Complementing an Interval Based Diagnosis Method With Sign Reasoning in the Automotive Domain

Xavier OLIVE[1,2] – Louise TRAVE-MASSUYES[1] – Jérôme THOMAS[2]

[1] LAAS-CNRS, 7 Avenue du Colonel ROCHE, F-31077 TOULOUSE Cedex 4
[2] ACTIA, 25 chemin de Pouvourville, BP4215, F-31432 TOULOUSE Cedex 4
xolive@laas.fr, louise@laas.fr, jthomas@actia.fr

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, interval based and sign consistency based, which compute diagnosis of electronic circuits are presented, their different features are outlined and it is shown how they can beneficially complement each other.

Keyword : diagnosis, sign and interval reasoning, electrical circuits, test sequencing problem

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) (Faure et al., 2001), 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 crosstable, and an AO* algorithm to obtain an optimal diagnosis tree in terms of measurement cost.

If this method is very powerful to diagnose extreme faults (i.e., faulty parameter value is zero or infinite), one of its main limitations is when the current fault is a parameter value deviation. The generated tree does not allow the garage mechanic to reach a diagnostic in this case.

Nevertheless, the study of the partial derivatives signs of each possible test w.r.t. the possible faulty parameters done during the interval based approach provides the signed influence of parameters on each possible test. This paper takes advantage of this knowledge to combine the interval based approach with a method using qualitative reasoning (Cascio et al., 1999). This combination allows us to diagnose extreme as well as deviation faults in an integrated approach.

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 (Dague et al., 1987).

Component Behavior Model

A behavioral model (Faure, 2001) is characterized by a set **Z** of n_Z mode variables $\{z_1,...,z_{nZ}\}$, a set **X** of n_X state variables $\{x_1,...,x_{nX}\}$ and a set **Y** of n_Y parameters $\{y_1,...,y_{nY}\}$.

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 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**. $\mathbf{R_k}$ is a set of couples (*logical condition*, *destination mode*) which describes mode switches caused by an action (ON/OFF for a switch), a cascaded fault (surintensity on a fuse) or an electrical constraint (voltage on a diode).

$$Z_k \Rightarrow \sigma_k = (b_k(X,Y), Y_k, R_k) \qquad (1)$$

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

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 (Darwiche, 1998).

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. This model is composed of both the behavioral models $BM\psi_i$ 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.

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^i_M possible modes. Consequently, the set **E** of possible configurations of the system ψ is composed of $n_E = \prod_{i=1,..,n\psi} n^i_M$ elements.

Example

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

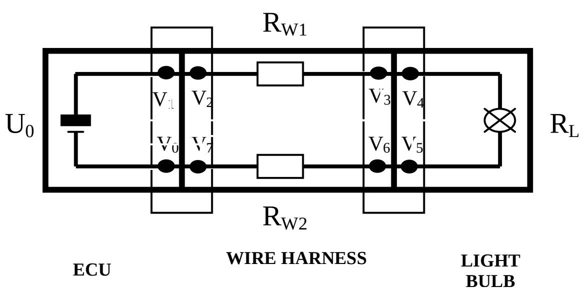

Figure 1 : Electrical circuit

We just give the (interval) values for the parameters in the nominal mode : $U_0 \in [4.9,5.1]$ V, $R_{W1} \in [0.001,0.001]\ \Omega$, $R_{W2} \in [0.001,0.001]\ \Omega$ and item $R_L \in [10,30]\ \Omega$.

Interval Based Method

The interval based method is a non interactive method which solves the Test Sequencing Problem (Pattipati and Dontamsetty, 1992). It first uses a prediction algorithm to anticipate the effects of a set of anticipated parameter faults to build a "cross table". Then an AO* algorithm is applied to obtain an optimal diagnosis tree in terms of measurement cost (Faure, 2001).

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 anticipated faults and **S** the set of the n_S anticipated 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**.**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 (Kole, 1996). The resulting test symbolic expression is proven to have a specific multi-variable homographic form. The partial derivative w.r.t. a parameter p of this formal expression have the property to keep the same sign on the domain of p (cf. proof in (Faure, 2001)).

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 (SIGLA/X, 1998). In order to find the corresponding interval outcome of a given test in the occurrence of a given fault, the extrema values of the symbolic expression are searched for on the parallelotop Δ defined by the parameter interval values.

Crosstable

A test-matrix, fault dictionary or crosstable A= $[m^j_i]$ is a matrix of dimension n_F x n_S where m^j_i 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^j_i is obtained from the test T_j outcome in the occurrence of fault F_i This crosstable is the input of an AO* algorithm (Bagchi and Mahanti, 1983) which outputs an optimal diagnosis tree (for details, see (Faure, 2001)).

Example

Due to practical considerations, the anticipated faults are extreme 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 the fault free case F_0 and 5 faults : F_1 an open circuit on U_0, F_2 an open circuit on R_{W1}, F_3 an open circuit R_{W2}, F_4 an open circuit on R_L and F_5 a short circuit on R_L.

Let us consider 9 tests : T_i the potential measurement on V_i, $\forall i \in \{0..7\}$ and T_8 the intensity measurement.

As an example, some results of prediction for test T_3 are given under assumption of fault occurrence : F_0 : [4.8,5.1], F_1 : [0,0], F_2 : [0,0], F_3 : [4.9,5.1], F_4 : [4.9,5.1], F_5 : [2.4,2.6].

Modalities are defined as : m^3_0([0,0]), m^3_1([2.4,2.6]), m^3_2([4.8,4.9]) et m^3_3([4.9,5.1]).

The crosstable corresponding to the example is given on the figure 2.

	T_1	T_2	T_3	T_4	T_5	T_6	T_8
F_0	m^1_1	m^2_1	m^3_2,m^3_3	m^4_2,m^4_3	m^5_0	m^6_0	m^8_1
F_1	m^1_0	m^2_0	m^3_0	m^4_0	m^5_0	m^6_0	m^8_0
F_2	m^1_1	m^2_1	m^3_0	m^4_0	m^5_0	m^6_0	m^8_0
F_F	m^1_1	m^2_1	m^3_3	m^4_3	m^5_2	m^6_2	m^8_0
F_4	m^1_1	m^2_1	m^3_3	m^4_3	m^5_0	m^6_0	m^8_0
F_5	m^1_1	m^2_1	m^3_1	m^4_1	m^5_1	m^6_1	m^8_2

Figure 2 : Crosstable example.

T_0 and T_7 have been removed because they have just one modality and therefore no discriminant power. They cannot be selected by the AO* algorithm.

Figure 3 provides the optimal diagnosis tree automatically generated by AGENDA.

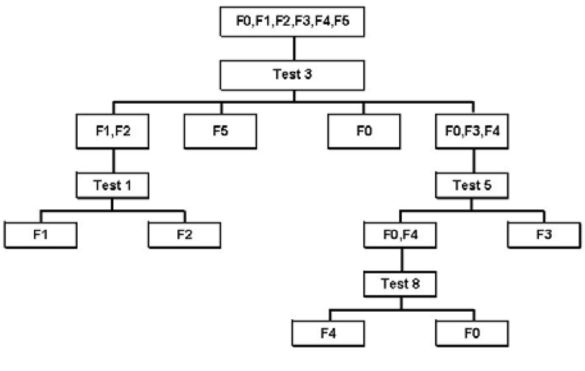

Figure 3 : Optimal diagnosis tree

Let us notice that if the system undergoes any fault which is not in the anticipated set, the diagnosis tree traversal stops and returns no result[1]. That is why we propose a method based on signs which beneficially complements this approach.

Sign Based Method

Fault and test sets anticipation

The test set is defined as for the interval based method. The sign based method is a pure consistency based method. The components that are considered as faulty candidates correspond to the parameters used in the interval based method for defining the fault set. So, let **F** be the set of n_F ($=n_Y$) faults and **S** be the set of n_S tests.

Sign Prediction

The interest of this method is that it uses as so the results coming from the interval based method for the fault free case. Indeed, during the interval computation and the global optimization, the sign of partial derivatives is analyzed (Faure, 2001). 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. But not all the partial derivative signs can be determined during the global optimization.

With the result of the optimization, it is however possible to deduce the not yet determined signs. Indeed, for any test, (Faure, 2001) proves that the test formal expression is monotone and that its extrema are on the edge of the parallelotop Δ .

So let p be a parameter of the system having the interval values $I_p = [a_p, b_p]$ and T_j a test formal expression. The minimum Y^- and the maximum Y^+ values of T_j on the paralleloptop Δ are respectively obtained for the vectors $X^-$and X^+ of Δ such that $X^-=(x^-_1, ..., x^-_n)^T$ and $X^+=(x^+_1, ..., x^+_n)^T$ where $x^s_i \in \{a_i, b_i\}$, $\forall s \in \{-,+\}$ and $\forall i \in \{1,...,n_Y\}$.

Then we can be deduced :

$$\left.\begin{array}{l} T_j \ is \ monotone \ w.r.t. \ p \\ a_p < b_p \\ x^-_p \left|\begin{array}{l} < \\ > \\ = \end{array}\right. x^+_p \end{array}\right| \Longrightarrow \frac{\delta T_j}{\delta p} \left|\begin{array}{l} \geq 0 \\ \leq 0 \\ \\ = 0 \end{array}\right. \quad (2)$$

From here, we consider that the partial derivatives are strictly increasing or strictly decreasing. Cases where the partial derivatives are null without having $T_j(a_p) = T_j(b_p)$ are not typical for our domain.

[1] Obviously, when a non anticipated fault has the same symptoms as an anticipated one, the tree leads to an incorrect diagnosis.

So all the partial derivative signs are determined.

Influence sign table

The influence sign table is a matrix composed of n_F x n_S entries. It provides the influence sign of every fault on every test formal expression (Travé et al., 1997).

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

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* (–), *nominal* (0) and *unknown* (?) used for initialization.

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

Now, if the value of the formal expression T_j is greater than the nominal value of the expression, i.e. $\partial T_j = +$, 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.

Influence sign table for the example :

	T_1	T_2	T_3	T_4	T_5	T_6	T_8
U_0	+	+	+	+	+	+	+
R_{W1}	0	0	–	–	–	–	–
R_{W2}	0	0	+	+	+	+	–
R_L	0	0	+	+	–	–	–

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"[2]. 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. $H(\partial T_j = +/–) = [H(p_1, \partial T_j = +/–), \ldots , H(p_{nY}, \partial T_j = +/–)]$.

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 (Travé et al., 1997) 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 hypothetical hypothesis $H'(\partial T_j)$ in the 2 possible cases $\partial T_j = +$ and $\partial T_j = –$, for every test to go.

$H'(\partial T_j=+)$ and $H'(\partial T_j=–)$ are obtained from the influence sign table as in the *making hypothesis* section. The corresponding new hypothesis $H'_{New}(\partial T_j=+)$ and $H'_{New}(\partial T_j=–)$ are computed as in the *hypothesis combination* section : $H'_{New}(\partial T_j=+) = H_C$ © $H'(\partial T_j=+)$ and $H'_{New}(\partial T_j=–) = H_C$ © $H'(\partial T_j=–)$.

The quantity of information provided by a test T_j is evaluated by the number n^{j+}_d and n^{j-}_d of syntactical differences of H_C w.r.t. the new hypotheses $H'_{New}(\partial T_j=+)$ and $H'_{New}(\partial T_j=–)$. Indeed, no syntactical difference just confirms the current hypothesis without providing new information.

The final criterion is the absolute value $|n^{j+}_d - n^{j-}_d|$ divided by the cost of the test.

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

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

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

Figure 4 : Successive hypotheses table.

Managing multiple configurations

As defined in the *Modeling System* section, the system may have several configurations, corresponding to components in different modes. The sign based method only uses fault free modes configurations. The "best" configuration to start with in the diagnostic process is intuitively the one which involves the maximum number of parameters.

Complementing Interval Based Method

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 crosstable on figure 2). 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 extreme faults and they make the test outcomes interpretation in terms of modalities more difficult. In this context, the tree traversal does not achieve a diagnosis.

Complementing the interval method with the sign based method

The process of tree generation is based on an anticipated faults set including the fault with the highest occurrence probability. 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.

Coming back to our example, let us assume that the fault to isolate is the increase of the parameter R_{W1}. The diagnosis tree traversal 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.

The sign based method allows to find the right diagnosis.

Application : The LED System

The circuit called the *LED system* represents a LED display circuit of a given ECU. We suppose that switches Sw_1 and Sw_2 cannot be closed simultaneously. It is composed by 6 components connected by wires. In our case, the wires are the faulty candidates. So the system has 14 resistive parameters that may be faulty. The test set is composed by 27 tests.

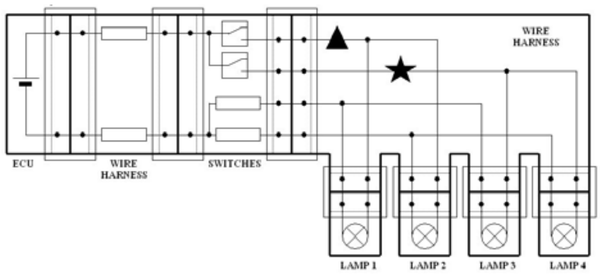

Figure 7 : The LED System

We consider 2 faulty scenarios visualized by the *star* and the *triangle* symbols (Figure 7) which define 2 faults due to the increase of the resistive parameter P_2 and P_5 respectively. Test sequences are given for both cases on figures 5 and 6.

On figure 5, the diagnosis is obtained after performing 4 tests (T_5, T_{16}, T_3 et T_{10}) in configuration 1 (i.e. Sw_2 opened, Sw_1 closed).

On figure 6, the diagnosis is obtained after performing 2 tests : T_5 in configuration 1 et T_5 in configuration 2 (i.e. inverting the 2 switch positions). Given that T_5^1 is nominal, and under exoneration assumption, all the parameters involved in this test are nominal. Consequently, the parameter P_5 keeps its suspect value (?) but is not proven guilty. Its variation w.r.t. its nominal value is confirmed by test T_5^2.

Conclusion

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

The interval based method uses an anticipated dictionary of faults giving faulty values to every parameter whereas in the qualitative one, faults are expressed in terms of deviations w.r.t. the nominal values : you can consider more easily a

	∂P_0	∂P_1	$\partial \mathbf{P}_2$	∂P_3	∂P_4	∂P_5	∂P_6	∂P_7	∂P_8	∂P_9	∂P_{10}	∂P_{11}	∂P_{12}	∂P_{13}
Inital hyp.	?	?	**?**	?	?	?	?	?	?	?	?	?	?	?
$\partial T^1_5 = +$	+	–	**+**	+	+	0	+	+	+	+	+	+	+	–
$\partial T^1_{16} = +$	+	–	**+**	+	+	0	0	0	0	0	+	0	0	0
$\partial T^1_3 = –$	0	0	**+**	+	+	0	0	0	0	0	+	0	0	0
$\partial T^1_{10} = +$	0	0	**+**	0	0	0	0	0	0	0	0	0	0	0

Figure 5 : Successive hypotheses table for *triangle* fault.

	∂P_0	∂P_1	∂P_2	∂P_3	∂P_4	$\partial \mathbf{P}_5$	∂P_6	∂P_7	∂P_8	∂P_9	∂P_{10}	∂P_{11}	∂P_{12}	∂P_{13}
Inital hyp.	?	?	?	?	?	**?**	?	?	?	?	?	?	?	?
$\partial T^1_5 = +$	0	0	0	0	0	**?**	0	0	0	0	0	0	0	0
$\partial T^1_{16} = +$	0	0	0	0	0	**+**	0	0	0	0	0	0	0	0

Figure 6 : Successive hypotheses table for *star* fault.

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 optimization of the next test to be performed and has no warranty in terms 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 scenarios.

References

Bagchi, A. and A. Mahanti (1983). *Admissible Heuristic Search in AND/OR Graphs*. Theoritical Computer Science 24, pp. 207--219.

Cascio, F., L. Console, M. Guagliumi, M. Osella, A. Panati, S. Sottano and D. Theseider-Dupré (1999). *Generating on-board Diagnostics of Dynamic Automotive Systems Based on Qualitative Models*. In: 10th International Workshop on Principles of Diagnosis, DX-99. Loch Awe (Scotland). pp. 78--89.

Dague, P., O. Raiman and P. Devès (1987). *Troubleshooting : when modeling is the trouble*. In: AAAI. Seattle (USA). pp. 600--605.

Darwiche, A. (1998). *Model-Based Diagnosis using Structured System Descriptions* Artificial Intelligence 8, 165--222.

Faure, P. P. (2001). *An Interval Model-Based Approach for Optimal Diagnosis Tree Generation : Application to the Automotive Domain*. PhD thesis. LAAS.

Faure, P. P., X. Olive, L. Travé-Massuyès and H. Poulard (2001). *AGENDA : Automatic GENeration of Diagnosis trees*. In: JDA'01. Toulouse (France).

Kole, M. E. (1996). *Algorithms for Symbolic Circuit Analysis Based on Determinant Calculations*. PhD thesis. Universiteit Twente.

Pattipati, K. R. and M. Dontamsetty (1992). *On a Generalized Test Sequencing Problem* IEEE Transactions on Systems, Man and Cybernetics 22(2), 392--396.

SIGLA/X, Group (1998). *Modal Intervals : draft for a basic tutorial*. Technical report. IMA Informatica y Matematica Aplicada department, University of Girona.

Travé-Massuyès, L., P. Dague and F. Guerrin (1997). *Le raisonnement qualitatif* Ed. HERMES.