Complementing an Interval Based Diagnosis Method With Sign Reasoning On an Automotive Subsystem

Xavier OLIVE[1,2], Louise TRAVE-MASSUYES[1], and Hervé POULARD[2]

[1] LAAS-CNRS, 7 Avenue du Colonel ROCHE F-31077 TOULOUSE Cedex 4
{xolive,louise}@laas.fr,
[2] ACTIA, 25 Chemin de Pouvourville - B.P. 4215 F-31432 TOULOUSE Cedex 4
{xolive,poulard}@actia.fr

Keywords : diagnosis, sign and interval reasoning, electrical circuits.

Abstract. In the automotive field, the use of an ECU (Electronic Control Unit) to control several functions (such as engine injection or ABS) increases. To diagnose such systems, diagnosis trees are built. These trees allow the garage mechanics to find the faulty component(s) by performing a set of tests (measurements) which has the lowest global cost as possible. Two methods which compute diagnosis of electronic circuits are presented, their different features are outlined and it is shown how they can beneficially complement each other. The first one is based on a dictionary of interval model based faults and allows one to generate off-line a diagnosis tree ; the second one uses a sign consistency based approach to derive a diagnosis hypothesis in case the fault is not among the anticipated ones.

1 Introduction

In the automotive field, the use of electronic systems to control several functions (fuel injection, ABS) has been widely developed during these last years.

These electronic systems are roughly composed of voltage supply, sensors and actuators linked to Electronic Control Units (ECU) by a wire harness.

The ECUs are equipped with an auto-diagnosis function that reliably detects the failing electronic circuit which is connected to this ECU. However, the ECU is not able to localize precisely the faulty components even if it is able to detect the failed electronic circuit.

In order to diagnose such electronic circuits, diagnosis trees are built. These trees allow the garage mechanic to find the faulty component(s) by performing a sequence of measurements which has the lowest global cost as possible. In order to automatically build these diagnosis trees from the design data supplied by the car manufacturer, AGENDA (Automatic GENeration of DiAgnosis trees) [1], an interval based approach has been developed. AGENDA is a non interactive method which uses a prediction algorithm to anticipate the effects of a set of anticipated parameter faults to build a "cross table" and an AO* algorithm to obtain an optimal diagnosis tree in term of measurements cost.

In our application domain, the performed measurements have bad precision. So the power of the interval based method cannot be fully exploited because the number of computed intervals is reduced during the "cross table" generation.

The other method uses sign reasoning. We apply an equivalent prediction algorithm to predict the fault free behavior. The study of the partial derivatives sign of each possible test w.r.t. the possible faulty parameters allow us to perform sign consistency based off-line or on-line diagnosis reasoning.

The next section describes the formalism used for modelling systems. The third section is about the interval based method and its following about the sign based method.

The fifth section outlines the different features of both methods and shows how they can beneficially complement each other.

2 System Modeling

Building a behavioral model of the system from the design data supplied by the car manufacturer is the first step according to a classical component-oriented approach [2].

2.1 Component Behavior Model

A behavioral model [3] is characterized by a set Z of n_Z mode variables $\{z_1, ..., z_{n_Z}\}$, a set X of n_X state variables $\{x_1, ..., x_{n_X}\}$ and a set Y of n_Y parameters $\{y_1, ..., y_{n_Y}\}$.

n_M different modes Z_k with $k \in \{1, ..., n_M\}$ are defined as vectors of n_Z values assigned to each of the mode variables z_i with $i \in \{1, ..., n_Z\}$. In the same way, n_L parameter initializations Y_k with $k \in \{1, ..., n_L\}$ are defined as vectors of n_Y values assigned to each of the parameters y_i with $i \in \{1, ..., n_Y\}$.

For any $k \in \{1, ..., n_M\}$, Z_k is associated with one triple σ_k defined by one behavior $b_k(X, Y)$, one parameter assignment vector Y_k and a mapping between Z_k and the other modes as shown in equation 1. The behavior $b_k(X, Y)$ is expressed as a system of equations involving state variables of the set X, parameters of the set Y and mode variables of the set Z. $\Re_k$ is a set of couples (*logical condition*, *associated mode*) which describes mode switch caused by an action (ON/OFF for a switch), a cascaded fault (surintensity on a fuse) or an electrical constraint. (voltage on a diode).

$$Z_k \Rightarrow \sigma_k = (b_k(X, Y), Y_k, \Re_k) \tag{1}$$

A component mode Z_k is said to be faulty if at least one of the mode variables in Z_F is assigned to a faulty mode ; it is fault-free otherwise.

2.2 System Model

Depending of the granularity, one may use a structural model expressing the connections between these modeled components as a system of equalities which equal two distinct state variables belonging to two distinct components [4].

Let Ψ be the system to be diagnosed defined as a set of n_Ψ individual components ψ_i with $i \in \{1, ..., n_\Psi\}$. The behavioral model of the system Ψ, called BM_Ψ, is built according to the above component oriented approach. As shown in equation 2, this model is composed of both the behavioral models BM_{ψ_i} corresponding to the components ψ_i with $i \in \{1, ..., n_\Psi\}$ and the structural model of the system Ψ, called SM_Ψ, which describes the way these components are interconnected.

$$BM_\Psi = SM_\Psi \cup BM_{\psi_1} \cup ... \cup BM_{\psi_{n_\Psi}} \tag{2}$$

A configuration of the system Ψ is defined as a n_Ψ dimension vector which associates to each component ψ_i one of its n_M^i possible modes. Consequently, the set E of possible configurations of the system Ψ is composed of $n_E = \prod_{i=1}^{n_\Psi} n_M^i$ elements.

2.3 Example

Consider the electrical circuit on figure 1. It is composed by 4 components (2 resistances, 1 voltage supply and 1 light bulb).

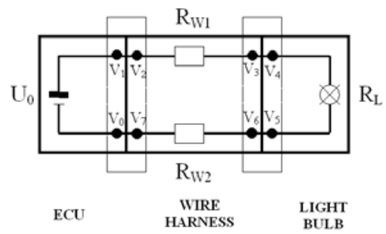

Fig. 1. Electrical circuit

We just give the (interval) values for the parameters in the nominal mode :

- $U_0 \in [4.9, 5.1]\ V$
- $R_{W1} \in [0.001, 0.001]\ \Omega$
- $R_{W2} \in [0.001, 0.001]\ \Omega$
- $R_L \in [10, 30]\ \Omega$

3 Interval Based Method

3.1 Fault and test sets anticipation

This subsection describes how the set F of faults and the set S of tests that can be performed on the system are anticipated from the behavioral model.

Let Ψ be the system to be diagnosed defined as the set of its n_Ψ individual components ψ_i with $i \in \{1, ..., n_\Psi\}$.

Fault set For each component ψ_i, let Φ^i be the set of the n^i_Φ possible modes. Let also $\Phi^i_{\neg AB}$ and Φ^i_{AB} be the set of the $n^i_{\neg AB}$ fault-free modes and the set of the n^i_{AB} faulty modes, respectively, such that $\Phi^i_{\neg AB} \cup \Phi^i_{AB} = \Phi^i$ and $\Phi^i_{\neg AB} \cap \Phi^i_{AB} = \emptyset$.

A faulty mode of the system Ψ, also called system fault, is defined as a n_Ψ dimension vector which associates to each component ψ_i one of its n^i_Φ possible faulty modes. Consequently, the set F of faults which may occur in the system Ψ is composed of $n_F = \prod_{i=1}^{n_\Psi} n^i_\Phi$ elements, called F_k with $k \in \{1, ..., n_F\}$. Note that only single faults have been considered with the definition of fault set.

Test set Let X be the set of the n_X system state variable defined by the structural model of the system Ψ.

A test is defined as a pair composed of a system state variable belonging to X and a system configuration belonging to E.

3.2 Prediction algorithm

This work assumes that the system to be diagnosed is an electronic circuit in the form of a resistive net supplied by one voltage source. For this system, let F be the set of the n_F considered faults and S the set of the n_S considered tests.

For any test in S and any fault in F, the aim of the prediction process is to provide the symbolic expression of the test outcome in the occurrence of the fault and ultimately its values interval.

Symbolic matrix expression of the system model The symbolic matrix expression of the system model is in the form $A \times X = B$ where A is a square matrix, X is a vector of state variables and B a vector of constants. This linear system represents the Ohm's and Kirchhoff's laws.

Test symbolic expression A test symbolic expression is then derived from the symbolic matrix expression of the system corresponding to the studied pairs (fault/test). This is performed by solving the symbolic matrix expression for the variables involved in the measurement corresponding to the test according to the Cramer's method [5]. The resulting test symbolic expression is proven to have a specific multi-variable homographic form [3].

Global Optimization The uncertainty that can be undertaken by the system parameter values is represented by intervals. An algorithm that optimizes the test symbolic expression is used [6]. In order to find the corresponding interval outcome of a given test in the occurrence of a given fault, the maximum and the minimum values of the symbolic expression of this test have to be found on the parallelotop defined by the parameter interval values.

Cross-table A test-matrix, fault dictionnary or "cross-table" $A = [m_i^j]$ is a matrix of dimension $n_F \times n_S$ where m_i^j represents a subset of the modalities of the test T_j.The whole set of modalities defines a partition of the test T_j domain value. m_i^j is obtained from the test T_j outcome in the occurrence of fault F_i.

3.3 AND/OR graph search

The problem of building an optimal diagnosis tree can be formulated as an ordered, best-first search on an AND/OR graph [7].

The explicit AND/OR search graph represents all the possible solutions of a given problem starting from the ground elements of this problem. For the test sequencing problem, the ground elements are the fault set, the test set and the corresponding cross-table. Obviously, the possible solutions are all the possible diagnosis trees that allow one to discriminate each fault of the fault set using any subset of the test set.

The principle of the AO* algorithm is to develop only parts of the explicit AND/OR search graph which correspond to the most interesting solutions of the problem, according to the objective function to optimize. The objective function is the function J to minimize (see equation 3).

$$J = \sum_{i=1}^{n_F} p_i \times \left(\sum_{j=1}^{n_S} d_{ij} \times c_j \right) \tag{3}$$

where n_F is the number of faults, p_i is the occurrence probability of the fault F_i, n_S is the number of tests, c_j is the cost of the test T_j and d_{ij} is 0 or 1 depending on the fact that the test T_j is on the path from the root tree to the leaf corresponding to the fault F_i.

At the end of the AO* algorithm, the optimal subgraph is a selected subgraph of the implicit AND/OR search graph that has been developed which corresponds to the optimal diagnosis tree T^*.

3.4 Example

Due to practical considerations, the anticipated faults are extrem faults (open and short circuit faults). Indeed, based on garage mechanics feedback report, their probabilities are an order of magnitude higher than deviation faults.

Let us consider 6 *faults*	*Let us consider* 9 *tests*
F_0 : *Fault free case*	T_0 : *Potential measurement on* V_0
F_1 : *Opencircuit on* U_0	T_1 : *Potential measurement on* V_1
F_2 : *Open circuit on* R_{W1}	T_2 : *Potential measurement on* V_2
F_3 : *Open circuit on* R_{W2}	T_3 : *Potential measurement on* V_3
F_4 : *Open circuit on* R_L	T_4 : *Potential measurement on* V_4
F_5 : *Short circuit on* R_L	T_5 : *Potential measurement on* V_5
	T_6 : *Potential measurement on* V_6
	T_7 : *Potential measurement on* V_7
	T_8 : *Intensity measurement*

Cross-table :

	T_1	T_2	T_3	T_5	T_6	T_7	T_8
F_0	m_1^1	m_1^2	m_2^3, m_3^3	m_0^5	m_2^6, m_3^6	m_0^7	m_1^8
F_1	m_0^1	m_0^2	m_0^3	m_0^5	m_0^6	m_0^7	m_0^8
F_2	m_1^1	m_1^2	m_0^3	m_0^5	m_0^6	m_0^7	m_0^8
F_3	m_1^1	m_1^2	m_3^3	m_2^5	m_3^6	m_2^7	m_0^8
F_4	m_1^1	m_1^2	m_3^3	m_0^5	m_3^6	m_0^7	m_0^8
F_5	m_1^1	m_1^2	m_1^3	m_1^5	m_1^6	m_1^7	m_2^8

T_0 and T_4 have been removed because they have just one modality and therefore no discriminant utility. They cannot be selected by the AO* algorithm.

As an example, the list of modalities for test T_3 is : $m_0^3([0,0])$, $m_1^3([2.4, 2.6])$, $m_2^3([4.8, 4.9])$ and $m_3^3([4.9, 5.1])$.

Figure 2 provides the optimal diagnosis tree automatically generated by AGENDA.

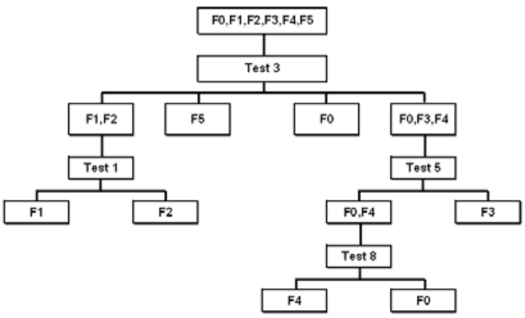

Fig. 2. Optimal diagnosis tree automatically generated

Let us notice that if the system undergoes any fault which is not in the anticipated set, the diagnosis tree traverse stops and returns no result. This is why we propose in the next section a method based on signs which can beneficially complement this approach.

4 Sign Based Method

4.1 Fault and test sets anticipation

The test set is defined as for the interval based method(see 3.1). The fault set is not relevant anymore since the sign based method is a pure consistency based method. However the components that are considered as faulty candidate corresponds to the parameters used in the interval base method for defining the fault set.

4.2 Prediction algorithm

The interest of this method is that it uses as so the results coming from the interval based method for fault free case. Indeed, during the interval computation and the global optimization, the sign of partial derivatives is studied [3]. The derivation operation is applied to the set of formal expressions obtained from the prediction. These expressions are derived w.r.t the parameter set Y.

4.3 Influence sign table

The influence sign table is a matrix composed of $n_F \times n_S$ entries where n_F is the number of faults and n_S the number of tests. It provides the influence sign of every fault on every test formal expression [8].

If the partial derivative of the formal expression is positive (+) (negative (-)), then the parameter and the expression vary in the same (opposite) direction.

If the influence sign is unknown (?), i.e. the sign of the partial derivative w.r.t. the faulty parameter is not the same on the paralleloptop defined by the parameter's interval values.

If there is no influence (0) of a fault on a formal expression, i.e. the partial derivative w.r.t. the faulty parameter is zero.

So a parameter may be associated with 4 qualitative values : *increases* (+), *decreases* (-), *unknown* (?) and *nominal* (0).

For a given quantity X, let us note its nominal value X^0 and its measured value X^m, and define a qualitative value $\partial X = Sign(X^m - X^0)$ where $\partial X \in \{+, -, ?, 0\}$.

Now, if the value of the formal expression F is greater than the nominal value of the expression, i.e. $\partial F = +$, one can deduce that one of the parameters p which have a positive sign influence has increased or one of the parameters p which have a negative sign influence has decreased (4).

$$\partial F = + \Longrightarrow (\delta F/\delta p = + \Longrightarrow \partial p = +) \ or \ (\delta F/\delta p = - \Longrightarrow \partial p = -) \tag{4}$$

Influence sign table for the example :

	T_1	T_2	T_3	T_5	T_6	T_7	T_8
U_0	$+$	$+$	?	$+$	?	$+$	$+$
R_{W1}	0	0	?	$-$	?	$-$	$-$
R_{W2}	0	0	$+$	?	$+$	?	$-$
R_L	0	0	$+$	$-$	$+$	$-$	$-$

4.4 Qualitative reasoning

Making hypothesis An hypothesis is a vector having n_F entries representing each parameter. The vector values are qualitative within the set $\{+, -, ?, 0\}$.

Hypothesis are directly computed from the influence sign table. An association between a parameter and its sign influence is defined.

The initial hypothesis is defined by initializing all parameters with *unknown*.

Given a test T_j, if $\partial T_j = 0$, all the parameters having a non zero influence sign on T_j are declared "not guilty"[3]. So only the parameters associated with influence sign 0 may be faulty.

If $\partial T_j = +$ or $\partial T_j = -$, an hypothesis related to T_j, noted $H_{\partial T_j=+}$ or $H_{\partial T_j=-}$, is derived from the influence sign table. $H_{\partial T_j=+/-}$ is a vector whose components refer to the parameters, i.e. $H_{\partial T_j=+/-} = [H^{P_1}_{\partial T_j=+/-}, \dots, H^{P_{n_Y}}_{\partial T_j=+/-}]$.

Hypothesis combination The aim of this part is to compute a new hypothesis from the current one and the one obtained from the recent test measurement.

Let us define the operator[8] applied to 2 hypothesis components and having for result the combination of both hypothesis components. We called it ©.

©	+	−	?	0
+	+	0	+	0
−	0	−	−	0
?	+	−	?	0
0	0	0	0	0

Criterion for test selection The selection of the next test to perform is done w.r.t. the current hypothesis and the influence sign table. The criterion must capture the additional information provided by every test.

Given a current hypothesis H_C, the idea is to check H_C against the hypothetic hypothesis $\hat{\mathrm{H}}_{\partial T_j}$ in the 2 possible cases $\partial T_j = +$ and $\partial T_j = -$, for every test to go.

$\hat{\mathrm{H}}_{\partial T_j=+}$ and $\hat{\mathrm{H}}_{\partial T_j=-}$ are obtained from the influence sign table as in the *making hypothesis* section. The corresponding new hypothesis $\hat{\mathrm{H}}_N^{\partial T_j=+}$ and $\hat{\mathrm{H}}_N^{\partial T_j=-}$ are computed as in the *hypothesis combination* section : $\hat{\mathrm{H}}_N^{\partial T_j=+} = \mathrm{H}_C © \hat{\mathrm{H}}_{\partial T_j=+}$ and $\hat{\mathrm{H}}_N^{\partial T_j=-} = \mathrm{H}_C © \hat{\mathrm{H}}_{\partial T_j=-}$.

The quantity of information provided by a test T_j is evaluated by the number n_d^{j+} and n_d^{j-} of syntactical differences of H_C w.r.t. the new hypotheses $\hat{\mathrm{H}}_N^{\partial T_j=+}$ and $\hat{\mathrm{H}}_N^{\partial T_j=-}$. Indeed, no syntactical difference just confirms the current hypothesis without providing new information.

The final criterion (5) is the absolute value $|n_d^{j+} - n_d^{j-}|$ divised by the cost of the test.

$$C_j = \frac{|n_d^{j+} - n_d^{j-}|}{cost(T_j)} \tag{5}$$

[3] This makes use of the exoneration assumption which is justified in the considered application domain (component models are functionally reversible)

As we want to build a balanced tree, the chosen test is the one obtaining the lowest criterion value.

5 Complementing Interval Based Method

5.1 Why ?

In automotive domain, resistance values may increase due to corrosion originated by humidity. Suppose in the previous example that the resistance value R_{W1} increases and takes a value about 50 Ω.

The first measurement to perform in the tree generated by the interval based method is the test T_3 (see detail of the cross table in section 3.4). The measure is about $3.5V$. This value is not a possible modality.

In general, this type of deviation faults are not anticipated in the interval based method because they are much less probable than extrem faults and they make the test outcomes interpretation in term of modalities more difficult. In this context, the tree traverse does not achieve a diagnosis.

5.2 Complementing the interval method with the sign based method

Sign based method may be used in complement of the interval based method. The process of tree generation may be based on an anticipated faults set including the fault with the highest occurrence probability. Then if the measurement performed for a test is not present in the modality set, the sign based method is called.

The already performed measurement set is used to generate the initial hypothesis using the same previously described combination mechanism. But it is important to notice that the faults remaining down in the diagnosis tree are not relevant. The sign method must consider the whole fault set again.

Coming back to our example, let us assume that the fault to isolate is the increase of the parameter R_{W1}. The diagnosis tree traverse gets stuck from the very first test (measurement is out of its modalities). The sign method determines (details are omitted due to the length restriction) that the fault can be isolated with 3 tests.

	∂U_0	∂R_{W1}	∂R_{W2}	∂R_L
Initial hypothesis	?	?	?	?
$After\ \partial\ T_1 = 0$	0	?	?	?
$After\ \partial\ T_8 = -$	0	+	+	+
$After\ \partial\ T_3 = -$	0	+	0	0

Fig. 3. Successive hypotheses table.

The sign based method allows to find the right diagnosis.

6 Conclusion

Both presented approaches are based on the same preprocessing method and allow to make diagnosis.

The interval based method uses an anticipated dictionnary of faults giving faulty values to every parameter whereas in the qualitative one, faults are expressed in term of deviations w.r.t. the nominal values : you can consider more easily a structural or topological fault in the first approach than in the second one.

The interval based method generates an optimal diagnosis tree in terms of cost whereas the sign based method is based on a local optimisation of the next test to be performed and has no warranty in term of global cost optimality.

But the main point of sign based method is the isolation of parameter deviation fault. That's why both methods are actually complementary as illustrated in the fault scenario example.

Acknowledgement.The authors thank Jerome THOMAS for his useful comments and interesting discussions.

References

1. P. P. Faure, X. Olive, L. Travé-Massuyès, and H. Poulard, "Agenda : Automatic GENeration of Diagnosis trees," in *JDA'01*, (Toulouse (France)), 2001.
2. P. Dague, O. Raiman, and P. Devès, "Troubleshooting : when modeling is the trouble," in *AAAI*, (Seattle (USA)), pp. 600–605, 1987.
3. P. P. Faure, *An Interval Model-Based Approach for Optimal Diagnosis Tree Generation : Application to the Automotive Domain.* PhD thesis, LAAS, 2001.
4. A. Darwiche, "Model-Based Diagnosis using Structured System Descriptions," *Artificial Intelligence*, vol. 8, pp. 165–222, 1998.
5. M. E. Kole, *Algorithms for Symbolic Circuit Analysis Based on Determinant Calculations.* PhD thesis, Universiteit Twente, 1996.
6. S. Group, "Modal Intervals : draft for a basic tutorial," tech. rep., IMA Informática y Matemática Aplicada department, University of Girona, Oct. 1998.
7. K. R. Pattipati and M. G. Alexandridis, "Application of Heuristic Search and Information Theory to Sequential Fault Diagnosis," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 20, pp. 872–887, July 1990.
8. L. Travé-Massuyès, P. Dague, and F. Guerrin, *Le raisonnement qualitatif.* HERMES, 1997.