The Probabilistic Interpretation of Model-based Diagnosis

Ildikó Flesch[1] and Peter Lucas[2]

[1]Tilburg Centre for Creative Computing, Tilburg University, The Netherlands
I.Flesch@uvt.nl

[2]Institute for Computing and Information Sciences, Radboud University Nijmegen, The Netherlands
peterl@cs.ru.nl

Tilburg University

Radboud University Nijmegen

Introduction

Bayesian networks are popular as formalisms to built model-based, diagnostic systems. An alternative theory of model-based diagnosis was developed at approximately the same time, founded on techniques from logical reasoning [5]. The General Diagnostic Engine, GDE for short, is a well-known implementation of the logical theory; however, it also includes a restricted form of uncertainty reasoning to focus the diagnostic reasoning process [2]. Previous research by Geffner and Pearl proved that the GDE approach to model-based diagnosis can be equally well dealt with using Bayesian networks [4].

This research shows that by adding probabilistic information to a model of a system, the predictions that can be made by the model can be decomposed into a logical and a probabilistic part. The logical specifications are determined by the system components that are assumed to behave normally, constituting part of the system behaviour. This is complemented by uncertainty about behaviour for components that are assumed to behave abnormally. In addition, it is shown that the Poisson-binomial distribution plays a central role in determining model-based diagnoses.

Poisson-binomial Distribution

Let $s = (s_1, \ldots, s_n)$ be a Boolean vector with elements $s_k \in \{0, 1\}$, where s_k is a Bernoulli discrete random variable with value success (1) or failure (0). Let the probability of success of trial k be indicated by $p_k \in [0, 1]$ and the probability of failure be set to $1 - p_k$. Then, the probability of obtaining vector s as outcome is equal to

$$P(s) = \prod_{k=1}^{n} p_k^{s_k}(1 - p_k)^{1-s_k}. \quad (1)$$

The Poisson-binomial distribution is employed to describe the outcomes of n independent Bernoulli distributed random variables, where only the number of successes and failures are counted [1]. The probability that there are m successful outcomes amongst the n outcomes is then defined as:

$$f(m; n) = \sum_{s_1+\cdots+s_n=m} \prod_{k=1}^{n} p_k^{s_k}(1 - p_k)^{1-s_k}, \quad (2)$$

where f is a probability function. Here, the summation means that we sum over all the possible values of elements of the vector s, where the sum must be equal to m.

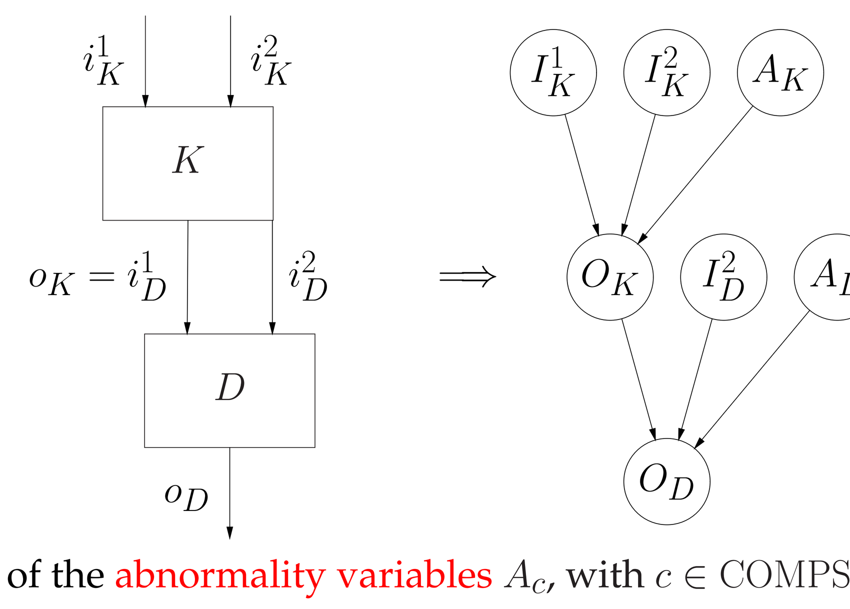
Furthermorre, suppose that we model interactions between the outcomes of the trials by means of a Boolean function b. The expectation or mean of this Boolean function is then equal to:

$$\mathcal{E}_P(b(S)) = \sum_{s} b(s)P(s). \quad (3)$$

Note that for $b(s) \equiv s_1 + \cdots + s_n = m$ (i.e., the Boolean function that checks whether the number of successful trials is equal to m), we have that $\mathcal{E}_P(b(S)) = f(m; n)$. Thus, Equation (3) can be looked on as a generic way to combine the outcomes of independent trials.

Bayesian Diagnostic Problems

A Bayesian diagnostic system is denoted by $\mathcal{S}_B = (G, P)$, where P is a joint probability distribution of the vertices of G, interpreted as random variables, and G is obtained by mapping a logical diagnostic system $\mathcal{S}_L$ to a Bayesian diagnostic system:



Let the set of values of the abnormality variables A_c, with $c \in \mathrm{COMPS}$, be denoted by

$$\delta_C = \{a_c \mid c \in C\} \cup \{\bar{a}_c \mid c \in \mathrm{COMPS} - C\},$$

and the set of observations by ω. The set of observed input and output variables are referred to as I_ω and O_ω, whereas the unobserved input and output variables will be referred to as I_u and O_u, respectively. The set of observations is, thus, decomposed as $\omega = i_\omega \cup o_\omega$. A Bayesian diagnostic problem is then defined as $\mathcal{P}_B = (\mathcal{S}_B, \omega)$.

Abductive Diagnosis

Determining the abductive diagnoses of a Bayesian diagnostic problem amounts to computing $P(\delta_C \mid \omega)$, and then finding the δ_C which maximises $P(\delta_C \mid \omega)$ [3], i.e.

$$\delta_C^* = \arg\max_{\delta_C} P(\delta_C \mid \omega).$$

The probability $P(\delta_C \mid \omega)$ can be computed by Bayes' rule, using the probabilities from the specification of a Bayesian diagnostic system:

$$P(\delta_C \mid \omega) = \frac{P(\omega \mid \delta_C)P(\delta_C)}{P(\omega)}. \quad (4)$$

Using the independence relations derived from a Bayesian diagnostic system, basic probability theory and the definition of a Bayesian diagnostic problem yield the following derivation:

$$P(\omega \mid \delta_C) = P(i_\omega) \sum_{i_u} P(i_u) \sum_{o_u} \prod_{c} P(O_c \mid \pi(O_c)). \quad (5)$$

Let the notation $S : v$ denote that v is a member of the set S. Then, the following additional assumptions can be made explicit:

- $P(O_c \mid \pi(O_c) : a_c) = P(O_c \mid a_c)$, i.e. the probabilistic behaviour of a component that is faulty is independent of its inputs;
- $P(O_c \mid \pi(O_c) : \bar{a}_c) \in \{0, 1\}$, i.e. normal components behave deterministically.

Decomposition

We start by distinguishing between various types of components, inputs and outputs, in order to make the necessary distinction:

- The sets of components assumed to function normally and abnormally will be denoted by $C^{\bar{a}}$ and C^a, respectively, with $C^{\bar{a}}, C^a \subseteq \mathrm{COMPS}$;
- The sets $C^{\bar{a}}$ and C^a are partitioned into sets of components, for observed and unobserved outputs, indicated by the sets $C_\omega^{\bar{a}}$, $C_u^{\bar{a}}$, C_ω^a, and C_u^a, respectively.

Thus, $C^{\bar{a}} = C_\omega^{\bar{a}} \cup C_u^{\bar{a}}$ and $C^a = C_\omega^a \cup C_u^a$.

It is now possible to decompose the product of the entire set of components, as follows:

$$\prod_{c} P(O_c \mid \pi(O_c)) = \prod_{c \in C_u^{\bar{a}}} P(O_c \mid \pi(O_c) : \bar{a}_c) \prod_{c \in C_\omega^{\bar{a}}} P(O_c \mid \pi(O_c) : \bar{a}_c) \times \prod_{c \in C_u^a} P(O_c \mid a_c) \prod_{c \in C_\omega^a} P(O_c \mid a_c).$$

As the probability of an output of a normally functioning component was assumed to be either 0 or 1, i.e. $P(O_c \mid \pi(O_c) : \bar{a}_c) \in \{0, 1\}$, these probabilities yield, when multiplied, Boolean functions. One of these Boolean functions, denoted by φ, is defined as follows: $\varphi(o_u^{\bar{a}}, o_u^a, i^{\bar{a}}) = \prod_{c \in C_u^{\bar{a}}} P(O_c \mid \pi(O_c) : \bar{a}_c)$, where the set of parents $\pi(O_c)$ may, but need not, contain random variables from the sets of random variables O_u^a and $I^{\bar{a}}$.

Theorem 1 *Let $\mathcal{P}_B = (\mathcal{S}_B, \omega)$ be a Bayesian diagnostic problem. Then, $P(\omega \mid \delta_C)$ can be expressed as follows:*

$$P(\omega \mid \delta_C) = P(i_\omega) \sum_{i_u^{\bar{a}}} P(i_u^{\bar{a}}) \sum_{o_u^a} b(o_u^a, i^{\bar{a}}) \prod_{c \in C^{a,o}} p_c \prod_{c \in C^{a,\bar{o}}} (1 - p_c),$$

where $b(o_u^a, i^{\bar{a}}) \in \{0, 1\}$ and $p_c = P(o_c \mid a_c)$.

An alternative version of the theorem can be obtained in terms of expectations using Equation (3) for the Poisson-binomial distribution:

$$P(i_\omega) \sum_{i_u^{\bar{a}}} P(i_u^{\bar{a}}) \sum_{o_u^a} b(o_u^a, i^{\bar{a}}) \prod_{c \in C^{a,o}} p_c \prod_{c \in C^{a,\bar{o}}} (1 - p_c) = P(i_\omega) \prod_{c \in C_\omega^a} P(O_c \mid a_c) \sum_{i_u^{\bar{a}}} P(i_u^{\bar{a}}) \mathcal{E}_P(b_{i^{\bar{a}}}(O_u^a)),$$

i.e. the sum of the mean of the Boolean functions $b_{i^{\bar{a}}}$, which are functions of the unobserved inputs $i_u^{\bar{a}}$, in terms of the probability function P (Equation (1)), weighed by the prior probability of unobserved inputs $i_u^{\bar{a}}$. Combining this with Equation (4) yields $P(\delta_C \mid \omega)$. Thus, to rank diagnoses δ_C probabilistically it is necessary to compute: (*i*) $\mathcal{E}_P(b_{i^{\bar{a}}}(O_u^a))$, the Poisson-binomial distribution mean of the behaviour of the normally assumed components, (*ii*) $P(i_u^{\bar{a}})$, (*iii*) $\prod_{c \in C_\omega^a} P(O_c \mid a_c)$, the observed abnormal components, and (*iv*) the prior $P(\delta_c)$. Note that $P(i_\omega)$ is just an invariant weight factor and $P(\omega)$ is a normalising factor; both can be ignored in computing the probabilistic ranking of the diagnoses δ_C.

Conclusions

We have shown that probabilistic model-based diagnosis can be decomposed into computation of various probabilities, in which a central role is played by the Poisson-binomial distribution. When all probabilities $p_c = P(o_c \mid a_c)$ are assumed to be equal, a common simplifying assumption in model-based diagnosis, the analysis reduces to the use of the standard binomial distribution.

So far, no attempts were made in related research to look inside what happens in the diagnostic process, as was done in this paper. We expect that it becomes thus possible to investigate further variations in probabilistic model-based diagnosis, for example, by adopting assumptions different from those in this paper with regard to fault behaviour in systems.

References

[1] L. Le Cam, 1960. An approximation theorem for the poisson binomial distribution. *Pacific Journal of Mathematics*, 10:1181–1197.

[2] J. de Kleer and B.C. Williams, 1987. Diagnosing multiple faults. *AI*, 32:97–130.

[3] J.A. Gámez, 2004. *Abductive inference in Bayesian Networks: a review*, pages 101–120. Springer, Berlin.

[4] H. Geffner and J. Pearl, 1987. Distributed diagnosis of systems with multiple faults. In *Proc. of the 3rd IEEE Conference on AI Applications*, pages 156–162. IEEE.

[5] R. Reiter, 1987. A theory of diagnosis from first principles. *AI*, 32:57–95.