Joint Probability Distribution.
Analysis with the HUGIN Decision Engine
Based on the HUGIN Training Course and HUGIN Manual

Prof. Dr. Galia Weidl

Basics of Decision-Making

TH Aschaffenburg

WS2021/2022

Functionality

The HUGIN Decision Engine (HDE) is the inference engine. The HDE supports functionality such as

- Construction (how to build a Bayesian Netowrk (BN) model?)
- Compilation (how to run/execute the developed BN model?)
- Inference (how to introduce evidence in BN and update the probabilities?)
- Analysis (how to get more insights on the variables in the BN model?)

Analysis

The HUGIN Decision Engine (HDE) is the inference engine. The HDE supports functionality such as Methods for Analysis

- Joint Analysis
- The joint Analysis can be used to identify the most probable configurations over a set of nodes when their joint is too large to be represented in main memory or if the joint cannot be computed.
-

https://download.hugin.com/webdocs/manuals/8.9/htmlhelp/pages/Manual/MethodsOfAnalysis/JointConfigurations.html

Joint Probability Distribution

Figure 1 shows the joint analysis pane appearing after activating the joint configurations dialog. In this pane the joint probability distribution over a set of nodes is shown (Figure 3). The set of nodes in the joint can be defined by selecting the "Variables" tab (Figure 2).
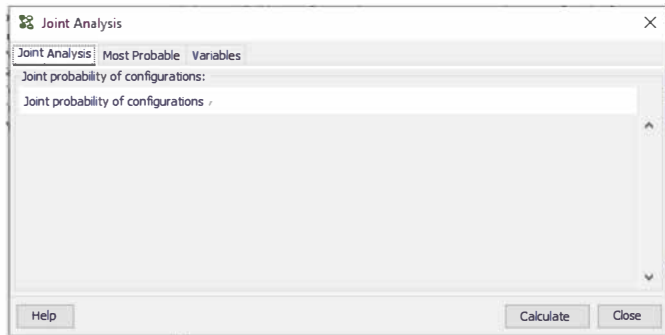

Figure 1: No joint distribution has been computed.

A joint probability distribution is computed over a set of nodes. The set of nodes can be selected as indicated in Figure 2.
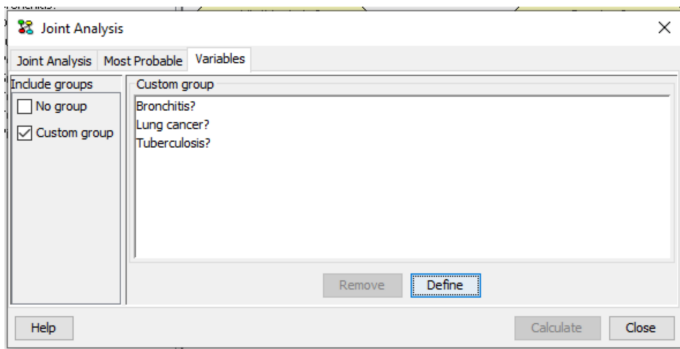

Figure 2: Selecting the set of nodes in the joint distribution.

Selecting nodes to appear in the joint proceeds in the same way as selecting *Target(s)* of *Instantiations* in the d-Separation pane.

After selecting the set of nodes their joint probability distribution is computed by selecting *Calculate*. The results for the example are shown below in Figure 3 where the probabilities are sorted.
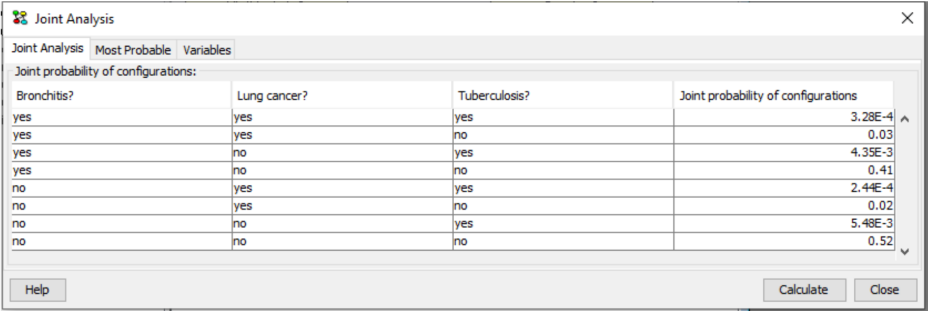

Figure 3: The joint distribution over the selected set of nodes.

The joint probability distribution over a set of nodes is computed by message passing in a junction tree. The feasibility of computing a joint distribution is in part determined by the size of the junction tree and the combined state space size of the nodes in the joint. The size of the joint distribution is equal to the product of the state space sizes of the nodes in the jiont. Thus, the size of the joint distribution grows exponentially with the number of nodes in the joint.

Notice that the most probable configurations can be identified by sorting the rightmost column. This do, however, require that the full joint is computed. Alternatively, the most probable configurations of a subset of nodes can be identified using the *Most probable* tab.
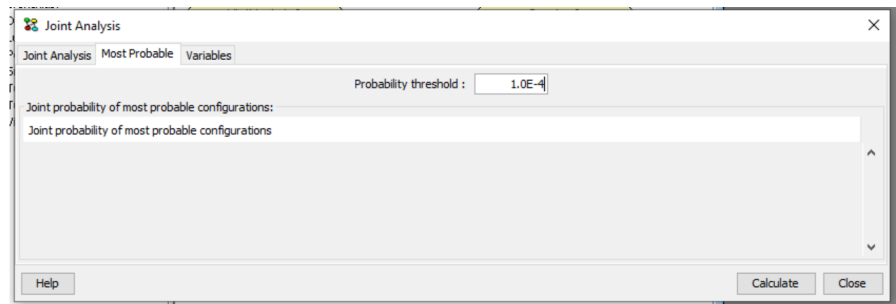

Figure 4: The most probable configurations over the selected set of nodes with a probability higher than the threshold.

This dialog can be used to identify the most probable configurations over a set of nodes when their joint is too large to be represented in main memory or if the joint cannot be computed. The user is asked to supply a threshold on the probability of the configurations to be included in the result. The lower the thredhold is, the longer the algorithm has to run.

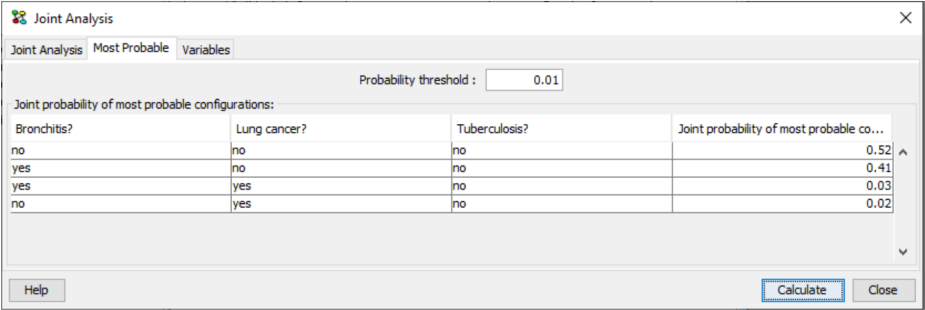The result for the example above with a threshold of 0.01 is shown below.

Figure 5: The most probable configurations with a probability higher than the defined threshold.

This method is not supported for networks with decisions.

Remark on Junction Tree and Cliques

The compilation of a HUGIN network involves a transformation of the network to a junction tree in which the propagations can take place - you need as a user, interested only in application, just to press the "run" button (called also: Compilation tab).

You can specify certain properties on the compilation in HUGIN from the Network Properties of the Network Menu. Figure 2 shows the Compilation tab of the Network Properties dialog box.

Most practical applications use default triangulation method, i.e. you do not need to specify anything, since triangulation is done at the compilation automatically. Generally, it is NP hard (mathematical way to say that it is impossible within reasonable time) to find the best internal representation of a HUGIN network (the junction tree representation). This is basically because it is NP hard to triangulate a graph optimally.

Figure 2. Compilation tab of the Network Properties dialog box.

In many cases, however, an optimal (or near-optimal) triangulation can in fact be found. The HUGIN Decision Engine offers a method for optimal triangulation, which in many cases results in junction trees of much lower complexity than those generated using the other (heuristic) triangulation methods provided. The optimal triangulation method works by searching for minimal separators. For some networks this results in extremely many possible optimal separators. You do not need (in this course) to spend time to understand how this works.
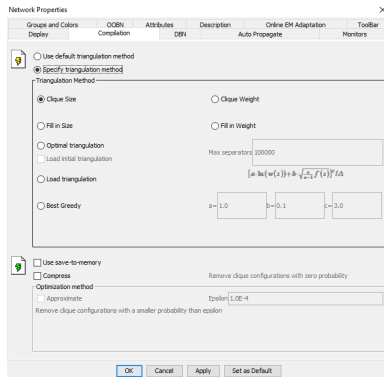
Remark on Junction Tree and Cliques

Some general information on the junction tree algorithm is given here:
https://en.wikipedia.org/wiki/Junction_tree_algorithm

The junction tree algorithm (also known as 'Clique Tree') is a method used in machine learning to extract marginalization in general graphs. The junction tree algorithm is executed automatically in the background when you switch from edit mode to run mode.

And **you Do Not Need to understand it when you are working with applications** of Bayesian Networks.
If you read a scientific paper, which mentions it shortly - just read the rest - you need also to learn how to read research papers.

The same comment is relevant about **Clique**, which is a notion in Graph theory (requires stong mathematics specialization).

We do not deal with cliques in ths cource which is focused just on **basics** of decision making and on **simple applications**.

https://en.wikipedia.org/wiki/Clique_(graph_theory)

How to see the Junction Tree ?

If you are curious what is a junction tree look like? - you can see it in HUGIN.
Inference in Bayesian networks and influence diagrams are performed in a secondary structure known as a junction tree. The junction tree for a network is generated when one switches from Edit Mode to Run Mode. The junction tree of the network can be displayed in a separate window, opened by selecting the Show Junction Tree item of the View Menu.

https://download.hugin.com/webdocs/manuals/8.9/htmlhelp/pages/Manual/Algorithms/JunctionTrees.html

Junction tree is used in Data Conflict Analysis

This belongs to the advanced topics of the algorithm and its applications.

http://download.hugin.com/webdocs/ manuals/8.9/htmlhelp/pages/Manual/ MethodsOfAnalysis/ DataConflictAnalysis.html#data-conflict- analysis