Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence

Large Scale Diagnosis Using Associations between System Outputs and Components

Ting Guo[1], Zhanshan Li[1], Ruizhi Guo[1], and Xingquan Zhu[2]

College of Computer Science and Technology, Jilin University, Changchun, 130012, China[1]
Centre for Quantum Computation & Intelligent Systems, FEIT, University of Technology, Sydney, NSW 2007, Australia[2]
nalanwutuo@gmail.com, zslizsli@163.com, guo_ruizhi@hotmail.com, xqzhu@it.uts.edu.au

Abstract

Model-based diagnosis (MBD) uses an abstraction of system to diagnose possible faulty functions of an underlying system. To improve the solution efficiency for multi-fault diagnosis problems, especially for large scale systems, this paper proposes a method to induce reasonable diagnosis solutions, under coarse diagnosis, by using the relationships between system outputs and components. Compared to existing diagnosis methods, the proposed framework only needs to consider associations between outputs and components by using an assumption-based truth maintenance system (ATMS) [de Kleer 1986] to obtain correlation components for every output node. As a result, our method significantly reduces the number of variables required for model diagnosis, which makes it suitable for large scale circuit systems.

Introduction

Model-based diagnosis, a sub-field in Artificial Intelligence, is a diagnostic method which uses internal system constructions and behaviour knowledge to diagnose possible faulty functions of a system [de Kleer et al. 1987]. To achieve the diagnostic goal, traditional methods require a large number of redundant behaviours of the system, which makes them inefficient in practice, especially for large scale systems. In addition, for many real-world systems, we may not know the exact faulty behaviours of the components, but do know exactly which components are subject to faulty functions. In this paper, we propose the concept of "coarse diagnosis", where a system is represented as (SD,COMPS,OBS) [de Kleer et al. 1992], with components (denoted by COMPS) only containing two types of behaviours, namely faulty behaviour or normal behaviour. In our definition, each behavior is denoted by an AB-literal which is AB(c) or $\neg$AB(c) for some $c \in$ COMPS. A component is an element $c \in$ COMPS, and an output indicates a measurement point. It is worth noting that an output is not a real terminal of the system. The value of an output is a result induced by values transferring though the components from the inputs to the output. An output is said to be abnormal if there is a difference between its predicted value and its observation value, and normal otherwise.

Based on the defined "coarse diagnosis", we propose a diagnostic framework by using association information between outputs and components to extract a Bi-layered model, which only maintains associations between components and outputs. Using the proposed Bi-layer model, the diagnosing process, as shown in Fig.1, is as follows: (1) when a fault appears in the system, we first extract predicted outputs and their support environments into a Bi-layer model by using ATMS, then (2) we collect all abnormal outputs based on the discrepancy between prediction and observation, and (3) process the diagnosis by the associations between outputs and components.

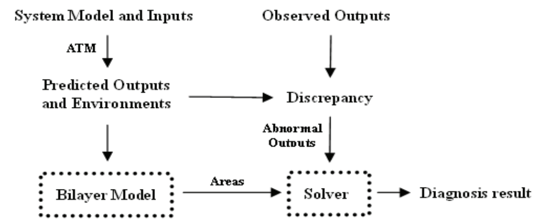

Fig. 1: Diagnosing process of the whole system (the dashed rectangle boxes represent work reported in this paper)

Diagnosis Based-on the Associations between Outputs and Components

ATMS is a truth maintenance system based on assumption. We propagate values, together with assumptions, by applying the thought of ATM which adds assumption to the support environment, and treats the observation as justification to find inconsistent environments. We use expression in GDE [de Kleer et al. 1987] and C(prev, env) to express the prediction and its support environment of node x. For instance, in Fig. 2, by adding input values through ATM, we have C(F=12, $\{A_1, M_1, M_2\}$) and C(G=12, $\{A_2, M_2, M_3\}$). For output G, the meaning of C(G=12, $\{A_2, M_2, M_3\}$) is that when there is no fault in $\{A_2, M_2, M_3\}$, the predicted value of G is 12. If our observations show that F=12 and G=10, it indicates that observations are inconsistent with the prediction, so G is an abnormal output.

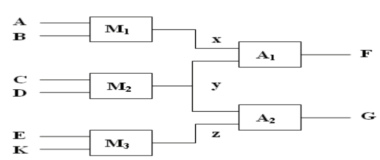

Fig. 2: An example of Davis Circuit (M_1, M_2, and M_3 are multipliers; A_1 and A_2 are adders; A=2, B=3, C=2, D=3, E=2, and K=3 denote measurements; and F and G are outputs)

Definition 1 (Output Constraint Component Set, Component Constraint Output Set) For a system (SD, COMPS, OBS), we call the set of components affecting output n the constraint component set, denoted by OCCS(n), of n. Similarly, we call the output set affected by component c ($c \in$ COMPS) the component constraint output set CCOS(c).

In Fig. 2, we have OCCS(F)={A_1, M_1, M_2} and CCOS(M_2) = {F, G}. The support environment of a predicted output is actually the constraint component set of the output. In Fig. 2, A_2, M_2, and M_3 are the only components having influences on the output G.

Definition 2 (Area, Association Outputs) An area is referred to as the set of components with the same constraint output set. The set of outputs affected by the components in area A is called association outputs, which is denoted by Outputs(A). In Fig. 2, CCOS(A_1) = CCOS(M_1) = {F}, so $Area_1$={A_1, M_1} and Outputs(Area1) = {F}.

Any component with a fault in the same area would cause the same abnormal outputs in the system.

Theorem 1 Given a system (SD, COMPS, OBS), output constraint component sets of any output y are the union of all areas which have influences on y, formally OCCS(y)={$\cup A \mid y \in$ Outputs(A)}.

According to Theorem 1, we can collect all components affecting output y, by observing association outputs of every area in the system, to find suspicious components. To achieve the goal, we propose a Bi-layer model, which only contains two parts: areas (all components of the original system are replaced by areas), and outputs. The connecting line represents the associations between areas and outputs. One area only affects the outputs, so we can easily find suspicious components with faulty functions through an inverse tracing of the connecting lines. Fig. 3 demonstrates the Bi-layer model of Fig.2 (using Definition 2). According to the associations between outputs and components, it is obvious that A_1 and M_1 only affect output F, A_2 and M_3 only affects output G, whereas M_2 affects both F and G.

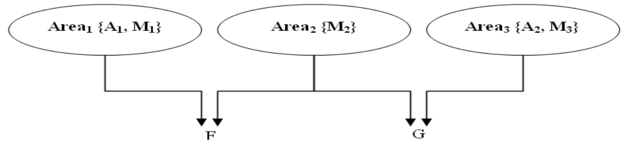

Fig. 3: Bi-layer Model of the Davis Circuit shown in Fig. 2

Predicted outputs and support environments acquired by using ATMS have provided all necessary information for generating Bi-layer models. At the same time, an abnormal output can be discovered by comparing the discrepancy between predictions and observations. Accordingly, we propose the following Diagnosis algorithm, where D is the diagnosis result (the initial value is nil), Areas is the set of areas, Ab_outputs is the set of abnormal outputs.

Algorithm Diagnosis (D, Areas, Ab_outputs)

```
For each area in Areas do
  If Outputs(area) = Ab_outputs
    return D = D × area
  If Outputs(area) ⊂ Ab_outputs
    D1 = D × area
    A1 = Ab_outputs - Outputs(area)
    Diagnosis(D1, Areas, A1)
  Else If Outputs(area) - Ab_outputs≠nil∧
                    (Outputs(area) ∩ Ab_outputs)≠nil
    D2 = D × area
    A2 = (Ab_outputs - Outputs(area)) ∩
                    (Outputs(area) - Ab_outputs)
    Diagnosis(D2, Areas, A2)
```

We explain the above algorithm by using the Davis Circuit in Fig.2. By using ATMS with F=12 and G=10, we know that G is an abnormal output, F is a normal output, and OCCS(G)= $Area_1 \cup Area_2$. Since $Area_1$ only affects G, any components with faulty functions in $Area_1$ are consistent with the observation, so we have diagnosis results{A_2, M_3}. Meanwhile, because $Area_2$ affects outputs G and F at the same time, if there are any components with faulty functions in $Area_2$, we should observe exceptions in both F and G. When we observe that F is normal, we should take account of the multi-component cooperative fault, and then find that F is affected by $Area_1$= {A_1, M_1} and $Area_2$. $Area_1$ only affects output F, and the observation is consistent with a cooperative fault in $Area_1$ and $Area_2$. As a result, the diagnosis results are the Cartesian product of $Area_1$ and $Area_2$: $Area_1 \times Area_2$ = {A_1, M_1}×{M_2}.

Using the Cartesian product to represent a diagnosis result not only saves diagnosis spaces, but is also beneficial for the selection of re-observing points. Finally, we obtain diagnosis results about the exception in G and normal in F under the current observation: {A_2, M_3}, {A_1, M_1}×{M_2}, which is identical to the results ({A_2},{M_3},{A_1, M_2},{M_1, M_2}) from GDE [de Kleer et al. 1987].

In Table 1, we compare efficiency between the proposed Bi-layer model and a model using Prime Implicate reduction rule (MFMCD) [de Kleer 2008], by using ISCAS-85 benchmark circuits [Brglez & Fujiwara 1985]. In Table 1, Original and Outputs denote the component and output numbers of the circuit system, respectively. The last two columns list the reduction, in component count, of two abstract methods on the circuit in each row. The results show significant component reduction of the Bi-layer model.

Table.1. Reduction in component count using the Bi-layer Model

Circuit	Original	Outputs	MFMCD Reduced	Bi-layer Model Reduced
C432	160	7	59	145
C499	202	32	58	159
C880	383	26	77	344
C1355	546	32	58	503
C1908	880	25	160	847
C2670	1193	140	169	1099
C3540	1669	22	353	1643
C6288	2416	32	1456	2354

Acknowledgement

Xingquan Zhu is sponsored by Australian Research Council Future Fellowship under Grant No. FT100100971.

References

[Brglez & Fujiwara 1985] Brglez, F., and Fujiwara, H. 1985. A neutral netlist of 10 combinational benchmark circuits and a target translator in fortran. In proc. IEEE Int. Symposium on Circuits and Systems: 695-698.

[de Kleer 1986] de Kleer, J. 1986. An assumption-based TMS. Artificial Intelligence 28:127-162.

[de Kleer et al. 1987] de Kleer, J., Williams, B.C. 1987. Diagnosing multiple faults. Artificial Intelligence 32 (1): 97-130.

[de Kleer et al. 1992] de Kleer, J., Mackworth, A., Reiter, R., Characterizing Diagnoses and Systems. Artificial Intelligence, 56(2-3):197–222.

[de Kleer 2008] de Kleer, J. 2008. An Improved Approach for Generating Max-Fault Min-Cardinality Diagnoses. In proc. DX'08: 247-252.

[Siddiq et al. 2007] Siddiq S., Huang J. 2007. Hierarchical Diagnosis of Multiple Faults. In proc. IJCAI'07: 581-586.