Proc. of Pacific Asian Conference on Expert Systems, pp.322-329, 1997

A Model-based Diagnosis with Fault Event Models

Yoshinobu KITAMURA[*], Masahiko UEDA[1**], Mitsuru IKEDA[*],
Satoshi KOBORI[**], Osamu KAKUSHO[***] and Riichiro MIZOGUCHI[*]
[*] *I.S.I.R., Osaka Univ., 567 Japan*
[**] *Faculty of Science and Technology, Ryukoku Univ., 520-21 Japan*
[***] *Faculty of Economics and Information Science, Hyogo Univ., 675-01 Japan*

Abstract: *The causal chains between faults and symptoms include abnormal phenomena such as breakage and corrosion which cannot be represented in terms of physical parameters. Most of model-based diagnosis systems based on physical parametric models cannot cope with such phenomena, and thus the deeper causes underlying the faults remain untouched. This article describes a diagnosis method based on the fault event models, which represent generic abnormal transitions which could commonly occur in components in terms of conceptual attributes such as existence and shape. A framework and an example of the reasoning integrated with the conventional methods are presented. The integrated reasoning system can uncover the deeper causes underlying the faults detected by the conventional methods.*

1. Introduction

One of the crucial performances of diagnostic systems is to infer causal chains between faults and symptoms. Such causal chains often include abnormal phenomena which cannot be represented in terms of physical parameters. Let us consider a causal chain of a fault in a super-heater and a turbine shown in Figure 1. First, in the super-heater, "corrosion" causes the existence of some fragments in the steam. Next, the fragments move into the turbine and collide with the turbine blade, which results in "breakage" of the blade. Eventually, the revolution of the shaft decreases. If one tries to represent "breakage" in terms of physical parameters such as width and height, it is too complicated to model because the shape will change to abnormal one in various ways. To cope with such a situation, we need to abstract the parameters to conceptual attributes such as shape.

Most of the conventional model-based diagnosis systems are based on physical parametric constraints. Systems based on the intended behavior models such as [2,4,12] define a fault by negation of the model. Fault modes used in the systems such as [13] specify possible faulty states of each component. As the consequences, the reasoning results of such systems are restricted to physical parametric faults. In this example, only misbehavior of the turbine could be detected. The fault of the super-heater could not be detected, because its phenomena "corrosion" and the causal chain between it and the symptom are outside the scope. In other words, no explanation why and how the fault of the turbine has occurred is generated.

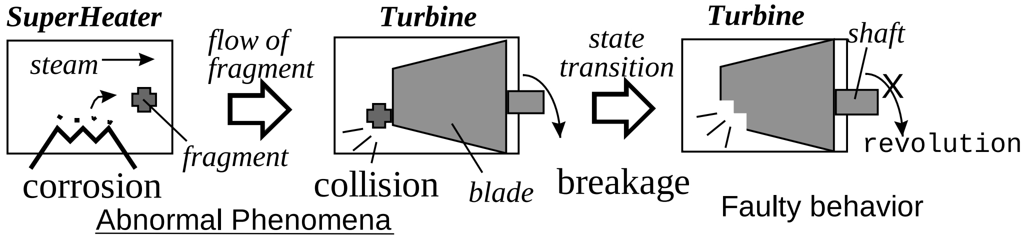

Figure 1: An example of causal chains including non-physical parametric faults

[1] Currently, Toko Seiki Co., Ltd.

On the other hand, heuristics-based diagnosis systems have a possibility to cope with non-parametric faults. However, because they represent direct symptom-fault associations, they are ad hoc and not reusable. We need to conceptualize a fault as an ontology in order to make them general and hence reusable.

The lack of abnormal phenomena model is because of the inadequate definition of the fault. The fault means not only misbehaviors as consequences of the faults but also the transitions from the intended behavior to the misbehavior. We call such transitions fault events. The fault of the target system can be viewed as the causal chains of the fault events among components. The fault event models are generalized fault events. For example, the fault event "adhesion" can be generally defined as the phenomena in which the existence of some small fragments in the neighborhood of a component causes the change of the shape of the component.

In this article, a diagnosis method based on the fault event models is proposed. Given observed symptoms, the method provides causal chains of fault events which enable explanations why and how the component has reached the faulty state. The reasoning based on the fault event model is complementary to the conventional methods whose basis is an intended behavior model. This article shows a framework of the reasoning system which consists of the conventional method and the reasoning based on the fault event models. The integrated reasoning system can uncover the deeper causes underlying the faults detected by the conventional methods.

2. Fault Events

2.1 A Fault Event

A fault event of a component is defined as a state transition from the normal state to a faulty state of the component. The state transitions are irreversible. In other words, the faulty state does not return to the intended state even if the influences which caused the fault event disappear. In cases of reversible transitions such as some kinds of thermal expansion, the abnormal states are called semi-faulty states.

A fault event consists of a causative state, a process and a resultant state. The causative state represents the influence which causes the state transition. The process represents the state transition between the causative state and the resultant state. The resultant state represents the faulty state caused by the state transition. For example, the causative state of the fault event "aberration" is the state in which the pressure exerted on the component is greater than the intended one. It causes the process, that is, the component moves from the intended position to the abnormal one. Then, a resultant state appears in which the position of the component is abnormal.

2.2 Causal Chains of Fault Events

A target system can be viewed as a set of components. The fault of the system can be represented by the causal chains of fault events which occur on each component. Figure 2 shows an example of a causal chain of fault events. The system consists of the three components, c_1, c_2 and c_3. The abnormal influence a_1 is propagated to the component c_1 and the state of c_1 becomes the state s_1. s_1 causes a fault event on c_1 in which the causative state s_1 changes into the faulty state s_2. Next, the influence of s_2 is propagated to c_2. In this case, no state transition occurs at c_2, and c_2 outputs the abnormal influence a_3. When a_3 is propagated to c_3, a normal state of c_3 changes into the faulty state s_5.

A fault in a system is represented by a fault event and a faulty component. In this example, the system has two faults and the faulty components are c_1 and c_3. A cause of a fault is defined as the abnormal influence to the causative state of a fault event. The abnormal influence of a fault event is called the *relative cause* of the fault event and the resultant state. The *absolute cause*

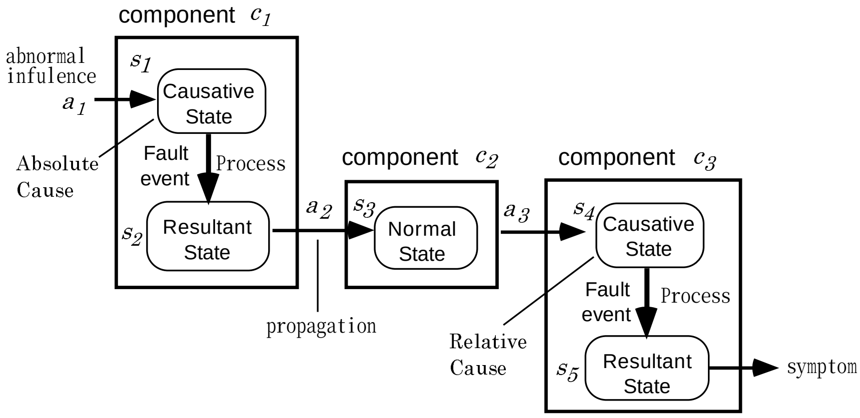

Figure 2: An example of a causal chain of fault events

represents such an abnormal influence that no fault event causes it in the target system. Such influences represent factors external to the target system under consideration such as abnormality of the environment of the system. In the example, the absolute cause is the abnormal influence a_1.

The diagnosis task can be defined as the process which detects all plausible causes of the fault events which have caused the given symptoms. There are, however, some cases where no cause of a fault is found by the diagnostic process. For example, a target system which has no fault event but an abnormal input influence also has abnormal symptoms.

3. Fault Event Models

A fault event model represents a generic fault event. It consists of a term representing the fault event, a causative state, a process and a resultant state. It means that if there is a component matching the condition of the causative state, then the process probably occurs, as the consequence, the status of the component changes into the resultant state. In other words, a fault event model represents general faulty phenomena which could occur commonly in components. The term is a noun representing the fault event for humans.

The causative state represents conditions for the fault event. A state is a set of four-tuples <object, attribute, relation, value>. The object is “component” or “environment”. The component can match all the components in the target system. The environment represents the neighborhood of a component. There are some conceptual attributes which do not directly correspond to physically measurable attributes. For example, the attribute shape of a component takes either “normal” or “abnormal” as its value. On the other hand, an environment has attributes such as temperature, pressure and fragment. The last attribute takes either “exists” or “not-exist” as its value which represents some small fragments of substance (e.g., dust) exist (or not) in the environment of a component.

The causative states are categorized into two types, namely necessary states and accelerative states. The fault event will not occur without the former states. The latter states just accelerate the event. The resultant states are categorized into two types, direct states and associated states. The latter states are not included in the meaning of the term of the fault event.

A process represents the state transition from the causative state to the resultant state. A description of a process consists of an object, a term and a time-interval. The term is a verb representing the process for humans. Most of them are identical with the verbalized terms of the noun term representing the fault events. The time-interval takes one of long, short and instant, which represents the time interval necessary for the event.

Table 1: Examples of the fault event models

Vocabulary	Causative States	Process	Resultant States
aberration	env:pressure > normal comp:shape = abnormal	comp:move (short)	comp:position = abnormal
abrasion	[A]comp:position=abnormal	comp:abrade (long)	comp:shape = abnormal [A]env:temperature > normal [A]env:fragment = exists
adhesion	env:fragment = exists	fragment: adhere (short)	comp:fragment = exists [A]comp:shape = abnormal [A]comp:fiction-resist > normal
breakage	comp:strength < normal or env:pressure > normal	comp:break (short)	comp:shape = abnormal [A]env:fragment = exists
collision	env:fragment = exists	fragment:coll ide(instant)	env:pressure > normal
corrosion	[A]env:fragment = exists	fragment:corr ode (long)	comp:strength < normal [A]comp:shape = abnormal
hit	comp:position = abnormal	comp:hit (instant)	env:pressure > normal
invasion	comp:shape = abnormal	fragment:inva de(instant)	in-env:fragment = exists

We have described 31 fault event models. Table 1 shows a sub-set of them. “object:attribute” denotes the attribute of the object. The “comp” and “env” are the abbreviation for the component and the environment, respectively. The mark [A] denotes accelerative or associative. See “adhesion” in Table 1, for example. When some fragments exist in the neighborhood of a component, the fragments probably adhere to the component in the short term. As associative consequences of the process, the shape of the component will be changed to abnormal and the resistance for friction becomes greater than that of normal.

4. Reasoning Based on the Fault Event Models

The fault event models make the reasoning on the causal chains of the fault events. There are two reasoning processes, the retrospective reasoning and the prospective reasoning. The former generates plausible causes which have probably caused the given state. The latter generates all resultant states to be caused by the given causative state.

Each reasoning process has two steps, that is, the matching step and the evaluation step. In the retrospective reasoning, the matching step searches for the fault events which have the resultant state matching the current state. The fault events found are events which have probably caused the current state and thus they are a part of plausible causal chains for the current state. The causative states of them are the relative causes of the current state. In cases where there are more than one plausible fault events for a current state, the relations among them are OR-relationship. In the evaluation step, the engine views the causative states of the events as new current states. Then, the further matching step is invoked for the new current states in order to detect deeper causes.

In the case of the prospective reasoning, the reasoning is done in the reverse direction. The engine searches for the events that have a causative state which matches the given state. Next, the evaluation step generates new current states according to the description of the resultant state of the events.

There are cases where no event is found in the matching step, because the attributes of the fault events are described in different grain sizes for generality. In such cases, the attributes are generalized according to the hierarchy of the attributes of the fault event. The knowledge

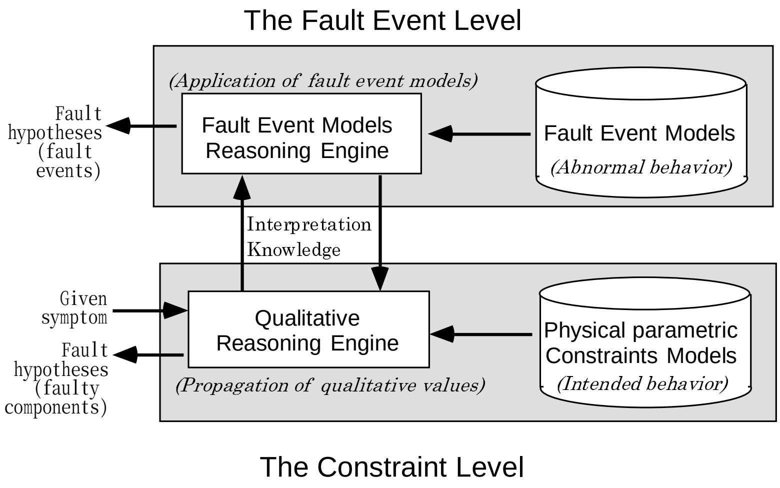

Figure 3: The framework of the integrated diagnosis method

representing such hierarchical relations between attributes is called the hierarchical attribute knowledge.

5. Integrated Diagnostic Method

We propose an integrated diagnostic method which has two reasoning methods. One is the method based on the fault event models discussed thus far, called the fault event level. The other is an ordinary model-based method using intended behavior models in terms of qualitative constraints, called the constraint level. Figure 3 shows the framework of the integrated diagnosis method.

5.1 The constraint level

At the constraint level, the knowledge base contains qualitative component models representing their intended behaviors. A component model consists of a set of physical parameters, qualitative constraints over parameters, ports for connections and causal properties of the parameters. The “physical parameters” means attributes which directly correspond to physically measurable quantities. A parameter takes one of the three qualitative values related to the deviation from a normal value. [+]([-]) represents a quantity greater(less) than the normal value, i.e. abnormal values. [0] represents a quantity equal to the normal value.

The reasoning engine at the constraint level [9,10] is based on qualitative reasoning techniques[3, 6, 7, 8]. The reasoning method is categorized as a type of the reasoning method proposed by de Kleer and Brown[3]. On the basis of the device modeling, it has two reasoning processes, that is, the intra-component reasoning and the inter-component reasoning. It can infer finer-grained causal ordering among physical phenomena in transition states and feedback loops on the basis of seven units of time resolution. For more detail, see [9, 10].

5.2 The Reasoning Processes

Our diagnostic method consists of two processes, the fault-hypotheses generation and the fault-hypotheses verification. The generation process derives plausible causes of the fault covering the given symptom. The verification process generates all the possible symptoms to be caused by the fault-hypotheses and checks predicted values against actual values. In this article, we concentrate on the fault-hypotheses generation process.

The integrated fault-hypotheses generation process consists of two level reasoning, the constraint level and the fault event level. Given the physically parametric symptom, the qualitative reasoning engine at the constraint level retrospectively propagates abnormal values according to the intended behavior model. First, the given symptom is propagated to the other parameters in the component. Next, the abnormal values are propagated to the neighboring components. The reasoning results at the constraint level are a set of abnormal parameters whose values have no deeper cause at the level, that is, the absolute causes at the level. Then, the causes are converted to the fault event model according to the interpretation knowledge in the subsection below. At the fault event level, causal chains of fault events are derived on the basis of the fault event models. In the case where causative states at the fault event level can be converted to the constraint level, the deeper causes are derived at the constraint level.

5.3 The Interpretation Knowledge

The conversions between physical parameters and faulty states are done according to the interpretation knowledge. The following relations are examples of the interpretation knowledge.

(1) IF humidity = [+] THEN environment:water-vapor = exists

(2) IF turbine-efficiency = [-] THEN component:shape = abnormal

The first one represents the equivalent relation between the physical parameter "humidity" and the existence of water vapor. This is independent of the component. On the other hand, the second one depends on the turbine and represents a causal relation.

6. An Example of Reasoning

In this section, we show an example of the integrated reasoning. The target component is a steam turbine of a power plant. The turbine is connected to a generator and a super-heater. The symptom given in this example is "the output power of the generator is lower than the normal value".

6.1 The Constraint Level

First, the engine focuses on the generator whose output is identical with the symptom and generates a relative cause, that is, the revolution of the shaft of the turbine is lower.

Next, the engine reasons about the turbine. Given the revolution of the shaft is lower, the following causes are generated:

- **Absolute Cause:** turbine-efficiency = [-]
- **Relative Cause:** flow-rate = [-], heat-inflow = [-], inlet-pressure = [-], outlet-pressure = [+]

The relative causes are associated with other components. Then, the abnormal values are propagated to the super-heater which is an upper component.

On the other hand, no deeper cause of the decreases of the turbine efficiency is derived at the constraint level. Then the engine converts the state to the fault event models according to the interpretation knowledge. In the paragraph below, we show an example in the case of an abnormal shape converted from the decreases of the efficiency.

6.2 The Fault Event Level

Figure 4 shows a part of a retrospective reasoning at the fault event level which starts with the abnormal shape. First, the engine generates the fault events which have resultant states matching "component:shape = abnormal". For example, "adhesion" is derived as one of the possible fault events. It means that the existence of some fragments has caused "adhesion" and then the shape becomes abnormal. Next, the engine views the causative states of the generated events as the resultant states and searches for the fault events to cause them. In this example, the existence of some fragments which is the causative state of the adhesion may be caused by the fault events

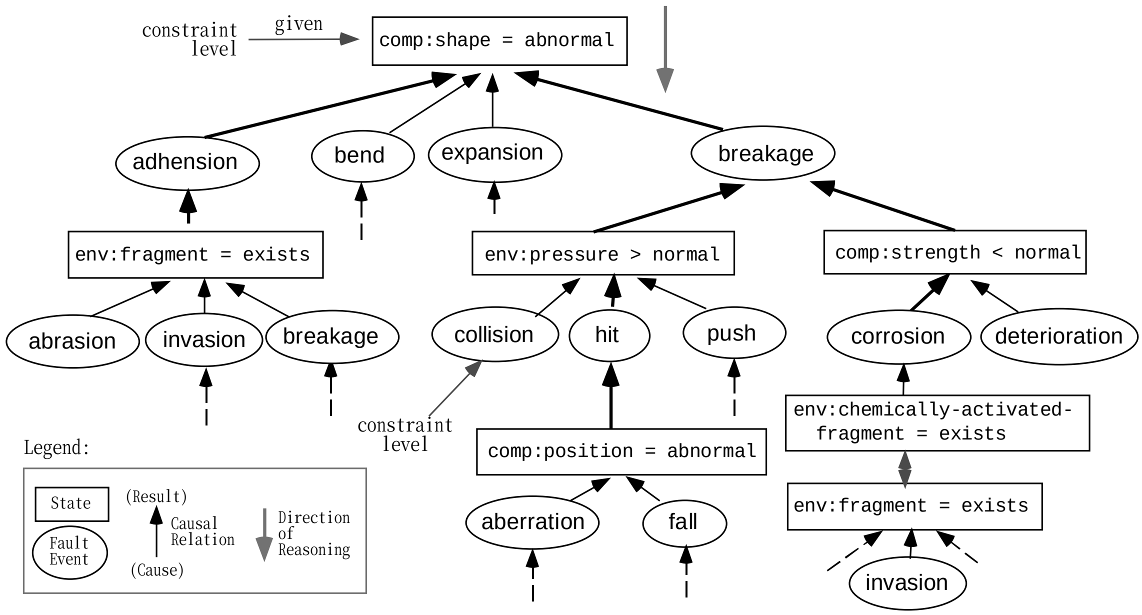

Figure 4: The causal chains generated for the abnormal shape

such as "invasion" and "abrasion". In the case where the turbine is new, the "abrasion" could not be plausible according to its description of the necessary time interval.

In the case of "breakage", the reasoner searches for the fault events which have the resultant state which matches "component:strength < normal" or "environment:pressure > normal". Then "corrosion", "collision" and so on are derived. Next, the engine searches for the event matching "environment:chemically-activated-fragment = exists" which is the causative state of the "corrosion". As there is no further event, the attribute is generalized to "fragment". Then, "invasion" of chemically-activated-fragments are generated.

The causative state of "collision" is an existence of fragments. According to the interpretation knowledge that "IF humidity > normal THEN environment:water-vapor = exists", the engine converts the state to the physical parameter. At the constraint level, the fault that outlet temperature of the super-heater is lower than that of normal is detected.

6.3 The Reasoning Results

Given the abnormal output power of the generator, 3 fault hypotheses at the constraint level and 28 fault hypotheses for the turbine at the fault event level are generated. Then, they are verified by the verification process. The reasoning results for the turbine cover all the faults of a turbine listed in a technical book on power plants.

7. Related work and Discussion

Our method based on the fault event models has wider scope of detectable faults than the conventional model-based systems[2,4,6,12,13]. Firstly, fault event models are represented in terms of conceptual attributes such as shape and existence which enable the reasoner to cope with complex phenomena, such as breakage, which is difficult for conventional systems based on physical parameters. Secondly, it uncovers unintended relations (interactions) underlying the intended behavior model. The faulty component which is the deeper cause of such a faulty component that the conventional systems could detect can be detected. On the other hand, the advantage of our method over the conventional heuristics-based systems is reusability. The fault event models represent primitive causal relations from the viewpoint of the fault event and they are described in terms of generalized parameters. Thus, they are generic and reusable. Composing the primitive fault events, the reasoning engine can generate symptom-fault associations which traditionally represented as the empirical heuristic rules.

The fault event models are, however, generic and coarser than the constraint models. The competence of this method with respect to resolution is not high. Our method enumerates fault events which have possibly caused the symptom. Therefore, this method is complementary to the conventional method.

In [1], abnormal inter-component interactions called hidden interactions are discussed. The hidden interaction models for the fluid domain such as leak can be viewed as a kind of generic fault events between components. Our fault event models include such cases as shown in Table 1.

The QP theory[5] introduces the concept of process. The fault event can be viewed as a special kind of the process. The fault events can represent discrete phenomena such as breakage. Moreover, our method deals with accelerative states and associative states for diagnosis.

Similar formulation is found in [11] where knowledge called naive physics represents general causal relations between naive concepts and phenomena causing them. For example, exist of liquid is caused by flow of liquid, melting of solid or condensation of gas.

8. Summary

A diagnosis based on the fault event models has been proposed. The integrated reasoning system has wider scope of detectable faults using generic fault event models.

Currently, the set of the fault event models are built on the plant domain including pumps and turbines. Extension of our work for other domains such as electronics and mechanics remains as future work. Furthermore, in order to develop a framework to diagnose unexpected faults, an investigation on control knowledge representing probabilities of fault events is also in progress.

Acknowledgments:
The authors would like to thank Takashi Washio and Munehiko Sasajima for their valuable comments.

References

[1] Bottcher, C., No Fault in Structure? How to Diagnose Hidden Interactions,Proc. of IJCAI-95,1728-1733, 1995
[2] Davis, R., Diagnostic Reasoning Based on Structure and Behavior, Artificial Intelligence, 24: 347-410, 1984.
[3] de Kleer, J., Brown, J. S., A Qualitative Physics Based on Confluences, Artificial Intelligence, 24: 7-83, 1984
[4] de Kleer, J., and Williams, B. C., Diagnosing Multiple Faults, Artificial Intelligence, 32: 97-130, 1987.
[5] Forbus, K. D., Qualitative Process Theory. Artificial Intelligence 24: 85-168, 1984
[6] Hamscher, W., et al. (Eds.), Readings in Model-based Diagnosis. Morgan Kaufmann, 1984.
[7] Iwasaki, Y., and Simon, H. A., Causality in Device Behavior, Artificial Intelligence, 29: 3-32, 1986
[8] Kuipers, B. J., Qualitative Reasoning, MIT Press, 1994.
[9] Kitamura, Y. et al., A Method of Qualitative Reasoning for Model-based Problem Solving and its Application to a Nuclear Plant, Expert Systems with Application, 10(3/4): 441-448, 1996.
[10] Kitamura, Y., Ikeda, M., and Mizoguch, R., A Qualitative Reasoning based on an Ontology of Fluid Systems and its Evaluation on a Power Plant. In Proc. of PRICAI-96, 288-299, 1996.
[11] Purna, Y. W., and Yamaguchi, T., Generating and Testing Fault Hypotheses with MODEST. In Proc. the Third World Congress on Expert Systems, 2:954-961, 1996.
[12] Reiter, R., A Theory of Diagnosis from First Principles, Artificial Intelligence, 32: 57-96, 1987.
[13] Struss, P., and Dressler, O., "Physical negation" - Integrating Fault Models into the General Diagnostic Engine. In Proc. of IJCAI-89, 1318-1323, 1989.