

Analysis of complex patterns of evidence in legal cases: Wigmore charts vs. Bayesian networks

V. Leucari*

October 2005

Typical features of the evidence arising from legal cases are its complex structure and ambiguity. Therefore, marshalling and evaluating evidence are two fundamental issues in forensic science, both for constructing arguments about questions of fact and for taking final decisions. The chart method for analysing evidence introduced by Wigmore is a technique which allows to organise and describe the available evidence, and to construct reasoning processes through sequential steps. Bayesian networks are a general statistical tool which can be applied in the legal context to model relationships between different sources of evidence, weigh the available evidence, and draw statistical inferences from it. Wigmore charts and Bayesian networks are clearly different in nature, but they are both an attempt to a rigorous and formal approach to the analysis of evidence. We illustrate and compare such techniques applied to a fictitious burglary example.

1 Introduction

In Dawid and Evett (1996) a hypothetical criminal case is presented, involving a complicated pattern of interactions between different items of evidence. Such interactions are represented in a Bayesian network, and the likelihood ratio for the hypothesis “is the person prosecuted for this crime truly the offender?” is then computed. In Schum (2005) the same example –and the

*University College London

same evidence— is analysed via the Wigmore chart method, and arguments are constructed in order to charge the defendant with the crime.

The aim of this paper is to compare and synthesize the two analyses: starting from the original Bayesian network in Dawid and Evett (1996), a more detailed network is constructed by taking into account the process followed in Schum (2005) of breaking down the evidence into single “units” and build a chain of reasoning by connecting them. The ultimate objective would be the development of a formal method to handle multiple sources of evidence in legal cases (and not only) by jointly exploiting Wigmore charts and Bayesian networks. This synthesis is largely at its very early stages and the model presented here is a first attempt at “combining” the two methods, and by no means the final product.

The example is briefly described in Section 2. Section 3 and 4 respectively contain the Wigmore chart and Bayesian network analyses. The two techniques are compared in Section 5. Some remarks are pointed out in Section 6.

2 A burglary example

We use a fictitious burglary example to illustrate the methods. The example is described in Dawid and Evett (1996) as well as in Schum (2005). An unknown number of offenders break in some commercial premises through a cut they make in a grille, likely with the intention of committing a burglary. Once inside, they face a guard who has the time of setting an alarm on before being punched in the face by one of the intruders. As the alarm sounds, the offenders run away without any money and escape on foot, since their getaway car has already left as a consequence of the alarm. Minutes later, a police officer driving nearby the crime scene finds the suspect trying to steal a car and apprehends him. A tuft of red fibres is found in the grille at the crime scene. A blood stain is found on the jumper worn by the suspect at the time of his arrest.

This example, though quite simple at first sight, possesses many of the features of legal cases that give raise to complex structures of the associated evidence. Most of them are obvious, but it may be useful to highlight them, as they recur in many criminal cases:

- There are multiple and different sources of evidence: an item found at the crime scene (the tuft of fibres), an item belonging to the suspect

(the jumper), a trace left by an unknown individual (the blood stain on the jumper), people involved with the crime (the suspect, the guard, the police officer).

- Further evidence can be (and will be) collected: laboratory analyses of the fibres found at the crime scene, tests on blood samples (taken from the suspect's jumper and possibly from various people.)
- There are no data, in the statistical meaning of repeated observations of an experiment: the “experiment” is unique and irreproducible.
- Uncertainty can be introduced at various levels: are we willing to assume that there was a crime (and take this for granted)? And many other such examples.

Notice that there are other classes of problems in other disciplines that exhibit some of these features or similar issues.

3 A Wigmore chart analysis

I will not go through the steps required for building a Wigmore chart, and will not describe how a Wigmore chart for the example above looks like, as this is well explained in Schum (2005). Instead, I will make some remarks that will be relevant to the successive Bayesian network analysis.

Something will be added here...

4 A Bayesian network analysis

Dawid and Evett (1996) show a possible Bayesian network for modeling relationships between various items of evidence. Here we try to improve the original network based on the Wigmore chart analysis in Schum (1996). Trying to translate from one technique to the other would be the most intuitive way of proceeding, but the two methods are too different in nature for this to be profitable. Rather, Wigmore charts could be usefully combined with Bayesian networks in order to provide a more satisfying model to be used to derive inference from the observed evidence. This is what we discuss in the remaining of this section.

4.1 The Hugin software for Bayesian networks

The Bayesian network analysis is performed in Hugin, a software for building Bayesian networks and evaluating the associated evidence. A brief description of how Hugin works:

1. Build the Bayesian network
 - 1.1 Define the random variables for the current problem (these are the nodes in the network).
 - 1.2 Specify the relationships among the variables, in terms of conditional independencies (these are the arrows in the network).
2. Enter prior knowledge
 - 2.1 Specify the values that each variables can take.
 - 2.2 Specify the conditional probability tables (conditional distributions of each variable given the parent variables).
3. Enter evidence: fix the value of the observed variables.
4. Propagate evidence: Hugin performs a propagation algorithm that updates the prior probabilities for the unobserved variables given the evidence entered for the observed variables.

The output for the burglary example will be, for instance, the value of the likelihood ratio for the hypothesis “the suspect is guilty”, after the evidence described in Section 2 is entered. Which decision will we take given this number?

4.2 Simple Bayesian network

The key list in Schum (2005), p. 5, describes all the items of evidence related to the burglary example and used to build the corresponding Wigmore chart. The Bayesian network in Figure 1 is an attempt to extend the model in Dawid and Evett (1996) in order to take into account all the elements in such a list. In particular, uncertainty about the crime (did it really happen?) is introduced, see variables C_1, C_2 in the network, as well as some details about the crime, see variables C_4, C_5, C_8 . Moreover, witnesses are implicitly represented, see variables in the bottom part of the table below. Table 1 is a list

of the variables and their possible states. The names in parenthesis are the labels in the Bayesian network in Dawid and Evett (1996). Notice that the graph obtained by removing all the new variables is the same as the one in Figure 1 in Dawid and Evett (1996). In the table below some of the names in Schum (2005) are used: BB stands for Blackbread Brewery, HS is the suspect, DI Leary is the detective inspector who apprehended the suspect. Finally, \mathcal{I} is the indicator function, that takes values 1 if the event is true and 0 otherwise. Once the evidence has been entered in the network, i.e. the observed variables have been fixed to their observed value, the probabilities of the nodes of interest, namely C_6 and C_8 , can be updated.

4.3 Object-oriented Bayesian network

It is often possible to find certain structures that are used repeatedly inside a single Bayesian network and in different networks constructed for different problems. An object oriented Bayesian network allows to define general single networks –or fragments– and to combine them together in a main network, having as nodes both random variables and Bayesian networks which are instances of the general fragments previously defined, see Dawid et al. (2005).

Figure 2 and Figure 3 represent the object-oriented network constructed from the network in Figure 1. Note that variables “cut”, “number”, “punch”, “scene”, “guilty”, “identity2” and “jumper true” here are called respectively $C_1, C_3, C_4, C_7, C_6, F_2$ and J_1 in Figure 4.2. The blue nodes are all instance of the fragment “report”, the green nodes are all instances of the fragment “matchevidence”, and the red nodes are all instances of the fragment “consequence”. Such fragments are shown in Figure 4, 5, 6 and are defined as follows:

report: this network represents a report (node “report”) regarding an item or event (variable “item A”). The report can be right, and hence equal to the original item or event, or wrong, and hence equal to a random item or event (variables “item B”). The variable “biased coin” allows for a non symmetric error in the report. Such fragment is used here for witnesses who testify about events related to the crime, and for laboratory analyses (e.g. microspectrofluorimetry) on items related to the crime.

match evidence: this fragment represents a match between a trace (variable “item cs”) and several possible sources of that trace (actually, only

| Label | Variable | States |
|---------------------|---|---------------------------------------|
| C_1 | \mathcal{I} {Someone made a cut in the grille in order to break in} | 0, 1 |
| C_2 | \mathcal{I} {An unknown number of persons entered BB} | 0, 1 |
| C_3 (N) | Number of offenders | 1, 2, 3, 4, 5, 6 |
| C_4 | \mathcal{I} {One of the offenders punched the guard} | 0, 1 |
| C_5 | Consequences of the punch | 0, 1 |
| C_6 (C) | \mathcal{I} {HS is guilty, i.e. he is one of the offenders} | 0, 1 |
| C_7 | \mathcal{I} {HS was nearby the scene of the crime} | 0, 1 |
| C_8 (B) | Identity of the person who punched the guard | 0, 1 |
| F_1 | \mathcal{I} {Fibre tuft shown at trial was found at crime scene} | 0, 1 |
| F_2 (A) | Identity of the person who left the fibre tuft in the grille | 0, 1 |
| F_3 (Y_1) | Properties of the fibre tuft | 0=other, 1=red acrylic |
| J_1 | \mathcal{I} {The jumper shown at trial is HS's jumper} | 0, 1 |
| J_2 (X_3) | Properties of HS's jumper | 0=other, 1=red acrylic |
| B_1 (X_1) | Blood sample from HS | A, B, AB, O |
| B_2 (X_2) | Blood sample from the guard | A, B, AB, O |
| B_3 (R) | Shape of blood stain on jumper | 0=no blood stain, 1=spray, 2=other |
| B_4 (Y_2) | Blood type on jumper | A, B, AB, O |
| PIC_{C_1} | Evidence from the photo of the hole in the grille | 0, 1 |
| PIC_{C_5} | Evidence from the photo of the guard's injury | 0, 1 |
| PIC_{C_7} | Evidence from the photo of HS after he was apprehended | 0, 1 |
| MSF | Result of microspectrofluorimetry | 0, 1 |
| TLC | Result of thin layer cromatography | 0, 1 |
| L_{C_1} | DI Leary's testimony about C_1 | 0, 1 |
| L_{C_2} | DI Leary's testimony about C_2 | 0, 1 |
| L_{C_7} | DI Leary's testimony about C_7 | 0, 1 |
| L_{F_1} | DI Leary's testimony about F_1 | 0, 1 |
| L_{F_3} | DI Leary's testimony about F_3 | 0, 1 |
| L_{J_1} | DI Leary's testimony about J_1 | 0, 1 |
| L_{J_2} | DI Leary's testimony about J_2 | 0, 1 |
| L_{B_1} | DI Leary's testimony about B_1 | A, B, AB, O |
| L_{B_2} | DI Leary's testimony about B_2 | A, B, AB, O |
| L_{B_3} | DI Leary's testimony about B_3 | 0, 1 |
| L_{B_4} | DI Leary's testimony about B_4 | 0, 1 |
| L_{MSF} | DI Leary's testimony about microspectrofluorimetry | 0, 1 |
| L_{TLC} | DI Leary's testimony about thin layer cromatography | 0, 1 |
| G_{C_2} | Guard's testimony about C_2 | 0, 1 |
| G_{C_3} (G_1) | Guard's testimony about C_3 | 1, 2, 3, 4, 5, 6 |
| G_{C_4} | Guard's testimony about C_4 | 0, 1 |
| G_{C_5} (G_2) | Guard's testimony about C_5 | 0, 1 |
| HS_{C_6} | HS's testimony about C_6 | 0, 1 |
| HS_{C_8} | HS's testimony about C_8 | 0, 1 |

Table 1: Variables in the network in Figure 1.

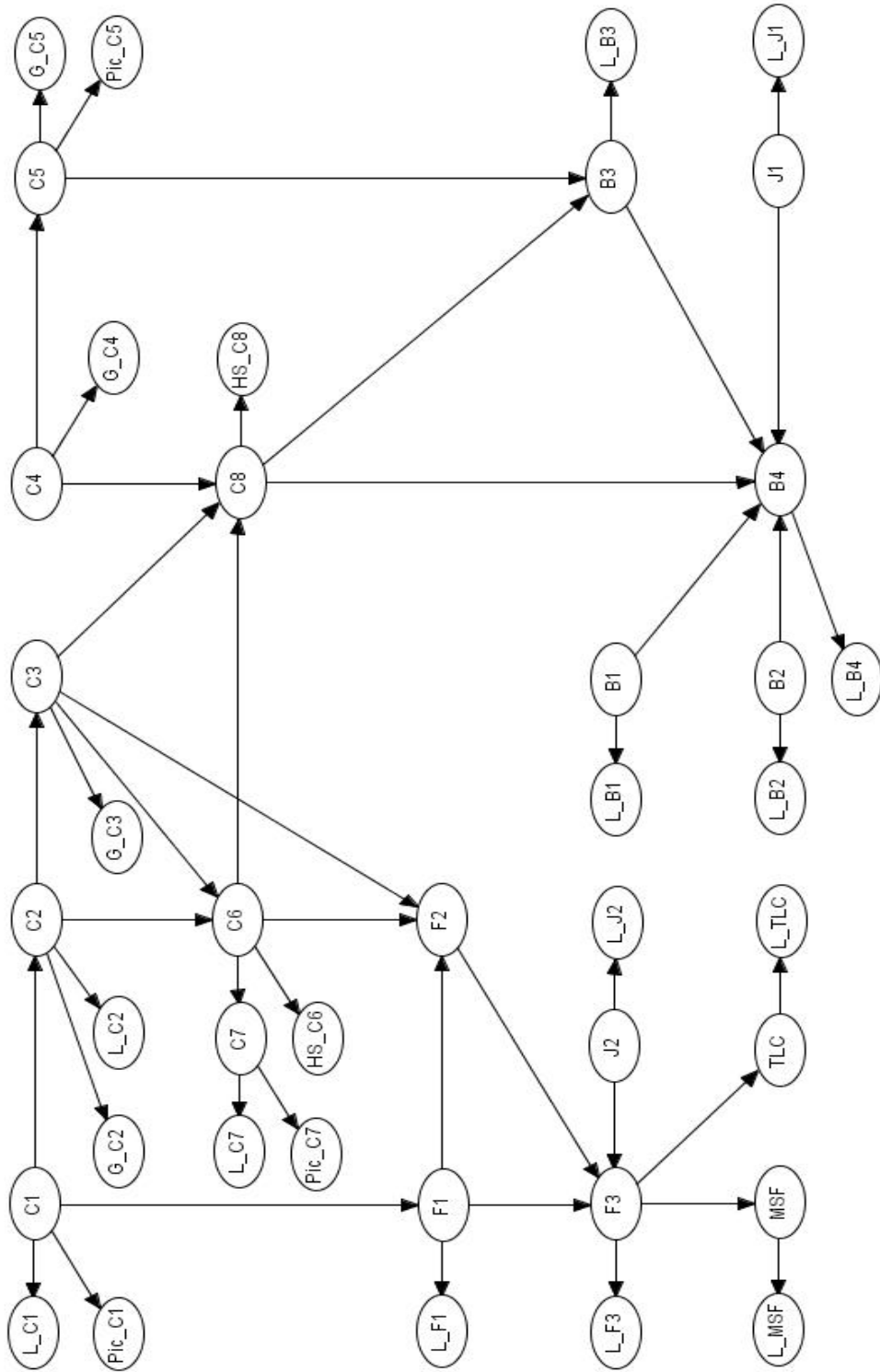


Figure 1: A non object oriented version of the Bayesian network for the burglary example.

one source is considered in this example, variable “item s”). The variable “H” is the hypothesis variables “who left the trace?”. Such fragment is used here for describing blood matches between the suspect, the guard, and the blood found on the suspect’s jumper, as well as the match between the fibres found at the crime scene and the suspect’s jumper.

consequence: this network represents events (variable “event”) that are consequences of some possibly false event (variable “crime”), typically the event “did the crime happen?”. Such fragment is used here to model meaningless conditioning, such as conditioning the break in variable on the fact that nobody cut the grille (and hence that nobody even tried to break in). The variable “bernoulli” is itself a network, represented in Figure 7.

Notice that most of the variables are binary, but everything can be done for polytomous variables as well.

Example with numbers to come...

5 Comparison

Both Wigmore charts and Bayesian networks are formal tools for handling evidence. What do they have in common? Which are the differences? Are they complementary methods in some sense? Some remarks/questions about comparing these two techniques are listed below.

- Wigmore charts are a technique for describing a reasoning process, and for constructing arguments. Bayesian networks are a statistical tool for deriving inferences given the available evidence. Wigmore charts are deterministic.
- Wigmore charts are constructed backward, after collecting the evidence, whereas Bayesian networks are meant to be a model for representing how things can happen (evidence is entered after the network has been built). Arrows in the two methods have opposite directions.
- No conditional independence in Wigmore charts, but still notion of relevance?

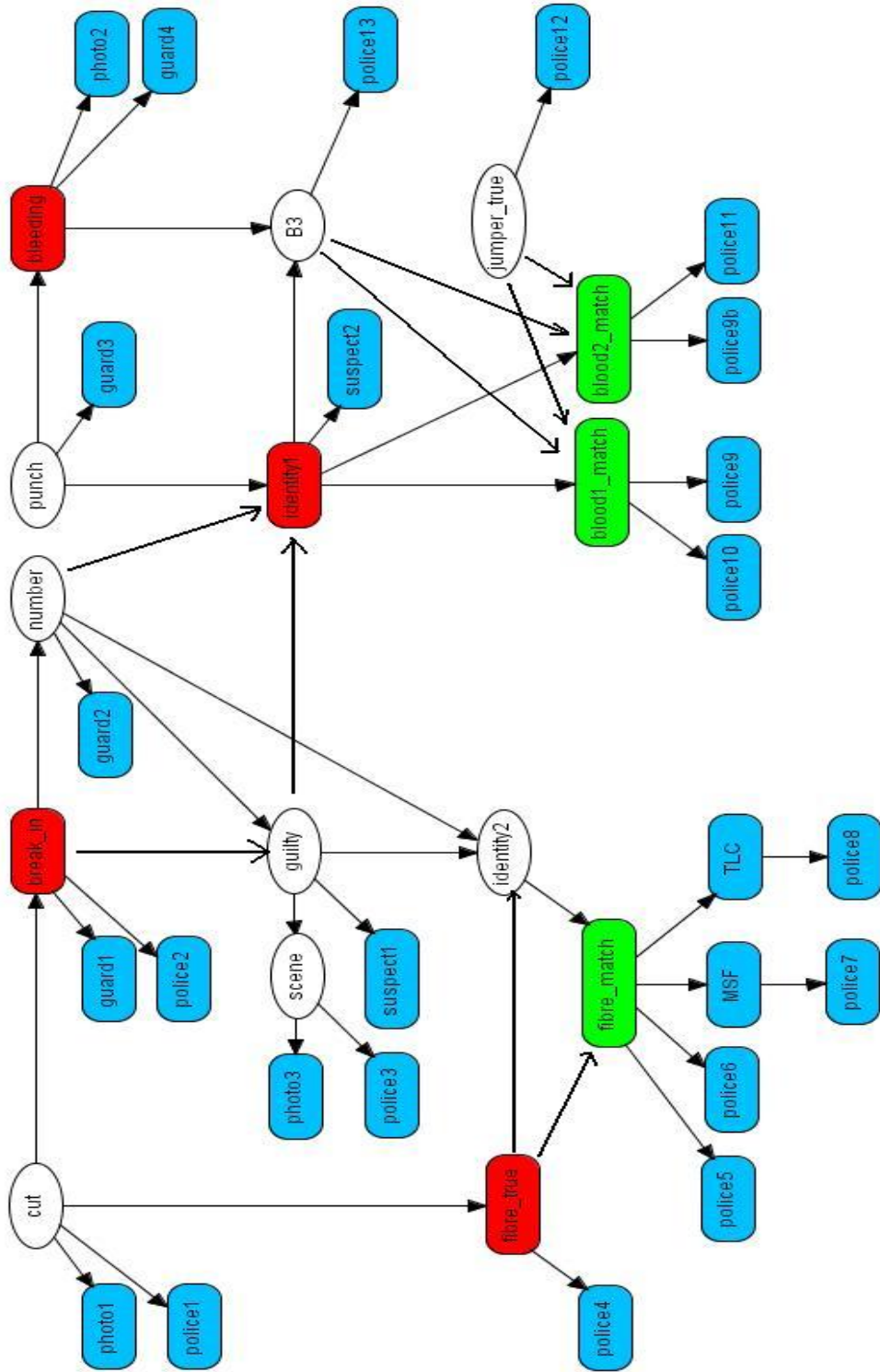


Figure 2: The object oriented Bayesian network for the burglary example.

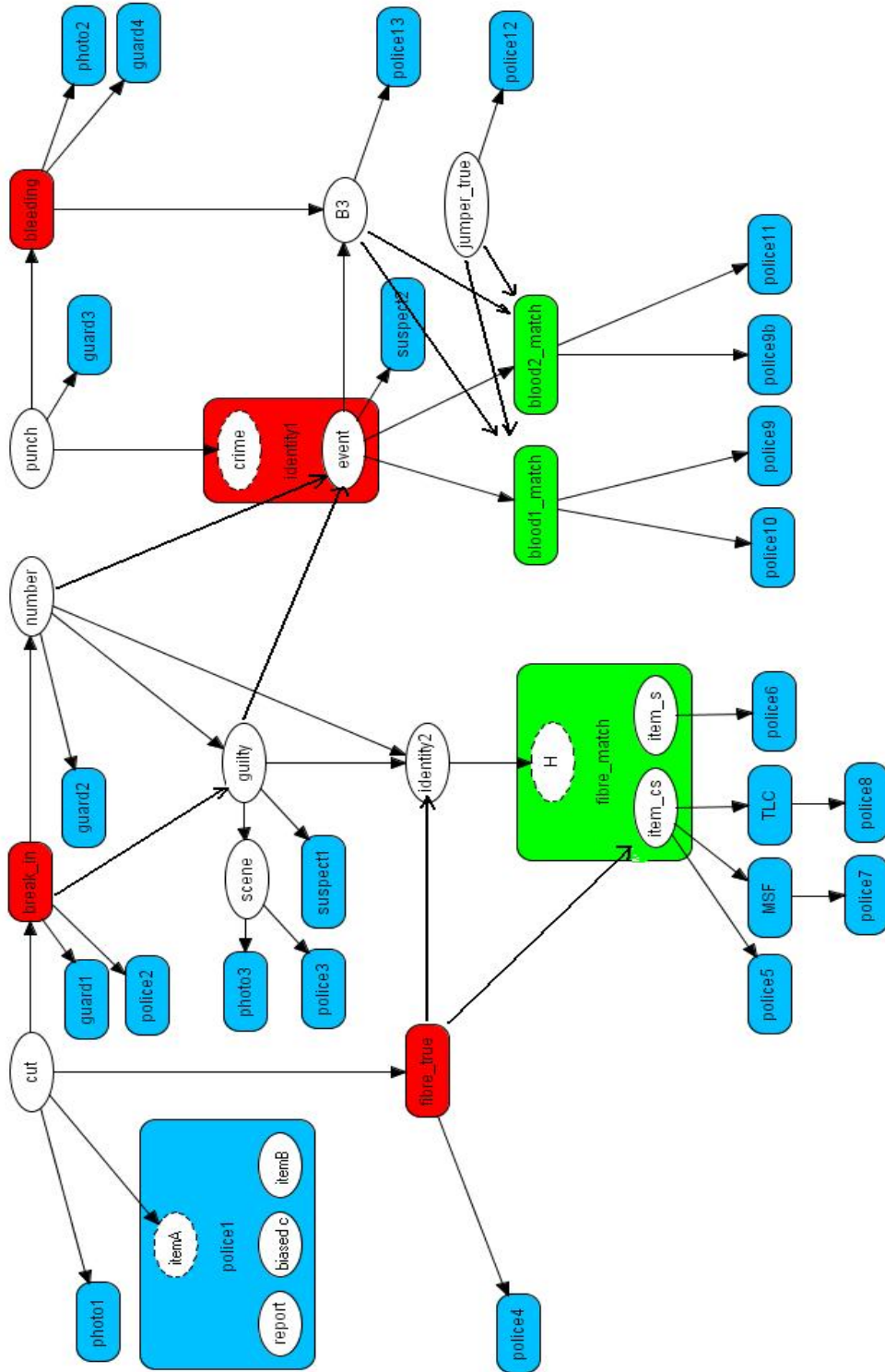


Figure 3: The object oriented Bayesian network for the burglary example.

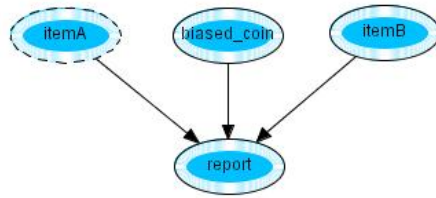


Figure 4: The “report” fragment.

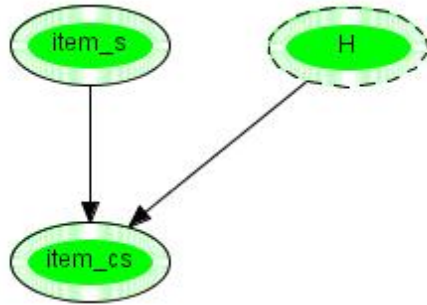


Figure 5: The “matchevidence” fragment.

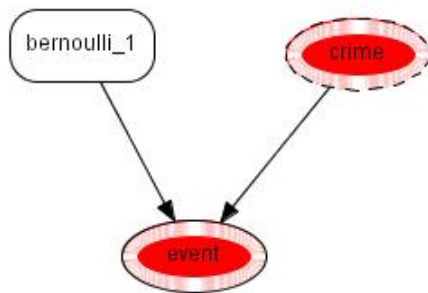


Figure 6: The “consequence” fragment.

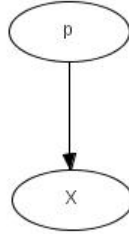


Figure 7: The “bernoulli” fragment.

- No quantification in Wigmore charts: how can one draw inferences from the observed evidence or formulate hypotheses?
- No construction of arguments in Bayesian networks?
- A lot of details vs. not so many details (take some of the elements for granted).
- In order to build a Bayesian network one has to choose a “starting point”, and hence take some of the events for granted and implicitly make assumptions which do not appear in the model, and thus decide what is evidence...
- A Wigmore chart describes different hypotheses at different levels (ultimate, penultimate and intermediate probanda). What is that correspond to this in Bayesian networks? Do we always want a single hypothesis (guilty/not guilty)?
- Construction of the network is a subjective matter, in both cases.
- Is the Wigmore chart method unnecessary complicated (lots of trifles) from a statistical point of view?

6 Discussion and further work

Many important issues have not been taken into account in the models described above. Further work is needed for the following:

- How to evaluate and compare different models, if needed (sensitivity analysis)

- What model when more than a suspect is involved.
- Manipulated evidence (e.g. testimonies of witnesses may not be genuine, or item presented at trial may not be those found at crime scene or could have been modified after being found at crime scene).
- Interactions between different witnesses.
- Credibility of witnesses.
- Limitations of Hugin when building a network.

References

Dawid, A. P. and Evett, I. W. (1997). Using a graphical method to assist the evaluation of complicated patterns of evidence. *Journal of Forensic Sciences* **42**, 226-231.

Dawid, A. P., Mortera, J. and Vicard, P. (2005). Object-oriented Bayesian networks for complex forensic DNA profiling problems. Technical Report 256, Department of Statistical Science, University College London.

Schum, D. A. (2005). A Wigmorean interpretation of the evaluation of a complicated pattern of evidence.