PERGAMON

Chemical Engineering Science 57 (2002) 2971–2985

Chemical Engineering Science

www.elsevier.com/locate/ces

A fuzzy diagnosis approach using dynamic fault trees

Sheng-Yung Chang, Cheng-Ren Lin, Chuei-Tin Chang *

Department of Chemical Engineering, National Cheng Kung University, Tainan 70101, Taiwan, Republic of China

Received 23 October 2001; received in revised form 2 May 2002; accepted 8 May 2002

Abstract

By incorporating digraph models, fault trees and fuzzy inference mechanisms in a unified framework, a novel approach for fault diagnosis is developed in this work. To relieve the on-line computation load, the fault origins considered in diagnosis are limited to the basic events in the cut sets of a given fault tree. The symptom occurrence order associated with each root cause is derived from system digraph with the qualitative simulation techniques. The implied candidate patterns are enumerated according to two proposed theorems and then encoded in the inference system with IF–THEN rules. The simulation results show that the proposed approach is not only feasible but also capable of identifying the most likely cause(s) of a hazardous event at the earliest possible time.

Keywords: Fault diagnosis; Fault tree; Digraph; Symptom occurrence order; Candidate pattern; Fuzzy logic

1. Introduction

Owing to the continuous increase in scale and complexity of the modern chemical processes, numerous computer-aided fault diagnosis methods have been developed in recent years to assist operator to locate fault origins which may lead to catastrophic consequences. A wide variety of different approaches were discussed in the literature, e.g. the expert systems (Rich & Venkatasubramanian, 1987; Petti, Kleni, & Dhurijati, 1991), the state observers (Chang, Mah, & Tsai, 1993; Chang & Chen, 1995; Chang & Hwang, 1998a, b) the neural networks (Venkatasubramanian & Chan, 1989; Watanabe, Matsuura, Kuboto, & Himmelblau, 1989; Venkatasubramanian, Vaidyanathan, & Yamamoto, 1990; Hoskins, Kalivur, & Himmeblau, 1991; Tsai, Chang, & Chen, 1996) and the signed directed graphs (SDG) (Iri, Aoki, O'Shima, & Matsuyama, 1979; Shiozaki, Matsuyama, O'Shima, & Iri, 1985; Kramer & Palowitch, 1987), etc. Ulerich and Powers (1988) reported the first attempt to perform fault diagnosis on the basis of fault trees. The most significant advantage of their approach is that the candidates of fault identification are restricted to only the causes of one or more given top events and, consequently, the diagnosis procedure can be greatly simplified. However, it should be noted that only a conceptual framework was proposed in this study. There is thus a need to develop systematic algorithms for its implementation. In addition, only the eventual symptoms in process measurements were utilized in Ulerich and Powers (1988). It is our belief that the *occurrence order* of these on-line symptoms should also be incorporated in an improved strategy for locating the fault origins.

In this work, the fuzzy set theory (Ross, 1995) is adopted to facilitate construction of a fault-tree based diagnosis system. This selection is mainly due to the following reasons:

- The fault tree is basically a *qualitative* model of all failure mechanisms leading to an eventual top event. Therefore, the on-line symptoms observed in process measurements must be interpreted accordingly in a consistent manner. In a traditional fault tree, the abnormal process conditions are often classified with linguistic terms, e.g. "too high" or "too low", and they can be appropriately characterized with the fuzzy membership functions.
- As suggested by Ulerich and Powers (1988), the occurrence probabilities of top event and also its causes should be computed in real time. To accomplish these tasks, it is necessary to produce the on-line estimates of (1) the reliability (or availability) of each hardware item in the given system and (2) the frequency of every significant external disturbance. In the former case, although it is possible to extract some of the reliability parameters, e.g. the failure rate and mean time to repair, from statistical data (if available), their accuracy is highly questionable.

* Corresponding author. Tel.: +886-6-275-7575 ext. 62663; fax: +886-6-234-4496.

E-mail address: ctchang@mail.ncku.edu.tw (C.-T. Chang).

0009-2509/02/$ - see front matter
PII: S0009-2509(02)00178-1

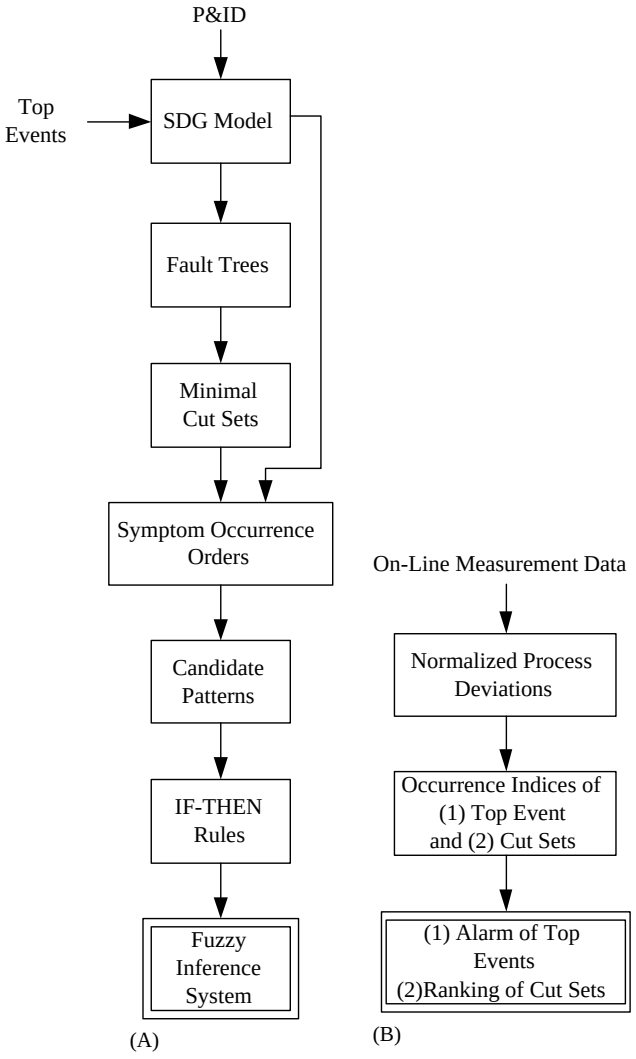

Fig. 1. The proposed on-line fault diagnosis procedure: (A) off-line preparation stage; (B) on-line implementation stage.

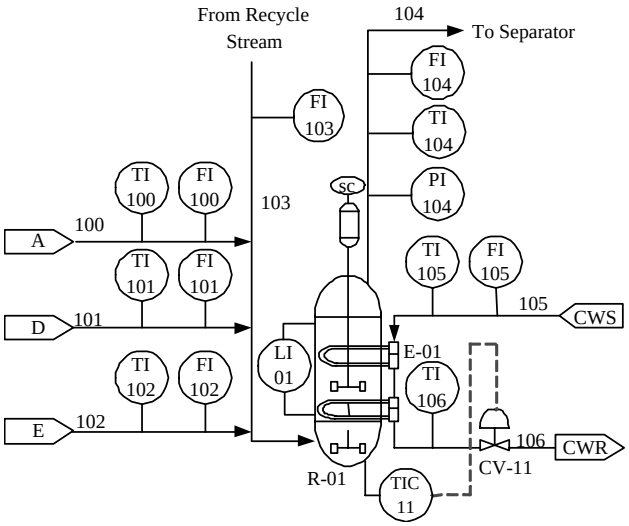

Fig. 2. Reactor section of TE process.

On the other hand, the latter information can only be obtained from operation experience and its nature is often imprecise. Since a rigorous statistical treatment is really not practical under this circumstances, a set of fuzzy occurrence measures are adopted instead in this study as the conclusions of fault diagnosis.

The proposed fault diagnosis procedure can be implemented in two stages: the off-line preparation stage and the on-line implementation stage (Fig. 1). In the former case, a SDG system model is first constructed and the fault trees corresponding to the given top events are then synthesized according to the Lapp-and-Powers algorithm (1977). The symptom occurrence order caused by the basic events in each cut set can be easily determined with the qualitative simulation techniques (Chang & Hwang, 1992) on the basis of the SDG model. In addition, two theorems are developed in this study to facilitate enumeration of all possible symptom patterns that may be observed at any time during operation. These patterns are then translated into a set of IF–THEN fuzzy inference rules for assessing the occurrence possibilities of the basic events in every cut set and also the top events. In the next stage, the on-line measurement data are first normalized. These normalized values are used as the inputs to a fuzzy inference system for computing the occurrence indices of top events and cut sets in real time. The Tennessee Eastman (TE) process simulation program (Downs & Vogel, 1993) is chosen as the platform for verifying the feasibility of the proposed approach. Extensive simulation results of the reactor section, the stripper section and also the entire process are presented at the end of this paper.

2. Qualitative system models

The products of the first three steps in the off-line preparation stage, i.e. the SDG, the fault trees and the minimal cut sets, can be considered as qualitative system models in different formats. It should be noted that the SDG construction methods have already been discussed extensively in the literature, e.g. Iri et al. (1979) This study follows basically the conventional approach to build digraph models. Before developing these digraphs, the *scope of fault diagnosis* (SCFD) must be established first so that the resulting SDG contains the nodes corresponding to the given top event and the fault origins *of interest*. For example, let us consider a SCFD restricted to the reactor section in TE process (see Fig. 2). The corresponding SDG model is presented in Fig. 3. The nodes in this digraph are mainly associated with the process variables, the measurement signals and the control signals within SCFD. Generally speaking, the process variables considered in this study are pressure, temperature, mass flow rate and liquid level.

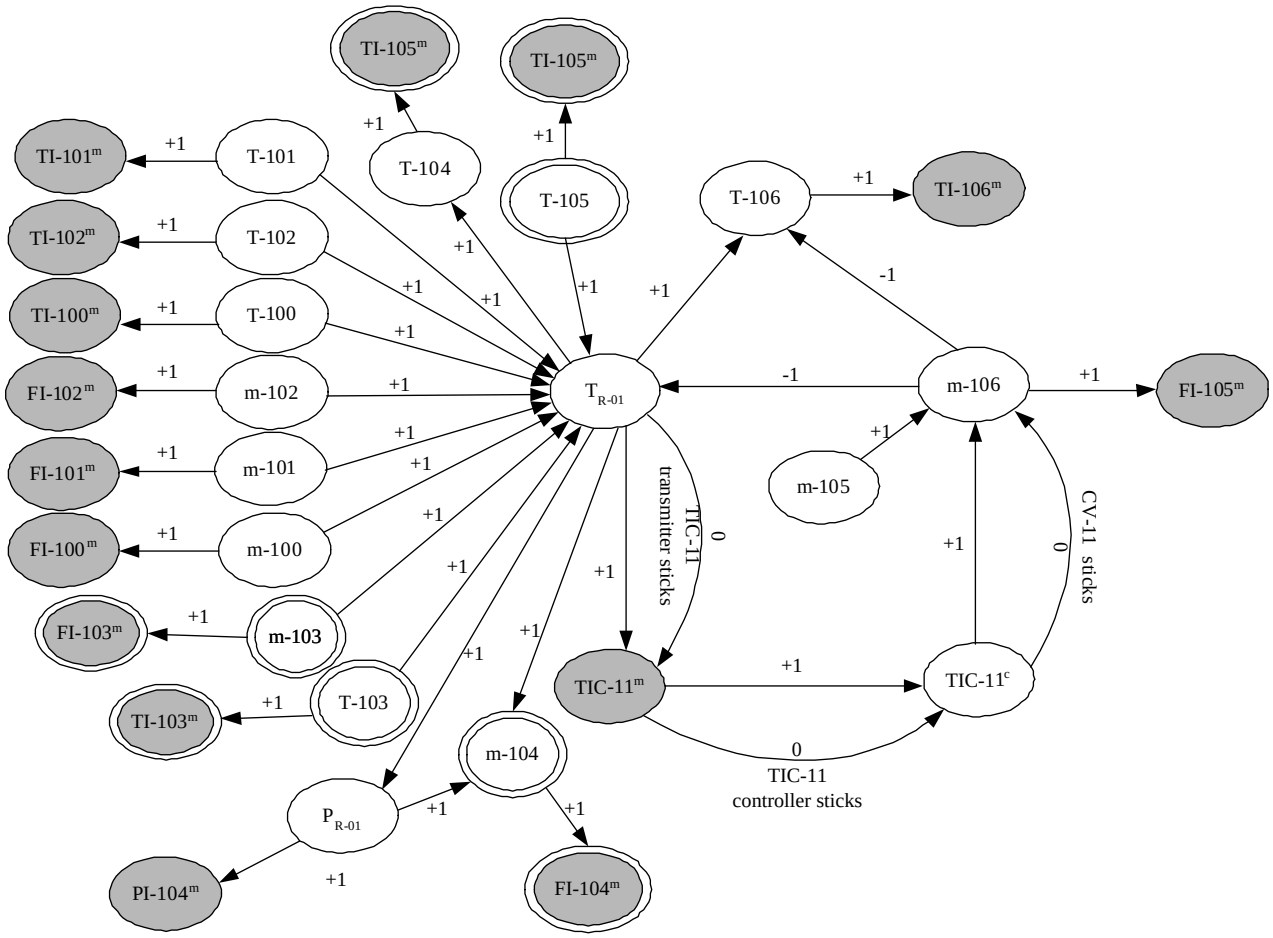

Fig. 3. SDG model of reactor section.

They are represented respectively in the digraph by combining an initial letter, i.e. P, T, m or L, with a numerical label or a subscript. The numerical labels are basically the stream numbers in the original process flow diagram, while the subscripts are always associated with process units. Thus, T-101 denotes the temperature of stream 101 and $T_{\text{R-01}}$ represents the temperature of reactor R-01. An on-line measurement instrument of a process variable is usually tagged in the process flow diagram by attaching the corresponding instrument number to two initial letters indicating the physical nature of the measured variable, i.e. PI, TI, FI or LI. Similarly, the controllers are usually identified according to its controlled variable and an equipment number. For example, TIC-11 is the tag number of the temperature controller for reactor R-01. In digraph, a measurement signal is represented by the tag of the corresponding sensor or controller with a superscript "m". For convenience in tracking the symptoms of fault propagation, only the nodes representing measurement signals are shaded in digraphs throughout this paper. The same approach is adopted to denote the controller outputs to drive the control valves, i.e. they are represented by the controller tags in process flow diagram with a superscript "c". Finally, notice that the process variables associated with the recycle stream are marked with double circles.

The conventional Lapp-and-Powers algorithm (1977) is adopted in this study to synthesize the fault trees for given top events. Let us select the event "high temperature in the reactor", i.e. $T_{\text{R-01}}(+1)$, as one of the top events for diagnosis purpose. The corresponding fault tree can be easily derived from the SDG in Fig. 3. Since the present SCFD is limited to the reactor section, fault origins originated from other sections of the TE process should be neglected. In other words, the disturbances originated from the double-circled nodes are not considered in constructing this fault tree. For the sake of brevity, the resulting fault tree is not presented in this paper. Instead, a complete list of its minimal cut sets is provided in Table 1. Notice that each row of this table represents a cut set. For example, if the basic events in the 11th row, i.e. "an increase in feed rate of reactant D" (m-101(+1)) and "cooling water control valve sticks" (CV-11 sticks), exist simultaneously, then the top event is bound to occur.

Table 1
The minimum cut sets (MCS) of the fault tree with top event $T_{\text{R-01}}(+1)$

MCS No.	Base event	
1	m-100(+10)	
2	m-101(+10)	
3	m-102(+10)	
4	T-100(+10)	
5	T-101(+10)	
6	T-102(+10)	
7	T-105(+10)	
8	m-100(+1)	CV-11 valve sticks
9	m-100(+1)	TI-01^c sticks
10	m-100(+1)	TI-01^m sticks
11	m-101(+1)	CV-11 valve sticks
12	m-101(+1)	TI-01^c sticks
13	m-101(+1)	TI-01^m sticks
14	m-102(+1)	CV-11 valve sticks
15	m-102(+1)	TI-01^c sticks
16	m-102(+1)	TI-01^m sticks
17	T-100(+1)	CV-11 valve sticks
18	T-100(+1)	TI-01^c sticks
19	T-100(+1)	TI-01^m sticks
20	T-101(+1)	CV-11 valve sticks
21	T-101(+1)	TI-01^c sticks
22	T-101(+1)	TI-01^m sticks
23	T-102(+1)	CV-11 valve sticks
24	T-102(+1)	TI-01^c sticks
25	T-102(+1)	TI-01^m sticks
26	T-105(+1)	CV-11 valve sticks
27	T-105(+1)	TI-01^c sticks
28	T-105(+1)	TI-01^m sticks
29	m-106(−10)	
30	m-106(−1)	TI-01^c sticks
31	m-106(−1)	TI-01^m sticks

3. Symptom occurrence order

The effects of base event(s) in a cut set usually propagate throughout the entire system sequentially. In general, a series of intermediate events may occur before the top event. Since the performance of a diagnosis scheme should be evaluated not only in terms of its correctness but also its timeliness, it is the intention of this research to develop a fault identification procedure taking both the eventual symptoms and also their *occurrence order* into consideration. To identify this *symptom occurrence order* (SOO) associated with a given fault origin, the following operations should be performed on the digraph:

(1) apply the techniques of qualitative simulation (Chang & Hwang, 1992) to identify the *fault propagation paths* (FPPs),
(2) merge every pair of measured variable and its measurement signal in the FPPs, and then
(3) eliminate the nodes representing the unmeasured variables.

Since only the nodes representing measurement signals remain in the resulting symptom occurrence order, the set containing all such nodes is referred to as the *range of influence* (ROI) of the given fault origin. For example, let us again consider the cut set {m-101(+1), CV-11 sticks} in the reactor section of TE process. Notice that m-101 is a *primal node*, i.e. a node without inputs. In fact, every cut set contains at least one event which is associated with a primal node. As mentioned above, the first step to identify the corresponding symptom occurrence order is to determine the FPPs with qualitative simulation techniques. A brief description of the required computation procedure is provided below.

It should be noted first that a digraph model explicitly describes the qualitative cause–effect relationships between deviations in process variables (represented by 0, ±1 and ±10) and component failures (represented by 0, 1 and 10). The effects of a disturbance originated from a primal node can thus be determined by first assigning a nonzero value (±1 or ±10) to the corresponding node variable, i.e. m-101, and then evaluating the values of all other affected variables. In a simple digraph, any of these variables can normally be obtained by multiplying its input value(s) with the corresponding edge gain(s). Since in this case an equipment failure, i.e. CV-11 sticks, occurs at the same time, it is thus necessary to remove the corresponding edge between TIC-11^c and m-106 before carrying out the above computation. The resulting fault propagation paths can be found in Fig. 4. The corresponding SOO can be produced by merging each pair of nodes representing the measured variable and its measurement signal in the FPPs, i.e. (m-101, FI-101^m), ($T_{\text{R-01}}$, TIC-01^m), (T-106, TI-106^m) and (m-106, FI-105^m), and then eliminating the unmeasured node TIC-11^c (see Fig. 5). The ROI in this case can be identified accordingly, i.e.

$$\mathbf{ROI}_{11} = \{\text{FI-101}^m, \text{TIC-11}^m, \text{TI-106}^m, \text{FI-105}^m\}, \qquad (1)$$

where $\mathbf{ROI}_{11}$ denotes the ROI of cut set 11 in Table 1.

4. Enumeration of candidate patterns

If all symptoms in a SOO can be observed on-line, then it is certainly reasonable to confirm the existence of corresponding fault origin(s). However, it is also possible to find that these symptoms are only partially developed during the incipient period of an eventual system hazard and, further, their pattern may vary at different times during operation. To facilitate later discussions, let us define the collection of on-line symptoms at any time after the introduction of basic event(s) in a cut set as a *candidate pattern*. It is obvious that any candidate pattern can be considered as an evidence for fault identification with a degree of confidence. Thus, it is important to enumerate all possible candidate patterns and evaluate their respective significance in the off-line preparation stage.

Let us first consider the simplest SOO, i.e. a single path. The total number of *candidate patterns* in this case can be determined with the following theorem:

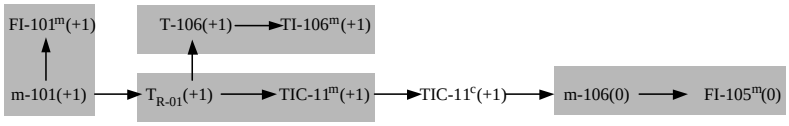

Fig. 4. The fault propagation paths (FPPs) of $\{m\text{-}101(+1), \text{CV-11 sticks}\}$.

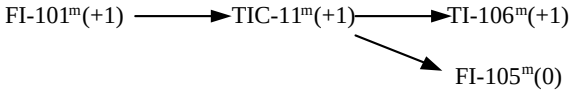

Fig. 5. The symptom occurrence order (SOO) of $\{m\text{-}101(+1), \text{CV-11 sticks}\}$.

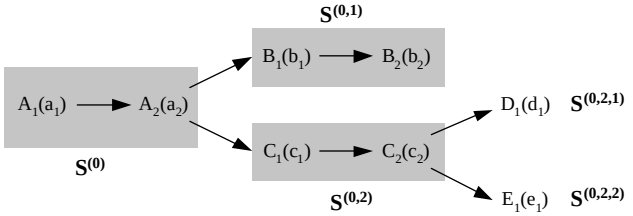

Fig. 6. The tree-shape symptom occurrence order in Example 2.

Table 2
The candidate patterns in Example 1

No.	Measurements		
	m_1	m_2	m_3
1	0	0	0
2	δ_1	0	0
3	δ_1	δ_2	0
4	δ_1	δ_2	δ_3

Theorem 1. *Let* $\mathbf{M}$ *be the set of all measured variables and* $\mathbf{X} = \{-1, +1\}$. *If* $\mathbf{S}(n) = m_1(\delta_1) \rightarrow m_2(\delta_2) \rightarrow \cdots \rightarrow m_n(\delta_n)$ (*where* $\forall m_i \in \mathbf{M}$ and $\forall \delta_i \in \mathbf{X}$) *denotes a* single-path *symptom occurrence order caused by a given fault origin, then the total number of* candidate patterns *is* $n + 1$.

The proof of this theorem is given in Appendix A. For illustration purpose, a simple example is provided below:

Example 1. Consider a single-path SOO: $m_1(\delta_1) \rightarrow m_2(\delta_2) \rightarrow m_3(\delta_3)$, $\forall \delta_i \in \{-1, +1\}$ and $i = 1, 2, 3$. All possible candidate patterns in this case can be found in Table 2. It is clear that Theorem 1 is correct here since the number of symptoms in SOO is 3 and the number of patterns is 4.

In general, the SOO assumes the form of a tree. The total number of candidate patterns in this case can be computed with the generalized version of the previous theorem.

Theorem 2. *Let* $\mathbf{M}$ *be the set of all measured variables and* $\mathbf{X} = \{-1, +1\}$. *If, in a tree-shape symptom occurrence order,* $\mathbf{S}^{(0)}(n_0)$ *denotes the initial path of length* n_0, $\mathbf{S}^{(0,i)}(n_{0,i})$ $(i = 1, 2, \ldots, N_0)$ *denotes the ith branch path of length* $n_{0,i}$ *connecting to the end of* $\mathbf{S}^{(0)}(n_0)$, $\mathbf{S}^{(0,i,j)}(n_{0,i,j})$ $(j = 1, 2, \ldots, N_{0,i})$ *denotes the jth branch path of length* $n_{0,i,j}$ *connecting to the end of* $\mathbf{S}^{(0,i)}(n_{0,i})$, *etc., then the total number of candidate patterns* N_{CP} *can be computed recursively according to*

$$N_{CP} = n_0 + \prod_{i=1}^{N_0} \left\{ n_{0,i} + \prod_{j=1}^{N_{0,i}} \left[n_{0,i,j} + \prod_{k=1}^{N_{0,i,j}} (n_{0,i,j,k} + \cdots) \right] \right\}, \tag{2}$$

where the convention $\prod_{l=1}^{0}(\bullet) = 1$ *must be followed if no further branches are attached to the branch path* $\mathbf{S}^{(0,i',j',\ldots)}(n_{0,i',j',\ldots})$, i.e. $N_{0,i',j',\ldots} = 0$.

Notice that Theorem 1 is simply a special case of Theorem 2 when $N_0 = 0$. The proof of this theorem is given in Appendix B. To demonstrate its correctness, let us consider another example:

Example 2. Consider the tree-shape SOO presented in Fig. 6. The total number of candidate patterns can be computed accordingly, i.e. $N_{CP} = 2 + \{(2+1) \times ([2 + (1+1) \times (1+1)]\} = 20$. Notice that the number of candidate patterns associated with each branch path can be determined individually by Theorem 1. The corresponding patterns can be expressed with the following logic statements:

$$\mathbf{P}^{(0)} = (A_1(a_1) \wedge A_2(a_2)) \vee (A_1(a_1) \wedge A_2(0)) \vee (A_1(0) \wedge A_2(0)),$$

$$\mathbf{P}^{(0,1)} = (B_1(b_1) \wedge B_2(b_2)) \vee (B_1(b_1) \wedge B_2(0)) \vee (B_1(0) \wedge B_2(0)),$$

$$\mathbf{P}^{(0,2)} = (C_1(c_1) \wedge C_2(c_2)) \vee (C_1(c_1) \wedge C_2(0)) \vee (C_1(0) \wedge C_2(0)),$$

$$\mathbf{P}^{(0,2,1)} = D_1(d_1) \vee D_1(0),$$

$$\mathbf{P}^{(0,2,2)} = E_1(e_1) \vee E_1(0).$$

Table 3
The candidate patterns in Example 2

No.	Measurements							
	A_1	A_2	B_1	B_2	C_1	C_2	D_1	E_1
1	0	0	0	0	0	0	0	0
2	a_1	0	0	0	0	0	0	0
3	a_1	a_2	0	0	0	0	0	0
4	a_1	a_2	0	0	c_1	0	0	0
5	a_1	a_2	0	0	c_1	c_2	0	0
6	a_1	a_2	0	0	c_1	c_2	0	e_1
7	a_1	a_2	0	0	c_1	c_2	d_1	0
8	a_1	a_2	0	0	0	0	0	0
9	a_1	a_2	b_1	0	c_1	0	0	0
10	a_1	a_2	b_1	0	c_1	0	0	0
11	a_1	a_2	b_1	0	c_1	c_2	0	0
12	a_1	a_2	b_1	0	c_1	c_2	0	e_1
13	a_1	a_2	b_1	0	c_1	c_2	d_1	0
14	a_1	a_2	b_1	0	c_1	c_2	d_1	e_1
15	a_1	a_2	b_1	b_2	0	0	0	0
16	a_1	a_2	b_1	b_2	c_1	0	0	0
17	a_1	a_2	b_1	b_2	c_1	c_2	0	0
18	a_1	a_2	b_1	b_2	c_1	c_2	0	e_1
19	a_1	a_2	b_1	b_2	c_1	c_2	d_1	0
20	a_1	a_2	b_1	b_2	c_1	c_2	d_1	e_1

It should be noted that the symptoms after a branching node in SOO cannot be observed until the symptom on this branching node is materialized. For example, if $\mathbf{P}^{(0)} = A_1(a_1) \wedge A_2(0)$, then the patterns on the two immediate branch paths after the initial branching node A_2 should be $B_1(0) \wedge B_2(0)$ and $C_1(0) \wedge C_2(0)$, respectively. On the other hand, if $\mathbf{P}^{(0)} = A_1(a_1) \wedge A_2(a_2)$, then all patterns on the branch paths $\mathbf{S}^{(0,1)}$ and $\mathbf{S}^{(0,2)}$ are possible. Finally notice that, from the SOO alone, it is not possible to determine the precedence order among the symptoms on parallel paths originated from the same branching node. A list of all possible candidate patterns is presented in Table 3.

5. Fuzzy inference system

In this work, the final product prepared off-line is a *fuzzy inference system* (FIS). A sketch of its framework is presented in Fig. 7. If this system is to be implemented on-line, the measurement data must be first converted to a set of normalized deviations with respect to the given reference values and, then, used as inputs to FIS. The core of FIS is a collection of IF–THEN rules, which can be further divided into three distinct classes. The outputs of FIS are the occurrence index of top event OITE and also that of every minimal cut set cs_i ($i = 1, 2, \ldots$). The former can be used to forecast a potential incident and the latter are adopted to rank all possible fault origins. The main elements of FIS are described in detail in the following subsections.

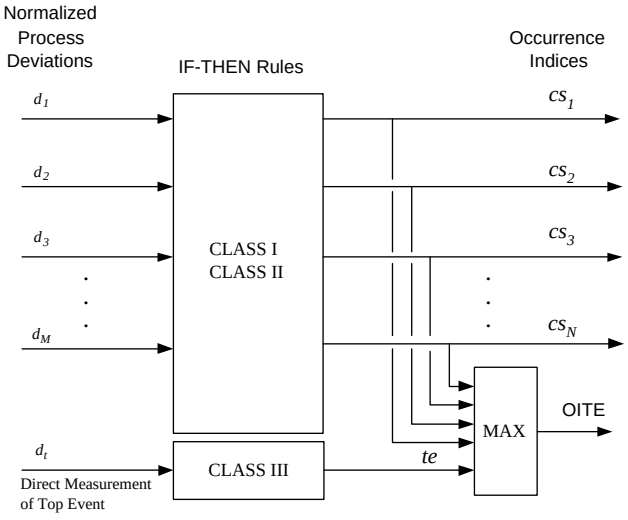

Fig. 7. The fuzzy inference system.

5.1. Classification of process deviations

Since the deviation values associated with disturbances and/or failures in a typical fault tree are *qualitative* in nature, it is necessary to introduce a more consistent description in fault diagnosis applications. In particular, a set of membership functions are used in this work for the purpose of classifying the on-line measurements into several fuzzy sets. Specifically, let us denote the ith process measurement as m_i and the same measurement at steady state as m_i^{ss}. The normalized process deviation of the ith measured variable, d_i, can be determined according to the

following equation:

$$d_i = \frac{m_i - \overline{m_i^{ss}}}{\sigma_i}. \quad (3)$$

where $\overline{m_i^{ss}}$ and σ_i denote, respectively, the mean and standard deviation of m_i^{ss}. In this study, three linguistic values, i.e. SN (small negative), ZE (zero) and SP (small positive), are assigned to the normalized process deviations to describe the qualitative concepts of -1, 0, and $+1$, respectively. They are represented with typical trapezoid membership functions in this study. It should be noted that, in practice, the appropriate shape and location of each membership function are dependent upon individual process considerations and thus should be determined on a case-by-case basis.

5.2. Classification of occurrence indices

As mentioned previously, an occurrence index can be viewed as the diagnosis concerning a top event or the basic events in a cut set. In general, three types of fuzzy sets, i.e. OCR, NOC and UCT_i are adopted in this work for its characterization. The symbol OCR is used to denote the belief that the corresponding event(s) will occur soon or have already occurred and, on the other hand, NOC is concerned with the opposite conclusion, i.e. the possibility of fault can be ruled out. Finally, UCT_i is essentially a less definite statement between OCR and NOC. The subscript i is used to reflect the degree of confidence toward confirmation of the corresponding event(s).

Notice that the membership functions of NOC and OCR used in this study are trapeziods. The locations of their four corners are (0, 0, 0.2, 0.4) and (0.6, 0.8, 1.0, 1.0), respectively. On the other hand, a total of $n-1$ triangular membership functions are used to characterize UCT_is. The locations of their apexes (V_i) are determined according to the following equation:

$$V_i = I_{\mathrm{NOC}} + \frac{I_{\mathrm{OCR}} - I_{\mathrm{NOC}}}{n} \times i \quad i = 1, 2, \ldots, n-1, \quad (4)$$

where n is the number of symptoms in SOO, and I_{NOC} and I_{OCR} denote, respectively, the locations of upper interior corners of the two trapezoids corresponding to NOC and OCR. The locations of bottom corners of each triangle are $V_k \pm \Delta$ and $\Delta = 0.1$ in this study.

5.3. Generation of inference rules

As mentioned previously, the IF–THEN inference rules can be divided into three different classes. The class I and class II rules are used for evaluating the existence potential of all basic events in a cut set, i.e. cs_i ($i = 1, 2, 3, \ldots$), and, on the other hand, the class III rules are used to compute a preliminary index for the top event, i.e. te.

The class I rules can be derived from the candidate patterns with the help of Theorem 1 or Theorem 2. To facilitate later discussions, let us define a *linguistic interpretation function F* as follows:

$$F(\delta_i) = \begin{cases} \mathrm{SN} & \text{if } \delta_i = -1, \\ \mathrm{ZE} & \text{if } \delta_i = 0, \\ \mathrm{SP} & \text{if } \delta_i = +1, \end{cases} \quad (5)$$

where $i = 1, 2, \ldots, n$. If the on-line symptoms are identical to those in a SOO, then it is highly possible that they are caused by the corresponding fault origin. Consequently, the following rule is incorporated in FIS to assert such a belief:

IF d_1 is $F(\delta_1)$ AND d_2 is $F(\delta_2)$ AND $\cdots$

AND d_n is $F(\delta_n)$

THEN cs_k is OCR. (6)

On the other hand, it is quite reasonable to disregard the possibility of a fault if none of the symptoms in the corresponding SOO can be observed. Thus, the following rule is also included:

IF d_1 is ZE AND d_2 is ZE AND $\cdots$

AND d_n is ZE

THEN cs_k is NOC. (7)

The remaining $N_{CP} - 2$ candidate patterns are translated into rules with uncertain conclusions, i.e. UCT_i and $i = 1, 2, \ldots, n-1$. The premises of each rule can be determined by substituting the qualitative deviation values in a candidate pattern into the linguistic interpretation function F. The number of matched symptoms in the candidate pattern is used as the degree of confidence i in conclusion UCT_i. Finally, it should be noted that, since three deviation values are used to characterize each symptom, the total number of *all* possible patterns should be 3^n. Obviously, there are $3^n - N_{CP}$ non-candidate patterns and the conclusions of the corresponding rules should all be NOC. These rules are excluded from FIS to reduce the computation load without causing significant impacts on the results of diagnosis. Three examples are provided in the sequel to illustrate the proposed rule generation procedure.

Table 4
The IF–THEN rules in Example 3

No.	IF			THEN
	d_{A_1}	d_{A_2}	d_{A_3}	
1	ZE	ZE	ZE	NOC
2	$F(\delta_1)$	ZE	ZE	UCT_1
3	$F(\delta_1)$	$F(\delta_2)$	ZE	UCT_2
4	$F(\delta_1)$	$F(\delta_2)$	$F(\delta_3)$	OCR

Example 3. Let us consider the SOO in Example 1. It has already been determined that the number of candidate patterns is 4 and these patterns are listed in Table 2. By implementing the procedure mentioned above, the corresponding inference rules can be generated in a straightforward fashion (Table 4).

Table 5
The IF–THEN rules in Example 4

No.	IF								THEN
	A_1	A_2	B_1	B_2	C_1	C_2	D_1	E_1	
1	ZE	ZE	ZE	ZE	ZE	ZE	ZE	ZE	NOC
2	F(a_1)	ZE	ZE	ZE	ZE	ZE	ZE	ZE	UCT_1
3	F(a_1)	F(a_2)	ZE	ZE	ZE	ZE	ZE	ZE	UCT_2
4	F(a_1)	F(a_2)	ZE	ZE	F(c_1)	ZE	ZE	ZE	UCT_3
5	F(a_1)	F(a_2)	ZE	ZE	F(c_1)	F(c_2)	ZE	ZE	UCT_4
6	F(a_1)	F(a_2)	ZE	ZE	F(c_1)	F(c_2)	ZE	F(e_1)	UCT_5
7	F(a_1)	F(a_2)	ZE	ZE	F(c_1)	F(c_2)	F(d_1)	ZE	UCT_5
8	F(a_1)	F(a_2)	ZE	ZE	F(c_1)	F(c_2)	F(d_1)	F(e_1)	UCT_6
9	F(a_1)	F(a_2)	F(b_1)	ZE	ZE	ZE	ZE	ZE	UCT_3
10	F(a_1)	F(a_2)	F(b_1)	ZE	F(c_1)	ZE	ZE	ZE	UCT_4
11	F(a_1)	F(a_2)	F(b_1)	ZE	F(c_1)	F(c_2)	ZE	ZE	UCT_5
12	F(a_1)	F(a_2)	F(b_1)	ZE	F(c_1)	F(c_2)	ZE	F(e_1)	UCT_6
13	F(a_1)	F(a_2)	F(b_1)	ZE	F(c_1)	F(c_2)	F(d_1)	ZE	UCT_6
14	F(a_1)	F(a_2)	F(b_1)	ZE	F(c_1)	F(c_2)	F(d_1)	F(e_1)	UCT_7
15	F(a_1)	F(a_2)	F(b_1)	F(b_2)	ZE	ZE	ZE	ZE	UCT_4
16	F(a_1)	F(a_2)	F(b_1)	F(b_2)	F(c_1)	ZE	ZE	ZE	UCT_5
17	F(a_1)	F(a_2)	F(b_1)	F(b_2)	F(c_1)	F(c_2)	ZE	ZE	UCT_6
18	F(a_1)	F(a_2)	F(b_1)	F(b_2)	F(c_1)	F(c_2)	ZE	F(e_1)	UCT_7
19	F(a_1)	F(a_2)	F(b_1)	F(b_2)	F(c_1)	F(c_2)	F(d_1)	ZE	UCT_7
20	F(a_1)	F(a_2)	F(b_1)	F(b_2)	F(c_1)	F(c_2)	F(d_1)	F(e_1)	OC

Example 4. Let us again consider the tree-shape SOO in Example 2. The number of candidate patterns has already been determined, i.e. $N_{CP}=20$. These candidate patterns are listed in Table 3. The corresponding inference rules should be the ones presented in Table 5.

Although the qualitative deviations in typical candidate patterns are restricted to $+1$ and -1 (see Theorems 1 and 2), the symptom on an end node in SOO may also assume the value of 0, e.g. see Fig. 5. This is mainly due to an equipment failure nullifying the causal relationship associated with an edge. In such cases, the number of candidate patterns can still be calculated with the proposed theorems if the value 0 is treated as a "deviation", i.e. the set $\mathbf{X}$ is replaced by $\mathbf{X}^*=\{-1,0,+1\}$. However, by adopting $\mathbf{X}^*$, one or more set of conflicting IF–THEN rules may be constructed according to the proposed rule-generation procedure. It should be stressed that, although these rules are inconsistent in terms of crisp logic, the resulting fuzzy inference calculations can still be carried out without difficulties. Furthermore, it is in fact quite reasonable to include these conflicting rules in FIS since they represent candidate patterns observed at different times during the fault propagation process. Following is an example.

Example 5. Consider the tree-shape SOO in Fig. 8. There are 6 candidate patterns as shown in Table 6. Suppose that $a_1=+1,\ a_2=+1,\ b_1=+1,\ c_1=0$ in this case. The corresponding inference rules can be found in Table 7. Notice first that the premises of the 3rd and 4th rules are the same. Row 3 is used to represent the scenario that the effects of fault and/or failure have not reached any of the branch paths.

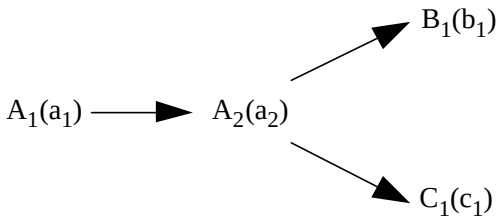

Fig. 8. The simple tree-shape symptom occurrence order in Example 5.

Table 6
The candidate patterns in Example 5

No.	A_1	A_2	B_1	C_1
1	0	0	0	0
2	a_1	0	0	0
3	a_1	a_2	0	0
4	a_1	a_2	0	c_1
5	a_1	a_2	b_1	0
6	a_1	a_2	b_1	c_1

On the other hand, an asterisk is used in row 4 to indicate the fact that the symptom associated with C_1 is *fully* developed. Thus, the degrees of diagnostic confidence in these two cases should be different and the conclusions of the IF–THEN rules must be specified accordingly. Notice also that the same can be observed from rules 5 and 6.

If a variable is not a member of $\mathbf{ROI}_k$, then it should remain unaffected by the corresponding basic event(s) in cut set k. The class II rules are mainly used to reduce the chance

Table 7
The IF–THEN rules in Example 5

No.	IF				THEN
	d_{A_1}	d_{A_2}	d_{B_1}	d_{C_1}	
1	ZE	ZE	ZE	ZE	NOC
2	SP	ZE	ZE	ZE	UCT_1
3	SP	SP	ZE	ZE	UCT_2
4	SP	SP	ZE	ZE*	UCT_3
5	SP	SP	SP	ZE	UCT_3
6	SP	SP	SP	ZE*	OCR

of misdiagnosis due to a different fault origin, say the basic events in cut set k', causing the same candidate patterns within $\mathbf{ROI}_k$. The typical class II rules can be written as follows:

IF d_u is SP THEN cs_k is NOC,

IF d_u is SN THEN cs_k is NOC, (8)

where cs_k is the occurrence index of given cut set and $u \notin \mathbf{ROI}_k$. If cut set k is indeed the correct root cause, then the rules in Eq. (8) are not supposed to be triggered. On the other hand, in the case of cut set k' occurring, these rules should definitely lower the value of cs_k. Thus, by assuming that the possibility of more than one fault origin occurring is negligible, the class II rules can enhance the diagnostic resolution. However, the drawback of including them in FIS is that the possibility of a nonminimal cut set or multiple cut sets may also be ignored. In other words, if the basic events in cut set k form a subset of all existing events, these rules may cause a decrease in the occurrence index cs_k. Therefore, this trade-off between resolution and comprehensiveness in diagnosis must be evaluated carefully before introducing the class II rules in practical applications. In this study, they are included on the ground that the probability of additional events occurring is extremely low.

Finally, if the top event is monitored directly with a sensor, then this information should be encoded with class III rules. Since there is only one measurement involved, three inference rules can be constructed to cover all possibilities. As an example, by assuming that "d_t is SP" is the event causing undesirable consequences, one can express the corresponding rules as

IF d_t is SP THEN te is OCR,

IF d_t is ZE THEN te is NOC,

IF d_t is SN THEN te is NOC. (9)

5.4. Forecast of top event

If the existence of the basic event(s) in any cut set is confirmed, then one can deduce logically that the top event is bound to occur. Thus, the outputs of class I and II rules, i.e. cs_i ($i = 1, 2, \ldots$), can also be treated as indications of system hazard. On the other hand, the class III rules exist only when the *direct* measurement of the variable associated with top event d_t is available. If this is the case, the final occurrence index of top event (OITE) should be obtained by taking the maximum of cs_is and te, i.e.

$$\mathrm{OITE} = \max(cs_1, cs_2, \ldots, \mathrm{te}). \quad (10)$$

6. The on-line implementation procedure

During operation, the process measurements are taken continuously on-line and then converted to the normalized deviations according to Eq. (3). These values are then used in the fuzzy inference system to calculate the occurrence indices of cut sets (cs_is) and also the top event (OITE). The fuzzy inference algorithm proposed by Mamdani and Assilian (1975) is adopted in this work. The index OITE can be utilized as the basis for alarm generation and cs_is are essentially the measures for ranking the possibilities of fault origins. Since the diffuzification procedure in the inference algorithm basically involves centroid calculation, the upper and lower limits of the output values are dependent upon the shapes and locations of the membership functions for OCR and NOC, respectively. In the present work, the outputs of FIS should be bounded between 0.153 and 0.847. For ease of interpretation, the raw output values are always re-scaled onto the interval $[0, 1]$ in our applications.

In this study, the proposed fault diagnosis approach is tested with on-line measurement data generated by the TE process simulator. Let us again consider the same scenario described in Sections 2 and 3, i.e. the incident associated with the 11th cut set in Table 1. In the simulation study, the failure "cooling water control valve CV-11 sticks" was first introduced at 500 s after the initiation time and the disturbance causing a 10% increase in the feed rate of reactant D was then created at 1000 s. The simulation results are presented in Fig. 9. The occurrence indices of top event and the cut set associated with the simulated fault origin are given in Fig. 9(A) and (B), respectively. It can be observed that the possibility of top event is detected earlier than that of the correct fault origin. More specifically, the occurrence of former is confirmed around 1150 s and the latter is identified around 1250 s. It is clear from this figure that the fuzzy inference approach is quite sensitive. In addition, the responses of the inference system corresponding to two different cut sets, cs_{11} and cs_{13}, are compared in Fig. 9(B) to demonstrate the superior diagnostic resolution achieved by the proposed strategy. It should be noted that the cut set cs_{11} represents the correct fault origin used in simulation. The basic events in cs_{13} are: "an increase in the feed rate of reactant A" and "reactor temperature sensor sticks". Since the initial symptoms of cs_{11} and cs_{13} are the same, the diagnosis results in both cases are identical temporarily in the beginning. However, the occurrence index of the former origin

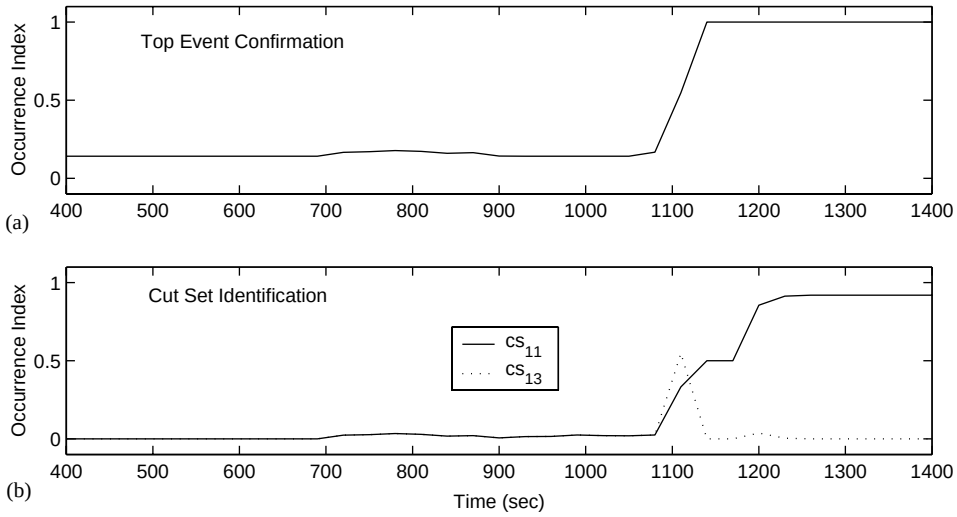

Fig. 9. Simulation results of reactor section.

increases continuously, but that of the latter drops to almost zero. This is apparently due to a mismatch between the later symptoms of cs_{13}.

7. Case studies

Extensive simulation studies have been carried out in this work to demonstrate the effectiveness of the proposed fault diagnosis approach. As mentioned before, the test data were generated with the TE process simulator. To execute the simulator successfully, it is necessary to specify a proper control configuration for the entire process. Thus, the following five control loops are included in all simulation runs:

(1) reactor temperature control loop,
(2) recycle flow control loop,
(3) separator temperature control loop,
(4) stripper temperature control loop and
(5) stripper level control loop.

The resulting process flow diagram can be found in Fig. 10. The simulation results of two scenarios are presented in the sequel.

7.1. Stripper section

Let us consider a SCFD restricted to the stripper section of Fig. 10. By selecting the top event "high liquid level in the stripper", i.e. $L_{S\text{-}01}(+1)$, a fault tree can be constructed accordingly. Let us study the fault propagation scenario associated with one of its cut sets, i.e. $\{m\text{-}500(+1)$, CV-08 sticks$\}$. The corresponding SOO can be identified with the proposed procedure (see Fig. 11).

In the simulation study, the failure "level control valve (CV-08) sticks" was first activated at 950 s and then a 10% increase was introduced in the feed rate of reactant A/C, i.e. m-500(+1), at 1000 s. A sample of the simulation results is presented in Fig. 12. From Fig. 12(A), it can be seen that the top event is confirmed almost immediately after the fault occurs. Initially, since the fault symptoms are ambiguous, the fuzzy inference system yields an occurrence index of 0.5 indicating the status of the given fault origin is uncertain. As more symptoms emerge later, the occurrence index gradually reaches its maximum value 1.0. The diagnosis results of two competing cut sets, cs_9 and cs_{11}, are presented in Fig. 12(B). The former is associated with the correct fault origin and the basic events in the latter are: "an increase in the feed rate of reactant A/C" and "level sensor sticks". Notice that the occurrence index cs_{11} is higher than cs_9 between 1000 and 2300 s. This is due to the facts that the early symptoms observed in the on-line measurement data match the SOO of the 9th cut set only partially but they are essentially identical to the symptoms of cs_{11}. However, as the full spectrum of fault propagation pattern develops, the fuzzy inference system eventually changes its conclusion around 2300 s.

7.2. The entire process

To test the feasibility of the fuzzy diagnosis approach in more complex systems, let us expand the SCFD to cover the entire TE process. For comparison purpose, the event "reactor temperature too high", i.e. $T_{R\text{-}01}(+1)$, is again chosen as the top event here. Notice that each event set listed in Table 1 should still be considered as a minimal cut set of the fault tree in the present case. However, the corresponding FPPs and also SOO must be extended from reactor section to other sections of the process.

Apart from the 31 cut sets listed in Table 1, there should be numerous causes of $T_{R\text{-}01}(+1)$ which can be attributed to disturbances and/or failures in units other than those in the reactor section. Let us study the fault propagation scenario

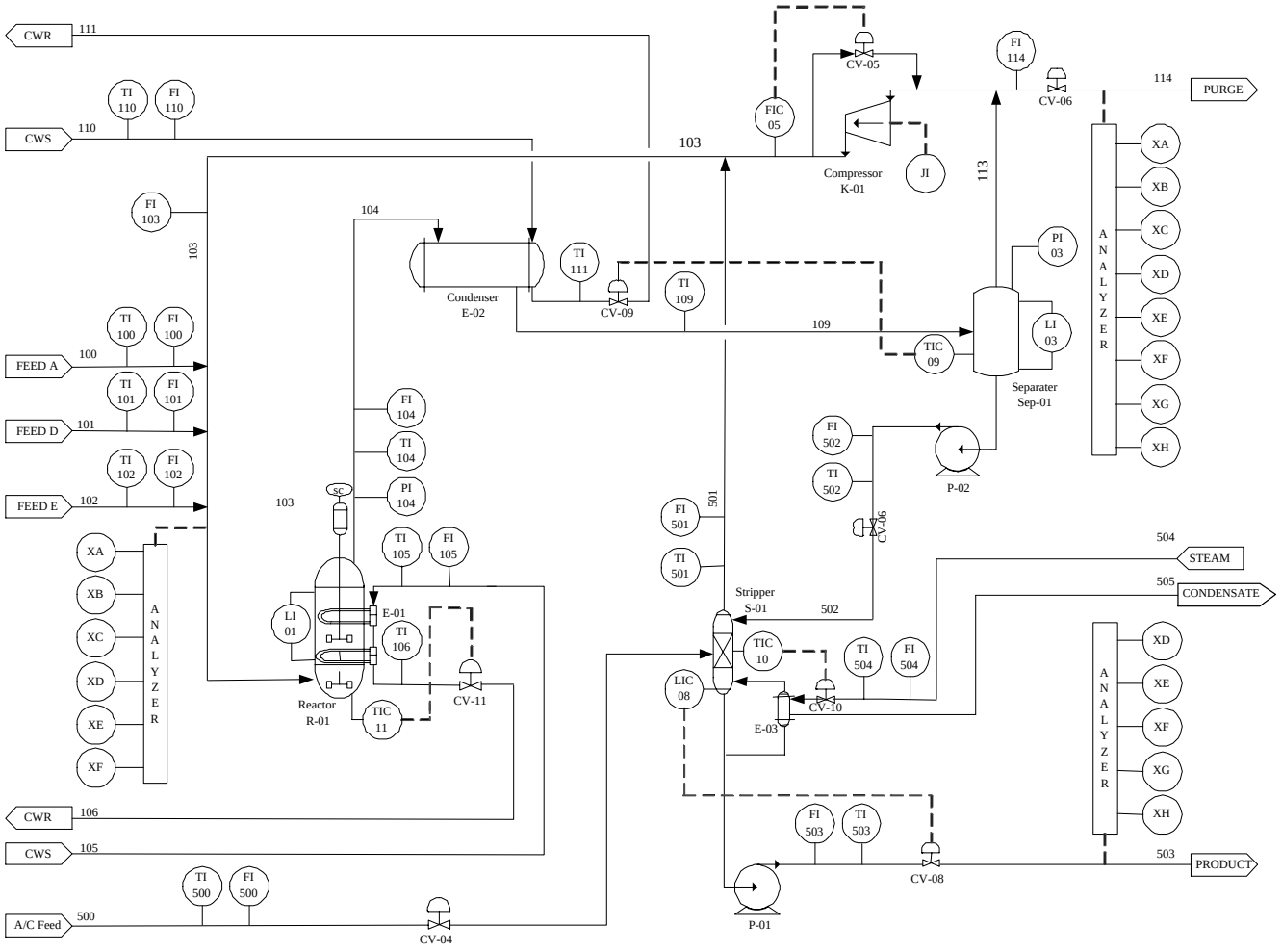

Fig. 10. TE process with control configuration.

associated with only one of them, i.e. the effects caused by

(1) an increase in the inlet temperature of condenser cooling water (T-110(+1)),
(2) failure of the recycle flow control valve (CV-05 sticks),
(3) failure of the separator temperature control valve (CV-09 sticks) and
(4) failure of the cooling water control valve in reactor section (CV-11 sticks).

The set of the above basic events, i.e. {T-110(+1), CV-05 sticks, CV-10 sticks, CV-11 sticks}, is labeled as the 32th cut set in this paper. The corresponding FPPs can then be determined with the qualitative simulation techniques. Since the number of symptoms embedded in these FPPs is quite large, it becomes necessary to produce a simpler version to relieve the implied computation load in fuzzy inference. In particular, the FPPs can be simplified by removing a large number of branch paths which are not critical to the diagnostic performance. Following is a list of possible candidates:

- the branch paths connected to the node representing top event or the feedback loop in which this node is located, and

$$\text{FI-500}^{m}(+1) \longrightarrow \text{LIC-03}^{m}(+1) \longrightarrow \text{FI-503}^{m}(+1)$$

Fig. 11. The symptom occurrence order (SOO) in {m-500(+1), CV-08 sticks}.

- the branch paths originated from the path between the nodes representing the fault origin and the top event.

Generally speaking, the first type of paths should be eliminated. This is due to the fact that the corresponding symptoms can be observed only *after* the occurrence of top event. Thus, a conclusive diagnosis cannot be obtained in a timely manner if confirmation of any of these symptoms is required in the inference rules. However, since a disturbance may propagate through the node corresponding to top event repeatedly if it is located in a feedback loop, all nodes in this loop must still be kept in the simplified FPPs.

The symptoms on the second type of paths may or may not appear before the top event. Thus, it is only necessary to delete the ones occurring later than the top event. In practice, their removal must be performed according to operation experience on a case-by-case basis. In this study, they

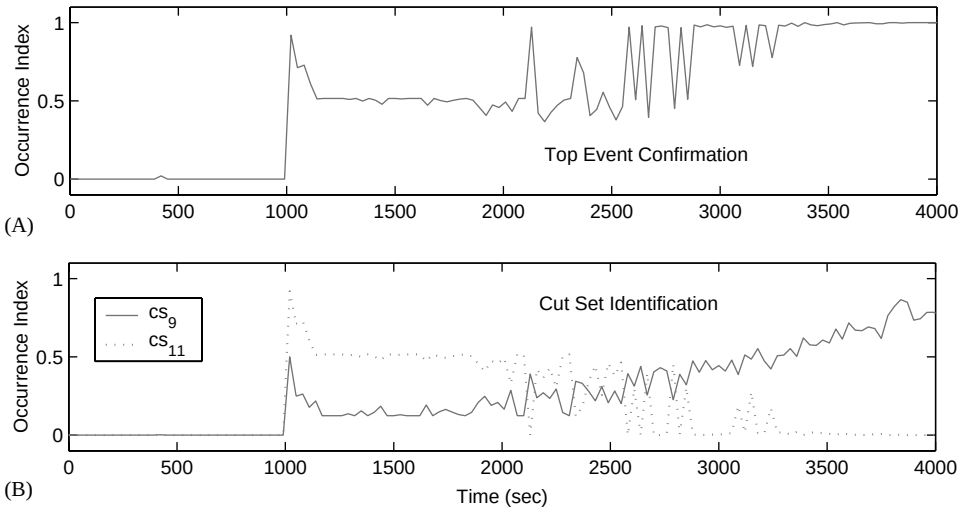

Fig. 12. Simulation results of stripper section.

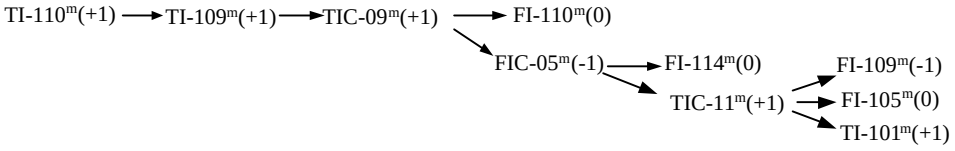

Fig. 13. The SOO of cs_{32} in the entire TE process.

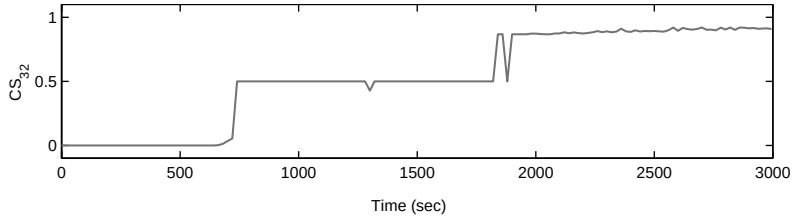

Fig. 14. The occurrence index of cs_{32}.

are severed on the basis of the observed fault propagation behavior in the simulation results. The resulting SOO is presented in Fig. 13. It can be observed that the number of symptoms is 10. The total number of candidate patterns can be determined with Eq. (2) and its value is 41 in this case.

In the simulation study, the process is at steady state in the beginning. The failures of control valves, CV-11, CV-10 and CV-05, are activated at 500, 600 and 700 s, respectively. The disturbance in the inlet temperature of condenser cooling water is introduced at 700 s. The simulation results for cs_{32} is presented in Fig. 14. Notice that, in this scenario, the basic events occur in sequence during the time period from 500 to 700 s. The output of FIS in Fig. 14 indicates that occurrence of the 32th cut set is detected with a confidence degree of 0.5 at about 750 s and its existence is confirmed at about 1900 s. Finally, it should be noted that the occurrence indices of all 31 cut sets in Table 1 remain very close to zero during the entire operation period, i.e. from 0 to 3000 s.

8. Conclusion

A novel fault diagnosis approach based on knowledge concerning all possible patterns in on-line symptoms is proposed in this paper. For simplification purpose, the candidates of fault identification is limited to the minimal cut sets of a given fault tree. The fault propagation paths (FPPs) and symptom occurrence order (SOO) associated with each cut set are generated from the system digraph with a systematic procedure. All possible symptom patterns of each fault origin can be enumerated according to its SOO with the help of two theorems developed in this study. These patterns are then encoded with IF–THEN inference rules and then incorporated in a fuzzy inference system. The proposed fault diagnosis approach has been tested in extensive simulation studies. The simulation results show that this approach is quite effective in detecting the occurrence of top event and also identifying the correct fault origin during the incipient stage of a serious accident.

Notation

cs_k	the occurrence index of kth cut set
d_i	the normalized process deviation of ith measured variable
F	the linguistic interpretation function
I_{NOC}	the location of upper interior corner of the membership function NOC
I_{OCR}	the location of upper interior corner of the membership function OCR
M	the set of all measured variables
m_i	the measurement value of ith measured variable
m_i^{ss}	the value of m_i at steady state
$\overline{m_i^{ss}}$	the average value of m_i^{ss}
N_{CP}	the number of candidate patterns
NOC	not occur, linguistic value of output fuzzy membership function
OCR	occurred, linguistic value of output fuzzy membership function
OITE	occurrence index of top event
$\mathbf{ROI}_k$	the range of influence of the kth cut set
$\mathbf{S}^{(0)}(n_0)$	the initial path of length n_0
$\mathbf{S}^{(0,i)}(n_{0,i})$	the ith branch path from $\mathbf{S}^{(0)}$ of length $n_{0,i}$
$\mathbf{S}^{(0,i,j)}(n_{0,i,j})$	the jth branch path of length $n_{0,i,j}$
SN	small negative—a linguistic value of the normalized process deviation d_i
SP	small positive—a linguistic value of the normalized process deviation d_i
te	preliminary occurrence index of top event
UCT_i	uncertain with confidence of degree i—a linguistic value of the performance index
V_k	the apex location of triangular membership function of UCT_i
$\mathbf{X}$	$\mathbf{X}=\{-1,+1\}$, the set of deviation state of measurements
$\mathbf{X}^*$	$\mathbf{X}^*=\{-1,0,+1\}$, an extended version of X
ZE	zero, linguistic value of input fuzzy membership function

Greek letters

δ_i	the qualitative deviation value of process variable m_i, $\delta_i \in \{-1, 0+1\}$
σ_i	the standard deviation of m_i^{ss}

Acronyms

FI	fuzzy inference
FIS	fuzzy inference system
FPP	fault propagation path
ROI	range of influence
SCFD	scope of fault diagnosis
SDG	signed directed graph
SOO	symptom occurrence order
TE	Tenneessee Eastman

Appendix A. Proof of Theorem 1

If $n=1$, then $\mathbf{S}(1)=m_1(\delta_1)$ and all possible candidate patterns can be described with a logic statement, i.e.

$$\mathbf{P}_1 = m_1(0) \vee m_1(\delta_1),$$

where the symbol $\vee$ denotes the logic operator "OR" and the number of terms connected by $\vee$ is the number of candidate patterns. If $n=2$, then $\mathbf{S}(2)=m_1(\delta_1)\rightarrow m_2(\delta_2)$ and all possible candidate patterns can be expressed as

$$\mathbf{P}_2 = (m_1(0)\wedge m_2(0)) \vee (m_1(\delta_1)\wedge m_2(0)) \vee (m_1(\delta_1)\wedge m_2(\delta_2)),$$

where the symbol $\wedge$ denotes the logic operator "AND". Thus, by counting the number of terms connected by "OR" in the logic statements, the theorem can be confirmed in both cases mentioned above.

Let us further assume that the theorem is valid for a particular case when $n=k$ $(k=1,2,3,\ldots)$ and then consider if it is also true when $n=k+1$. The corresponding propagation path in the symptom occurrence order can be written as

$$\mathbf{S}(k+1)=\mathbf{S}(k)\rightarrow m_{k+1}(\delta_{k+1}),$$

where $\mathbf{S}(k)=m_1(\delta_1)\rightarrow m_2(\delta_2)\rightarrow\cdots\rightarrow m_k(\delta_k)$. Thus, all possible candidate patterns in this case can be expressed as

$$\mathbf{P}_{k+1} = (\mathbf{P}_k \wedge m_{k+1}(0)) \vee (m_1(\delta_1)\wedge m_2(\delta_2)\wedge\cdots \wedge m_k(\delta_k)\wedge m_{k+1}(\delta_{k+1})).$$

Since the number of terms in $\mathbf{P}_k\wedge m_{k+1}(0)$ is the same as that in $\mathbf{P}_k$, the total pattern number in $\mathbf{P}_{k+1}$ should be $k+2$. The validity of the theorem is thus confirmed. □

Appendix B. Proof of Theorem 2

If $N_0=0$, then the symptom occurrence order is a single path and $N_{CP}=n_0+1$. The proof for this case has already been given in Theorem 1.

Let us consider the simplest tree-shape symptom occurrence order in which $N_0>1$ and $N_{0,i}=0\ \forall i$. It is clear that the symptoms in the branch paths $\mathbf{S}^{(0,i)}(n_{0,i})$ $(i=1,2,\ldots,N_0)$ can only be observed after the last symptom in $\mathbf{S}^{(0)}(n_0)$ occurs. Notice also that only one out of the n_0+1 possible patterns derived from $\mathbf{S}^{(0)}(n_0)$ satisfies this requirement. By Theorem 1, the number of candidate patterns associated with each branch path $\mathbf{S}^{(0,i)}(n_{0,i})$ $(i=1,2,\ldots,N_0)$ should be $n_{0,i}+1$ in this situation. Since, after the occurrence of the last symptom in the initial path, the ensuing effects propagate along the branches *in parallel*, the total number of the corresponding patterns should be $\prod_{i=1}^{N_0}(n_{0,i}+1)$. Therefore, the number of all possible candidate patterns for the entire symptom occurrence order should be

$$N_{CP} = n_0 + \prod_{i=1}^{N_0}(n_{0,i}+1).$$

This formula can be rewritten by following the convention adopted in Theorem 2, i.e.

$$N_{CP} = n_0 + \prod_{i=1}^{N_0} \left[n_{0,i} + \prod_{j=1}^{N_{0,i}} (\bullet) \right],$$

where the first term is associated with the first n_0 patterns derived from $\mathbf{S}^{(0)}(n_0)$ and the second is corresponding to the last pattern when all symptoms in the initial path materialize.

Let us next consider a symptom occurrence order in which $N_0 > 1$, $N_{0,i} > 1\ \exists i$ and $N_{0,i,j} = 0\ \forall j$. Without loss of generality, let us further assume that

$$N_{0,i_1} > 1 \quad i_1 = 1, 2, \ldots, N_0^*,$$

$$N_{0,i_2} = 0 \quad i_2 = N_0^* + 1, \ldots, N_0.$$

In the former case, on the basis of the observations that (1) the symptoms in the branch paths can be observed only after the symptom corresponding to the branching node occurs and (2) the resulting effects propagate along the branches *in parallel*, one can see clearly that the number of candidate patterns associated with $\mathbf{S}^{(0,i_1)}(n_{0,i_1})$ $(1 \leqslant i_1 \leqslant N_0^*)$ and its branch paths $\mathbf{S}^{(0,i_1,j)}(n_{0,i_1,j})$ $(j = 1, 2, \ldots, N_{0,i_1})$ can be determined with the same approach described previously. More specifically, this number should be $n_{0,i_1} + \prod_{j=1}^{N_{0,i_1}} (n_{0,i_1,j} + 1)$. In the latter case, the number of candidate patterns associated with $\mathbf{S}^{(0,i_2)}(n_{0,i_2})$ $(N_0^* + 1 \leqslant i_2 \leqslant N_0)$ has already been shown to be $n_{0,i_2} + 1$. Thus, by repeatedly applying the computation procedure for branching nodes, the total number of candidate patterns can be determined, i.e.

$$N_{CP} = n_0 + \left\{ \prod_{i=1}^{N_0^*} \left[n_{0,i} + \prod_{j=1}^{N_{0,i}} (n_{0,i,j} + 1) \right] \right\} \times \left\{ \prod_{i=N_0^*+1}^{N_0} [n_{0,i} + 1] \right\}.$$

Again, let us adopt the proposed convention to produce a more concise formulation

$$N_{CP} = n_0 + \left\{ \prod_{i=1}^{N_0^*} \left[n_{0,i} + \prod_{j=1}^{N_{0,i}} \left(n_{0,i,j} + \prod_{k=1}^{N_{0,i,j}} (\bullet) \right) \right] \right\} \times \left\{ \prod_{i=N_0^*+1}^{N_0} \left[n_{0,i} + \prod_{j=1}^{N_{0,i}} (\bullet) \right] \right\} = n_0 + \prod_{i=1}^{N_0} \left[n_{0,i} + \prod_{j=1}^{N_{0,i}} \left(n_{0,i,j} + \prod_{k=1}^{N_{0,i,j}} (\bullet) \right) \right].$$

The formula for computing the total number of candidate patterns can be extended to symptom occurrence orders with additional branches. Specifically, let us consider the case in which $N_0 > 1$, $N_{0,i} > 1\ \exists i$, $N_{0,i,j} > 1\ \exists j$ and $N_{0,i,j,k} = 0\ \forall k$. Let us assume that

$$N_{0,i_1,j_1} > 1 \quad j_1 = 1, 2, \ldots, N_{0,i_1}^*,$$

$$N_{0,i_1,j_2} = 0 \quad j_2 = N_{0,i_1}^* + 1, \ldots, N_{0,i_1}.$$

By the same reasoning process, the following equation can be derived

$$N_{CP} = n_0 + \left\{ \prod_{i=1}^{N_0^*} \left[n_{0,i} + \left(\prod_{j=1}^{N_{0,i}^*} \left\langle n_{0,i,j} + \prod_{k=1}^{N_{0,i,j}} (n_{0,i,j,k} + 1) \right\rangle \right) \times \left(\prod_{j=N_{0,i}^*+1}^{N_{0,i}} \langle n_{0,i,j} + 1 \rangle \right) \right] \right\} \left\{ \prod_{i=N_0^*+1}^{N_0} [n_{0,i} + 1] \right\} = n_0 + \prod_{i=1}^{N_0} \left\{ n_{0,i} + \prod_{j=1}^{N_{0,i}} \left[n_{0,i,j} + \prod_{k=1}^{N_{0,i,j}} \left(n_{0,i,j,k} + \prod_{l=1}^{N_{0,i,j,k}} (\bullet) \right) \right] \right\}.$$

Thus, the validity of the theorem can be confirmed by this recursive deduction procedure for a tree-shape symptom occurrence order of arbitrary shape and depth. □

References

Chang, C. T., & Chen, J. W. (1995). Implementation issues concerning the ekf-based fault diagnosis technique. *Chemical Engineering Science*, *50*(18), 2861.

Chang, C. T., & Hwang, H. C. (1992). New developments of the digraph-based techniques for fault-tree synthesis. *I & EC Research*, *31*, 1490.

Chang, C. T., & Hwang, J. I. (1998a). Simplification of techniques for ekf computations in fault diagnosis-suboptimal gains. *Chemical Engineering Science*, *53*, 3853.

Chang, C. T., & Hwang, J. I. (1998b). Simplification techniques for ekf computations in fault diagnosis-model decomposition. *AIChE Journal*, *44*(6), 1392.

Chang, C. T., Mah, K. N., & Tsai, C. S. (1993). A simple design strategy for fault monitoring systems. *AIChE Journal*, *39*(7), 1146.

Downs, J. J., & Vogel, E. F. (1993). A plant-wide industrial process control problem. *Computers & Chemical Engineering*, *17*(3), 245.

Hoskins, J. C., Kalivur, K. M., & Himmeblau, D. M. (1991). Fault diagnosis in complex chemical-plants using artificial neural networks. *AIChE Journal*, *37*, 137.

Iri, M., Aoki, K., O'Shima, E., & Matsuyama, H. (1979). An algorithm for diagnosis of system failure in the chemical process. *Computers & Chemical Engineering*, *3*, 489.

Kramer, M. A., & Palowitch, B. L. (1987). A rule-based approach to fault diagnosis using the signed directed graph. *AIChE Journal*, *33*, 1067.

Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, *7*, 1.

Petti, T. F., Kleni, J., & Dhurijati, P. S. (1991). Diagnostic model processor: Using deep knowledge for process fault diagnosis of system failures in the chemical process *AIChE Journal*, *36*, 565.

Rich, S. H., & Venkatasubramanian, V. (1987). Model-based ressoning in diagnostic expert system for chemical process plants. *Computers & Chemical Engineering*, *11*, 111.

Ross, T. J. (1995). *Fuzzy logic with engineering applications*. New York: Mc-Graw-Hill.

Shiozaki, J., Matsuyama, H., O'Shima, E., & Iri, M. (1985). An improved algorithm for diagnosis of system failures in the chemical process. *Computers & Chemical Engineering*, *9*(3), 285.

Tsai, C. S., Chang, C. T., & Chen, C. S. (1996). Fault detection and diagnosis in batch and semi-batch processes using artificial neural networks. *Chemical Engineering Communications*, *143*, 39.

Ulerich, N. H., & Powers, G. J. (1988). On-line hazard aversion and fault diagnosis in chemical processes: The diagraph + fault tree method *IEEE Transactions on Reliability*, *37*, 171.

Venkatasubramanian, V., & Chan, K. (1989). A neural network methodology for process diagnosis. *AIChE Journal*, *35*, 1993.

Venkatasubramanian, V., Vaidyanathan, R., & Yamamoto, Y. (1990). Process fault detection and diagnosis using neural network-i. steady-state Process. *Computers & Chemical Engineering*, *14*, 699.

Watanabe, K., Matsuura, M. A., Kuboto, M., & Himmelblau, D. M. (1989). Incipient fault diagnosis of chemical process via artificial neural networks. *AIChE Journal*, *35*, 1803.