BAYESIAN NETWORKS AND THE EVALUATION OF SCIENTIFIC EVIDENCE: A THEORETICAL APPROACH

Franco TARONI[1, 2], Paolo GARBOLINO[3]

[1] Institut de Police Scientifique et de Criminologie, University of Lausanne, Lausanne, Switzerland

[2] Institut de Médecine Légale, University of Lausanne, Lausanne, Switzerland

[3] Dipartimento di Filosofia, Scuola Normale Superiore, Pisa, Italy

ABSTRACT: Recently, methods that deal with formal analysis of decision making have been developed. Bayesian networks – also known as belief networks and causal probabilistic networks – provide a method of representing relationships between characteristics even if the relationships involve uncertainty, unpredictability or imprecision. This quantitative method assists the scientist not only in describing a problem and communicating information about its structure but also in calculating the effect of the truth of one proposition or piece of evidence on the plausibility of others. Notably, Bayesian networks are a network-based framework for representing and analysing situations involving uncertainty (i.e. evidence evaluation, criminal investigation, etc.).

Information is presented in a graph as a set of nodes (representing the variables) linked by directed arcs (or edges) and the direction of the arc represents an influential relationship. The absence of an arc between two nodes implies that the two variables associated with these nodes are independent of each other, conditional on knowledge of the values of the other variables.

The aim of this paper is to show how such a methodology could facilitate the representation and the evaluation of the scientific evidence. A simple scenario involving a transfer evidence is developed to show the role of different variables.

INTRODUCTION

Forensic literature has pointed out the utility of methods that deal with formal analysis of decision making. Notably, it has been underlined that complex framework of circumstances – situations involving many variables – requires a logical assistance [2]. Methods of formal reasoning has been proposed to assist the forensic scientist to understand all of the depend-

encies which may exist between different aspects of the evidence [6]. Notably, it has been shown that Bayesian networks provide a valuable aid for representing relationships between characteristics in situation of uncertainty, unpredictability or imprecision. They assist the user not only in describing a complex problem and communicating information about its structure but also in calculating the effect of the truth of one proposition or piece of evidence on the plausibility of others.

Using a simple case examples, we would like to demonstrate how such a graphical method could be used to assess scientific evidence.

BAYESIAN NETWORKS

Bayesian networks are a method for discovering valid, novel and potentially useful patterns in data where uncertainty is handled in mathematically rigorous but simple and logical way.

Imagine a directed arc linking nodes A and B in a graph. A and B represent events and the arc directed from the node A to node B expresses that the event A is a "cause" of event B. For the arc (A to B) the scientist could give a matrix expressing the probability , that specifies for each state of A the probability of each state of B. In other words, each node represents a random variable that can assume the alternative values "occurs" and "does not occur" (or "true" and "false").

Such a pictorial scheme allows the users to directly translate qualitative structure into dependence (or causal relationship), independence and conditional independence relations of one variable to another under specified circumstances. This specifies the relevance of the variables in the context of the case.

Links between variables can then be established deterministically or probabilistically using data coming from surveys or expert opinions and estimations [4, 8].

Practically, Bayes networks allow the forensic scientist to:
1. learn a way of thinking about the problem involving uncertain information;
2. learn how to apply these methods to drawn inferences about the world of interest;
3. learn how to act rationally under uncertainty;
4. learn how the model can be induced from data.

Scientists have to cope with uncertainty; they have to take into account:
1. theoretical ignorance, because no complete theory is known about the problem domain (e.g. police investigation or crime analysis);

2. laziness, because the space of relevant factors is very large and would require too much work to list the complete;
3. a practical ignorance, because there is uncertainty about a particular individual in the domain of interest;
4. decision-making under uncertainty, because it is necessary to make rational decisions able to decide even when there is not enough information to prove that an action will work.

Bayesian decision theory could be a useful tool in the hand of forensic scientists essentially because humans often fail to follows a logical framework in complex situations; they need a support as it has been already suggested in forensic literature [3].

The graphical nature of the approach facilitates formal discussion of the system structure. Ability to specify the relationships between variables in uncertain terms is ideal to describe the relationship between dynamics which may not be well understood. On a quantitative level, the approach enables the user's knowledge to be incorporated into the model on the same basis as more objectively derived data (e.g. data coming from surveys). Such features allow the creation of a model which may contain mathematical relationships as well as subjective elements corresponding to the experience of the people [11].

HOW TO WORK WITH PROBABILITIES IN BAYESIAN NETWORKS: THE SIMPLE "ONE-TRACE" CASE EXAMPLE

Suppose a crime has been committed and a blood stain has been found at the scene of the crime. The DNA profile of the stain is different from the victim's blood profile and may have been left by the offender. Let's suppose, for the sake of simplicity, that it's part of the background knowledge that there was only one offender. A blood sample has been taken from a suspect; this sample and the recovered stain share the same DNA profile.

Let H be the *probandum*: "The suspect is the offender", F the proposition: "The crime blood stain came from the suspect", G: "The crime stain came from the offender" and E: "The suspect's blood sample and the blood stain found at the crime scene share the same DNA profile".

The appropriate graphical representation of these assertions is presented in Figure 1 where the arc from the node (variable) H to node F expresses that the event denoted by H is a "cause" of the event denoted by F, or "brings about", or in some way "influences" it. We don't have to be committed to a particular causal theory: the arcs represent probabilistic relevance relations which might be understood, but not necessarily, as causal relations. For the arc (H to F) the scientist could give a matrix expressing the probabil-

ity $P(F\,|\,H)$, to be read as "the probability of F given H", that specifies for each truth-value of H the probability of each truth-value of F: for example the node H of Figure 1 correspond to a random variable with the values $h = 1$ (h = "true") and $h = 0$ (or h = "false"). For sake of simplicity, we shall write H and $\overline{H}$ (meaning "not-H") from now on.

Schemes as Figure 1 allow the users to directly translate qualitative structure into dependence (or causal relationship), independence and conditional independence statements of one variable to another under specified circumstances. This specifies the relevance of the variables in the context of the case.

Variables H and G are probabilistically independent (there is not a direct arc between H and G), but they are conditionally dependent given F, because they have a common descendant, namely, F. Moreover, the variable F "screens off" E from H and G: if F is true (or false), then E is independent from H and G, that is, $P(E\,|\,F,H) = P(E\,|\,F)$.

H : The suspect is the offender

G : The crime stain came from the offender

F : The crime stain came from the suspect

E : The suspect's blood sample and the bloodstain found at the scene of the crime share the same DNA profile
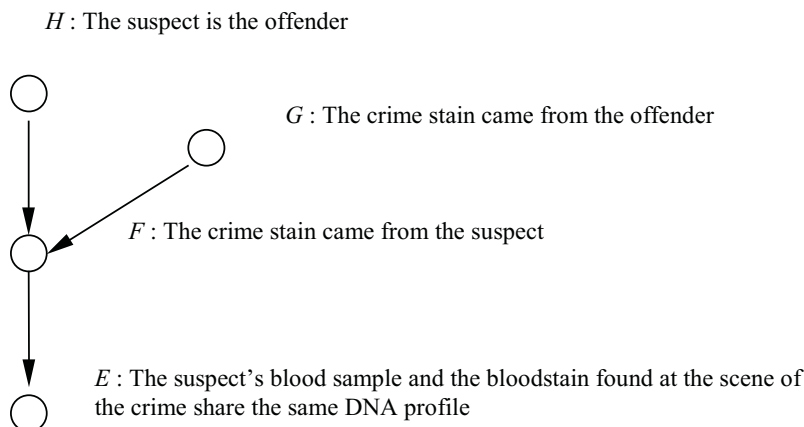
Fig. 1. Bayesian network describing "one-trace" scenario.

In addition to the probabilistic dependencies represented in Figure 1, the following assumptions seem quite straightforward:

1. If is true, then certainly E: $P(E\,|\,F) = 1$.
2. If $\overline{F}$ is false, then the probability of E is given by the random match probability f of the DNA profile among members of the relevant population: $P(E\,|\,\overline{F}) = f$.
3. If the suspect is the offender and the crime stain came from the offender, then certainly the stain came from the suspect: $P(F\,|\,G,H) = 1$.

4. If the suspect is the offender and the stain did not come from the offender, then certainly the crime stain didn't come from the suspect: $P(\mathrm{F}|\overline{G},H) = 0$.
5. If the suspect is not the offender and the stain came from the offender, then certainly the crime stain didn't come from the suspect: $P(F|G,\overline{H}) = 0$.

By the assumptions {1} and {2}, and by the fact that F "screens off" E from H, we obtain that the probability of E given H is given by the following formula:

$$P(E\,|\,H) = P(F\,|\,H) + P(\overline{F}\,|\,H)\cdot f. \qquad \{1\}$$

Then we calculate the probability of F given H by "extension of the conversation" [9] to G:

$$P(F\,|\,H) = P(F\,|\,G,H)\cdot P(G\,|\,H) + P(F\,|\,\overline{G}\,|\,H)\cdot P(\overline{G}\,|\,H). \qquad \{2\}$$

Given assumption {4}, the probability of F given H is reduced to:

$$P(F\,|\,H) = P(F\,|\,G,H)\cdot P(G\,|\,H) + \overset{0}{\overbrace{P(F|\overline{G},H)}}\cdot P(\overline{G}\,|\,H) = P(F\,|\,G,H)\cdot P(G\,|\,H). \quad \{3\}$$

By assumption {3}, and by the hypothesis that H and G are independent (see Figure 1), that simplifies into:

$$P(F\,|\,H) = \overset{1}{\overbrace{P(F|G,H)}}\cdot (P(G\,|\,H) = P(G). \qquad \{4\}$$

By the same assumption, and by assumption {4}, we get immediately that the probability of $\overline{F}$ given H is equal to the probability of $\overline{G}$:

$$P(\overline{F}|\,H) = P(\overline{G}). \qquad \{5\}$$

Scientist is concerned with the assessment of the likelihood ratio, $P(E|H)/\,P(E|\overline{H})$, and now we know that the numerator (3.1) of the LR can be written as:

$$P(E\,|\,H) = P(G) + P(\overline{G})\cdot f. \qquad \{6\}$$

Given the same hypotheses about probabilistic independence, and using assumption {5}, the formula for the denominator, $P(E\,|\,\overline{H})$, can also be written as:

$$P(E\,|\,\overline{H}) = P(F\,|\,\overline{H}) + P(\overline{F}\,|\,\overline{H})\cdot f\,. \qquad \{7\}$$

where $P(F|\overline{H}) = P(F|\overline{G},\overline{H})\cdot P(\overline{G})$ and $P(\overline{F}|\overline{H}) = P(G) + P(\overline{F}|\overline{G},\overline{H})\cdot P(\overline{G})$.

The likelihood ratio then becomes, from {6} and {7}:

$$LR = \frac{P(E|H)}{P(E|\overline{H})} = \frac{P(G) + P(\overline{G})\cdot f}{P(F|\overline{G},\overline{H})\cdot P(\overline{G}) + \left[P(G) + P(\overline{F}|\overline{G},\overline{H})\cdot P(\overline{G})\right]\cdot f}\,. \qquad \{8\}$$

Assume that $P(G) = r$ (the relevance term [10]), and $P(F|\overline{G},\overline{H}) = p$, the latter is the probability that the stain would have been left by the suspect even though he was innocent of the offence. It is the probability the crime stain was left innocently by someone who is now a suspect. Here is assumed that the propensity to leave a stain is independent of the blood profile of the person leaving the stain [1]. We then obtain a (simplified) version of Evett's formula [7]:

$$LR = \frac{\overbrace{P(G)}^{r} + \overbrace{P(\overline{G})}^{(1-r)} \cdot f}{\underbrace{P(F|\overline{G},\overline{H})}_{p} \cdot \underbrace{P(\overline{G})}_{(1-r)} + \underbrace{P(G)}_{r} \cdot f + \underbrace{P(\overline{F}|\overline{G},\overline{H})}_{(1-p)} \cdot \underbrace{P(\overline{G})}_{(1-r)} \cdot f} \; ;$$

$$LR = \frac{r + (1-r) \cdot f}{p \cdot (1-r) + r \cdot f + (1-p) \cdot (1-r) \cdot f} ;$$

$$LR = \frac{r + (1-r) \cdot f}{r \cdot f + (1-r) \cdot \left[p + (1-p) \cdot f \right]} \; . \qquad \{9\}$$

If the assessment of the relevance reaches its maximum, $r = 1$, then the likelihood ratio is reduced to its simplest form, *1/f*.

The Bayes network in Figure 1 can be helpful in understanding where the formula {9} comes from: it shows graphically the dependence, and independence, assumptions made by the scientist, and the conditional probability assessments needed to evaluate the Evett's formula. Furthermore, it is a standard model that can be used in any "one-trace" case: for instance, the trace could be fibres or fingerprints.

CONCLUSION

The use of a graphical model is appealing because it allows the scientist to concentrate on the structure of the problem before having to deal with the assessment of quantitative issues [5].

Likelihood ratio formula like {9} might look complicated to people who are not trained into probabilistic analysis. Bayesian networks like the one depicted in Figure 1 represent in an economic and intuitive way the hypotheses about the probabilistic relations existing among the variables of interest in formula {9}.

References:

1. Aitken C. G. G., Statistics and the evaluation of evidence for forensic scientists, John Wiley & Sons, Chichester 1995.

2. Aitken C. G. G., Gammerman A., Probabilistic reasoning in evidential assessment, *Journal of the Forensic Science Society* 1989, vol. 29, pp. 303–319.
3. Aitken C. G. G., Gammerman A., Zhang G. [et al.], Bayesian belief networks with an application in specific case analysis, [in:] Computational learning and probabilistic reasoning, Gammerman A. [ed.], John Wiley & Sons, London 1996.
4. Cowell R. C., Dawid A. P., Lauritzen S. L. [et al.], Probabilistic Networks and Expert Systems, Springer, New York 1999.
5. Curran J., Hicks T., Buckleton J., Forensic Interpretation of Glass Evidence, CRC Press, Boca Raton 2000.
6. Dawid A. P., Evett I. W., Using a graphical method to assist the evaluation of complicated patterns of evidence, *Journal of Forensic Sciences* 1997, vol. 42, pp. 226–231.
7. Evett I. W., Establishing the evidential value of a small quantity of material found at a crime scene, *Journal of the Forensic Science Society* 1993, vol. 33, pp. 83–86.
8. Jensen F. V., An introduction to Bayesian networks, UCL Press, London 1996.
9. Lindley D. V., Probability theory, [in:] The use of statistics in forensic science, Aitken C. G. G., Stoney D. A. [eds.], Ellis Horwood, New York 1991.
10. Stoney D. A., Transfer evidence, [in:] The use of statistics in forensic science, Aitken C. G. G., Stoney D. A. [eds.], Ellis Horwood, New York 1991.
11. Taroni F., Aitken C. G. G., Garbolino P., De Finetti's subjectivism, the assessment of probabilities and the evaluation of evidence: a commentary for forensic scientist, *Science & Justice* 2001 [submitted for publication].