Solution notes

Artificial Intelligence II 2005 – Paper 8 Question 2 (SBH)

This addresses the section of the course devoted to the use of probability for dealing with uncertainty, and specifically the use of basic Bayesian networks.

(*a*) The equation is

$$\Pr(Q|e) = \frac{1}{Z}\Pr(Q,e) = \frac{1}{Z}\sum_u \Pr(Q,e,u)$$

and so the process is merely one of summing over the joint distribution. The problems are that the joint distribution in any reasonable problem will be a huge table meaning the computational complexity in terms of both space and time is a limiting factor. Also actually populating the table with numbers is not feasible.

(*b*) A Bayesian network is a directed acyclic graph in which each node is a RV with associated distribution $\Pr(N_i|\text{parents}(N_i))$. It represents the joint distribution of all the RVs as

$$\Pr(N_1 = n_1, \ldots, N_n = n_n) = \prod_{i=1}^{n} \Pr(N_i = n_i|\text{parents}(N_i))$$

(*c*) Conditional independence: $\Pr(AB|C) = \Pr(A|C)\Pr(B|C)$.

The simplification is then best illustrated by writing the joint distribution as a product of conditional distributions and using the fact that conditional independence allows us to make simplifications along the lines of $\Pr(A|BC) = \Pr(A|C)$, thus allowing us to obtain a complete representation of the joint distribution using a produce of relatively simple conditional distributions. Naive Bayes takes this to an extreme using

$$\Pr(A, B_1, \ldots, B_n) = \Pr(A)\prod_{i=1}^{n}\Pr(B_i|A)$$

(*d*) Again, any reasonable material is acceptable. For example: the need to include continuous RVs by introducing usable standard distributions, and the fact that computational complexity remains an issue that requires the introduction of more subtle exact algorithms or of approximation algorithms such as Markov chain Monte Carlo.