Transactions on Information and Communications Technologies vol 8, © 1995 WIT Press, www.witpress.com, ISSN 1743-3517

Fuzzy fault evaluation in causal diagnostic reasoning

P.F. Parra[a], A. Ligeza[b]
[a]*University of Balearic Islands, Cra. de Valldemossa, km 7.5, E-07071 Palma, Spain*
[b]*Institute of Automatics AGH, al. Mickiewicza 30, 30-059 Kraków, Poland*

Abstract

Causal reasoning is a practical support for model-based diagnosis [1]. The paper addresses the issues of extending diagnostic reasoning based on abductive analysis of causal structures. The extension is aimed at admitting fuzzy characterization of faults; binary evaluation (i.e. faulty/correct) is no longer necessary. The degree of faultyness is expressed with use of basic fuzzy notions. This extension uses a uniform model for representing causal behaviour of diagnosed systems: it has the form of an AND/OR/NOT causal graph allowing for specification of causality types reflecting the basic logical operations [3]. The graph can be used to search for potential (possible) diagnoses. Validation of generated diagnoses is performed by propagation of fuzzy faults upwards the graph [3]. A motivational discussion introducing the presented ideas at an intuitionistic level is presented in brief. Finally, possibilities for further extensions and related work are pointed out.

1 Introduction

Automated diagnosis constitutes an important area of applied Artificial Intelligence (AI) and control theory. The practical importance of diagnostic methods consists in supporting human operator with means for symbolic representation and reasoning incorporating large bodies of knowledge so as to efficiently detect faults (causes) in complex systems, being given only a set of symptoms of abnormal behaviour. The use of automated diagnostic systems saves both time and effort and, what is more important allows for easy copying and wide spread of such systems (the majority of expenditures are spent on first working system), and certain logical properties (consistency, correctness) are kept constant over time.

Present-day supervision, diagnostic and control systems are implemented mainly as rule-based expert systems based on shallow diagnostic knowledge of the experts [10, 11]. The implementation and debugging of such systems is a time-consuming, very tedious task, and the performance of such systems is strictly limited to the class of problems described with the expert-acquired rules.

In order to overcome the above difficulties of the mentioned above first-generation expert diagnostic systems another approach is put forward by AI [2, 7, 8, 9]. The very basic idea of this novel approach consists in representing the internal structure and causal behaviour of systems rather than the shallow diagnostic knowledge. It is assumed, that a sufficiently complex reasoning mechanism will allow then for determination of faults based on deep knowledge and causal reasoning about the systems behaviour, provided that a set of symptoms is given. This kind of approach is also referred to as *model-based diagnosis* or *diagnosis from first principles*. In order to do it in an efficient manner an appropriate knowledge representation scheme is needed. Moreover, a powerful reasoning paradigm incorporating causal reasoning and backward graph traversing for determining the set of potential failures, capable of focusing on specific problem area and dealing with incomplete and uncertain knowledge is necessary. The advantages of the new approaches include also a significant degree of device independence of diagnostic procedures, the ability to deal with novel, not yet encountered problems, and the ability to reason at different levels of hierarchy, i.e. with regard to different degree of details.

2 Motivational discussion

What strikes in most of the literature on diagnosis, is that almost all diagnostic approaches assume just binary model of potentialy faulty components, i.e. a component $c \in \mathbf{C}$ can take two states – it can work correctly or it can be faulty. Although some authors admit the discussion of slightly extended models (in [7, 4] it is shown that several possible faulty states of a component can be taken into account), the consideration usually seems not to go beyond the stage of general discussion. Although the simplest binary fault model usually fits well the class of mostly considered applications (i.e. digital circuits diagnosis), there are areas where the overal failure of a system can be caused by "partial" faulty behaviour of its components.

Let us consider for example a "slightly" damaged tube; it is working "almost" O.K., but there is a small leakage of the transported medium. Let V_0 denote the volume of the leakage while V_m the volume of the "correct" flow. One can define $d = V_0/(V_0 + V_m)$ to be a fuzzy value of the faulty

Transactions on Information and Communications Technologies vol 8, © 1995 WIT Press, www.witpress.com, ISSN 1743-3517

behaviour of the tube. Of course, there is $0 \leq d \leq 1$; further, $d = 1$ can be interpreted as total fault of the tube and $d = 0$ as the correct behaviour.

A fault similar in nature is typically caused by sediment blocking the flow of media. In this case one can define the "degree" of faultyness d as $d = (A - a)/A$, where a is the current area still opened for flow, A is the potentially (initially) maximal area, and where $0 \leq a \leq A$; typically a decreases as a function of time, i.e. the longer the system works, the smaller a becomes.

Similar considerations apply in a natural way to dynamic systems. If the observed trajectory is not the one desired (e.g. optimal), one can apply a normalised "measure" of the distance between the desired trajectory and the observed one. Evaluation of such a fuzzy failure can constitute a basic factor for determining the quality of the process and invoking diagnostic procedure if the observed quality is too low (see also [5]).

Any symptom of fault can be described with use of natural language. This is usually done by a domain expert or system user. For diagnostic reasoning we may need more precise definitions. In case of technical systems symptoms can usually be defined with use of facts or more complex formulae of propositional or predicate calculus. For example, a symptom like `battery_low` may be defined by an atom `voltage(battery)` being assigned fuzzy truth value btween 0 and 1 depending on the required normal voltage and the currently observed one. More complex definitions may also be possible; a symptom like `overheating` can be defined by logical conjunction of facts, for example `temperature(coolant)` and `time(observation)`. A symptom may have more than one definition, i.e. its definition has alternative possibilities of occuring.

A more general fuzzy model of component fault can be considered as follows. Let us consider a system component with output variable denoted by y. Assume that y can take the values from some interval; let us put $y \in [y_{min}, y_{max}]$. Assume that in the ideal case y_{max} is achieved, while y_{min} denotes the worst value. Further, let f be any continuous, monotonic function defined on the considered interval (the choice of f is arbitrary; it may depend on domain expert knowledge). Now, one can define fuzzy measure of component fault d as:

$$d = \frac{f(y_{max}) - f(y)}{f(y_{max}) - f(y_{min})} \tag{1}$$

with $d = 1$ denoting total failure (maximal fault), and $d = 0$ denoting perferctly correct work. Any intermediate value denotes some strength of the observed fault.

2.1 Motivation for this paper

In this paper we would like to present and extension of diagnostic reasoning concerning "degree of faultyness" of components in the diagnoses system. Most of the current approaches admits only binary evaluation of faults, i.e. a component can be faulty or correct. In some papers, the possibility of diagnosing several different fault states of components is pointed out (e.g. [7]), however the practical considerations are limited to the mentioned above binary case. The reasons for that seem obvious – most of the practical examples concern diagnosis digital (binary) circuits and in such systems an approach consisting in binary classification faulty/O.K. is satisfactory; the same applies to simplified considerations concerning most simple electrical circuits. However, for more complex systems incorporating large number of diversified components (e.g. pneumatic, hydraulic, mechanical, chemical, biochemical, etc.) the above approach is not necessarily correct. Moreover, it seems unreasonable to attempt to build somewhat general theory of diagnosis basing only on one application domain. The point is that in such systems certain components can be regarded as "faulty to certain degree", not just *faulty* or *correct.*

The sources of such "partial" faults usually lies in the characteristics of the components – they perform *continuous* processes with some boundary conditions determined by the state of elementary parts of them. This type of processes include ones based on different sorts of flows, concentration, exchange of energy, etc., where the state parameters change in a continuous way. After some time of working (as a natural process) or due to some accidential changes the process becomes one not a one hundred percent correct

3 AND/OR/NOT causal graphs

The key issue for explaining faulty behaviour in technical systems is the knowledge of *causal* relationship among *symptoms* occurring in the system. In order to model formally causal relationship among symptoms we assume that whenever there is such a relation between nodes n and n', there is a directed arc pointing from n to n'. An arc pointing from n to n' says that n *may cause* n' to occur. Let $\mathbf{E}^2$ denote all such simple dependencies (i.e. be a binary relation). In case symptoms $n_1, n_2, \ldots, n_{i-1}$ *may* cause n_i only when occurring simultaneously we shall say that conjunction of $n_1, n_2, \ldots, n_{i-1}$ causes n_i. To denote the set (i.e. an i-ary relation) of all such dependencies we shall use the symbol $\mathbf{E}^i$, where $i \geq 3$. For simplifying the notation let us put $\mathbf{E}^* = \mathbf{E}^2 \cup \mathbf{E}^3 \cup \ldots \cup \mathbf{E}^l$, where $l-1$ is the maximal number of symptoms causing simultaneously some symptom to occur. Further, let $\mathbf{E}^-$ denote the set of binary negative influences, i.e. $(n, n') \in \mathbf{E}^-$ iff the lack of n causes n' to occur and vice versa.

Transactions on Information and Communications Technologies vol 8, © 1995 WIT Press, www.witpress.com, ISSN 1743-3517

Further, we shall refer to nodes from which no arc points to another node as to *final* or *terminal nodes* (roots); the set of final nodes will be denoted as $\overline{\mathbf{F}}$. Similarly, the set of nodes such that no arc is pointing to any of them will be referred to as *initial* or *starting nodes* (leafs); such a set will be denoted as $\underline{\mathbf{I}}$. We can define the basic causal structure for further considerations, i.e. an AND/OR/NOT causal graph [3].

Definition 1 *Let* $\mathbf{N}$ *denote the set of considered symptoms,* $\mathbf{N} = \mathbf{D} \cup \overline{\mathbf{D}} \cup \mathbf{V} \cup \mathbf{F}$, *where* $\mathbf{F}$ *is a set of fault symptoms,* $\mathbf{D}$ *is a set of elementary diagnoses,* $\overline{\mathbf{D}}$ *is is the set of complementary elementary diagnoses,* $\mathbf{V}$ *is a set of pre-specified intermediate symptoms. Further, let* $\mathbf{E}^*$ *denote a set of relations defining causal dependencies, and* $\mathbf{E}^-$ *a set of relations defining binary negative dependencies. The AND/OR/NOT causal graph is a structure* $(\mathbf{N}, \mathbf{E}^*, \mathbf{E}^-)$ *satisfying the following conditions:*

- *the set of nodes of the graph is the set* $\mathbf{N} = \mathbf{D} \cup \overline{\mathbf{D}} \cup \mathbf{V} \cup \mathbf{F}$ *(for simplicity we do not distinguish between the graph nodes and labelling them symptoms),*
- *the causal relations given by* $\mathbf{E}^*$ *and* $\mathbf{E}^-$ *define the arcs in the graph,*
- *there is* $\mathbf{F} \subseteq \overline{\mathbf{F}}$ *and* $\underline{\mathbf{I}} \subseteq \mathbf{D} \cup \overline{\mathbf{D}}$ *in the graph.*
- $(\mathbf{D} \cup \overline{\mathbf{D}}) \cap \mathbf{V} = \emptyset$.
- $(\mathbf{D} \cup \overline{\mathbf{D}}) \cap \mathbf{F} = \emptyset$
- $\mathbf{F} \cap \mathbf{V} = \emptyset$,
- *there are no loops in the graph.*

The interpretation of the above definition is simple. There is a directed arc from a node n to node n' whenever there is direct causal relation between n and n', i.e. n is a cause or reason for n' (possible one of a set of such causes).

Definition 2 *Let* $\mathbf{N} = \mathbf{D} \cup \overline{\mathbf{D}} \cup \mathbf{V} \cup \mathbf{F}$ *be a set of considered symptoms, and let* $\mathbf{G} = (\mathbf{N}, \mathbf{E}^*, \mathbf{E}^-)$ *be an AND/OR/NOT causal graph. A diagnostic problem is a pair* $(\mathbf{G}, \mathbf{F}^0)$, *where* $\mathbf{F}^0$ *is a set of observed failure symptoms,* $\mathbf{F}^0 \neq \emptyset$.

; but the system can still work. However, its overall performance decreases, as it is influenced by the state of components.

Our main point is to put forward a fuzzy characterization of component faults. Further, we present a simple model for representation of causality in a form of AND/OR/NOT causal graph. This kind of graph reflects logical connectors most commonly used in reasoning. We present an approach to

Transactions on Information and Communications Technologies vol 8, © 1995 WIT Press, www.witpress.com, ISSN 1743-3517

propagate the fuzzy measures of faults through the causal graph. As a result we can estimate the influence of fuzzy faults on the observed fuzzy failures. This approach provides a tool for evaluation (validation) of fuzzy diagnoses in case of incomplete and imprecise model of the diagnosed system.

4 An approach to continuous fault diagnosis with fuzzy fault measures

Let us present a model of search of diagnosis enchanced with fuzzy coefficients of faulty behaviour. We assume that any component (possibly control action and operation condition) has assigned some rough fuzzy estimation of its faultyness, i.e. for any $d_i \in \mathbf{D}$ there is known some value α_i denoting the "state of faultyness; there is $\alpha_i \in [0,1]$. Similar assumption, preserving consistency, concerns the elements of $\overline{\mathbf{D}}$, and the respective values are denoted with $\overline{\alpha}_i$ ($\alpha_i + \overline{\alpha}_i = 1$).

Following this line of reasoning, we can further assume that also the "degree" or "strength" of influence of certain fuzzy faults can be characterized with fuzzy coefficients; thus we assume that the arcs of the causal graph are assigned some numbers β_i, where $\beta_i \in [0,1]$.

For performing operations on the above fuzzy coefficients one can apply just min and max operations, or, more generally, any $T-norm$ and $S-norms$ ($T-conorms$) [6]. In the following let T denote any selected $T-norm$, and S - a $S-norm$. Let $\mathbf{G} = (\mathbf{N}, \mathbf{E}^*, \mathbf{E}^-)$ be an AND/OR/NOT causal graph. Assume that the fuzzy coefficients of fault occurrence for elements of $\mathbf{D} \cup \overline{\mathbf{D}}$ are assigned to the considered elements. The following rules define the possibility to propagate the fuzzy coefficients of faultyness upward graph $\mathbf{G}$:

- The case of an OR node: let $(n_1, n_i), \ldots, (n_{i-1}, n_i) \in \mathbf{E}^2$, and let α_j denote the fuzzy coefficients of n_j, $j = 1, 2, \ldots, i-1$. Further, let β_j denote the coefficients of fault propagation assigned to arcs. The calculated coefficient α_i of n_i is obtained as

$$\alpha_i = S(T(\alpha_1, \beta_1), T(\alpha_2, \beta_2), \ldots, T(\alpha_{i-1}, \beta_{i-1})). \tag{2}$$

- The case of an AND node: let $(n_1, n_2, \ldots n_{i-1}, n_i) \in \mathbf{E}^i$, and let α_j denote the fuzzy coefficients of n_j, $j = 1, 2, \ldots, i-1$. Further, let β_j denote the coefficients of fault propagation assigned to arcs. The calculated coefficient α_i of n_i is obtained as

$$\alpha_i = T(T(\alpha_1, \beta_1), T(\alpha_2, \beta_2), \ldots, T(\alpha_{i-1}, \beta_{i-1})) \tag{3}$$

- The case of a NOT arc: let $(n_1, n_2) \in \mathbf{E}^-$, and let α_1 denote the fuzzy coefficient of n_1; the coefficient α_2 of n_2 is calculated as $s_2 = 1 - \alpha_1$.

According to the above rules one can inductively assign fuzzy coefficients to a maximal subset of $\mathbf{N}$. This kind of extension of the basic model can be applied for three different reasons.

First, the search for diagnoses can be performed assuming the classical binary model; then at the evaluation stage, after obtaining the exact values of α_i the overall degree of failure with respect to different diagnoses can be evaluated. In this way preference among diagnoses can be established.

Second, under assumption that the degree of faultyness is evaluable (e.g. it can be estimated as a function of time, as in the case of sediment in a pipe), this approach can be used to order the search for diagnoses; by propagating the expected fuzzy values of faults, most likely symptoms (nodes) can be searched first.

Third, the whole model can be used for simulation of the influence of fuzzy faults of system components on the observed failures and estimation of expected fuzzy values of them.

5 Conclusions

In this paper the problem of fuzzy faults has been discussed. A fuzzy model of component faults and a motivational discussion have been presented. With respect to the accepted definition of causal structure in the form of AND/OR/NOT causal graphs, the rules for fault propagation have been given. Our work is related to the abductive approaches based on causal reasoning; our graph serves as a direct tool for search (backwards) of diagnosis. Thus fuzzy evaluation of diagnoses allows for rough estimation of fuzzy diagnoses and their influence on the observed fuzzy failure and for ordering among generated possible diagnoses. The complete model of the diagnosed system is not required.

Acknowledgements

The research was supported within the Grant No. 8 T11A 013 08 of KBN Formal Methods and Tools Supporting Design and Analysis of Real Time Software for Process Control (Poland).

And the research was supported by the *Comisión Interministerial de Ciencia y Tecnología* trough the TAP93-0596-C04-04 project (Spain).

References

[1] Aguilar, J. & Fuster, P. Knowledge-based approach of model-based diagnosis of complex processes, in Model-Based Fault Diagnosis, pp.

94 to 103, *Proceedings of the Workshop of Tempus Impact Jep 2011-93/3*, Zakopane, Poland, 1994.

[2] DeKleer, J. & Williams, B.C.: Diagnosing multiple faults. *Artificial Intelligence*, 1987, **32**, 97-130.

[3] Ligęza, A. & Fuster, P.: Qualitative knowledge representation and processing for causal diagnostic reasoning. Extended reasoning modes and efficiency-related issues. *Research Report of Institute of Automatics AGH*, **38**, Cracow, 1994.

[4] Mozetic, I.: A polynomial-time algorithm for model-based diagnosis, in ECAI92, (ed B. Neumann), pp. 729 to 733, *Proceedings of 10th European Conference on Artificial Intelligence*, John Wiley & Sons, 1992.

[5] Ng, H.T.: Model-based, multiple fault diagnosis of time-varying, continuous physical devices. in IEEE/90, pp. 9 to 15, *Proceedings of the IEEE*, 1990.

[6] Pedrycz, W.: *Fuzzy Control and Fuzzy Systems.* Research Studies Press Ltd., Taunton, Somerest, England, John Wiley & Sons Inc., New York, 1989.

[7] Reiter, R.: A theory of diagnosis from first principles. *Artificial Intelligence*, 1987, **32** 57-95.

[8] Saucier, G., A.Ambler & M.A.Breuer, *Knowledge Based Systems for Test and Diagnosis*, North-Holland, Amsterdam, New York, 1989.

[9] Torasso, P. and Console, L. *Diagnostic Problem Solving. Combining heuristic Approximate and Causal Reasoning.* North Oxford Academic, A Division of Kogan Page, London, 1989.

[10] Tzafestas, S. AI techniques in control: an overview, in Reprints of the IMACS, pp. 55 to 67, *Int. Symposium on AI, Expert Systems and Languages in Modelling and Simulation*, Barcelona, Spain, 1987.

[11] Tzafestas, S.G., Ata-Doss, S.A.E. & Papakonstantinou, G., Expert system methodology in process supervision and control, Chapter 11, *Knowledge-Base System Diagnosis Supervision and Control*, ed Tzafestas, S.G., pp 181-215, Plenum Press, New York and London, 1989.