[4] Entropy and Mutual Information

Entropy

Entropy is a measure of unpredictability in a system Bedford and Cooke (2001). In a Bayesian Network, higher entropy means the node states are more randomly distributed, therefore more uncertain. Zero entropy means the state is known exactly (with a probability of 1). Maximum entropy corresponds to all states having equal probability - maximum unpredictability. For a variable (node) X, with n discrete states, the entropy $H(X)$ is defined as follows:

$$H(X)=\sum_{i=1}^{n} P(x_i)\log_2\left(\frac{1}{P(x_i)}\right) \qquad (1)$$

Where $P(x_i)$ is the probability of X being in state x_i. The maximum entropy for any given node will therefore depend on the number of possible states.

Mutual Information

The mutual information (MI) of nodes X and Y, $MI(X,Y)$, is the reduction of uncertainty in Y by knowing X. MI can be computed for any pair of nodes, regardless of whether they are directly connected. Mutual information is used in algorithms for estimating network structure from data. High MI between a pair of nodes indicates that an arc exists (in some direction). Low MI, typically defined by some cut-off point, implies conditional independence (Kane et al., 2003). As the BN presented here has been developed using expert opinion (rather than observational data), MI gives a measure of the perceived strength of connection between nodes.

For two variables X and Y with discrete states, mutual information $MI(X,Y)$ is defined as:

$$MI(X,Y)=H(Y)-H(Y|X)=H(X)-H(X|Y) \qquad (2)$$

Where $H(X)$ denotes the entropy of node X (Ebert-Uphoff , 2007). The function is symmetric in X and Y, and does not indicate the direction of influence (does not distinguish X as a parent of Y or vice versa).

Mutual information percentage (MI%) is the percentage reduction of uncertainty in target node Y due to knowledge of X, and is not symmetric in X and Y:

$$MI\%(X,Y)=\frac{H(Y)-H(Y|X)}{H(Y)}100 \qquad (3)$$

Bedford, T., Cooke, R., 2001. Probabilistic Risk Analysis, Foundations and Methods. Cambridge University Press. 393pp.

Ebert-Uphoff, I., 2007. Users Guide for the LinkConnectionStrength Package (Version 1.0) – A PNL Package for the Calculation and Visualization of Entropy, Link Strengths and Connection Strengths in Discrete Bayesian Networks. http://www.dataonstage.com/PNL/PACKAGES/PTG_LCS/PTG_LCS.html (accessed 5/6/13)

Kane, M., Sahin, F., Savakis, A., 2003. A two phase approach to Bayesian network model selection and comparison between the MDL and DGM scoring heuristics. International Conference on Systems, Man and Cybernetics. IEEE 5, :4601–4606.