B)$ corresponding to the summation over C. [6 marks]

(d) Complete the computation to find the distribution $\Pr(A|\overline{e})$. [4 marks]