Computing Manifestations of Max-Size Min-Cardinality Ambiguity Groups

Alexander Feldman and **Johan de Kleer** and **Gregory Provan** [1]

Abstract. The application of Model-Based Diagnosis to systems that are under-observed (e.g., sensor-lean systems) is severely hindered by the ambiguity of the diagnostic result. In the worst-case, even in very restricted frameworks such as the one presented in this paper, an observation may lead to an exponential number of diagnoses. This is the case even if we impose a minimality criterion such as cardinality-minimal diagnoses. To solve this problem researchers have proposed a number of information gathering approaches such as probing and active-testing. There is little literature however, on evaluating the performance of these information-gathering algorithms. In this paper we analyze a new class of observations that maximize the size of the minimal-cardinality (MSMC) ambiguity group. We show a probing framework for which these observation lead to worst-case probing sessions. We exhaustively compute these MSMC initial observations for a benchmark of 74XXX digital circuits.

1 Introduction

Model-Based Diagnosis (MBD) aims to compute, given a model SD and an observation α, diagnoses, minimal under some minimality criterion, e.g., the minimal-cardinality set of faulty components. Since the model SD is known *a priori*, much work has been devoted to optimizing the inference process by pre-processing SD. However, little work has focused on how α affects the inference process. This paper focuses on how α affects the cardinality and minimality properties of diagnoses. We define a metrics, MSMC, which addresses diagnostic ambiguity (or indistinguishability of diagnoses).

Prior work on probing [3] is multi-valued in it does not restrict the variables domains to Boolean values. This makes the proposed approach very useful in practical situations, however it makes the analysis of algorithmic performance more difficult. In this paper we propose a strictly propositional framework that allows more intuitive presentation of assumptions, and analysis of algorithms and properties.

The contributions of this paper are as follows. (1) We define a new class of MSMC observation vectors that are worst-case scenarios for a class of information gathering algorithms and probing in particular. (2) We provide a formal framework for the evaluation of probing algorithms. (3) In the proposed framework, we show a probing algorithm that uses a myopic one-step look-ahead to compute optimal probe variables. We show that MSMC observation vectors are worst-case scenarios for this probing algorithm. (4) We compute MSMC properties for a class of 74XXX benchmarks.

The majority of existing MBD research as well as this paper consider only realistic cases in which the original "injected" faults do not change their location. An interesting alternative to this is a case proposed by de Kleer that we call "The MBD Game". The board of this game is a digital circuit. In the beginning of the game the antagonist "injects" a fault that stays hidden from the protagonist. In each turn of the game, the protagonist proposes a measurement (a probe), the antagonist gives the value of this measurement and changes the location of the fault so the protagonist does not find it. The goal of the protagonist is to minimize the number of probes before finding the fault while the antagonist aims at the opposite (maximizing the number of probes). In a subsequent game the protagonist and antagonist change roles and the winner is the one who uses a smaller number of probes to uniquely determine all faults.

[1] University College Cork, email: a.feldman@ucc.ie, dekleer@parc.com, g.provan@cs.ucc.ie

2 Related Work

To the best of our knowledge, we are the first to define the notion of an MSMC observation vector.

Early work [3] aimed at diagnostic convergence by computing a probe sequence for reducing diagnostic entropy using a myopic search strategy. This paper complements this work by providing a strict probing framework in which we can show worst-case scenarios.

Probing is not the only way to perform diagnostic information gathering. Another approach is active testing [8] in active testing one computes a set of optimal control settings that lead to observations of small cardinality-minimal ambiguity groups. The MSMC framework presented in this paper is also a bound on the performance of this class of algorithms.

The material presented in this paper shows an approach to evaluating the performance of the information gathering part of diagnostic algorithms in worst-case scenarios. The international diagnostic competition DXC [4] has similarly evaluated algorithms. The main goal there, however, is the comparison of different diagnostic algorithms and not stress-testing under worst-case conditions. This work may facilitate similar diagnostic competitions in the future by allowing algorithms to compete on difficult observation vectors.

A problem that is related to MSMC is that of computing observations leading to cardinality-minimal diagnoses of maximal cardinality. These observations are called MFMC observations and are studied in [6].

3 Concepts and Definitions

Our discussion continues by formalizing some MBD notions. This article uses the traditional diagnostic definitions [3], except that we use propositional logic terms (conjunctions of literals) instead of sets of failing components.

Central to MBD, a *model* of an artifact is represented as a Well-Formed Propositional Formula (**Wff**) over some set of variables. We discern subsets of these variables as *assumable* and *observable*.[2]

Definition 1 (Diagnostic System)**.** A diagnostic system DS is defined as the quadruple DS $=\langle$SD, COMPS, IN, OUT, INT$\rangle$, where SD is a propositional theory over a set of variables V, COMPS $\cup$ IN $\cup$ OUT $\subseteq V$, COMPS is the set of assumables, IN is the set of primary inputs, OUT is the set of primary outputs, and INT $= V \setminus \{$COMPS $\cup$ IN $\cup$ OUT$\cup\}$. The set of observables OBS is defined as OBS $=$ IN $\cup$ OUT.

Throughout this article we restrict SD to propositional theories derived from Boolean circuits. We assume that SD $\not\models\bot$, i.e., SD is not faulty (does not lead to diagnoses) when there is no observation. We also assume that the sets IN, OUT, and COMPS are disjoint. Further, we assume that the model SD is acyclic, testable, and connected, i.e., starting from a primary input in IN we can always reach a primary output in OUT, thus defining direction of each connection (we will illustrate this with an example).

The internal variables of SD are all variables in V that are neither assumables nor primary inputs nor primary outputs, i.e., $V \setminus \{$IN $\cup$ OUT $\cup$ COMPS$\}$.

Let COMPS $= \{h_i\}$ for $i = 1, 2, \ldots, n$. We use positive assignments $h_i = \mathbf{True}$, or simply positive literals h_i, to denote healthy components; conversely, we use negative assignments $h_i = \mathbf{False}$, or negative literals $\neg h_i$, to denote faulty components. Other authors use different mnemonics for this: some denote faulty components with "ab" for abnormal, while others denote healthy components using "ok".

Not all propositional theories used as system descriptions are of interest to MBD. Diagnostic systems can be characterized by a restricted set of models, the restriction making the problem of computing diagnosis amenable to algorithms like the ones presented in this article. We consider two main classes of models.

Definition 2 (Weak-Fault Model)**.** A diagnostic system DS $=$ $\langle$SD, COMPS, IN, OUT, INT$\rangle$, OBS $=$ IN $\cup$ OUT, belongs to the class **WFM** iff for COMPS $= \{h_1, h_2, \ldots, h_n\}$, SD is equivalent to $(h_1 \Rightarrow F_1) \wedge (h_2 \Rightarrow F_2) \wedge \ldots \wedge (h_n \Rightarrow F_n)$ and COMPS $\cap\, V' = \emptyset$, where V' is the set of all variables appearing in the propositional formulae $F_1, F_2, \ldots, F_n$.

Weak-fault models are sometimes referred to as models with *ignorance of abnormal behavior* [2], or *implicit fault systems*. Alternatively, a model may specify the faulty behavior for its components [14]. In the following definition, with the aim of simplifying the formalism throughout this article, we adopt a slightly restrictive representation of faults, allowing only a single fault mode per assumable variable. This can be easily generalized by introducing multi-valued logic or suitable encodings [12, 5].

Definition 3 (Strong-Fault Model)**.** A diagnostic system DS $=$ $\langle$SD, COMPS, IN, OUT, INT$\rangle$, OBS $=$ IN $\cup$ OUT, belongs to the class **SFM** iff SD is equivalent to $(h_1 \Rightarrow F_{1,1}) \wedge (\neg h_1 \Rightarrow F_{1,2}) \wedge \cdots \wedge (h_n \Rightarrow F_{n,1}) \wedge (\neg h_n \Rightarrow F_{n,2})$ such that $1 \leq i, j \leq n, k \in \{1, 2\}$, $\{h_i\} \subseteq$ COMPS, $F_{j,k} \in$ **Wff**, and none of h_i appears in $F_{j,k}$.

Membership testing for the **WFM** and **SFM** classes can be performed efficiently in many cases, for example, when a model is represented explicitly as in Def. 2 or Def. 3.

[2] In the MBD literature the assumable variables are also referred to as "component", "failure-mode", or "health" variables. Observable variables are also called "measurable", or "control" variables.

3.1 A Running Example

We use the Boolean circuit shown in Fig. 1 to illustrate many notions and algorithms in this article. The subtractor, shown there, consists of seven components: an inverter, two or-gates, two xor-gates, and two and-gates. The expression $h \Rightarrow (o \Leftrightarrow \neg i)$ models the normative (healthy) behavior of an inverter, where the variables i, o, and h represent input, output and health respectively. Similarly, an and-gate is modeled as $h \Rightarrow (o \Leftrightarrow i_1 \wedge i_2)$ and an or-gate by $h \Rightarrow (o \Leftrightarrow i_1 \vee i_2)$. Finally, an xor-gate is specified as $h \Rightarrow [o \Leftrightarrow \neg\,(i_1 \Leftrightarrow i_2)]$.

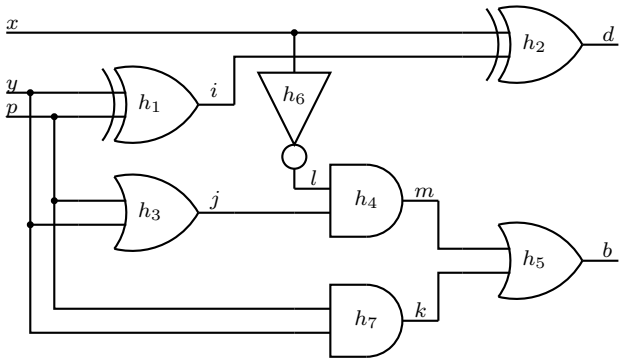


Figure 1. A subtractor circuit

The above propositional formulae are copied for each gate in Fig. 1 and their variables renamed in such a way as to properly connect the circuit and disambiguate the assumables, thus obtaining a propositional formula for the Boolean subtractor, given by:

$$\begin{aligned} \mathrm{SD}_w = &\{h_1 \Rightarrow [i \Leftrightarrow \neg\,(y \Leftrightarrow p)]\} \wedge \\ &\{h_2 \Rightarrow [d \Leftrightarrow \neg\,(x \Leftrightarrow i)]\} \wedge \\ &[h_3 \Rightarrow (j \Leftrightarrow y \vee p)] \wedge [h_4 \Rightarrow (m \Leftrightarrow l \wedge j)] \wedge \\ &[h_5 \Rightarrow (b \Leftrightarrow m \vee k)] \wedge [h_6 \Rightarrow (x \Leftrightarrow \neg l)] \wedge \\ &[h_7 \Rightarrow (k \Leftrightarrow y \wedge p)] \end{aligned} \tag{1}$$

A strong-fault model for the Boolean circuit shown in Fig. 1 is constructed by assigning fault-modes to the different gate types. We will assume that, when malfunctioning, the output of an xor-gate has the value of one of its inputs, an or-gate can be stuck-at-one, an and-gate can be stuck-at-zero, and an inverter behaves like a buffer. This gives us the following strong-fault model formula for the Boolean subtractor circuit:

$$\begin{aligned} \mathrm{SD}_s = \mathrm{SD}_w \wedge &[\neg h_1 \Rightarrow (i \Leftrightarrow y)] \wedge [\neg h_2 \Rightarrow (d \Leftrightarrow x)] \\ &(\neg h_3 \Rightarrow j) \wedge (\neg h_4 \Rightarrow \neg m) \wedge (\neg h_5 \Rightarrow b) \\ &[\neg h_6 \Rightarrow (x \Leftrightarrow l)] \wedge (\neg h_7 \Rightarrow \neg k) \end{aligned} \tag{2}$$

For both models (SD_s and SD_w), the set of assumable variables is COMPS $= \{h_1, h_2, \ldots, h_7\}$ and the set of observable variables is OBS $= \{x, y, p, d, b\}$, where IN $= \{x, y, p\}$ are the primary inputs and OUT $= \{d, b\}$ are the primary outputs.

Note that each component in SD_w or SD_s has inputs and an output. For example, the inverter which is associated with h_6 has x as its input and its output is l. The and-gate h_4 has two inputs: l and j and one output m.

3.2 Diagnosis and Minimal Diagnosis

The traditional query in MBD computes terms of assumable variables which are explanations for the system description and an observation.

Definition 4 (Health Assignment)**.** Given a system $\mathrm{DS} = \langle \mathrm{SD}, \mathrm{COMPS}, \mathrm{IN}, \mathrm{OUT}, \mathrm{INT} \rangle$, an assignment ω to all variables in COMPS is defined as a health assignment.

A health assignment ω is a conjunction of propositional literals. In some cases it is convenient to use the set of negative or positive literals in ω. These two sets are denoted as $Lit^{-}(\omega)$ and $Lit^{+}(\omega)$, respectively.

In our example, the "all nominal" assignment is $\omega_1 = h_1 \wedge h_2 \wedge \cdots \wedge h_7$. The health assignment $\omega_2 = h_1 \wedge h_2 \wedge h_3 \wedge \neg h_4 \wedge h_5 \wedge h_6 \wedge \neg h_7$ means that the two and-gates from Fig. 1 are malfunctioning.

What follows is a formal definition of consistency-based diagnosis.

Definition 5 (Diagnosis)**.** Given a diagnostic system $\mathrm{DS} = \langle \mathrm{SD}, \mathrm{COMPS}, \mathrm{IN}, \mathrm{OUT}, \mathrm{INT} \rangle$, $\mathrm{OBS} = \mathrm{IN} \cup \mathrm{OUT}$, an observation α over some variables in OBS, and a health assignment ω, ω is a diagnosis iff $\mathrm{SD} \wedge \alpha \wedge \omega \not\models \bot$.

There is a total of 96 possible diagnoses given SD_w and an observation $\alpha_1 = x \wedge y \wedge p \wedge b \wedge \neg d$. Example diagnoses are $\omega_3 = \neg h_1 \wedge h_2 \wedge \cdots \wedge h_7$ and $\omega_4 = h_1 \wedge \neg h_2 \wedge h_3 \wedge \cdots \wedge h_7$. Trivially, given a weak-fault model, the "all faulty" health assignment (in our example $\omega_a = \neg h_1 \wedge \cdots \wedge \neg h_7$) is a diagnosis for any instantiation of the observable variables in OBS (see Def. 2).

In the MBD literature, a range of types of "preferred" diagnosis has been proposed. This turns the MBD problem into an optimization problem. In the following definition we consider the common subset-ordering.

Definition 6 (Minimal Diagnosis)**.** A diagnosis $\omega^{\subseteq}$ is defined as minimal, if no diagnosis $\tilde{\omega}^{\subseteq}$ exists such that $Lit^{-}(\tilde{\omega}^{\subseteq}) \subset Lit^{-}(\omega^{\subseteq})$.

Consider the weak-fault model SD_w of the circuit shown in Fig. 1 and an observation $\alpha_2 = \neg x \wedge y \wedge p \wedge \neg b \wedge d$. In this example, two of the minimal diagnoses are $\omega_5^{\subseteq} = \neg h_1 \wedge h_2 \wedge h_3 \wedge h_4 \wedge \neg h_5 \wedge h_6 \wedge h_7$ and $\omega_6^{\subseteq} = \neg h_1 \wedge h_2 \wedge \cdots \wedge h_5 \wedge \neg h_6 \wedge \neg h_7$. The diagnosis $\omega_7 = \neg h_1 \wedge \neg h_2 \wedge h_3 \wedge h_4 \wedge \neg h_5 \wedge h_6 \wedge h_7$ is non-minimal as the negative literals in $\omega_5^{\subseteq}$ form a subset of the negative literals in ω_7.

Definition 7 (Subset-Minimal Ambiguity Group)**.** The subset-minimal ambiguity group of a system description SD and an observation α, denoted as $\Omega^{\subseteq}(\mathrm{SD} \wedge \alpha)$, is defined as the set of all minimal diagnoses of $\mathrm{SD} \wedge \alpha$.

Note that the set of all minimal diagnoses characterizes all diagnoses for a weak-fault model, but that does not hold in general for strong-fault models [2]. In the latter case, faulty components may "exonerate" each other, resulting in a health assignment containing a proper superset of the negative literals of another diagnosis not to be a diagnosis. In our example, given SD_s and $\alpha_3 = \neg x \wedge \neg y \wedge \neg p \wedge b \wedge \neg d$, it follows that $\omega_8^{\subseteq} = h_1 \wedge h_2 \wedge \neg h_3 \wedge h_4 \wedge \cdots \wedge h_7$ is a diagnosis, but $\omega_9 = h_1 \wedge h_2 \wedge \neg h_3 \wedge \neg h_4 \wedge \cdots \wedge h_7$ is not a diagnosis, despite the fact that the negative literals in ω_9 ($\{\neg h_3, \neg h_4\}$) form a superset of the negative literals in $\omega_8^{\subseteq}$ ($\{\neg h_3\}$).

Definition 8 (Number of Minimal Diagnoses)**.** Given a system description SD and an observation α, the number of minimal diagnoses, denoted as $|\Omega^{\subseteq}(\mathrm{SD} \wedge \alpha)|$, is defined as the size of the subset-minimal ambiguity group $\Omega^{\subseteq}(\mathrm{SD} \wedge \alpha)$.

Continuing our running example, $|\Omega^{\subseteq}(\mathrm{SD}_w \wedge \alpha_2)| = 8$ and $|\Omega^{\subseteq}(\mathrm{SD}_s \wedge \alpha_3)| = 2$. The number of non-minimal diagnoses of $\mathrm{SD}_w \wedge \alpha_2$ is 61.

Definition 9 (Cardinality of a Diagnosis)**.** The cardinality of a diagnosis ω, denoted as $|\omega|$, is defined as the number of negative literals in ω.

Diagnosis cardinality gives us another partial ordering: a diagnosis is defined as *minimal cardinality* iff it minimizes the number of negative literals.

Definition 10 (Minimal-Cardinality Diagnosis)**.** A diagnosis $\omega^{\leq}$ is defined as minimal-cardinality if no diagnosis $\tilde{\omega}^{\leq}$ exists such that $|\tilde{\omega}^{\leq}| < |\omega^{\leq}|$.

The cardinality of a minimal-cardinality diagnosis computed from a system description SD and an observation α is denoted as $MinCard(\mathrm{SD} \wedge \alpha)$. For our example model SD_w and an observation $\alpha_4 = x \wedge y \wedge p \wedge \neg b \wedge \neg d$, it follows that $MinCard(\mathrm{SD}_w \wedge \alpha_4) = 2$. Note that in this case all minimal diagnoses are also minimal-cardinality diagnoses.

A minimal cardinality diagnosis is a minimal diagnosis, but the opposite does not hold. There are minimal diagnoses that are not minimal-cardinality diagnoses. Consider the example SD_w and α_2 given earlier in this section, and the two resulting minimal diagnoses $\omega_5^{\subseteq}$ and $\omega_6^{\subseteq}$. From these two, only $\omega_5^{\subseteq}$ is a minimal-cardinality diagnosis.

Definition 11 (Minimal-Cardinality Ambiguity Group)**.** The minimal-cardinality ambiguity group of a system description SD and an observation α, denoted as $\Omega^{\leq}(\mathrm{SD} \wedge \alpha)$, is defined as the set of all minimal-cardinality diagnoses of $\mathrm{SD} \wedge \alpha$.

Counting the number of diagnoses in $\Omega^{\leq}(\mathrm{SD} \wedge \alpha)$ gives us the final definition for this section.

Definition 12 (Number of Minimal-Cardinality Diagnoses)**.** The number of minimal-cardinality diagnoses, denoted as $|\Omega^{\leq}(\mathrm{SD} \wedge \alpha)|$, is defined as the cardinality of $\Omega^{\leq}(\mathrm{SD} \wedge \alpha)$.

Computing the number of minimal-cardinality diagnoses for the running example results in $|\Omega^{\leq}(\mathrm{SD}_w \wedge \alpha_2)| = 2$, $|\Omega^{\leq}(\mathrm{SD}_s \wedge \alpha_3)| = 2$, and $|\Omega^{\leq}(\mathrm{SD}_w \wedge \alpha_4)| = 4$.

4 Observation Vector Optimization Problems

Consider the set of diagnoses in a subset-minimal ambiguity group $\Omega^{\subseteq}(\mathrm{SD} \wedge \alpha)$. We can construct a distribution of the subset-minimal diagnoses in $\Omega^{\subseteq}(\mathrm{SD} \wedge \alpha)$ by counting the number of diagnoses with cardinality $0, 1, 2, \ldots$ and computing how frequently each cardinality appears in $\Omega^{\subseteq}(\mathrm{SD} \wedge \alpha)$. The distribution of the diagnosis cardinalities in $\Omega^{\subseteq}(\mathrm{SD} \wedge \alpha)$ is denoted as $\widetilde{\Omega}^{\subseteq}(\mathrm{SD} \wedge \alpha)$. Note that $\widetilde{\Omega}^{\subseteq}(\mathrm{SD} \wedge \alpha)$ can be arbitrary, i.e., we can construct a system description SD and an observation α resulting in any $\widetilde{\Omega}^{\subseteq}(\mathrm{SD} \wedge \alpha)$. In this paper, SD is fixed, and the main focus of our work is how $\widetilde{\Omega}^{\subseteq}(\mathrm{SD} \wedge \alpha)$ changes for various instantiations of the observation α. In particular we are interested in computing observations α that optimize certain parameters defined on the distribution $\widetilde{\Omega}^{\subseteq}(\mathrm{SD} \wedge \alpha)$.

Figure 2 shows $\widetilde{\Omega}^{\subseteq}(\mathrm{SD} \wedge \alpha)$ for a weak-fault model of the 74182 combinatorial circuit (part of the 74XXX/ISCAS85 benchmark, see

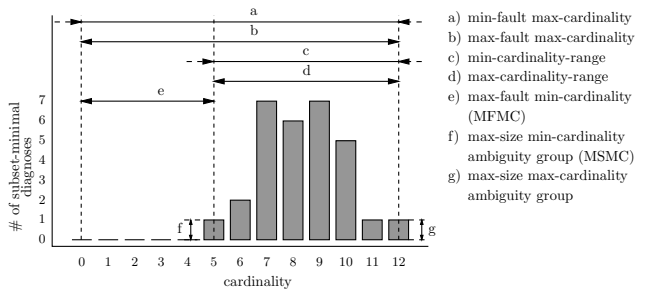


Figure 2. An example distribution of the cardinalities of the subset-minimal ambiguity group for a given observation α. Other observations lead to different distributions. All problems are defined as computing an observation vector (showing all possible observation vectors for this example would add another dimension to the figure) that optimizes certain properties of this distribution. These properties are indicated by arrows.

Sec. 6) and an arbitrary observation α. In addition to that, Fig. 2 illustrates a number of observation vector optimization problems.

From the seven observation vector optimization problems shown in Fig. 2, two are of practical significance to MBD: MFMC and MSMC. We next formally define those.

Problem 1 (MFMC Observation)**.** Given a system $\mathrm{DS} = \langle \mathrm{SD}, \mathrm{COMPS}, \mathrm{IN}, \mathrm{OUT}, \mathrm{INT} \rangle$, compute an observation α (defined as Max-Fault Min-Cardinality (MFMC) observation) such that ω is a minimal-cardinality diagnosis of $\mathrm{SD} \wedge \alpha$ and $|\omega|$ is maximized.

In addition to an MFMC observation, we also refer to an MFMC diagnosis of a model SD. This refers to any of the minimal-cardinality diagnoses $\omega^{\leq}$ of $\mathrm{SD} \wedge \alpha$ where α is an MFMC observation. The cardinality of this diagnosis is denoted as $MFMC(\mathrm{SD})$ and, next to the associated MFMC observations, this is a key model property we seek to compute.

Problem 2 (MSMC Observation)**.** Given a system $\mathrm{DS} = \langle \mathrm{SD}, \mathrm{COMPS}, \mathrm{IN}, \mathrm{OUT}, \mathrm{INT} \rangle$, compute an observation α (defined as Max-Size Min-Cardinality ambiguity group (MSMC) observation) such that $|\Omega^{\leq}(\mathrm{SD} \wedge \alpha)|$ is maximized.

We denote $|\Omega^{\leq}(\mathrm{SD} \wedge \alpha)|$ where α is an MSMC observation as $MSMC(\mathrm{SD})$.

Fig. 2 also illustrates some MBD problems that are less often encountered in practice. The min-fault max-cardinality problem, for example, is to compute the following observation. First, consider the subset-minimal ambiguity group of each different observation (there are $2^{|\mathrm{OBS}|}$ different observations). Second, take the observation that minimizes the number of faults in the maximum-cardinality diagnosis in each subset-minimal ambiguity group.

A related problem that is not illustrated in Fig. 2 is the max-size subset-minimal ambiguity group. The problem is to compute an observation α that maximizes the size of the subset-minimal ambiguity group.

5 Probing

Probing aims to minimize the expected number of diagnoses that result from the possible set of outputs that may occur from the measurement of a given internal (probe) variable.

5.1 Computing the Expected Number of MC Diagnoses

We will compute the expected number of diagnoses for a set of observable variables M ($M \subseteq \mathrm{OBS}$). The initial observation α and the set of MC diagnoses $D = \Omega^{\leq}(\mathrm{SD}, \alpha)$ modify the probability density function of subsequent outputs (observations), i.e., a subsequent observation α' changes its likelihood. The (non-normalized) a posteriori probability of an observation α', given a function $\Omega^{\leq}$ that computes the set of MC diagnoses and an initial observation α, is:

$$\Pr(\alpha'|\mathrm{SD}, \alpha) = \frac{|\Omega^{\cap}(\Omega^{\leq}(\mathrm{SD}, \alpha), \alpha')|}{|\Omega^{\leq}(\mathrm{SD}, \alpha)|} \tag{3}$$

The above formula computes the probability of a given a priori set of diagnoses restricting the possible outputs, i.e., we assume that the probability is the ratio of the number of remaining diagnoses to the number of initial diagnoses. In practice, there are many α for which $\Pr(\alpha'|\mathrm{SD}, \alpha) = 0$, because a certain fault heavily restricts the possible outputs of a system (i.e., the set of the remaining diagnoses in the numerator is empty).

The expected number of remaining MC diagnoses for a variable set M, given an initial observation α, is then the weighted average of the intersection sizes of all possible instantiations over the variables in M (the weight is the probability of an output):

$$E^{\leq}(\mathrm{SD}, M|\alpha) = \frac{\sum_{\alpha' \in M^*} |\Omega^{\cap}(D, \alpha')| \cdot \Pr(\alpha'|\mathrm{SD}, \alpha)}{\sum_{\alpha' \in M^*} \Pr(\alpha'|\mathrm{SD}, \alpha)} \tag{4}$$

where $D = \Omega^{\leq}(\mathrm{SD}, \alpha)$ and M^* is the set of all possible assignments to the variables in M. Replacing (3) in (4) and simplifying gives us the following definition:

Definition 13 (Expected Minimal-Cardinality Diagnoses Intersection Size)**.** Given a system ATS and an initial observation α, the expected remaining number of MC diagnoses $E^{\leq}(\mathrm{SD}, \mathrm{OBS}|\alpha)$ is defined as:

$$E^{\leq}(\mathrm{SD}, \mathrm{OBS}|\alpha) = \frac{\sum_{\alpha' \in \mathrm{OBS}^*} |\Omega^{\cap}(\Omega^{\leq}(\mathrm{SD}, \alpha), \alpha')|^2}{\sum_{\alpha' \in \mathrm{OBS}^*} |\Omega^{\cap}(\Omega^{\leq}(\mathrm{SD}, \alpha), \alpha')|} \tag{5}$$

where OBS^* is the set of all possible assignments to all variables in OBS.

The expected number of remaining MC diagnoses for one variable simplifies the expression in (5) to:

$$E^{\leq}(\mathrm{SD}, v|\alpha) = \frac{p(\mathrm{SD}, v, \alpha)^2 + q(\mathrm{SD}, v, \alpha)^2}{p(\mathrm{SD}, v, \alpha) + q(\mathrm{SD}, v, \alpha)} \tag{6}$$

where

$$p(\mathrm{SD}, v, \alpha) = |\Omega^{\cap}(\Omega^{\leq}(\mathrm{SD}, \alpha), v)| \tag{7}$$

and

$$q(\mathrm{SD}, v, \alpha) = |\Omega^{\cap}(\Omega^{\leq}(\mathrm{SD}, \alpha), \neg v)| \tag{8}$$

Algorithm 1: Probing framework

Input: DS, a diagnostic system,
$\text{DS} = \langle \text{SD}, \text{IN}, \text{OUT}, \text{COMPS}, \text{INT} \rangle$
Result: p, number of probes, $R \in \mathbb{Z}$

Local variables: α, observation term
$\omega^{\leq}$, cardinality-minimal diagnosis
z, probe variable
l, literal

1 $\langle \alpha, \omega^{\leq} \rangle \leftarrow$ INJECTFAULT(DS)
2 $p \leftarrow 0$
3 **while** $|\Omega^{\leq}(\text{SD}, \alpha)| \neq 1$ **do**
4 $\quad z \leftarrow$ COMPUTEPROBE(SD, α, INT)
5 $\quad l \leftarrow$ EVALUATEPROBE(SD, $\alpha, \omega^{\leq}, z$)
6 $\quad \alpha \leftarrow \alpha \wedge l$
7 $\quad \text{INT} \leftarrow \text{INT} \setminus z$
8 $\quad p \leftarrow p + 1$
9 **return** p

5.2 Probing Algorithm

Algorithm 1 shows a generalized procedure for the evaluation of the performance of probing algorithms. It can be generalized to evaluate the performance of any information gathering procedures (such as active testing [7]), to include probing costs, etc.

Algorithm 1 starts by generating an observation α that leads to a cardinality-minimal diagnosis $\omega^{\leq}$. This is done by a call to the INJECTFAULT subroutine in line 1. Algorithm 1 needs a diagnostic engine that can count the number of cardinality-minimal diagnoses (line 3). The probing algorithm is called in line 4. The probing algorithm returns a variable (probe) that will be "measured". The "measured" values of probe z is computed by the EVALUATEPROBE auxiliary subroutine in line 5. Methods such as Binary Constraint Propagation (BCP) [10] or SAT solvers are suitable for calculating the value of z given the observation and the injected cardinality-minimal fault. Algorithm 1 evaluates the performance of probing algorithms in terms of the number of probes p.

The following assumptions are made when designing Alg. 1:

Monotone $|\Omega^{\leq}(\text{SD}, \alpha)|$: We restrict ourselves to such system descriptions SD such that if α and β are two observations such that $\alpha \supseteq \beta$ then it holds that $|\Omega^{\leq}(\text{SD}, \alpha)| \geq |\Omega^{\leq}(\text{SD}, \alpha)|$. We can proof that this holds for "well-formed" system descriptions and weak-fault models. The idea is to construct a system of Boolean equations B in the following manner. First, the propositional **Wff** in SD is converted to a Boolean equation in a straightforward manner and the latter is added to B. Second, for each literal $l_i \in \alpha$, an equation of the form $l_i = 1$ or $l_i = 0$ (depending on the polarity of l_i) is appended to B. A system of Boolean equations B' is constructed from SD and β in an analogous way. The solutions of B and B' are the implicants of $\text{SD} \wedge \alpha$ and $\text{SD} \wedge \beta$, respectively. Observe, that, due to the fact that $\alpha \supseteq \beta$, the equations in B' are a superset of these in B and both are over the same set of variables. But $S(B') \leq S(B)$, where $S(X)$ denotes the number of solutions in a system X. The above holds also when the solutions of B and B' are ordered according to their cardinality. Hence, if a diagnosis with a cardinality smaller than the smallest cardinality diagnosis in B' exists, it is in B.

Non-ambiguous fault: Given a diagnostic system $\text{DS} = \langle \text{SD}, \text{IN}, \text{OUT}, \text{COMPS}, \text{INT} \rangle$ we assume that there exists an observation α and an instantiation over a set of variables $P \subseteq \text{INT}$ such that $|\Omega^{\leq}(\text{SD}, \alpha)| = 1$. This is easily achievable if $\text{SD} \in \mathbf{WFM}$ and if $\text{INT} = V \setminus \{\text{IN} \cup \text{OUT} \cup \text{COMPS}\}$.

No "don't cares" and well-formed SD: We require all SD to be models of well-formed digital circuits. A well-formed digital circuits is constructed from standard AND, OR, NAND, or NOR gates of two or more inputs, from XOR gates, buffers, and inverters. There are no "hanging" wires, each output is connected to the input of another gate or two a primary output. A well-formed circuit does not use any feedback.

Algorithm 2 shows a simple greedy approach to compute the optimal probe variable based on the expected cardinality-minimal intersection size.

The computational performance of Alg. 2 is dominated by the complexity of the diagnostic engine that counts the remaining number of cardinality minimal diagnoses in lines 2 and 3. Assuming that this number decreases monotonically improves the complexity significantly.

Algorithm 2: Probing algorithm

Input: SD, a system description
Input: α, an observation
Input: INT, a set of probe variables
Result: an optimal probe variable $z \in \text{INT}$

Local variables: p, q, number of diagnoses
$E, E^{\star}$, reals, expected number of diagnoses
v, candidate probe variable

1 **foreach** $v \in \text{INT}$ **do**
2 $\quad p \leftarrow |\Omega^{\cap}(\Omega^{\leq}(\text{SD}, \alpha), v)|$
3 $\quad q \leftarrow |\Omega^{\cap}(\Omega^{\leq}(\text{SD}, \alpha), \neg v)|$
4 $\quad E = \frac{p^2+q^2}{p+q}$
5 $\quad$ **if** $E^{\star} < E$ **then**
6 $\quad\quad E^{\star} \leftarrow E$
7 $\quad\quad z \leftarrow v$
8 **return** z

We can see that the number of probes k required for the uniquely (non-ambiguous) isolation of a fault can be an arbitrary value $0 \leq k \leq |\text{INT}|$. There are circuits such as n chained buffers (or inverters) for which Alg. 2 can isolate a single-fault in $k = \log n$ calls. In the worst-case, Alg. 2 needs to probe each probe variable.

We can see that the performance of Alg. 2 is determined by SD and the injected fault $\omega^{\leq}$ injected by Alg. 1. As SD is given, the only variable that Alg. 1 can modify is the initial set of cardinality-minimal diagnoses. It is straightforward to show then that a worst-case scenario for Alg. 2 is when INJECTFAULT(DS) returns an MSMC fault.

6 Experimental Results

This section discusses some results from an implementation of the algorithms described in the previous sections.

6.1 Experimental Setup, Simplification Results and Bounds

We have experimented on the medium-sized circuits from the 74XXX family [11]. Table 1 provides a summary of the 74XXX

circuits. The number of inputs, outputs and components are given in the third, fourth, and fifth column of Table 1, respectively.

Table 1. 74XXX circuits

Name	Description	|IN|	|OUT|	|COMPS|
74182	4-bit CLA	9	5	19
74L85	4-bit comparator	11	3	33
74283	4-bit adder	9	5	36
74181	4-bit ALU	14	8	65
c6288	32-bit multiplier	32	32	2 416

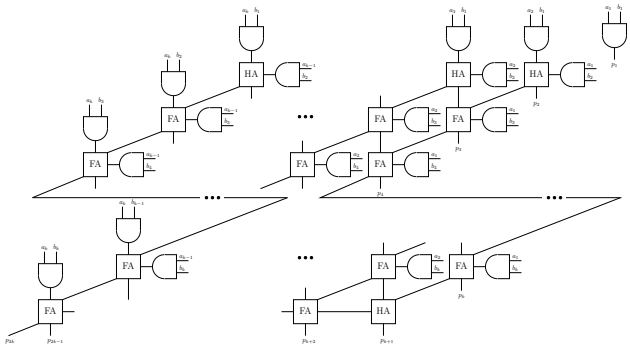


Figure 3. n-bit parallel multiplier ($n = 2k$). For c6288, $n = 32$.

In addition to the 74XXX circuits we have also considered a variation of the c6288 multiplier, part of the ISCAS85 benchmark. Despite the large number of components, c6288 has very regular structure: it is composed entirely of Boolean adders and and-gates as shown in Fig. 3 (the high-level structure of c6288 and Fig. 3 are due to the reverse-engineering efforts of [11]). Inspecting the reverse-engineered c6288 allowed us to construct similar smaller multipliers that have between 1 and 32 outputs. The smallest of them is amenable to an exhaustive approach. The regular structure of c6288 allows us to analytically hypothesize about the MFMC/MSMC properties of c6288 and the whole family of multipliers that have the same high-level structure. For example, one can show experimentally that for an n-bit multiplier having the structure of Fig. 3, it always holds that $MFMC^{\leq} = n$.

6.2 Solving the 74XXX Models Exhaustively

We first tried to exhaustively enumerate the space of all input/output assignments. For 74182, 74L85, and 74283 the size of this space is 16 384 (14 observable variables), while for 74181, it is 4 194 304 (22 observable variables). We used two state-of-the-art complete diagnostic solvers: HA* [9] and NGDE [1].

By using HA* in combination with cones [13] we computed all minimal-cardinality ambiguity groups for the 74XXX models. 74182 was the only circuit for which we could compute all subset-minimal ambiguity groups (these are different from the cardinality-minimal ambiguity groups). We recomputed all diagnoses with NGDE, which is a completely independent implementation by one of the authors of this paper, and the HA* and NGDE results match. Furthermore, NGDE did not use cones for 74XXX, while HA* did, thus independently verifying the correctness of the MFMC/MSMC values, and the correct implementation of the algorithms for computing minimal diagnoses.

The exhaustive search results of the small circuits are shown in Table 2. We can see that for 74182, 74L85, and 74283, $MFMC(\mathrm{SD}) = |\mathrm{OUT}|$, and for 74181 the MFMC value is smaller than the number of outputs $|\mathrm{OUT}|$. The MSMC value for the 74XXX models grows quickly with increasing model size ($|\mathrm{COMPS}|$).

Table 2. Properties of 74XXX subset-minimal diagnoses

Optimization Problem	74182	74L85	74283	74181
min-fault max-cardinality	1	1	–	–
max-fault max-cardinality	14	10	–	–
min-cardinality-range	0	0	–	–
max-cardinality-range	9	8	–	–
MFMC	5	3	5	7
MSMC	400	468	9 132	42 112

Figure 4 is a two-dimensional histogram of the minimal-cardinality ambiguity groups of the 74XXX models. Figure 4 plots on the z-axis the number of observation vectors leading to a minimal-cardinality ambiguity group of size $|\Omega^{\leq}(\mathrm{SD} \wedge \alpha)|$ (y-axis) and minimal cardinality $|\omega|$ (x-axis). We can see that there are no observations leading to low minimal-cardinality and high ambiguity group size and vice versa. We can also see that, in general, an increase in MFMC leads to an increase in MSMC. Furthermore, MFMC/MSMC observation vectors are relatively rare and the MSMC observation vectors are not always MFMC observation vectors (consider, for example, the histogram of 74181 in Figure 4) and vice-versa.

There are 36 MSMC observation vectors for 74182, for example. Of those, 18 observations lead to a minimal-cardinality diagnosis of cardinality 4 and 18 observations lead to a minimal-cardinality diagnosis of cardinality 5. All MSMC observations lead to nearly MFMC diagnoses. As it is visible from Fig. 4, MFMC observations lead to multiple values for the sizes of the minimal-cardinality ambiguity groups. In 74182, for example, there are 7 MFMC observations that lead to a unique minimal-cardinality diagnosis.

Given a system DS, we denote as $g(\mathrm{DS})$ the probability density function of the minimal-cardinalities of the diagnoses of all observations in DS. Figure 5 shows a histogram of the true minimal-diagnosis cardinalities for the four 74XXX circuits for which we have exhaustively determined $g(\mathrm{DS})$, fitted by a normal distribution.

Figure 5 shows the number of observations per minimal-cardinality. We noticed that a normal distribution fits the empirical data well in Fig. 5 (the standard error for 74182, 74L85, 74283, and 74283 is 154, 244, 100, and 18 955, respectively). This is explained as follows. Given an observation α leading to a k-fault minimal diagnosis, we associate a nominal-diagnosis observation α_n, which may differ from α only in the OUT sub-vector. The number of OUT-values in which α and α_n differ is called the *distance* of α, $D(\mathrm{SD}, \alpha)$. If $n = |\mathrm{OUT}|$ is the number of output variables in SD, then starting from any nominal observation α_n, there are ${}_nC_k$ ways to select a distance-k vector α, each of which corresponds to a diagnosis. In the case where each such diagnosis is a minimum cardinality diagnosis, $g(\mathrm{SD})$ is binomially-distributed.

To understand better why the distribution of the minimal-cardi-

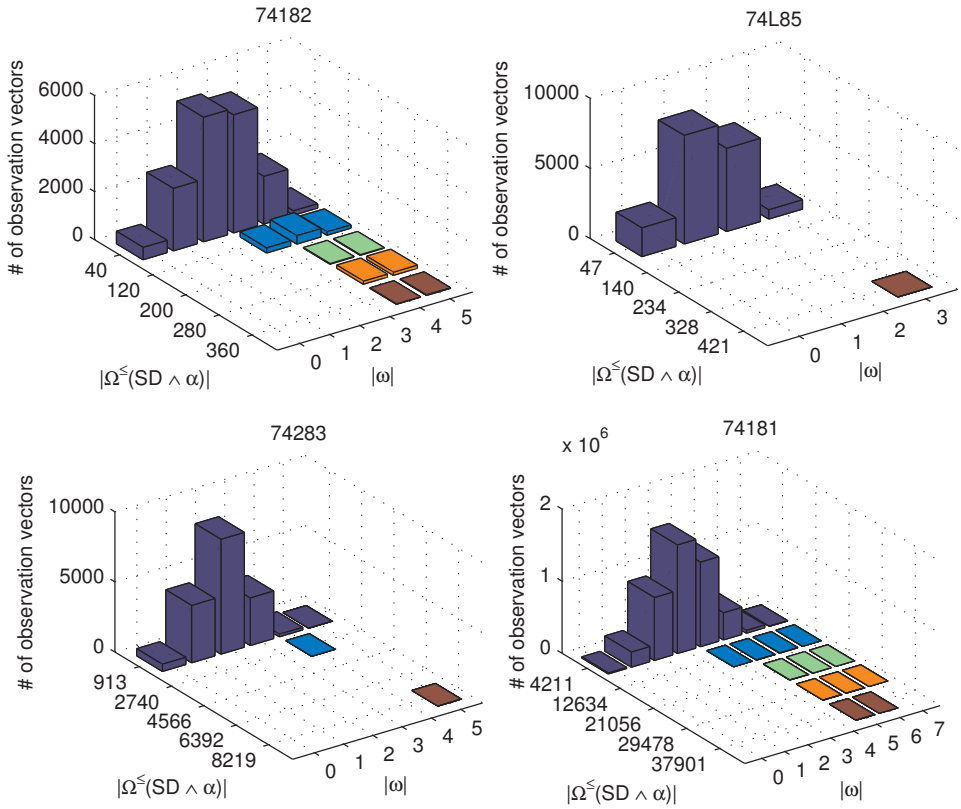


Figure 4. Number of observation vectors vs. cardinality and number of minimal-cardinality diagnoses bivariate histograms for 74XXX

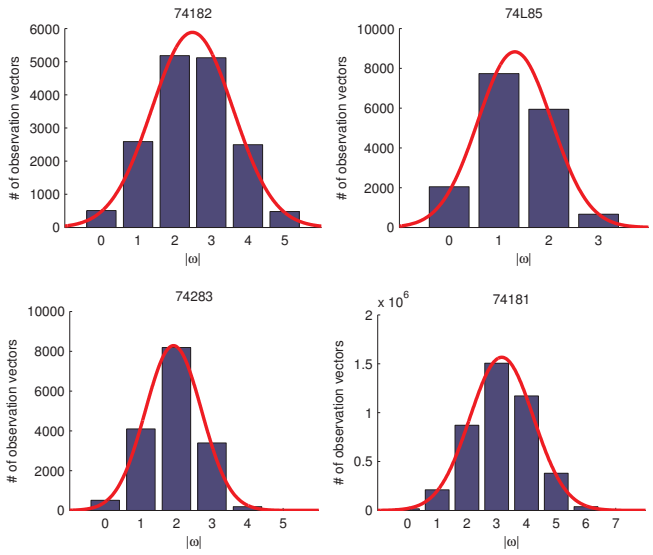
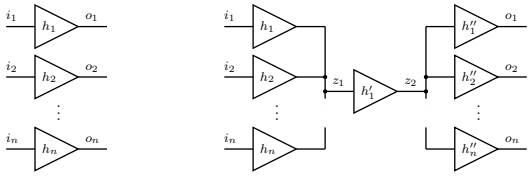


Figure 5. 74XXX minimal-cardinalities distribution



(a) Circuit containing n separate buffers

(b) Converging and diverging circuit

Figure 6. A model with a binomial minimal-cardinality distribution (left) and a model with one nominal minimal-cardinality diagnosis and one single-fault minimal-cardinality diagnosis (right)

nality diagnoses of many circuits can be approximated with a binomial distribution, consider **WFM** of the two synthetic circuits shown in Fig. 6. Both circuits consists of buffers only, where each buffer is modeled as $h \Rightarrow (o \Leftrightarrow i)$. Both circuits have the same input and output variables (IN $= \{i_1, i_2, \ldots, i_n\}$, and OUT $= \{o_1, o_2, \ldots, o_n\}$). The distributions of the minimal-cardinality diagnoses, however, are very different. The model of the circuit shown in Fig. 6(a) has one nominal behavior (health assignment in which all health literals are positive) and n single faults (health assignments in which there is exactly one negative literal). The same circuit has $\frac{n(n-1)}{2}$ double faults, $\frac{n(n-1)(n-2)}{6}$ triple faults, etc. Any observation for the model of the circuit shown in Fig. 6(b), however, leads either to nominal behavior or to the single fault $\neg h'_1$. As we have seen in Fig. 5, the distributions of the 74XXX circuits resemble more the distribution associated with Fig. 6(a). The density mass of all distributions shown in Fig. 5 are skewed to the left and the amount with which a distribution is skewed to the left depends on the masking

phenomenon demonstrated in Fig. 6(b).

Although the above model is an approximation, it provides useful bounds on MFMC errors. Let m be the number of bits in the output assignment that differ from the nominal output value. For the 74XXX and ISCAS85 benchmarks, the fraction of "m-flips" resulting in minimal-cardinality diagnoses of cardinality smaller than m is relatively small and does not vary significantly for different m.

Figure 7 shows a histogram of the minimal-cardinality ambiguity group sizes for all 74XXX circuits. We can see that when increasing the minimal-cardinality ambiguity group size, the number of observation vectors decreases rapidly. This depends on the model topology, and is less prevalent in 74182. The MSMC value of 74181, for example, is 42 112, and as is visible from Fig. 7, there are relatively few observations leading to such large ambiguity groups.

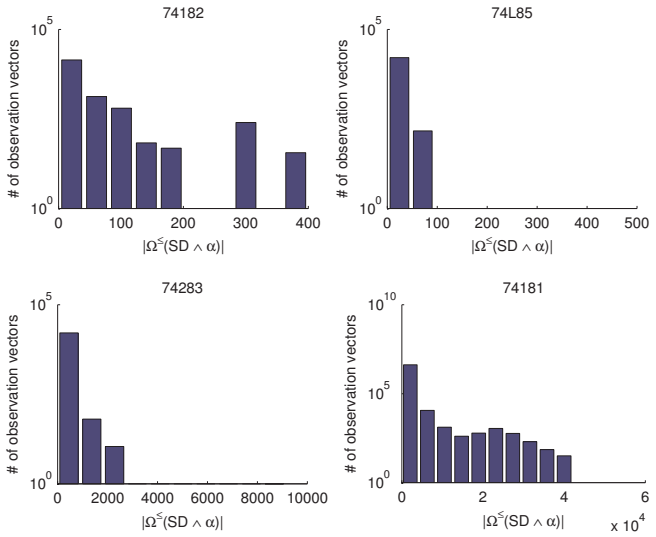


Figure 7. 74XXX minimal-cardinality ambiguity group sizes distribution

Table 3 shows the MFMC and MSMC values of several small multipliers (see Fig. 3). We have created two types of fault-models: in Type I models we have assigned only one health variable to each half-adder or full-adder (i.e., all the gates in an adder fail simultaneously), and in Type II fault-models we have associated an assumable with each logic gate (as everywhere else in this paper). The 2-bit multiplier consists only of a single and-gate, hence all MFMC and MSMC values are trivially 1. For Type I fault-models we can see that the MFMC value of an n-bit multiplier is $n/2+1$ ($n = 2k, k \in \mathbb{N}^+$). For Type II multipliers the MFMC value of an n-bit multiplier is n.

Table 3. MFMC and MSMC values of small multipliers

	Type I		Type II	
bits ($\|\mathrm{OUT}\|$)	MFMC	MSMC	MFMC	MSMC
2	1	1	1	1
4	3	6	4	9
6	4	58	6	3 969
8	5	845	–	–

7 Conclusions

This paper has defined a class of MSMC observation vectors which are the worst-case for the fault ambiguity (or indistinguishability). The MSMC of real-world systems is an important property quantifying the diagnosability of a model, as it shows the maximum number of cardinality-minimal diagnoses that can be returned by observing a set of variables.

We have shown a probing algorithm for which an MSMC observation vector results in the largest number of steps for reducing the initial set of cardinality-minimal diagnoses to a single candidate.

Computing MSMC-related properties of models of real-world artifacts is important for (1) assessing the performance of MBD and information gathering algorithms and (2) better understanding the diagnosability properties of the design.

Computing MSMC is a difficult counting problem and its complexity is hypothesized to be at least the complexity of counting the number of cardinality-minimal diagnoses entailed by a system description and an observation. As a result algorithms that can compute MSMC must utilize properties of the model, such as structure and hierarchy, in order to provide results for systems of practical size.

We have computed MSMC values for the 74XXX models. As a future work we plan to design more efficient MSMC algorithms and to apply them to a class of larger benchmarks.

REFERENCES

[1] Johan de Kleer, 'Minimum cardinality candidate generation', in *Proc. DX'09*, pp. 397–402, (2009).

[2] Johan de Kleer, Alan Mackworth, and Raymond Reiter, 'Characterizing diagnoses and systems', *Artificial Intelligence*, **56**(2-3), 197–222, (1992).

[3] Johan de Kleer and Brian Williams, 'Diagnosing multiple faults', *Artificial Intelligence*, **32**(1), 97–130, (1987).

[4] Alexander Feldman, Tolga Kurtoglu, Sriram Narasimhan, Scott Poll, David Garcia, Johan de Kleer, Lukas Kuhn, and Arjan van Gemund, 'Empirical evaluation of diagnostic algorithm performance using a generic framework', *International Journal of Prognostics and Health Management*, 1–28, (2010).

[5] Alexander Feldman, Jurryt Pietersma, and Arjan van Gemund, 'A multi-valued SAT-based algorithm for faster model-based diagnosis', in *Proc. DX'06*, (2006).

[6] Alexander Feldman, Gregory Provan, and Arjan van Gemund, 'Computing minimal diagnoses by greedy stochastic search', in *Proc. AAAI'08*, pp. 919–924, (2008).

[7] Alexander Feldman, Gregory Provan, and Arjan van Gemund, 'FRACTAL: Efficient fault isolation using active testing', in *Proc. IJCAI'09*, (2009).

[8] Alexander Feldman, Gregory Provan, and Arjan van Gemund, 'Stochastic algorithms for sequential model-based diagnosis', *Journal of Artificial Intelligence Research*, **39**, 301–334, (2010).

[9] Alexander Feldman and Arjan van Gemund, 'A two-step hierarchical algorithm for model-based diagnosis', in *Proc. AAAI'06*, (2006).

[10] Kenneth Forbus and Johan de Kleer, *Building Problem Solvers*, MIT Press, 1993.

[11] Mark Hansen, Hakan Yalcin, and John Hayes, 'Unveiling the ISCAS-85 benchmarks: A case study in reverse engineering', *IEEE Design & Test*, **16**(3), 72–80, (1999).

[12] Holger Hoos, 'SAT-encodings, search space structure, and local search performance', in *Proc. IJCAI'99*, pp. 296–303, (1999).

[13] Sajjad Siddiqi and Jinbo Huang, 'Hierarchical diagnosis of multiple faults', in *Proc. IJCAI'07*, pp. 581–586, (2007).

[14] Peter Struss and Oskar Dressler, 'Physical negation: Integrating fault models into the general diagnosis engine', in *Proc. IJCAI'89*, pp. 1318–1323, (1989).