LA-UR-12-21818
Approved for public release; distribution is unlimited.

Title: Uncertainty Quantification of Hypothesis Testing for the Integrated Knowledge Engine

Author(s): Cuellar, Leticia

Intended for: Documentation for sponsor

Uncertainty Quantification of Hypothesis Testing for the Integrated Knowledge Engine[1]

Leticia Cuéllar
D-6 Risk Analysis and Decision Support
Los Alamos National Laboratory
March, 2012

1. Introduction

The Integrated Knowledge Engine (IKE) is a tool of Bayesian analysis, based on *Bayesian Belief Networks* or *Bayesian networks* for short. A Bayesian network is a graphical model (directed acyclic graph) that allows representing the probabilistic structure[2] of many variables assuming a localized type of dependency called the Markov property. The Markov property in this instance makes any node or random variable to be independent of any non-descendant node given information about its parent. A direct consequence of this property is that it is relatively easy to incorporate new evidence and derive the appropriate consequences, which in general is not an easy or feasible task.

Typically we use Bayesian networks as predictive models for a small subset of the variables, either the leave nodes or the root nodes. In IKE, since most applications deal with diagnostics, we are interested in predicting the likelihood of the root nodes given new observations on any of the children nodes. The root nodes represent the various possible outcomes of the analysis, and an important problem is to determine when we have gathered enough evidence to lean toward one of these particular outcomes.

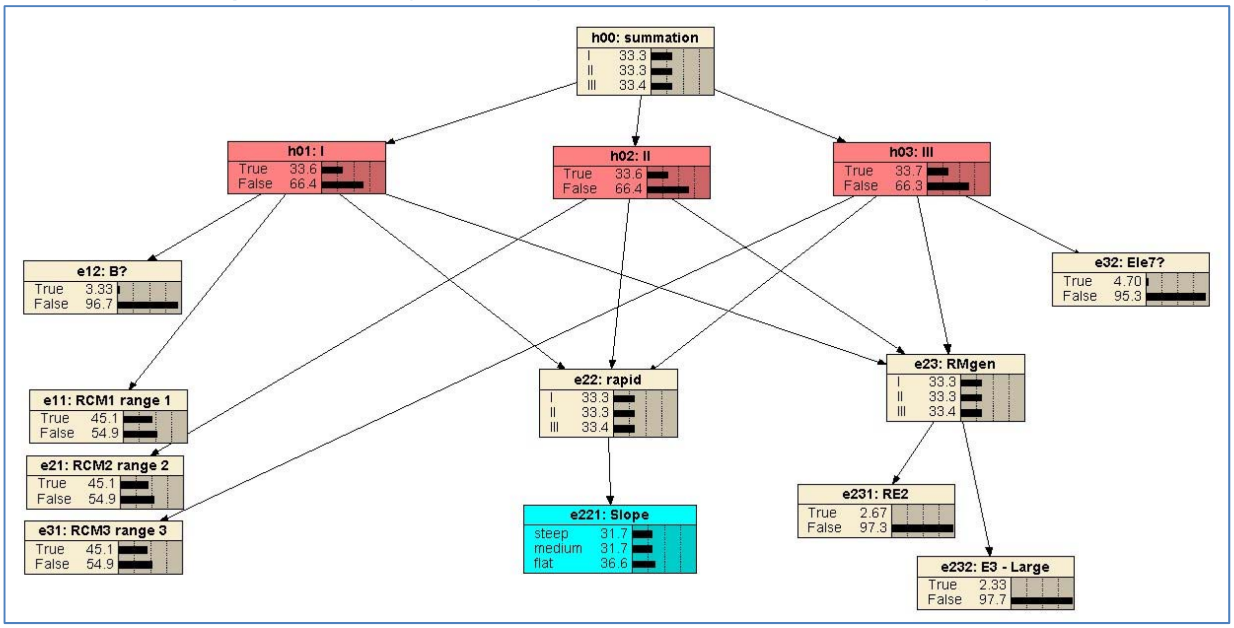
This document presents criteria to decide when the evidence gathered is sufficient to draw a particular conclusion or decide in favor of a particular outcome by quantifying the uncertainty in the conclusions that are drawn from the data. The material in this document is organized as follows: Section 2 presents briefly a forensics Bayesian network, and we explore evaluating the information provided by new evidence by looking first at the posterior distribution of the nodes of interest, and then at the corresponding posterior odds ratios. Section 3 presents a third alternative: Bayes Factors. In section 4 we finalize by showing the relation between the posterior odds ratios and Bayes factors and showing examples these cases, and in section 5 we conclude by providing clear guidelines of how to use these for the type of Bayesian networks used in IKE.

[1] LA-UR 12-xxxxx
[2] The joint probability distributions of all the variables represented in the network.

2. A Forensics Bayesian Network and Posterior Probabilities

We use IKE in many different situations that involve large data streams where important decisions need to be made based upon these data. Particular instances include monitoring, surveillance and forensics. To illustrate, the typical Bayesian network that IKE employs for forensics analysis, uses evidence variables to try to determine the type of device. The N devices under consideration are represented as the parent nodes and are binary variables taking a value of 1 if the device was detonated, and 0 otherwise (equivalently they could take the values *true* or *false*). These variables are named Hypothesis 1 through Hypothesis N. Figure 1 depicts a simplified version of such network where the hypotheses are labeled h_{01}, h_{02}, and h_{03}.

Figure 1. A simplified Bayesian network for Forensics Analysis



Note that node h_{00} is an artifact node to ensure that the Bayesian network captures the fact that hypotheses 1 through N are mutually exclusive to account for the fact that only one device was detonated. A direct consequence of this assumption is that the individual probabilities of each device having been the one detonated have to add up to one.

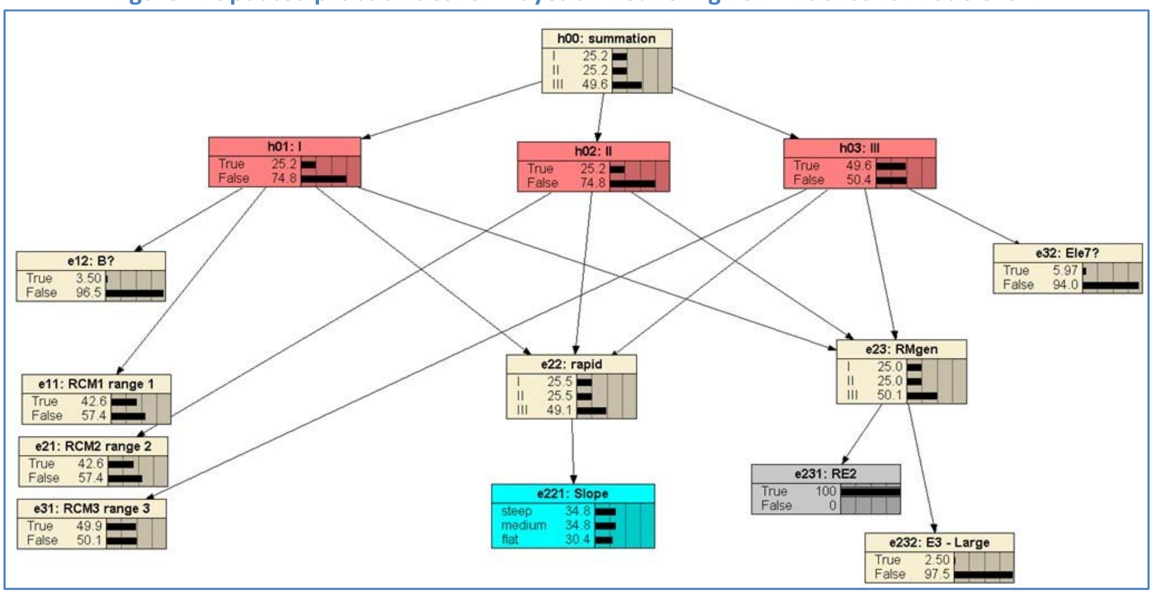
2.1 Comparing Hypotheses given New Evidence using Posterior Probabilities

This construction allows comparing individual hypotheses directly. For example, *a priori*, i.e. before getting any specific information about any particular detonation or event, we typically assume that all devices are equally likely, but once evidence starts coming in, and the Bayesian network's probabilities are suitably updated, one can compare the individual hypothesis posterior probabilities. For example, assume that evidence e_{231} is determined to be true, then the posterior probability that hypotheses 1, 2 and 3 are true become 0.25, 0.25, and 0.50 respectively (see Figure 2) making hypothesis 3 twice as likely to be true than either hypothesis 1 or 2. Formally, we are considering the *posterior* probability ratio of hypothesis 3 versus hypothesis 1 (or 2) given the evidence,

$$\frac{\mathbf{P}(H_3 = 1 |\, e_{231} = 1)}{\mathbf{P}\left(H_j = 1 \middle|\, e_{231} = 1\right)} = \frac{0.5}{0.25} = 2 \qquad for\ j = 1, 2. \qquad \textbf{(1)}$$

Similarly, the *posterior* probability ratios of hypothesis 1 versus hypothesis 3 or versus hypothesis 2 given the evidence are ½ and 1 respectively: ($\mathbf{P}(H_1 = 1 |\, e_{231} = 1)/\mathbf{P}(H_3 = 1 |\, e_{231} = 1) = 1/2$ and $\mathbf{P}(H_1 = 1 |\, e_{231} = 1)/\mathbf{P}(H_2 = 1 |\, e_{231} = 1) = 1$.

Figure 2. Updated probabilities for Bayesian network given Evidence for node e231.



In this example, hypothesis 3 is the most likely. Note that since the posterior probabilities depend on the data, they are random variables. The question then becomes: how sure are we that hypothesis 3 is true?

2.2 Comparing Hypotheses given New Evidence using Posterior Odds Ratios

An alternative to the posterior probability ratio in expression **(1)** that compares H_3 to say H_1, is to consider the *posterior* odds ratio of hypothesis H_3 given the evidence e_{231}, namely

$$\frac{\mathbf{P}(H_3 = 1 |\, e_{231} = 1)}{\mathbf{P}(H_3 = 0 |\, e_{231} = 1)} = \frac{0.5}{0.5} = 1$$

Similarly, we could compute odds ratios for the other two hypotheses, for *j*=1, 2

$$\frac{\mathbf{P}(H_j{=}1|\ e_{231}{=}1)}{\mathbf{P}(H_j{=}0|\ e_{231}{=}1)} = \frac{0.25}{0.75} = \frac{1}{3}.$$

These simple comparisons can easily be obtained as more evidence becomes available by looking at the posterior probabilities of H_j given all the evidence. To look up these numbers in the Bayesian network requires the appropriate belief updating of the marginal probabilities of the network given the evidence, i.e. the individual distributions of the nodes: $\mathbf{P}(H_j = x \mid e_k)$ and consequently $\mathbf{P}(e_j = x \mid e_k)$.

A third method to decide which hypothesis is the most likely to be true (in this instance, deciding which device was actually detonated), is to calculate Bayes Factors which are presented next.

3. Bayes Factors

3.1 Definition and Selection Criteria

Definition

Bayes factors provide ways of incorporating external information into the evaluation of evidence about a hypothesis.

> Given a model selection problem in which we have to choose between two models, on the basis of observed data *D*, the plausibility of the two different models H_0 and H_1, parametrized by model parameter vectors θ_0 and θ_1 is assessed by the Bayes factor $B_{0,1}$ given by
>
> $$B_{0,1} = \frac{\mathbf{P}(D|H_0)}{\mathbf{P}(D|H_1)} = \frac{\int \mathbf{P}(\theta_0|H_0)\mathbf{P}(D|\ \theta_0, H_0)\boldsymbol{d}\theta_0}{\int \mathbf{P}(\theta_1|H_1)\mathbf{P}(D|\ \theta_1, H_1)\boldsymbol{d}\theta_1} \qquad (2)$$
>
> where $\mathbf{P}(D|H_i)$ is called the marginal likelihood for model *i* (see [2] and [4]).

Selection Criteria

How to decide in favor of hypothesis H_0 versus H_1? There are various empirical guidelines for choosing one hypothesis over another depending on the actual values of the Bayes factors. Table 1 shows guidelines provided by Jeffrey [2] and Kass & Rafftery [4]. The intention of the latter one is to provide the same scale as the familiar deviance and likelihood ratio test statistics.

Table 1 Criteria for selecting hypothesis H_1 over hypothesis H_0.

Jeffrey's Criterion		Kass & Raftery		Interpretation
$\log_{10} B_{1,0}$	$B_{1,0}$	$2\log_e B_{1,0}$	$B_{1,0}$	Evidence against H_0
[0, ½]	[1, 3.2]	[0, 2]	[1, 3]	*Not worth than a bare mention*
(½ , 1]	(3.2, 10]	(2, 6]	(3, 20]	*Substantial*
(1, 2]	(10, 100]	(6, 10]	(20, 150]	*Strong*
> 2	>100	>10	>150	*Decisive*

The accuracy of these guidelines can be tested using existing forensics data, and can be adjusted appropriately if needed.

3.2 Bayes Factors for Forensics Bayesian Networks

In our context, the model H_i corresponds to the hypothesis that device *i* was detonated, and the data D corresponds to the evidence gathered. Since Bayes factors compare two possible models or hypotheses at a time, we can either compare

i) H_j *versus* H_k *for all* $j \neq k$ or
ii) H_j *versus* $\overline{H_j}$, where $\overline{H_j}$ stands for the hypothesis H_j being false.

The first option would require N choose 2 comparisons (i.e. N(N-1)/2), while the second option would require just N comparisons. Since, the aim is to identify the only hypothesis that is true, a reasonable criteria would be to calculate the N Bayes factors B_j that compare H_j with $\overline{H_j}$ *,* and chose the one with the largest Bayes factor, namely max$\{B_j\}$.

3.2.1 Calculating the Bayesian Factor given New Evidence

To test hypothesis H_j being true versus it being false, given the first piece of evidence $\boldsymbol{E_1}$ we compute the corresponding Bayes factor as,

$$B_j = \frac{\mathbf{P}(E_1|H_j)}{\mathbf{P}(E_1|\overline{H}_j)} \quad (3)$$

The Bayes factor B_j represents the odds of observing E_1 under hypothesis H_j versus under hypothesis $\overline{H_j}$, i.e. the ratio of the likelihood that evidence E_1 is observed given that H_j is true and the likelihood that evidence E_1 is observed given that H_j is false.

The quantity in the nominator in equation (1) can easily be computed using Netica by setting hypothesis H_1 to true, and all other hypotheses H_2, H_3 and H_N to false, and by subsequently propagating the appropriate probabilities. The probability in the denominator can be re-written as:

$$\mathbf{P}(E_1|\overline{H_j}) = \frac{\mathbf{P}(E_1 \cap \overline{H_j})}{\mathbf{P}(\overline{H_j})} = \frac{\sum_{k\neq j} \mathbf{P}(E_1 \,|H_k)\mathbf{P}(H_k)}{1-\mathbf{P}(H_j)}.$$

Since H_j is false if and only if at least one of the other *N-1* hypotheses is true ($\overline{H_j} \; = \; \cup_{k\neq j} H_k$), and these are mutually exclusive, we can rewrite,

$$\mathbf{P}(E_1|\overline{H_j}) = \frac{\mathbf{P}(E_1 \cap \cup_{k\neq j} H_k)}{1-\mathbf{P}(H_j)} = \frac{\sum_{k\neq j} \mathbf{P}(E_1 \cap H_k)}{1-\mathbf{P}(H_j)}.$$

Consequently,

$$\mathbf{P}(E_1|\overline{H_j}) = \frac{\sum_{k\neq j} \mathbf{P}(E_1 \,|H_k)\mathbf{P}(H_k)}{1-\mathbf{P}(H_j)}.$$

Further assuming that *a priori* all hypotheses are equally likely (i.e. $P(H_i) = 1/N$), we can rewrite the Bayes Factor B_j from expression **(3)** to choose between hypothesis H_j versus $\overline{H_j}$ as,

$$B_j \;=\; \frac{\mathbf{P}(E_1|H_j)}{\frac{1}{N-1}\sum_{k\neq j} \mathbf{P}(E_1|H_k)}. \qquad \textbf{(4)}$$

Forensics Example

Consider the forensics network provided in Figure 1, and assume that the first piece of evidence is that e_{231} is true. The Bayes factor for testing H_3 against all other hypotheses, namely $\overline{H_3}$, is (see Figure 3 and Figure 4),

$$B_3 = \frac{\mathbf{P}(e_{231} = 1|\; H_3 = 1)}{\mathbf{P}(e_{231} = 1|\overline{H_3} = 1)} = \frac{3.96}{\frac{1}{2}\;(2.02 + 2.02)} = 1.96$$

The Bayesian network given H_1 also produces $\mathbf{P}(e_{231} = 1|\; H_1 = 1) = \; 0.0202$. Similarly, *for j=1,2*

$$B_j = \frac{\mathbf{P}(e_{231} = 1|\; H_j = 1)}{\mathbf{P}(e_{231} = 1|\overline{H_j} = 1)} = \frac{2.02}{\frac{1}{2}\;(3.96 + 2.02)} = 0.67$$

According to Table 1 none of these provides enough evidence against (or in favor) of any hypothesis.

Figure 3. Bayesian network assuming H_3 is true.

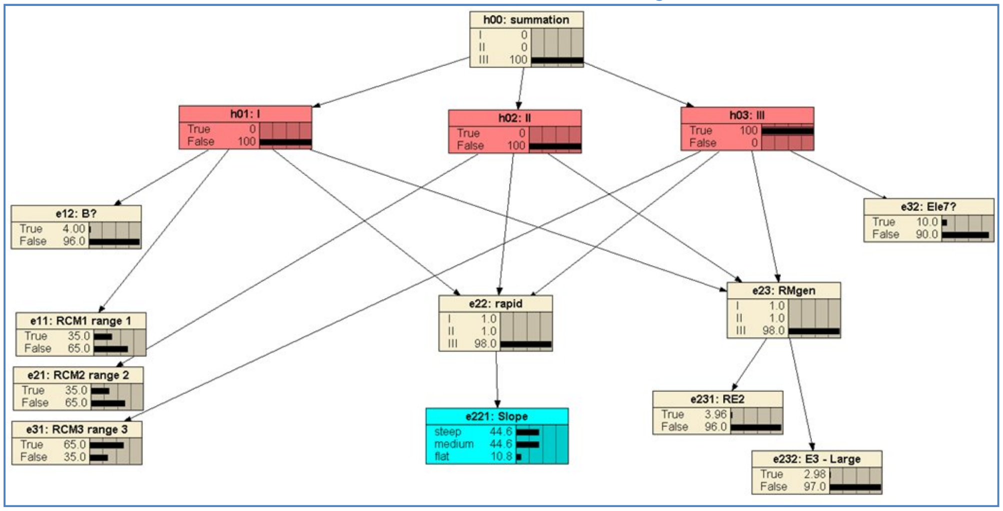
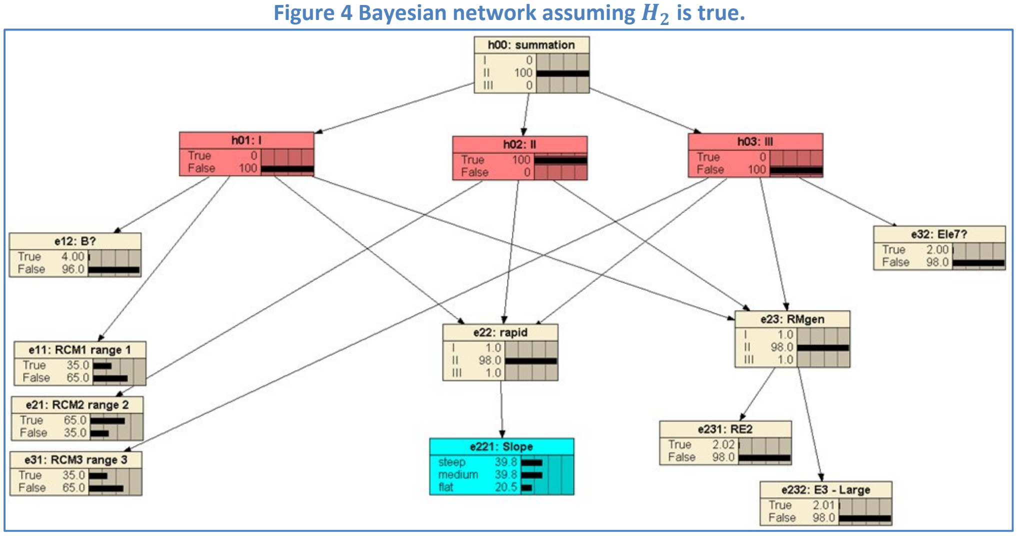


Figure 4 Bayesian network assuming H_2 is true.



3.2.2 Bayes Factor given M pieces of Evidence $E_1, \ E_2, \ldots, E_M$

As more evidence (more data) becomes available, we can compute the new Bayes factor that considers all the data available so far. Assuming M new pieces of evidence, say E_1, $E_2, \ldots, E_M$, the corresponding Bayes factor $B_{1,2}^M$ for testing hypothesis H_1 versus H_2 is given by,

$$B_{1,2}^M = \frac{\mathbf{P}(E_1, \ E_2, \ldots, E_M | H_1)}{\mathbf{P}(E_1, \ E_2, \ldots, E_M | H_2)}.$$

Fortunately, this quantity can be decomposed as the product of the odds ratios of the evidence given all prior data, namely

$$B_{1,2}^{M} = \frac{\mathbf{P}(E_1|H_1)}{\mathbf{P}(E_1|H_2)} \frac{\mathbf{P}(E_2|E_1,H_1)}{\mathbf{P}(E_2|E_1,H_2)} \times \ldots \times \frac{\mathbf{P}(E_M|E_{M-1}\ldots E_2,E_1,\ H_1)}{\mathbf{P}(E_M|E_{M-1}\ldots,E_2,E_1,\ H_2)}$$

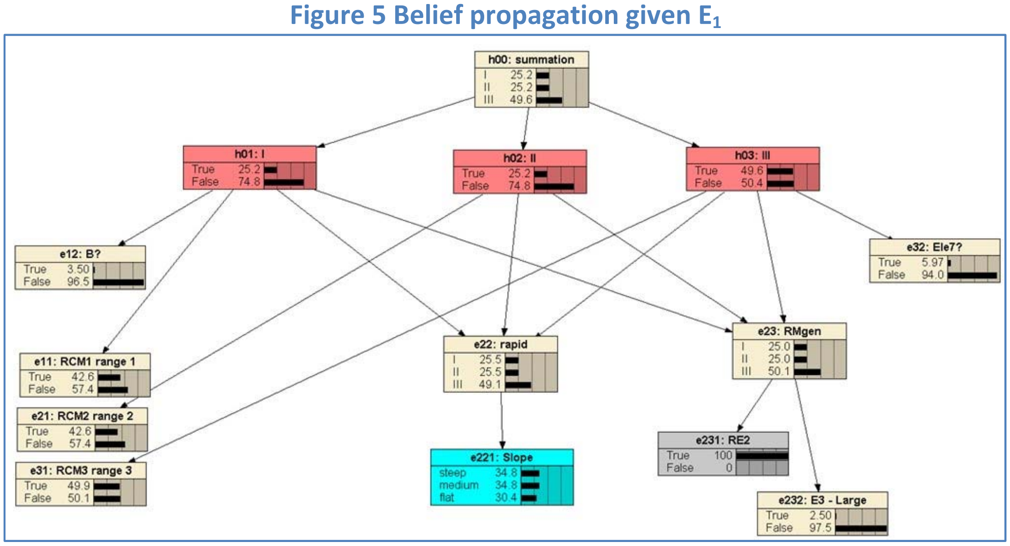
$$= \ B_{1,2}^{M-1} \times R_{1,2}^{M} \, . \qquad (5)$$

This is convenient, since as a new piece of evidence emerges, we only need to compute the ratio $R_{1,2}^{M}$ and then multiply it by the Bayes factor computed given all prior evidence (E_1, $E_2, \ldots, E_{M-1}$). Note that after computing the factor B_M, we need to update the probabilities throughout the Bayesian network by entering the new evidence (E_M) and propagating the appropriate probabilities, allowing us, in some sense, to perform sequential testing. It also is important to notice that because of the Bayesian network probabilistic structure, the probabilities $\mathbf{P}(E_M|E_{M-1} \ldots E_2,\ E_1,\ H_1)$ can be simplified in many cases, since E_M depends on $E_{M-1} \ldots E_2,\ E_1$, and H_1 only through their most common ancestor (common ancestors of E_M and $E_{M-1} \ldots E_2,\ E_1,\ H_1$).

Forensics Example

Continuing with our example, Figure 5 displays the updated Bayesian network given E_1 = {e_{231}=1}. Now, suppose that the second piece of evidence E_2 is that e_{221} has a medium slope.

Figure 5 Belief propagation given E_1



To compute the Bayes factors for say H_3 against $\overline{H_3}$ given E_1 and E_2, find first $\mathbf{P}(E_2|E_1, H_3) = 0.446$,[3] $\mathbf{P}(E_2|E_1, H_2) = 0.398$, and $\mathbf{P}(E_2|E_1, H_1) = 0.106$ (e.g. see Figure 6), and then compute,

$$R_3^2 = \frac{\mathbf{P}(E_2|E_1,H_3)}{\mathbf{P}(E_2|E_1,\overline{H_3})} = \frac{\mathbf{P}(E_2|E_1,H_3)}{\mathbf{P}(E_2|E_1,H_1)\frac{\mathbf{P}(E_1|H_1)}{\mathbf{P}(E_1|H_1)+\mathbf{P}(E_1|H_3)} + \mathbf{P}(E_2|E_1,H_3)\frac{\mathbf{P}(E_1|H_3)}{\mathbf{P}(E_1|H_1)+\mathbf{P}(E_1|H_3)}}$$

$$= \frac{44.6}{10.6\,\frac{202}{0.02.02 + 3.96} + 39.8\,\frac{3.96}{2.02 + 3.96}} = 1.77\,.$$

Finally, multiply by B_3^1, so that

$$B_3^2 = \quad R_3^2 \times B_3^1 = 1.77 \times 1.96 = 3.47. \qquad (6)$$

Similarly, one can find

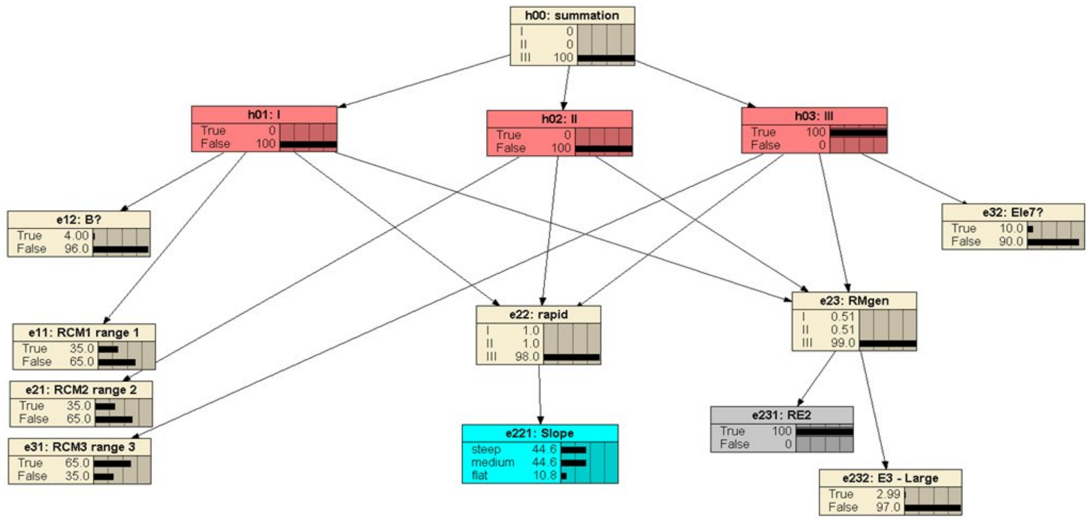
$$B_2^2 = 1.2 \times 0.68 = 0.82 \quad \text{and} \quad B_1^2 = 0.25 \times 0.68 = 0.17. \qquad (7)$$

According to Table 1, equations (6) and (7) provide substantial evidence in favor of hypothesis H_3.

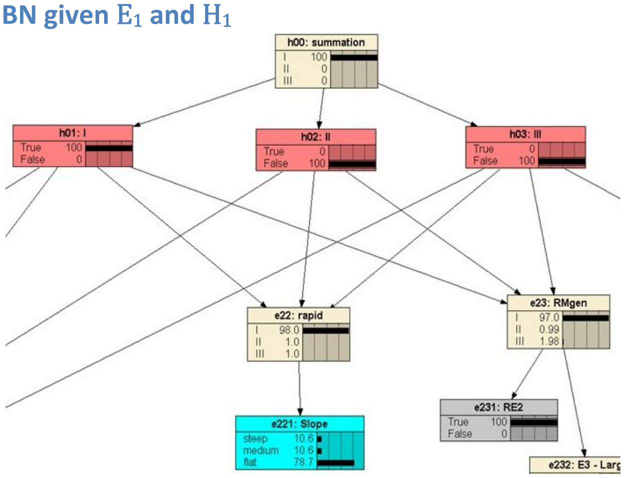
[3] Note that $\mathbf{P}(E_2|E_1, H_3) = \mathbf{P}(E_2|H_3)$, $\mathbf{P}(E_2|E_1, H_2) = \mathbf{P}(E_2|H_2)$, and $\mathbf{P}(E_2|E_1, H_1) = \mathbf{P}(E_2|H_1)$.

Figure 6 Belief propagation given E_1 and H_j

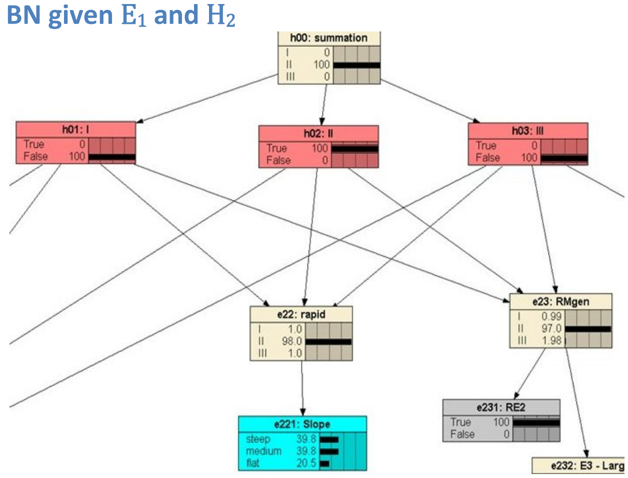
BN given E_1 and H_3



BN given E_1 and H_1



BN given E_1 and H_2



4. Relating Bayes Factors and Posterior Probabilities for IKE

As seen in the previous sections, *posterior odds ratios* and *Bayes factors* provide mechanisms to compare various hypotheses. A couple of comments: First, although we did not provide any particular criteria for what values posterior odds ratios need to achieve in order to decide in favor of a particular hypothesis, they are easier to calculate than Bayes factors (at least in IKE), but we can relate Bayes factors to posterior odds ratios as follows.

Bayes Factors. Bayes factors can be expressed as the *ratio of the posterior odds of H to its prior odds*, namely

$$B = \frac{\mathbf{P}(D|H)}{\mathbf{P}(D|\bar{H})} = \frac{\frac{\mathbf{P}(H|D)}{\mathbf{P}(\bar{H}|D)}}{\frac{\mathbf{P}(H)}{\mathbf{P}(\bar{H})}}.$$

This follows from,

$$B = \frac{\mathbf{P}(D|H)}{\mathbf{P}(D|\bar{H})} = \frac{\frac{\mathbf{P}(H,D)}{\mathbf{P}(H)}}{\frac{\mathbf{P}(\bar{H},D)}{\mathbf{P}(\bar{H})}} = \frac{\frac{\mathbf{P}(H|D)\mathbf{P}(D)}{\mathbf{P}(H)}}{\frac{\mathbf{P}(\bar{H}|D)\,\mathbf{P}(D)}{\mathbf{P}(\bar{H})}} = \frac{\mathbf{P}(H|D)}{\mathbf{P}(\bar{H}|D)}\;\frac{\mathbf{P}(\bar{H})}{\mathbf{P}(H)}.$$

Assuming N hypotheses with uniform priors, the Bayes factor for testing hypothesis H against $\bar{H}$ can be calculated simply as

$$B = \frac{\mathbf{P}(D|H)}{\mathbf{P}(D|\bar{H})} = (N-1)\frac{\mathbf{P}(H|D)}{\mathbf{P}(\bar{H}|D)}. \tag{8}$$

Under an uniform prior, the Bayes factor for hypothesis H becomes the posterior odds ratio weighted by the number of hypothesis that H is being compared to.

Thus, equation (8) provides a fairly easy way to calculate Bayes factors for IKE. This can be done by hand for now, but it would be nice to have IKE automatically calculate these values each time new evidence is entered in the network.

Posterior Probabilities. Similarly, the posterior probabilities of an hypothesis given new evidence, can be expressed in terms of the Bayes factors for all the hypotheses. We include this just for completeness, but we think that for now, this is not of particular use in IKE.

Assuming N+1 mutually exclusive hypotheses $(H_0, H_1 \dots, H_N)$, the posterior of hypothesis H_k can be related to the Bayes factors $\{B_{j,0}\}_{j=0}^{N}$ comparing all hypotheses to H_0 as follows,

$$\mathbf{P}(H_k|D) = \frac{\alpha_k B_{k,0}}{\sum_{j=0}^{N} \alpha_j B_{j,0}} \qquad \text{where} \quad \alpha_j = \frac{\mathbf{P}(H_j)}{\mathbf{P}(H_0)}$$

This easily follows from Baye's rule and by multiplying and dividing by $\mathbf{P}(D|H_\mathbf{0})\mathbf{P}(H_\mathbf{0})$**,**

$$\mathbf{P}(H_k|D) = \frac{\mathbf{P}(D|H_\mathbf{k})\mathbf{P}(H_\mathbf{k})}{\sum_{\mathbf{j=0}}^{\mathbf{N}} \mathbf{P}(D|H_\mathbf{j})\mathbf{P}(H_\mathbf{j})}$$

Thus the posterior of H_k can be seen as the proportion that its Bayes factor (that compares it to H_0) contributes to the overall weighted sum of all Bayes factors that compare all hypotheses to H_0.

Examples. For the examples presented throughout the document, the following table contains all the calculations for the posterior odds ratios, the Bayes Factors calculated directly and calculated using the posterior odds ratios.

	Hypothesis H_1	Hypothesis H_2	Hypothesis H_3
Posterior Odds Ratios			
Given evidence			
E_1={e_{231}=1}	$\frac{P(H_1=1\mid E_1)}{P(H_1=0\mid E_1)} = \frac{0.252}{(1-0.252)} = 0.34$	$\frac{P(H_2=1\mid E_1)}{P(H_2=0\mid E_1)} = \frac{0.252}{(1-0.252)} = 0.34$	$\frac{P(H_3=1\mid E_1)}{P(H_3=0\mid E_1)} = \frac{0.496}{1-0.496} = 0.98$
E_1={e_{231}=1}, E_2={e_{221}=medium}	$\frac{P(H_1=1\mid E_1,E_2)}{P(H_1=0\mid E_1,E_2)} = \frac{0.0771}{0.923} = 0.084$	$\frac{P(H_2=1\mid E_1,E_2)}{P(H_2=0\mid E_1,E_2)} = \frac{0.288}{0.712} = 0.41$	$\frac{P(H_3=1\mid E_1,E_2)}{P(H_3=0\mid E_1,E_2)} = \frac{0.635}{0.365} = 1.74$
Bayes Factors			
Given evidence			
E_1={e_{231}=1}	$B_1^1 = \frac{P(E_1\mid H_1=1)}{P(E_1\mid \overline{H_1}=1)} = \frac{2.02}{\frac{1}{2}(3.96+2.02)} = 0.68$	$B_2^1 = \frac{P(E_1\mid H_2=1)}{P(E_1\mid \overline{H_2}=1)} = \frac{2.02}{\frac{1}{2}(3.96+2.02)} = 0.68$	$B_3^1 = \frac{P(E_1\mid H_3=1)}{P(E_1\mid \overline{H_3}=1)} = \frac{3.96}{\frac{1}{2}(2.02+2.02)} = 1.96$
E_1={e_{231}=1}, E_2={e_{221}=medium}	$R_1^2 = \frac{P(E_2\mid E_1,H_1)}{P(E_2\mid E_1,\overline{H_1})} = \frac{10.6}{39.8\frac{2.02}{0.02.02+3.96} + 44.6\frac{3.96}{0.02.02+2.02}} = 0.25$	$R_2^2 = \frac{P(E_2\mid E_1,H_2)}{P(E_2\mid E_1,\overline{H_2})} = \frac{39.8}{10.6\frac{2.02}{0.02.02+3.96} + 44.6\frac{3.96}{2.02+3.96}} = 1.2$	$R_3^2 = \frac{P(E_2\mid E_1,H_3)}{P(E_2\mid E_1,\overline{H_3})} = \frac{44.6}{10.6\frac{2.02}{0.02.02+3.96} + 39.8\frac{3.96}{2.02+3.96}} = 1.77$
	$B_1^2 = R_1^2 \times B_1^1 = 0.25 \times 0.68 = 0.17$	$B_2^2 = R_2^2 \times B_2^1 = 1.2 \times 0.68 = 0.82$	$B_3^2 = R_3^2 \times B_3^1 = 1.77 \times 1.96 = 3.47$
Bayes Factors via Posterior Odds Ratios			
Given evidence			
E_1={e_{231}=1}	$B_1^1 = 2 \times \frac{P(H_1=1\mid E_1)}{P(H_1=0\mid E_1)} = 2 \times 0.34 = 0.68$	$B_2^1 = 2 \times \frac{P(H_2=1\mid E_1)}{P(H_2=0\mid E_1)} = 2 \times 0.34 = 0.68$	$B_3^1 = 2 \times \frac{P(H_3=1\mid E_1)}{P(H_3=0\mid E_1)} = 2 \times 0.98 = 1.96$
E_1={e_{231}=1}, E_2={e_{221}=medium}	$B_1^2 = 2 \times \frac{P(H_1=1\mid E_1,E_2)}{P(H_1=0\mid E_1,E_2)} = 2 \times 0.084 = 0.17$	$B_2^2 = 2 \times \frac{P(H_2=1\mid E_1,E_2)}{P(H_2=0\mid E_1,E_2)} = 2 \times 0.41 = 0.82$	$B_3^2 = 2 \times \frac{P(H_3=1\mid E_1,E_2)}{P(H_3=0\mid E_1,E_2)} = 2 \times 1.74 = 3.48$

5. Conclusion

We presented three criteria to compare hypotheses given new hard evidence in a Bayesian network: comparing posterior probabilities, looking at posterior odds ratios, and Bayes factors. Comparing posterior probabilities and posterior odds ratios given the new evidence is fairly simple, but *a priori* we have no specific guidelines as when we have enough evidence to decide that one of the hypothesis is true or not. In contrast, Bayes factors are not hard to calculate but require extra calculations then the normal belief updating that is done when entering new evidence. But we have empirical guidelines (see Table 1) for deciding when enough evidence has been gathered to support a particular hypothesis (see [2] and [4]). Fortunately, we have an easy alternative way to compute Bayes factors based only on the posterior odds ratios. The advantage of this procedure, is twofold, easy calculations and available empirical guidelines that are commonly used in many arenas.

Future work would include testing/validating these guidelines for the various IKE applications, and in principle, if enough data is available, we may even decide to adjust thresholds appropriately. Finally, it would be great to implement calculating Bayes factors in IKE and reporting how much the new evidence supports the different hypotheses.

Bibliography

[1] Gelman, A., J. Carlin, H. Stern, and D. Rubin (1995). Bayesian Data Analysis. Chapman and Hall/CRC.

[2] Jeffrey, H. (1961). Theory of Probability, 3rd edition, Oxford, U.K.: Oxford University Press.

[3] Kass, R. E. (1993). Bayes Factors in Practice, Journal of the Royal Statistical Society. Series D The Statistician, Vol. 42, No. 5, pp 551-560.

[4] Kass, R. E. and A. Raftery (1995). Bayes Factors, Journal of the American Statistical Association, Vol. 90, No. 430.