Accounting for False Indication in a Bayesian Diagnostics Framework

John W. Sheppard, PhD
ARINC Engineering Services, LLC
2551 Riva Road
Annapolis, MD 21401
John.Sheppard@arinc.com

Abstract—**Accounting for the effects of test uncertainty is a significant problem in test and diagnosis. Specifically, how does one assess the level of uncertainty and then utilize that assessment to improve diagnostics? One approach, based on measurement science, is to treat the probability of a false indication (false alarm or missed detection) as the measure of uncertainty. Given the ability to determine such probabilities, a Bayesian approach to diagnosis suggests itself. In the following, we present a mathematical derivation for false indication and apply it to the specification of Bayesian diagnosis. We draw from measurement science, reliability theory, and the theory of Bayesian networks to provide an end-to-end probabilistic treatment of the fault diagnosis problem.**

Index Terms—**Bayesian inference, diagnosis, false indication, measurement science.**

I. Introduction

WHAT is a false indication? How can false indications be predicted? How can one use this prediction to aid diagnosis rather than hinder diagnosis? These are the questions that concern us in this paper. Further, given recent efforts of the IEEE to standardize metrics for testability and diagnosability, answers to these questions are particularly relevant. The long-time bane of embedded diagnostics (such as BIT), it is becoming increasingly important to be able to predict and track false indications in test and diagnosis. Unfortunately, it is generally accepted that, even if we are able to predict false indications, verifying these predictions with field data is problematic. At best, we might be able to measure false-removal rates, cannot duplicate rates, or retest-OK rates, but these rates can be attributed to multiple causes that include false indications.

MIL-STD-2165 (now MIL-HDBK 2165A) [5] defines a false alarm as "a fault indicated by built-in test (BIT) or other monitoring circuitry where no fault exists." The former MIL-STD-1309C [4] defines false alarm the same way, limiting the definition to BIT. The IEEE standard dictionary is in agreement, defining false alarm more generally as "an indicated fault where no fault exists" [2]. A missed detection can be defined similarly as "an indication of no fault where a fault exists." Combining, we can then define a false indication simply as "an erroneous test result."

Work from the measurement science community provides a means for predicting uncertainty in measurement that can be used as the basis for predicting false indications. As long as we rely on calibration laboratories to validate these uncertainties and we focus on applying the uncertainties to the test results, we can control the impact of our limited ability to verify false alarms. Specifically, we look to apply measurement uncertainty as a component in the overall model for predicting false indications, combined with probability of failure and probability of observing failure. We then use these probabilities in the context of a Bayesian Belief Network to perform diagnosis. We recognize that there are other possible causes for false indications, including environmental factors, human error, and systems operating near their functional limits; however, we limit the discussion to measurement uncertainty as a starting point in our exploration.

The remainder of this paper is organized as follows. Section II provides basic background material from measurement theory. Sections III–V provide the main tools used in this paper—predicting probability of false indication, predicting probability of failure, and predicting instrument uncertainty respectively. We then provide the main result in Section VI, showing how to combine these predictions in Bayesian diagnosis. We provide concluding remarks in Section VII.

II. Measurement Theory

Under ideal conditions, with "accurate" measurement devices and full information about the system under test, diagnostic errors will still occur. These are mathematical residues that cannot be eliminated completely. It is important to understand that some complicated elements of diagnosis involve chasing these residues [9].

A prominent issue in testing is the impact of *precision* and *accuracy* of test resources on the certainty in the resulting test outcome. Formalizing the impact of precision and accuracy on test confidence comes from work in measurement theory. Given the need to formalize these factors, we can consider each of them in terms of probability distributions. The precision of a resource characterizes the amount of "scatter" one can expect when repeated measurements are made of the same signal. Typically, precision is applied to the measurement process or resource rather than the measurement itself. Thus the precision of a resource can be characterized by considering a probability distribution of measurements. A precise resource would yield a narrow distribution of measured values, while an imprecise resource would yield a wide distribution of measured values.

0-7803-7837-7/03/$17.00 ©2003 IEEE

On the other hand, the accuracy of a measurement corresponds to the level of deviation from some *reference value*. As with precision, accuracy can be modeled as a probability distribution. Typically, accuracy is applied to the measured value rather than the process or instrument used in taking the measurement. Frequently, it is determined through taking several independent measurements and computing the mean deviation from the reference value.

Statistically, inaccuracy is evaluated in terms of two potential sources—referred to as random error and systematic error. Systematic error is also called *bias*. Many consider accuracy only in terms of random error in which bias is ignored or calibrated out (in which case the estimate becomes an unbiased estimate of error). Others consider total error in which the bias is included.

To determine the appropriate tolerances for a particular test, and thereby determine relevant test outcomes, one must consider the precision of the required resources. The precision will be used to determine the associated bias of the instrument. To determine the precision, typically, several independent measurements are taken for that instrument under known conditions and the resulting distribution determined. Then the width of the distribution is determined by using, for example, variance (defined as the average deviation from the mean of the distribution).

Next, measurement error is considered by examining the distribution characterizing accuracy. Given the system to be tested and a measurement to be made, a distribution of "nominal values" for that measurement can be determined. From this distribution, in the simplest case, PASS/FAIL criteria are established based on the probability of a measurement occurring within some set of defined limits applied to that measurement.

Generally the PASS/FAIL criteria are determined by considering expected values for a fault-free unit. "Significant" deviation from these expected values results in the FAIL outcome for that test. The limits define what is meant by "significant." Ideally, conditions are considered in which a similar distribution is determined in the presence of faulty conditions, and the limits are set such that maximum separation between the distributions (in terms of probability density) is obtained. During testing, a measurement value is typically mapped into a discrete outcome determined based upon which side of these limits the measurement falls.

Once the measured value is mapped into a discrete outcome, interpreting the test results becomes problematic. Specifically, when considering a single measurement, the probability of that measurement being within the set limits will be $\Pr(x)$, and the probability of it being out of the limits will be $1 - \Pr(x)$. (Actually, these are conditional probabilities since the actual probability depends on whether or not the unit has failed, as we will see below.)

In determining the base confidence (e.g., either the expected confidence or the maximum confidence) in the outcome, we want to consider both the probability of observing a PASS outcome given the test should pass and the probability of observing a FAIL outcome given the test should fail, i.e., $\Pr(o(P) \mid P)$ and $\Pr(o(F) \mid F)$ respectively. Problems occur when a measurement for a good unit is out of tolerance (referred to as a "Type I error") or a measurement for a bad unit is in tolerance (referred to as a "Type II error"). Given that we consider a test outcome of FAIL to be a "positive" indication of a problem, Type I errors correspond to false positives while Type II errors correspond to false negatives. These two types of errors are depicted in Fig. 1.

III. Determining Probability of False Indication

Based on the fact Type I and Type II errors occur in practice, we want to know the rate at which we will be faced with such errors. Repeating the definitions of Section I, we define a *false alarm* to be an indication of a fault where no fault exists. Similarly, we define a *missed detection* as an indication of no fault when a fault does exist. Finally, a *false indication* is either a false alarm or a missed detection. Each of these concepts will be defined mathematically in this section.

Formally, a false indication occurs whenever a test result is inconsistent with the underlying state of the unit being tested. When determining the probability of a false indication occurring, we must consider these joint effects. Consider the situation where we only have a single test, and it is possible for that test to either PASS or FAIL. Further, assume that we consider the possibility of a fault truly being present or not being present. Thus we need to consider four combinations as follows:

- True pass: We observe the test passing ($o(P)$) and there is no fault (P).
- Missed detection: We observe the test passing ($o(P)$) and there is a fault (F).
- False alarm: We observe the test failing ($o(F)$) and there is no fault (P).
- True fail: We observe the test failing ($o(F)$) and there is a fault (F).

Now we want to consider the probabilities of these joint events occurring. This corresponds to $\Pr(o(P) \wedge P)$, $\Pr(o(P) \wedge F)$, $\Pr(o(F) \wedge P)$, and $\Pr(o(F) \wedge F)$ respectively. Note that these four situations cover all possible combinations, so we can define the probability of false indication, $\Pr(FI^u)$, as follows:

$$\Pr(FI^u) = \frac{\Pr(o(P),F) + \Pr(o(F),P)}{\Pr(o(P),P) + \Pr(o(F),F) + \Pr(o(P),F) + \Pr(o(F),P)}. \quad (1)$$

Note that, from a diagnostic perspective, this definition is not particularly useful. This is because the process of performing diagnosis corresponds to observing test results and then inferring diagnoses (i.e., possible faults) within the system. Thus, in reality we are interested in the conditional probabilities of the state of the unit given the test results. Recall the definition of conditional probability that says

$$\Pr(A \mid B) = \frac{\Pr(A,B)}{\Pr(B)}. \quad (2)$$

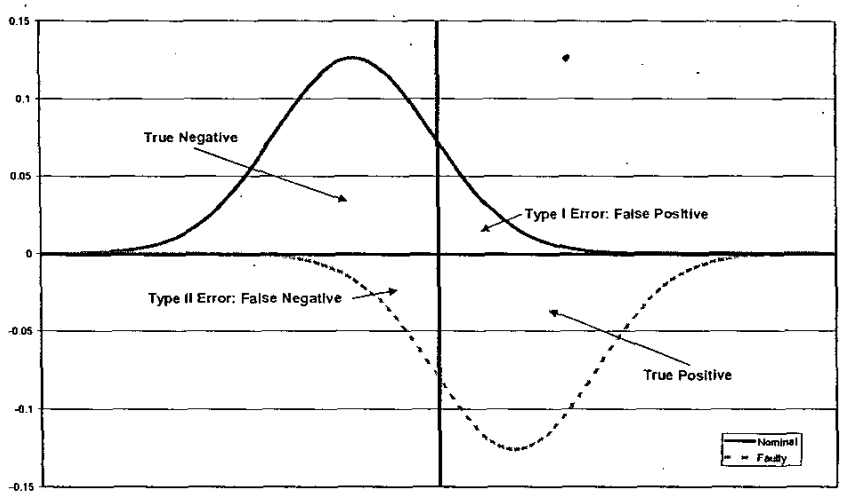

Fig. 1. Distributional sources of Type I and Type II errors, respectively.

Using this definition, we can rewrite Equation 1 as follows:

$$\Pr(FI^{u}) = \Pr(F \mid o(P))\Pr(o(P)) + \Pr(P \mid o(F))\Pr(o(F)) \quad (3)$$

Unfortunately, it is very difficult to determine the probabilities, $\Pr(o(P))$ and $\Pr(o(F))$ respectively, so it is easier to consider the equivalent form:

$$\Pr(FI^{u}) = \Pr(o(P) \mid F)\Pr(F) + \Pr(o(F) \mid P)\Pr(P) \quad (4)$$

This form will be particularly useful in diagnosis in that $\Pr(F)$ is derivable directly from the failure rate of the unit and $\Pr(o(F) \mid F) = 1 - \Pr(o(P) \mid F)$ is a measure of the ability of a test to detect the associated failure which we will derive below.

Now we will consider these alternative conditional probabilities. Specifically,

- $\Pr(o(F) \mid F)$ is the probability we observe a failure, given the unit is faulty.
- $\Pr(o(P) \mid P)$ is the probability we observe no failure, given the unit is not faulty.
- $\Pr(o(F) \mid P)$ is the probability we do not observe a failure, given the unit is not faulty.
- $\Pr(o(P) \mid F)$ is the probability we do not observe a failure, given the unit is faulty.

Then we can consider the confidence (or actually, the lack of confidence) in our test, independent of the failure rate of the unit tested, to be the following.

$$\Pr(FI^{c}) = \frac{\Pr(o(F) \mid P) + \Pr(o(P) \mid F)}{\Pr(o(F) \mid P) + \Pr(o(F) \mid F) + \Pr(o(P) \mid P) + \Pr(o(P) \mid F)} \quad (5)$$

This can be simplified by observing that the sum of the conditional probabilities $\Pr(o(F) \mid P) + \Pr(o(P) \mid P) = 1$ and $\Pr(o(P) \mid F) + \Pr(o(F) \mid F) = 1$. Thus

$$\Pr(FI^{c}) = \frac{1}{2}[\Pr(o(F) \mid P) + \Pr(o(P) \mid F)]. \quad (6)$$

To make sure that we have not violated the axioms of probability (specifically, the axiom that states that the sum of the probabilities over a domain must be 1), we note the following. First, $\Pr(o(P) \mid P) = 1 - \Pr(o(F) \mid P)$ and $\Pr(o(F) \mid F) = 1 - \Pr(o(P) \mid F)$. So the probability of a true indication ($\Pr(TI^{c})$) can be computed similarly:

$$\Pr(TI^{c}) = \frac{\Pr(o(F) \mid F) + \Pr(o(P) \mid P)}{\Pr(o(F) \mid P) + \Pr(o(F) \mid F) + \Pr(o(P) \mid P) + \Pr(o(P) \mid F)}. \quad (7)$$

Observe that, given the above axiom of probability, we must have $\Pr(TI^{c}) = 1 - \Pr(FI^{c})$. Using the complementary of probabilities defined above we have,

$$\begin{aligned}\Pr(TI^{c}) &= \frac{(1 - \Pr(o(F) \mid P)) + (1 - \Pr(o(P) \mid F))}{\Pr(o(F) \mid P) + \Pr(o(F) \mid F) + \Pr(o(P) \mid P) + \Pr(o(P) \mid F)} \\ &= 1 - \frac{1}{2}[\Pr(o(F) \mid P) + \Pr(o(P) \mid F)] \\ &= 1 - \Pr(FI^{c})\end{aligned}$$

So we see that the axioms of probability hold as desired.

Thus, assuming independence in testing, we can derive corresponding test accuracy probabilities from instrument accuracy and reliability measures. Assuming T independent tests, this can be given as

$$\Pr(\mathbf{FI}) = \frac{\sum_{i=1}^{T} \Pr(FI_i^c)}{\sum_{i=1}^{T} \Pr(FI_i^c) + \sum_{i=1}^{T} \Pr(TI_i^c)} = \frac{1}{T} \sum_{i=1}^{T} \Pr(FI_i^c) \quad (8)$$

When wanting to predict $\Pr(FI^c)$, we can apply Baye's Rule to each of the respective terms where,

$$\Pr(o(F) \mid P) = \frac{\Pr(P \mid o(F)) \Pr(o(F))}{\Pr(P)}$$
$$\Pr(o(P) \mid F) = \frac{\Pr(F \mid o(P)) \Pr(o(P))}{\Pr(F)} \quad (9)$$

Notice that $\Pr(F)$ is simply a failure probability, which can be derived from reliability models. Further, $\Pr(P \mid o(P))$ can be derived from the instrument uncertainty, and $\Pr(o(P))$ corresponds to the prior probability of observing no fault (i.e., $1 - \Pr(o(F))$, i.e., the confidence in the test result.

IV. Determining Probability of Failure

Traditional reliability theory determines probability of failure based on the failure rate of an item. Failure rates are determined using combinations of historical data and physical models and are typically expressed in terms of numbers of failures per million hours (or some other unit). Traditionally, the failure rate is denoted λ, and the mean time between failures (MTBF) is $1/\lambda$.

Usually, failure probability is computed under the assumption of an exponential probability distribution, based on a further assumption of constant failure rate [3].[1] In other words, if we let D_i denote failure (or diagnosis) i, the probability of failure D_i is computed as

$$\Pr(D_i) = 1 - \exp[-\lambda_i t] \quad (10)$$

In performing diagnosis under uncertainty, it is common to compute probability of failure over members of an ambiguity group as a relative measure directly from the failure rates. Specifically, we define an ambiguity group to be a set of diagnoses for which no tests have been performed that differentiate the set (perhaps because they do not exist). Assume (for simplicity) that a diagnosis corresponds to a single fault mode. Assume also that each fault mode has an associated failure rate. Then given a set of diagnoses D_i in an ambiguity group **A**,

$$\Pr(D_i \mid \mathbf{A}) = \frac{\lambda_{D_i}}{\sum_{D_j \in \mathbf{A}} \lambda_{D_j}} \quad (11)$$

However, the failure probability is time-dependent and Markov in nature. In other words, the actual probability of a fault depends upon the time since the last evaluation of that unit. Since ambiguity group **D** corresponds (as defined above) to a group of possible diagnoses that have not been differentiated, they have not yet been evaluated; therefore, the more accurate method for computing the relative probabilities of the diagnoses in **A** is,

$$\Pr(D_i) = \frac{1 - \exp[-\lambda_{D_i} t]}{\sum_{D_j \in \mathbf{A}} (1 - \exp[-\lambda_{D_j} t])} \quad (12)$$

Suppose we have three possible diagnoses with failure rates (assuming common units such as per million hours) set at 0.001, 0.005, 0.01, 0.05, and 0.1 respectively. Using Equation 11, we find the relative failure probabilities for these are approximately 0.006, 0.03, 0.06, 0.3, and 0.6 respectively. However, if we use Equation 12, we find that due to the dependence on time, these relative probabilities drastically change, eventually converging with time to uniform. We can then use this time-dependent view in the calculation of probability of false indication, as defined above.

V. Determining Instrument Uncertainty

Consider now the task of determining uncertainty in a measurement from some instrument. Typically, a measurement μ is mapped to a PASS or FAIL outcome when testing; however, based on the uncertainty of the measurement, u_μ, the probability of observing the PASS or FAIL outcome can be determined.

Note that when measuring the state of a unit under test, we are not directly measuring that state but are measuring some signal that depends on a number of independent variables (inputs, internal state, fault conditions, environmental conditions, etc.). Thus we can consider the measurement process as one of estimating a function $y = f(x_1, \ldots x_n)$. Uncertainty in the dependent variable y ultimately depends on the uncertainty in the independent variables x_i, (denoted $u(x_i)$ respectively).

Typically, the uncertainty is determined through one of two methods: Type A uncertainty is based on statistical sampling, and Type B uncertainty is based on the assumption of some underlying distribution [10]. If we assume a Gaussian distribution, we find that, frequently, the Type A and Type B uncertainties are approximately the same. Either way, the objective is to identify *standard uncertainty*, which is typically treated as a standard deviation of the underlying distribution.

More specifically, consider some measurement, x_m that is intended to observe some "true value" x. The uncertainty associated with this measurement is then u_x, and we say that $x = x_m \pm u_x$. Recognizing that the estimation process involves combining multiple measurements, we get that we are seeking some $y = f(x_1, \ldots, x_n)$. Given that the uncertainty u_i of independent variable x_i can be determined either statistically or analytically, the goal is to determine the uncertainty in the dependent variable y. The "combined" standard uncertainty of y, denoted $u_c(y)$, represents the estimated standard deviation in the result, and is computed as

[1] When the failure rate is not constant, it is common for the Weibull distribution to be used instead. It is relevant to observe that, for the Weibull distribution, when the shape parameter β is set to 0, the distribution reduces to exponential.

$$u_c(y) = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial y}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}u(x_i,x_j)}. \quad (13)$$

and is called the *law of uncertainty propagation* [1], [11]. Here, $u(x_i)$ denotes the estimated variance of x_i, and $u(x_i,x_j)$ denotes the estimated covariance associated with x_i and x_j. This law is derived from a first-order Taylor series approximation of the function $y = f(x_1, \ldots, x_n)$. For many functional forms, Equation 13 reduces to simple linear forms. For complex cases where there is significant non-linearity, higher-order approximations are required.

Interpreting the result of Equation 13, since we are treating standard uncertainty as the standard deviation of the estimate y and assuming a Gaussian (i.e., normal) distribution, we note that the actual value $Y = y \pm u_c$ approximately 68% of the time and $Y = y \pm 2u_c$ approximately 95% of the time. Note that this uncertainty can also be interpreted as the amount of deviation off of the true value.

Consider now a test that has been "shifted" such that its distribution is centered on zero with a nominal range of $\pm z$. The probability of the test actually PASSing given the measurement y is given as,

$$\Pr(P \mid y) = \Pr(P \mid o(P))\Pr(o(P) \mid y) + \Pr(P \mid o(F))\Pr(o(F) \mid y). \quad (14)$$

However, when conditioning on observing a PASS, we note that the second term drops out since $\Pr(o(F) \mid y) = 0$. Also, since $\Pr(o(P) \mid y) = 1$, we have,

$$\Pr(P \mid y) = \Pr(P \mid o(P)). \quad (15)$$

Similarly,

$$\Pr(F \mid y) = \Pr(F \mid o(F))\Pr(o(F) \mid y) + \Pr(F \mid o(P))\Pr(o(P) \mid y) \quad (16)$$

So by the same argument, when conditioning only on observing a FAIL, we find that

$$\Pr(F \mid y) = \Pr(F \mid o(F)). \quad (17)$$

Without loss of generality, focus on $\Pr(P \mid y) = \Pr(P \mid o(P))$. If we apply Bayes rule, we get,

$$\Pr(P \mid y) = \frac{\Pr(y \mid P)\Pr(P)}{\Pr(y)}.$$

After examining the underlying distributions for each of the probabilities on the right hand side (where Ψ corresponds to $\Pr(P)$, which is a constant), we get

$$\Pr(P \mid o(P)) = \Psi\left[\frac{1}{\sqrt{2\pi}u_c}\int_{-z}^{z}\exp\left[-\frac{1}{2}\left(\frac{x-y}{u_c}\right)\right]dx\right] \quad (18)$$

and

$$\Pr(F \mid o(F)) = 1 - \Psi + (1 - \Psi)\Pr(P \mid o(P)). \quad (19)$$

We now have all of the pieces required to determine test confidence for Bayesian diagnosis.

It is interesting to note that repeated measurements enable revision of uncertainty and can tend to reduce uncertainty. The idea of "repeat polling" has been proposed as a technique for reducing the probability of a false indication by allowing transient signals and other sources of measurement uncertainty to "dampen out" [7]. Specifically, using this idea of repeat polling, uncertainty can be revised as

$$u_c(t) = \left(\frac{1}{u_c^2(0)} + \frac{1}{u_c^2(t-1)}\right)^{-\frac{1}{2}}. \quad (20)$$

Given constant uncertainty (for a particular measurement type), we see that

$$u_c(t) = \left(\sum_{i=1}^{t}\frac{1}{u_c^2}\right)^{-\frac{1}{2}} = u_c/\sqrt{t}. \quad (21)$$

Further, the best estimate of the measurand based on repeated measurement can be determined by computing the weighted sum,

$$y = x(t)\left(\frac{\gamma^2}{1+\gamma^2}\right) + y(t-1)\left(\frac{1}{1+\gamma^2}\right). \quad (22)$$

where $x(t)$ is the current measurement, $y(t-1)$ is the best estimate through $t-1$ measurements, and $\gamma = u_c(t-1)/u_c(0)$ represents the ratio between uncertainty at the previous time step and initial uncertainty (thus showing the amount of relative reduction at this time step).

VI. Diagnosis with Bayesian Networks

Recall that $\Pr(TI) = 1 - \Pr(FI)$. We will use this as a measure of confidence in our test results. Note that, if we have measures for $\Pr(o(F) \mid P)$ and $\Pr(o(P) \mid F)$ for each of the measured tests, then we can use Bayes Rule and break the confidence into its respective parts for PASS or FAIL. Specifically, our confidence in a PASS result becomes $\Pr(P \mid o(P)) = 1 - \Pr(F \mid o(P))$. Similarly, our confidence in a FAIL result becomes $\Pr(F \mid o(F)) = 1 - \Pr(P \mid o(F))$.

Once we have formulated probabilities such as these, we are in a position to use these probabilities as confidence values in a diagnostic system. There are many diagnostic systems that allow for reasoning under uncertainty; however, the natural approach that uses probabilities directly is the *Bayesian Belief Network* (BBN) [6]. Formally, a BBN = $\langle \mathbf{V}, \mathbf{E}, \mathbf{C} \rangle$, where

- $\mathbf{V}$ is a set of vertices corresponding to random variables $V_i \in \mathbf{V}$,
- $\mathbf{E}$ is a set of directed edges $e_{ij} \in \mathbf{E}$ where the source of the edge corresponds to V_i, the destination of the edge

corresponds to V_j, and the edge represents a conditional dependence relationship of V_j on V_i,

- **C** is a set of conditional probability tables $C_{Pa(i)} \in \mathbf{C}$ where each entry provides the probability of V_i given the set of parents of V_i ($Pa(V_i)$).

The structure of the BBN depends on the concept of conditional independence. Given two random variables, V_1 and V_2, V_1 is *conditionally independent* of V_2 iff $Pr(V_1 \mid V_2) = Pr(V_1)$. Further, given random variables V_1, V_2, and V_3, V_1 is *conditionally independent* of V_2 given V_3 iff $Pr(V_1 \mid V_2, V_3) = Pr(V_1 \mid V_3)$ and $Pr(V_2 \mid V_1, V_3) = Pr(V_2 \mid V_3)$. Within the BBN, we say that a random variable V_i is conditionally independent of all other variables not connected to it given its parents $Pa(V_i)$.

If we continue to assume that tests are performed independently from one another, then we connect tests (as random variables) to possible diagnoses (as random variables). We can also assume that the diagnoses are independent from one another (i.e., the existence of one fault does not cause another fault to occur). Thus the only dependence relationships modeled are between tests and diagnoses.[2] In addition, we must find a way to handle the relationships between the intended states of the tests and the observations of those tests. Thus, we can define a causal model such as the one illustrated in Fig. 2.

The diagnostic problem consists of inferring the probability of each of the diagnoses in the BBN given the test results. Note that the joint probability distribution over all of the variables in the BBN, $Pr(\mathbf{V})$, is given as the product of the probability distributions of each over each of the vertices (random variables) conditioned on their parents, i.e.,

$$Pr(\mathbf{V}) = \prod_{V_i \in \mathbf{V}} Pr(V_i \mid Pa(V_i)) \tag{23}$$

Assume we subdivide the set of random variables **V** into two subsets, **T** and **D** corresponding to tests and diagnoses respectively, as discussed above. Further, assume that $\mathbf{T} \cup \mathbf{D} = \mathbf{V}$ and $\mathbf{T} \cap \mathbf{D} = \varnothing$ (i.e., **T** and **D** are disjoint but define the entire set of random variables **V**). Finally, assume **T** contains the set of observations (i.e., the evidence variables) and **D** contains everything else (including, if needed, the true states of the tests). Given a target set of test results, τ, we calculate $Pr(\mathbf{T}' = \tau)$ (where $\mathbf{T}' \subseteq \mathbf{T}$) by marginalizing out the remaining variables given by $\mathbf{V} \setminus \mathbf{T}'$. This is done by summing over all $Pr(\mathbf{V} \setminus \mathbf{T}', \mathbf{T}' = \tau)$, where "\" denotes set difference:

$$Pr(\mathbf{T}' = \tau) = \sum_{\mathbf{V} \setminus \mathbf{T}} Pr(\mathbf{V} \setminus \mathbf{T}', \mathbf{T}' = \tau) \tag{24}$$

When evaluating BBNs, there are several techniques for deriving probabilities based on evidence values. For the restricted set of BBNs corresponding to polytrees, exact solutions exist. For general BBNs, computing the joint probability distribution is *NP*-hard; therefore, a variety of techniques existing ranging from constructing join-trees to applying various sampling techniques such as Markov Chain Monte Carlo (MCMC) and importance sampling.

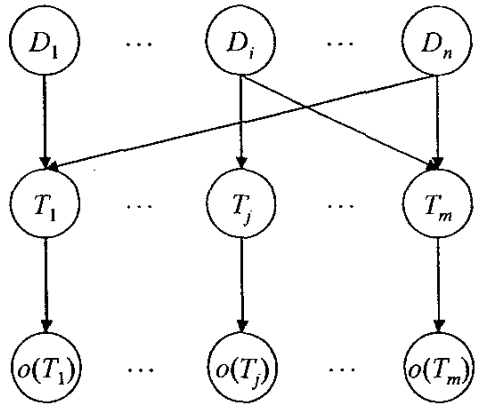

Fig. 2. Diagnostic Bayesian Network Structure

Applying the method described above for determining probability of false alarm or probability of missed detection, we can then use these probabilities directly without the need for stochastic sampling methods. By assuming independence among the random variables in **D** (i.e., the diagnoses) as well as independence among the random variables in **T** (i.e., the test observations), the characteristics of direction-dependent separation (i.e., *d*-separation) allow for simple propagation of the probabilities to the diagnoses. Specifically, we say that a set of evidence nodes E in a BBN (i.e., nodes that can be observed) *d*-separates two sets of nodes X and Y if every undirected path from a node $x \in X$ to a node $y \in Y$ is "blocked" given E. A path is blocked given E if there exists some node z on the path for which one of the following three conditions holds.

1. If $z \in E$, z has one parent that is on the path, and z has one child that is also on the path.
2. If $z \in E$ and z has two children, both of which are on the path.
3. If $z \notin E$, $\forall x\ pa(x) = z \notin E$, and z has two parents, both of which are on the path.

Then, we can assert the diagnoses are conditionally independent of each other given the test results because of the third case, even though the diagnoses have no parents. Specifically, taking the graph transpose of the network (i.e., reversing the directions of the arcs) results in the evidence variables being the parents of the diagnoses, and none of the diagnoses being evidence variables.

Given the conditional independence of the diagnoses, we can then compute the posterior probabilities of each of the diagnoses given the test results as follows. First, we will

[2] These two assumptions are not necessarily true, and in general such dependence relationships, when known, can be modeled directly by the BBN by inserting appropriate edges between pairs of tests or between pairs of diagnoses.

partition the random variables explicitly into three sets: **D** (the diagnoses), **T** (the true test states), and **O** (the test observations). The evidence variables will be restricted to **O**.

$$\begin{aligned} \Pr(D_i \mid \mathbf{O}) &= \alpha \Pr(\mathbf{O} \mid D_i)\Pr(D_i) \\ &= \alpha[\Pr(\mathbf{O} \mid \mathbf{T})\Pr(\mathbf{T} \mid D_i)\Pr(D_i)] \\ &= \alpha \Pr(D_i)\sum_{T_j \in \mathbf{T}} \Pr(o(T_j) \mid T_j)\Pr(T_j \mid D_i) \end{aligned} \tag{25}$$

Here, α is a normalizer over the set **D**, equal to

$$\alpha = \sum_{D_i \in \mathbf{D}} \Pr(D_i)\sum_{T_j \in \mathbf{T}} \Pr(o(T_j) \mid T_j)\Pr(T_j \mid D_i).$$

Observe that $\Pr(T_j \mid D_i) \in \{0, 1\}$, so the members of the sum are restricted only to those tests that observe D_i. Then we only need to consider $\Pr(D_i)$, which corresponds to the prior probability for D_i, and $\Pr(o(T_j) \mid T_j)$, which corresponds to the confidence value assigned to the observed test result. Using the Baye's maximum *a posteriori* hypothesis, we determine the most likely diagnosis simply as

$$D_{MAP} = \arg\max_{D_i \in \mathbf{D}} \{\Pr(D_i \mid \mathbf{T})\}. \tag{26}$$

In other words, we provide the most probable diagnosis as a means of minimizing expected error in the diagnostic process.

VII. Summary

In this paper, we discussed a formal approach to predicting probability of false indication and using this information as a means of setting test uncertainty in a diagnostic and prognostic system. Diagnosis was performed based on Bayesian Belief Networks and incorporated information on failure probability, instrument uncertainty, and the predictions for false indication. This is the first time such a unified, Bayesian view of the test and diagnosis problems has been presented. The advantage to the discussed approach is that it provides a formally consistent and theoretically sound approach to diagnosis that can be adapted and matured as better estimates of the associated probabilities become available.

Acknowledgment

We would like to thank the members of the IEEE SCC20 Diagnostic and Maintenance Control Subcommittee for reviewing several critical parts of this paper. Specifically, we thank Tim Wilmering, Mark Kaufman, Greg Bowman, Les Orlidge, and Li Pi Su for their comments as the ideas were formulated. This work was a direct outgrowth of the standardization work being performed under the DMC for testability/diagnosability prediction and diagnostic maturation.

References

[1] D'Agostini, Giulo, *Probability and Measurement Uncertainty in Physics—A Bayesian Primer*, DESY 95-242, Deutsches Electronen Synchnotron (DESY), December 1995.

[2] IEEE Std 100-2000, *The Authoritative Dictionary of IEEE Standard Terms*, Seventh Edition, IEEE Standards Association Press, Piscataway, NJ, 2000.

[3] Klion, Jerome, *Practical Electronic Reliability Engineering*, Van Nostrand Reinhold, New York, 1992.

[4] MIL-STD 1309C, *Definition of Terms*, November 18, 1989.

[5] MIL-STD 2165, *Testability Program for Electronic Systems and Equipments*, January 26, 1985.

[6] Pearl, Judea, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1991.

[7] Phillips, S.D., W. T. Estler, M. S. Levenson, and K. R. Eberhardt, "Calculation of Measurement Uncertainty Using Prior Information," *Journal of Research of the National Institute of Standards and Technology*, Vol 103, No. 6, November–December 1998, pp. 625–632.

[8] Sheppard, John W. and Mark Kaufman, "IEEE 1232 and P1522 Standards, *AUTOTESTCON 2000 Conference Record*, IEEE Press, September 2000.

[9] Sheppard, John W. and William R. Simpson, "Managing Conflict in System Diagnosis," in *Research Perspectives and Case Studies in System Test and Diagnosis*, Kluwer Academic Publishers, 1998.

[10] Taylor, Barry N. and Chris E. Kuyatt, "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results," NIST Technical Report 1297, 1994.

[11] Tripp, John S. and Ping Tcheng, "Uncertainty Analysis of Instrument Calibration and Application," NASA/TP-1999-209545, October 1999.