Hierarchical Bayesian Networks

Elias Gyftodimos Peter Flach

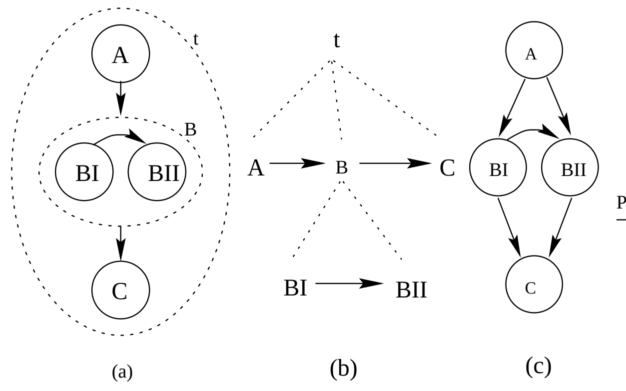
June 14, 2002

Bayesian Networks are a useful tool for probabilistic inference, but are unable to deal with non-propositional domains. Aggregate data types can be represented either using a flattened representation, which results in loss of expressivity, or by representing the aggregate type as a single node, disregarding possible probabilistic dependencies between the components of the structure.

Hierarchical Bayesian Networks are a generalisation of standard Bayesian Networks, where a node in the network may be an aggregate data type. This allows the random variables of the network to represent arbitrarily structured types. Within a single node, there may also be links between components, representing probabilistic dependencies among parts of the structure.

Hierarchical Bayesian Networks encode conditional probability dependencies the same way as standard Bayesian Networks. Hierarchical Bayesian Networks can express further knowledge about variable structure and use that knowledge to build more realistic probabilistic models.

A Hierarchical Bayesian Network consists of two parts. The *structural part* contains the variables of the network and describes the *part-of* relationships and the *probabilistic dependencies* between them. The *part-of* relationships in a structural part may be illustrated either as nested nodes (Fig. 1(a)) or as a tree hierarchy (Fig. 1(b)). The second part of a Hierarchical Bayesian Network, the *probabilistic part*, contains the conditional probability tables that quantify the links introduced at the structural part.



P(A)	a1	a2
	0.3	0.7

P(BII|A,BI)	bII1	bII2	bII3
a1bI1	0.3	0.4	0.3
a1bI2	0.1	0.5	0.4
a2bI1	0.7	0.1	0.2
a2bI2	0.9	0.1	0

P(BI|A)	bI1	bI2
a1	0.4	0.6
a2	0.8	0.2

P(C|BI,BII)	c1	c2	c3
bI1bII1	0.2	0.6	0.2
bI1bII2	0.3	0.4	0.3
bI1bII3	0.2	0.2	0.6
bI2bII1	0.5	0.3	0.2
bI2bII2	0.6	0.2	0.2
bI2bII3	0.7	0.2	0.1

(d)

Figure 1: A simple Hierarchical Bayesian Network. (a) Nested representation of the network structure. (b) Tree representation of the network structure. (c) Standard Bayesian Network expressing the same dependencies. (d) Probabilistic part.

Currently, the only aggregation our model supports for composite nodes is a tuple. We are planning to introduce more aggregation operators for types, such as lists and sets. The calculation of the joint probability will in that case involve the assumption of a probability distribution over variable length structures, as are lists, sets and multisets.

Our present work is focused on designing a Bayesian method for learning the structure of a Hierarchical Bayesian Network given a set of data. Assuming that a type structure is known, the problem lies in identifying the network that could most likely have generated the observed data.