FAULT ISOLABILITY WITH DIFFERENT FORMS OF THE FAULTS–SYMPTOMS RELATION

JAN MACIEJ KOŚCIELNY a,∗, MICHał SYFERT a, KORNEL ROSTEK a, ANNA SZTYBER a

aInstitute of Automatic Control and Robotics
Warsaw University of Technology, ul. Św. Andrzeja Boboli 8, 02-525 Warsaw, Poland
e-mail: {jmk,m.syfert,k.rostek,a.sztyber}@mchtr.pw.edu.pl

The definitions and conditions for fault isolability of single faults for various forms of the diagnostic relation are reviewed. Fault isolability and unisolability on the basis of a binary diagnostic matrix are analyzed. Definitions for conditional and unconditional isolability and unisolability on the basis of a fault information system (FIS), symptom sequences and directional residuals are formulated. General definitions for conditional and unconditional isolability and unisolability in the cases of simultaneous evaluation of diagnostic signal values and a sequence of symptoms are provided. A comprehensive example is discussed.

Keywords: diagnosis, fault diagnosis, fault isolation, sequences, information systems.

1. Introduction

With the increasing complexity of contemporary technical systems, ensuring safety and reliability of industrial processes has become increasingly important. Methods of fault detection and isolation (FDI) allow us to swiftly detect and isolate faults, thus enabling us to prevent undesirable consequences.

Undeniably there have been published numerous papers considering various issues of FDI methods. Two major approaches have emerged: model-based and model-free. In this work we focus on model-based techniques. We obtain the residuals using balance equations and by comparing modeled signals with the corresponding measured process outputs. Models can include information about faults or not. Both the situations are covered in this paper.

Fault isolability obtained in a diagnostic system is closely related to the nature of diagnostic signals, i.e., outputs of detection algorithms, and the form of notation used for describing the relation between diagnostic signal values and faults. Three types of diagnostic signals may be distinguished: binary, multi-valued (e.g., three-valued: −1, 0, +1) and continuous.

In order to formulate diagnosis, the knowledge of the relationship between faults and diagnostic signal values is necessary. A symptom is the appearance of such a value of a diagnostic signal which indicates the presence of a fault in the monitored part of a diagnostic process. It is usually assumed that the zero value corresponds to a faultless state, while other values indicate faults. Therefore, the relation between diagnostic signal values and faults will be referred to as the faults–symptoms relation. This relationship takes various forms depending on the available type of diagnostic signals. The binary diagnostic matrix (BDM) is primarily used (Chen and Patton, 1999; Gertler, 1998; Isermann, 2006; Korbicz et al., 2004; Patton et al., 2000; Bartyś, 2013). Another solution is a fault isolation system (FIS) (Korbicz et al., 2004; Kościelny, 1999; Kościelny et al., 2006), which assumes the use of multi-valued diagnostic signals. From the BDM, other forms of notation can be derived, such as logic functions, IF–THEN rules, fault trees (Korbicz et al., 2004). They correspond to rows and columns of the initial matrix. Analogous rules can be derived from the FIS. Those secondary forms of notation do not change fault isolability. The relation between faults and continuous diagnostic signals is described by regions in the residual space (Isermann, 2006; Korbicz et al., 2004), vectors of directions in the residual space (Chen and Patton, 1999; Gertler, 1998; Patton et al., 2000) and sequences of symptoms (Kościelny et al., 2013).

The above relations are determined based on
1. process modeling with fault influence, i.e., built on derived residual equations in internal form (Gertler, 1998);

2. learning, i.e., creating regions in the diagnostic signal space that correspond to individual faults, applying process data registered during faulty process states (Koivo, 1994; Patton et al., 1999);

3. expert knowledge (Kościelny, 1999; Syfert and Kościelny, 2009).

There are also successive works in the area of data-driven fault diagnosis (e.g., Yin and Huang, 2015; Yin et al., 2016; 2015). In the work of He et al. (2013) a novel method of residual matching (RM) is proposed to isolate and estimate faults.

The model of a diagnosed process, utilizing the relation between inputs, outputs and faults (and possibly disturbances and measurement noise), is available only in the first case. Also, the majority of publications in the field of fault isolation relate only to the first case (Ossmann and Varga, 2015). Isolability definitions given in the overview works by Basseville (1997; 1999) and Ding (2008) also refer to the first case. They assume the knowledge of the process model accommodating fault influence.

Fault isolability has been usually defined in the context of the adopted diagnostic method, in particular, the form of the notation of faults–symptoms relation. Most often, fault isolability was analyzed in the case of the BDM (incidence matrix) derived from the structure of linear equations of residuals in internal form (Gertler, 1998).

Faults isolability obtained on the basis of the BDM and the FIS derived from expert knowledge was analyzed in the works of Korbicz et al. (2004) and Kościelny et al. (2006). In the case of multiple-valued and continuous diagnostic signals, it is not always possible to determine subsets of isolable faults. Then, a conditional isolability is considered. Also, when analyzing isolability in the residual space, conditional isolability of regions corresponding to given faults is used (Korbicz et al., 2004).

There is a parallel approach to fault diagnosis developed by the computer science and artificial intelligence community called DX. This approach uses consistency-based logical methods derived from Reiter’s theory (De Kleer et al., 1992; Reiter, 1987). In recent years there has been a substantial effort to combine FDI and DX techniques (Travé-Massuyès, 2014).

In the DX approach, diagnoses are deduced from sets of conflicts. The conflict is a set of system components where the assumption that all these components are healthy is inconsistent with the system description. The diagnoses are hitting sets of the conflicts sets. It should be noted that diagnosis is defined as a set of components, so multiple faults are considered.

In the original approach the conflicts were calculated on-line (off-line extensions were proposed by Görny and Ligeza (2002) or Pulido and González (2004)). Therefore, in the classic DX approach there is no form of the notation for the faults–symptoms relation and the definition of isolability.

To combine and compare DX and FDI results, new definitions were necessary. Bridging techniques are based on the concept of an analytical redundancy relation (ARR), which is a consistency relation. The set of all ARRs leads to the signature matrix (FS) which is an equivalent of the BDM used in this paper. In the work of Cordier et al. (2004) definitions of ARR-d-completeness and ARR-i-completeness were proposed. The former is related to fault detectability. The latter means that all multiple faults are isolable, which is a very strong requirement. When these properties are fulfilled, the equivalence of FDI and DX approaches can be shown.

In FDI works, ARR-exoneration (Cordier et al., 2004) is often assumed. Under this assumption, if some ARR is satisfied by observation, then any component in its support is considered normal. We will also use that assumption in this paper.

Other important definitions are related to the structural system description. Structural detectability and isolability are widely used in diagnosability analysis (Düşteğör et al., 2006; Frisk et al., 2012) and sensor placement algorithms (Krysander and Frisk, 2008). The main advantage of these definitions is their close relation to dedicated computational schemes. On the other hand, they are limited to structural or analytical system descriptions. Multi-valued diagnostic signals and symptom sequences cannot be handled.

A definition of isolability was also proposed by Travé-Massuyès et al. (2006). It is related to definitions using the residual space (Isermann, 2006). It should be noted that in this approach conditional isolability is not symmetric. This definition allows us to consider multi-valued diagnostic signals but not symptom sequences.

There are also methods to account for noise and uncertainty in a diagnostic process. For example, in the work of Eriksson et al. (2013) the Kullback–Leibler divergence is used to measure fault distinguishability. Unfortunately, it requires an analytical model of the diagnosed process.

The definition of isolability is a crucial concept in fault diagnosis, but there exists no solution covering all different approaches. The presented definitions are not complete, since distinct faults may be characterized by a specific sequence of symptom occurrence.

The purpose of this paper is to show the dependence between the source of knowledge about
the faults–symptoms relation, its notation form and the achieved fault isolability. The results of different approaches to the diagnosis process are compared. Moreover, new and more general definitions of fault isolability are presented. The extension comes from accommodating not only the signatures specified by the values of individual diagnostic signals, but also the symptom sequence. In this work, only single faults are analyzed. The results of our research can be used during the design of the FDI system. The designer of such a system needs to be able to compare results obtained with different forms of notation of the faults–symptoms relation to choose the one most suitable for his or her needs.

In Section 2 definitions of fault isolability for a binary diagnostic matrix are given. In Section 3 analysis of fault isolability with sequential residuals is discussed. Fault isolability in respect of this method is defined. In Section 4 the FIS is introduced along with corresponding generalizations of the presented isolability definitions. Section 5 gives definitions of isolability for diagnostic signal sets $S = \{ s_j : j = 1, 2, \ldots, J \}$ and faults $F = \{ f_k : k = 1, 2, \ldots, K \}$:

$$R_{F,S} \subset S \times F.$$  \hfill (3)

An example of a BDM is shown in Table 1.

A fault signature is defined as a vector of values of diagnostic signals corresponding to this fault:

$$V(f_k) = [v_{1,k} \ v_{2,k} \ \ldots \ v_{J,k}]^T.$$ \hfill (4)

Therefore, the columns of the binary diagnostic matrix are the signatures of the corresponding faults.

Other forms used for this purpose are logic functions and rules. They correspond to the columns or rows of the BDM and are an alternative notation of the same knowledge about the faults–symptoms relation. There is a rule for each signature:

$$\text{If } (s_1 = v_{1,k}) \land \cdots \land (s_j = v_{j,k}) \land (s_j = v_{j,k}) \text{ then } f_k, \quad (5)$$

while the rows are represented by the following rule:

$$\text{If } (s_j = 1) \text{ then } f_a \lor \cdots \lor f_n. \quad (6)$$

The logic functions are created analogously.

2. Fault isolability based on a binary diagnostic matrix

2.1. Binary diagnostic matrix. A model of a linear object often used for fault diagnostics is a set of equations defining the dependence of outputs on object inputs and faults (Gertler, 1998):

$$y(s) = G(s)u(s) + H(s)f(s), \quad (1)$$

where $y$ is the vector of inputs, $u$ is the vector of inputs, $f$ is the vector of faults, $G(s)$ is the matrix of input-output transfer functions, $H(s)$ is the matrix of fault-output transfer functions.

This model describes both the linear and non-linear systems in the neighborhood of a selected operating point on the static characteristic (this point usually corresponds to the nominal or average operating conditions of the system).

A BDM is the most widely used form of description of the faults–symptoms relation. It can be obtained with all methods.

If matrix $H(s)$ is known then elements of the BDM can be obtained as

$$v_{j,k} = \begin{cases} 0 & \text{if } H_{j,k}(s) = 0, \\ 1 & \text{if } H_{j,k}(s) \neq 0. \end{cases} \quad (2)$$

The BDM defines the relation between faults and binary diagnostic signals. It is a form of notation of a relationship specified as a subset of the Cartesian product of diagnostic signal sets $S = \{ s_j : j = 1, 2, \ldots, J \}$ and faults $F = \{ f_k : k = 1, 2, \ldots, K \}$:

$$R_{F,S} \subset S \times F.$$  \hfill (3)

The logic functions are created analogously.

2.2. Definition of isolability. The following definitions of unisolability and isolability of faults based on the BDM can be given.

Definition 1. Faults $f_k, f_m \in F$ are unisolable based on the BDM if, and only if, their signatures are identical, $f_k R_{U|BDM} f_m \Leftrightarrow \forall s_j \in S \left[ v_{j,k} = v_{j,m} \right]$. \hfill (7)

where $R_{U|BDM}$ denotes the unisolability relation based on BDM. Subsequent isolability and unisolability relations, for other forms of notation, are denoted analogously.

Relation $R_{U|BDM}$ divides the set of faults $F$ into subsets of unisolable faults.

Definition 2. Faults $f_k, f_m \in F$ are isolable based on the BDM if, and only if, their signatures are different, $f_k R_{I|BDM} f_m \Leftrightarrow \exists s_j \in S \left[ v_{j,k} \neq v_{j,m} \right]$. \hfill (8)

This condition is analogous to the weakly isolating structure defined by Gertler (1998) as a structure where each fault response is nonzero and different. Gertler defined also strong fault isolability for structured residuals. However, in this work we focus on the basic understanding of fault isolability.

The maximum number of faults that can be isolated on the basis of the set of $J$ binary diagnostic signals is $2^J - 1$. When the process is in a faultless state, values
Table 1. Binary diagnostic matrix for the two-tanks system with primary nonlinear residuals.

<table>
<thead>
<tr>
<th></th>
<th>f₁</th>
<th>f₂</th>
<th>f₃</th>
<th>f₄</th>
<th>f₅</th>
<th>f₆</th>
<th>f₇</th>
</tr>
</thead>
<tbody>
<tr>
<td>s₁</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>s₂</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>s₃</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Diagnosed faults in an example object.

<table>
<thead>
<tr>
<th>Fault</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₁</td>
<td>Clog between tanks 1 and 2</td>
</tr>
<tr>
<td>f₂</td>
<td>Clog in the output of tank 2</td>
</tr>
<tr>
<td>f₃</td>
<td>Leak in tank 1</td>
</tr>
<tr>
<td>f₄</td>
<td>Leak in tank 2</td>
</tr>
<tr>
<td>f₅</td>
<td>Faulty measurement of F</td>
</tr>
<tr>
<td>f₆</td>
<td>Faulty measurement of L₁</td>
</tr>
<tr>
<td>f₇</td>
<td>Faulty measurement of L₂</td>
</tr>
</tbody>
</table>

Fig. 1. Two-tank system. F: input flow to tank 1, L₁, L₂; water levels in tank A₁, A₂; surface areas in tanks 1 and 2.

of diagnostic signals are zero. In practice, the number of faults distinguished on the basis of such a set is usually smaller.

The basic method of increasing fault isolability is the generation of secondary residuals. The method of designing secondary residuals depends on whether or not the internal form of primary residuals is known (Gertler, 1998).

**Example 1.** As an example, we shall analyze the two-tank system (Fig. 1). It can be described by the following equations:

\[
A₁ \frac{d(L₁ + f₆)}{dt} = F + f₅ - f₃ - A₁ \frac{d(L₁ + f₆)}{dt} \\
- \alpha₁₂(S₁₂ - f₁) √2g(L₁ + f₆ - (L₂ + f₇)), \tag{9}
\]

\[
A₂ \frac{d(L₂ + f₇)}{dt} = \alpha₁₂(S₁₂ - f₁) √2g(L₁ + f₆ - (L₂ + f₇)) \\
- \alpha₂(S₂ - f₂) √2g(L₂ + f₇), \tag{10}
\]

where \(\alpha₁₂ S₁₂\) denotes a flow constant between tanks 1 and 2.

There are three measurements: the input flow to tank 1, \(F\), and the water levels in two tanks, \(L₁, L₂\). The faults considered are presented in Table 2.

From Eqns. (9) and (10), we conclude that the residuals are

\[
r₁ = F + f₅ - f₃ - A₁ \frac{d(L₁ + f₆)}{dt} \\
- \alpha₁₂(S₁₂ - f₁) √2g(L₁ + f₆ - (L₂ + f₇)) \tag{11}
\]

and

\[
r₂ = \alpha₁₂(S₁₂ - f₁) √2g(L₁ + f₆ - (L₂ + f₇)) \\
- \alpha₂(S₂ - f₂) √2g(L₂ + f₇) \\
- f₄ - A₂ \frac{d(L₂ + f₇)}{dt}. \tag{12}
\]

We can also obtain the third residual by combining \(r₁\) with \(r₂\):

\[
r₃ = r₁ + r₂ \\
= F + f₅ - \alpha₂(S₂ - f₂) √2g(L₂ + f₇) \\
- f₄ - A₂ \frac{d(L₂ + f₇)}{dt} \\
- A₁ \frac{d(L₁ + f₆)}{dt}. \tag{13}
\]

The binary diagnostic matrix describing the system is shown in Table 1. We assume that a binarized value of residual \(r_i\) is a diagnostic signal \(s_i\).

Using residuals \(r₁, r₂\) and \(r₃\) from Table 1 and Definitions 1 and 2, the following sets of unisolable faults can be distinguished: \(\{f₁, f₂, f₄\}, \{f₃, f₅\}, \{f₆, f₇\}\). Faults from different sets are isolable.

An alternative form of those residuals can be obtained by linearization and the Laplace transform:

\[
r₁(s) = -L₁(s) + \frac{k₁}{T₁s + 1}F(s) + \frac{k₂}{T₁s + 1}L₂(s) \\
- f₆(s) + \frac{k₁}{T₁s + 1}f₅(s) + \frac{k₂}{T₁s + 1}f₇(s) \\
+ \frac{k₃}{T₁s + 1}f₁(s) - \frac{k₄}{T₁s + 1}f₃(s), \tag{14}
\]
Table 3. Binary diagnostic matrix for the two-tank system with primary non-linear residuals and secondary residual $r_4$.

<table>
<thead>
<tr>
<th>$r_1$</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>$f_3$</th>
<th>$f_4$</th>
<th>$f_5$</th>
<th>$f_6$</th>
<th>$f_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_1$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$r_2$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$r_3$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$r_4$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

\[
r_2(s) = -L_2(s) + \frac{k_5}{T_{2s} + 1} L_1(s) = -f_7(s) + \frac{k_5}{T_{2s} + 1} f_6(s) - \frac{k_6}{T_{2s} + 1} f_1(s) + \frac{k_5 s}{T_{2s} + 1} f_2(s) - \frac{k_8}{T_{2s} + 1} f_4(s),
\]

\[
r_3(s) = -L_2(s) - \frac{k_9 s}{T_{2s} + 1} L_1(s) + \frac{k_{10}}{T_{2s} + 1} F(s) = -f_7(s) - \frac{k_9 s}{T_{2s} + 1} f_6(s) + \frac{k_7}{T_{2s} + 1} f_2(s) - \frac{k_8}{T_{2s} + 1} f_4(s) + \frac{k_{10}}{T_{2s} + 1} f_3(s) - \frac{k_{11}}{T_{2s} + 1} f_3(s).
\]

The binary diagnostic matrix and the fault information system do not require linearization and can be used as well with non-linear residuals.

By knowing the inner form of primary residuals, secondary residuals insensitive to chosen faults can be easily obtained. For example, the following residual is insensitive to $f_6$:

\[
r_4(s) = r_1(s) + \frac{T_{2s} + 1}{k_3} r_2(s) = \left( \frac{k_3}{T_{1s} + 1} - \frac{k_6}{k_5} \right) f_1(s) + \frac{k_7}{k_5} f_2(s) - \frac{k_4}{T_{1s} + 1} f_3(s) - \frac{k_8}{k_5} f_4(s) + \frac{k_1}{T_{1s} + 1} f_5(s) + \frac{k_2}{k_5} \left( \frac{T_{2s} + 1}{T_{1s} + 1} - \frac{k_6}{k_5} \right) f_7(s).
\]

This allows us to isolate $f_6$ from $r_7$, which results in the following sets of unisolable faults (Table 3): \{ $f_1$, \{ $f_2$, $f_4$ \}, \{ $f_3$, $f_5$ \}, \{ $f_6$ \}, \{ $f_7$ \}.

Unfortunately, faults \{ $f_2$, $f_4$ \} and \{ $f_3$, $f_5$ \} cannot be separated with this technique. Methods described later in this paper need to be used for that purpose.

3. Fault isolability based on sequential residuals

3.1. Sequences of symptoms and methods of their determination. The order of symptom emergence is important information, worth using in the diagnostic process. A different order of symptom emergence may be, in fact, the basis for fault isolation. The appearance of a sequence of symptoms can be determined for each fault on the basis of the transmittance $H_{j,k}$, where

\[
H_{j,k}(s) = \frac{y_j(s)}{f_k(s)}, \quad k = 1, \ldots, K.
\]

If the transmittance $H_{j,k}$ is unknown, it is often possible to use expert knowledge to determine differences in symptom sequences for some pairs of diagnostic signals for particular faults.

The times of symptom emergence depend on the dynamic properties of the tested object, the nature of the fault (abrupt, linearly increasing, etc.), and on the method and parameters of the detection algorithm.

Let us assume the occurrence of a single fault $f_k$.

The residual equation becomes

\[
r_j(s)|_{f_k} = H_{j,k}(s) f_k(s), \quad f_m = 0, \quad m = 1, 2, \ldots, K, \quad m \neq k.
\]

The time series of the residual may be determined with the knowledge of the function $f_k(t)$, based on the inverse Laplace transform:

\[
r_j(t)|_{f_k} = L^{-1} r_j(s)|_{f_k} = L^{-1}[H_{j,k}(s) f_k(s)].
\]

Assuming a form of the function $f_k(t)$ (the step function being the simplest) and a threshold value for the residual $A_j$, one can determine the time after which the $j$-th symptom of the $k$-th fault will appear. This calculation should be performed for all residuals that are sensitive to fault $f_k$. For fault functions other than the simple step function, times of symptoms appearance will be different, but the order will not change for given $A_j$.

Let us denote by $e_{S,j,p}(f_k)$ the elementary sequence, i.e., a sequence of two symptoms $j$ and $p$ for the fault $f_k$. The sequences are equal if the order of symptoms in both the sequences is the same. Sequences of some symptom couples may be defined on the basis of expert knowledge, without modeling the influence of faults on the system (Sywert and Koscielny, 2009) or using causal GP graphs (Sztyber et al., 2015). An example of such a graph is shown in Fig. 2.

3.2. Definition of isolability.

Definition 3. Faults $f_k, f_m \in F$ are unisolable (unconditionally unisolable) on the basis of elementary
residuals for two or more faults detectable by the primary residuals. In such a case, different time delays between symptoms should be chosen. Particularly, it can be assumed that \( \tau_{k,j,p} = 0 \), which means simultaneous occurrence of symptoms \( s_j \) and \( s_p \).

For delay-designed elementary sequences \( es_{j,p}^d(f_k, \tau_{k,j,p}) \), analogous definitions of unisolability and isolability of faults apply to the sequence \( es_{j,p}^d(f_k) \) (Definitions 3 and 4).

**Definition 6.** Faults \( f_k, f_m \in F \) are **unisolable** on the basis of delay-designed elementary sequences of symptoms \( es_{j,p}^d(f_k, \tau_{k,j,p}) \) if, and only if, the corresponding delay-designed elementary sequences of symptoms are identical,

\[
\forall_{s_j, s_p \in S} es_{j,p}^d(f_k) = es_{j,p}^d(f_m). \tag{24}
\]

**Definition 7.** Faults \( f_k, f_m \in F \) are **isolable** on the basis of delay-designed elementary symptom sequences \( es_{j,p}^d(f_k, \tau_{k,j,p}) \) if, and only if, there exists a pair of symptoms with a different sequence for those faults,

\[
\exists_{s_j, s_p \in S} es_{j,p}^d(f_k, \tau_{k,j,p}) \neq es_{j,p}^d(f_m, \tau_{m,j,p}). \tag{25}
\]

It is worth noting that (25) holds when \( \tau_{k,j,p} \neq \tau_{m,j,p} \) and the order of symptoms is the same, or when the order of symptoms is different.

Between Definitions 4 and 7 there is a significant difference, associated with the design of delay of symptoms. In the case of Definition 7 isolability is also obtained with an identical order of pair of symptoms but different delays. Design freedom is therefore greater. The knowledge of the internal form of residuals is, however, required.

**Example 2.** Applying a method of sequential residuals to the two-tank system given in Section 11 with a known internal form of primary residuals, we obtain the results in symptom sequences presented in Table 4.

In this example, it is assumed that a time constant of the second tank is greater than that of the first tank (\( T_1 < T_2 \)). Numbers indicate the order of symptoms. Two identical numbers in a signature mean that it is impossible to determine which symptom would be first.
Using Definitions 3 and 4, the following sets of unisolable faults can be distinguished: \( \{ f_1 \}, \{ f_2, f_4 \}, \{ f_3, f_5 \}, \{ f_6 \}, \{ f_7 \} \). Therefore, faults \( \{ f_6, f_7 \} \), which are unisolable using the BDM with primary residuals, can be isolated using sequences of symptoms.

It is possible to further improve isolability by designing secondary sequential residuals. In the discussed two-tank system, it is beneficial to introduce the following residuals:

\[
\begin{align*}
r_{1,3/5} &= \frac{k_{10}}{(T_2 s + 1)} r_1, \\
r_{3,1/5} &= \frac{k_1}{(T_1 s + 1)} r_3 e^{-\tau_{3,1}s}.
\end{align*}
\]

Residuals \( r_{1,3/5} \) and \( r_{3,1/5} \) allow us to isolate \( f_5 \) from \( f_3 \). If fault \( f_5 \) occurs, both residuals will react in the same way, delayed by \( \tau_{3,1} \).

4. Fault isolability based on the information system

4.1. Information system (FIS). Definitions of the information system and the rough information system are very helpful in defining the faults–symptoms relation. The FIS has been defined (Korbicz et al., 2004; Kościelny, 1999) as an information system in the form of the following quadruple:

\[
\text{FIS} = (F, S, V_S, q),
\]

where \( F \) denotes faults, \( S \) denotes diagnostic signals, \( V_S = \bigcup_{j \in S} V_j \) is the set of the values of diagnostic signals, \( V_j \) is the set of the values of the \( j \)-th diagnostic signal, \( q \) is the projection

\[
q : F \times S \rightarrow \Phi(V_S),
\]

assigning to each element of the Cartesian product a subset of values of diagnostic signals:

\[
q(f_k, s_j) \equiv V_{j,k} \subset V_j
\]

which can appear during the occurrence of fault \( f_k \).

For example, \( V_j = \{-1, 0, +1\} \) and \( V_{j,k} = \{-1, +1\} \subset V_j \). Furthermore, we assume that the value of the diagnostic signal \( v_j = 0 \) corresponds to a state without faults, and other values are symptoms of faults.

Then, the FIS is a form of notation determining the pattern of diagnostic signals for individual faults. It is a generalization of the BDM. If a set of all diagnostic signal values is identical and equal to \( V_s = \{0, 1\} \) and \( V_{j,k} \) is a single element set, then the FIS simplifies to the BDM. Significant extensions of the FIS towards binary diagnostic matrix are follows:

(a) every diagnostic signal may have assigned its individual set of values;
(b) a set \( V_j \) of values of the \( j \)-th diagnostic signal contains a finite (possibly greater than 2) number of elements;
(c) any set \( V_{j,k} \) of the FIS may contain one value of the diagnostic signal as well as a subset of them.

The signature of the \( k \)-th fault corresponds to the column of the FIS. It is a generalization of the signature defined by the formula 4.

\[
Q(f_k) = [ V_{1,k} \ V_{2,k} \ \ldots \ V_{J,k} ]^T.
\]

A complex signature is equivalent to the rule in the form

\[
\text{If } (s_1 \in V_{1,k}) \land \cdots \land (s_j \in V_{j,k}) \land (s_j \in V_{j,k}) \text{ then } f_k.
\]

We can define the rules corresponding to the rows of the FIS. The number of the rules corresponding to one row of the FIS is equal to the number of the values of the diagnostic signal \( v \in V_j \) different from zero.

\[
\text{If } (s_j = v_{j,k} \neq 0), \text{ then } f_n \lor \cdots \lor f_n.
\]

4.2. Definition of isolability. In the work of Kościelny et al. (2006), definitions of unconditional and conditional unisolability in the FIS were given.

Definition 8. Faults \( f_k, f_m \in F \) are unisolable (unconditionally unisolable) based on the FIS, if and only if, their signatures are identical.

\[
f_k R_{U/FIS} f_m \iff \forall s_j \in S V_{j,k} = V_{j,m}.
\]

The signatures of unconditionally unisolable faults are identical. Therefore, it is impossible to isolate those faults. In the case of the FIS and other polyvalent forms of notation, conditional unisolability is possible. Then faults are unisolable only for some, not all, values of diagnostic signals. Also, other values of diagnostic signals are possible, where the same faults are isolable.

Definition 9. Faults \( f_k, f_m \in F \) are conditionally unisolable based on the FIS if, and only if, for every signal, subsets of its values corresponding to faults \( f_k \) and \( f_m \) have a nonempty intersection and those faults are not unconditionally unisolable.

\[
f_k R_{CU/FIS} f_m \iff \forall s_j \in S V_{j,k} \cap V_{j,m} \neq \emptyset \land \exists s_j \in S V_{j,k} \neq V_{j,m}.
\]

Definition 10. Faults \( f_k, f_m \in F \) are unconditionally isolable based on the FIS if, and only if, there is a diagnostic signal for which subsets of values corresponding to those faults are disjoint

\[
f_k R_{I/FIS} f_m \iff \exists s_j \in S V_{j,k} \cap V_{j,m} = \emptyset.
\]
In the case of the polyvalent classification of diagnostic signals, it is not possible to determine isolability in an explicit way. It depends on a combination of values of those signals. In some combinations the obtained isolability is bigger, in others it is smaller.

The maximum number of failures that can be isolated on the basis of a set of trivalent diagnostic signals is $3^3 - 1$. In the general case, the maximum number of isolable faults does not exceed the product of the powers of sets $V_j$ for all diagnostic signals, i.e., $\prod_{j \in S} |V_j|$. In practice, the number of isolated faults is usually much lower. This results from physical dependencies in the process.

Polyvalent evaluation of residual values may, but does not need to, lead to increased fault isolability in comparison with binary evaluation. Usually, taking the sign of a residual (three-valued evaluation) into account increases the obtained fault isolability.

In the case of continuous diagnostic signals, described by regions in the residual space, analogous definitions for conditional and unconditional isolability can be given.

**Example 3.** For the two-tank system described in Section 1 and by considering the direction of a change for the two-tank system described in Example 3. Continuation of the example from Section 1, was analyzed.

In practice, it is not necessary to examine directions in the space of all residuals. A sufficient condition of isolability for any pair of faults is that corresponding directions in the plane defined by any two residuals differ by more than a predetermined angle.

5. Fault isolability based on directional residuals

5.1. Vectors of fault directions in the residual space. This method is referred to as directional residuals. In order to obtain fault isolation, a set of residuals is designed in such a way that the occurrence of individual faults is characterized by a particular position in the residual space (called the parity space). Each fault therefore corresponds to an individually designed directional vector (Chen and Patton, 1999; Gertler, 1998). This is illustrated in Fig. 3.

Primary directional residuals are derived from the system of equations of residuals in the internal form (29). After replacing the transfer functions $H_{j,k}(s)$ by gains of the particular residuals for faults $c_{j,k}(s)$

$$c_{j,k}(s) = \begin{cases} 0 & \text{if } H_{j,k}(s) = 0, \\ \lim_{s \to 0} H_{j,k}(s) & \text{if } H_{j,k}(s) \neq 0, \end{cases}$$

(37)

a vector of gains of residuals corresponding to a given fault $f_k$ $k = 1, \ldots, K$ defines a direction $w_k = [c_{1,k}, \ldots, c_{j,k}, \ldots, c_{f,k}]$ in the residual space, characteristic for the fault. A given fault, after the disappearance of transient states, always manifests itself in this direction. This is the basis of the directional residuals method (Chen and Patton, 1999; Gertler, 1998; Patton et al., 2000).

Fault isolation occurs after the assessment of the coincidence of the direction of the residual vector with the direction specific to individual faults as in Fig. 3.

Vectors of fault directions in the residual space are designed on the basis of object modeling, including fault influence. In theory, vectors can also be determined with learning, but this requires testing faulty objects and is almost feasible. Experts are also usually not able to determine the directional vectors for the individual faults.

5.2. Definition of isolability. The following definition of fault isolability can be formulated based on the vectors of fault directions in the residual space: Faults are isolable if directions in the residual space corresponding to them are different. Such a definition is not relevant in practice, because a small difference in directions makes faults impossible to be isolated.

In practice, it is not necessary to examine directions in the space of all residuals. A sufficient condition of isolability for any pair of faults is that corresponding directions in the plane defined by any two residuals differ by more than a predetermined angle.

Directional fault vectors $f_k$ and $f_m$ in the plane defined by the residuals $r_j$ and $r_p$ have the form $w_k = [c_{j,k}, c_{p,k}]$ and $w_m = [c_{j,m}, c_{p,m}]$. The angle between them is determined by the formula

$$\alpha = \arccos \left( \frac{C_{j,k}C_{j,m} + C_{p,k}C_{p,m}}{|w_k||w_m|} \right).$$

(38)

In the work of Kościelny and Łabęda-Grudziak (2013) double fault distinguishability for linear systems was analyzed.

**Example 4.** Continuing the example from Section 1 the direction corresponding to faults is the residual space spanned by primary residuals (29). From the analysis
Faults (internal and computational) are known. A definition for the case where both forms of residuals diagnostic signals (BDM, FIS) with signatures in the form of a relation. In particular, it is appropriate to combine two or more forms of notation for the faults–symptoms of the inner form of residuals, Table 6 can be obtained. Here transient states are omitted and only static behavior is considered.

Each fault has a different direction in the residual space. For example, using this previously unisolable method, faults $f_2$ and $f_4$ can be easily isolated since they have opposite directions.

6. Generalization of the definitions of fault isolation

Generalization of the above definitions of fault isolability can be achieved after putting together adequate conditions regarding the values of diagnostic signals (residuals) with the conditions determining symptom sequences. The condition for fault unisolability while using simultaneously two or more faults–symptoms relations may be formed as a logic product of unisolability conditions for individual forms of notation of the faults–symptoms relation, whereas the condition of isolability is a logic alternative to partitive conditions.

One can distinguish two main cases of knowledge about the diagnosed object. First, when models of the object with fault influence are known, i.e., when the internal form of residuals is known. The second is more common in practice and relates to the case when only the computational form of residuals is known. For both of these cases, fault unisolability and isolability conditions can be formulated by using simultaneously two or more forms of notation for the faults–symptoms relation. In particular, it is appropriate to combine signature-based methods specified by the model values of diagnostic signals (BDM, FIS) with signatures in the form of sequences of symptoms.

Below one can find examples of such a generalized definition for the case where both forms of residuals (internal and computational) are known.

Definition 11. Faults $f_k, f_m \in F$ are unisolable (unconditionally unisolable) based on delay-designed sequences of symptoms and the FIS if, and only if, the conditions which are the conjunction of conditions (34) and (24) are met:

$$f_k R^U_{|FIS,SEQ} f_m \iff \exists V_{j,k} \cap V_{j,m} = \emptyset$$

$$\land \left[ \forall_{s_j \in S} es^d_{j,p}(f_k, \tau_{k,j,p}) = es^d_{j,p}(f_m, \tau_{m,j,p}) \right].$$

(39)

Definition 12. Faults $f_k, f_m \in F$ are isolable (unconditionally isolable) based on delay-designed sequences of symptoms and the FIS if, and only if, the conditions which are the alternative conditions (40) and (21) are met:

$$f_k R^U_{|FIS,SEQ} f_m \iff \exists V_{j,k} \cup V_{j,m} = \emptyset$$

$$\land \left[ \forall_{s_j \in S} es^d_{j,p}(f_k, \tau_{k,j,p}) \neq es^d_{j,p}(f_m, \tau_{m,j,p}) \right].$$

(40)

For an analogous case, when only residuals in computational form are known, the following definitions can be given.

Definition 13. Faults $f_k, f_m \in F$ are unisolable (unconditionally unisolable) on the basis of the FIS and symptom sequences if, and only if, signatures and sequences of those faults are identical. Thus the conditions (34) and (21) are met:

$$f_k R^U_{|FIS,SEQ} f_m \iff \exists V_{j,k} \cap V_{j,m} = \emptyset$$

$$\land \left[ \forall_{s_j \in S} es_{j,p}(f_k) = es_{j,p}(f_m) \right].$$

(41)

Definition 14. Faults $f_k, f_m \in F$ are (unconditionally) isolable on the basis of the FIS and symptom sequences if, and only if, there is a diagnostic signal for which subsets of values corresponding to those faults are disjoint or their elementary symptom sequences are different. This means

<table>
<thead>
<tr>
<th>$f_1$</th>
<th>$f_2$</th>
<th>$f_3$</th>
<th>$f_4$</th>
<th>$f_5$</th>
<th>$f_6$</th>
<th>$f_7$</th>
<th>$V_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_1$</td>
<td>+1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>-1,0, +1</td>
</tr>
<tr>
<td>$r_2$</td>
<td>-1</td>
<td>+1</td>
<td>0</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>-1,0, +1</td>
</tr>
<tr>
<td>$r_3$</td>
<td>0</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>-1,0, +1</td>
</tr>
</tbody>
</table>

Table 6. Fault directions in the three-dimensional space of primary residuals.
combining the conditions \((36)\) and \((22)\).

\[
\begin{align*}
&f_k R_{\text{FIS,SEQ}} f_m \\
&\iff \left[ \exists s_j \in S, V_{j,k} \cap V_{j,m} = \emptyset \right] \\
&\vee \left\{ \exists s_j, s_p \in S, [es_{j,p}(f_k) = \langle s_j, s_p \rangle] \\
&\land [es_{j,p}(f_m) = \langle s_p, s_j \rangle] \right\}.
\end{align*}
\]

(42)

What results from the above definitions is the fact that utilizing knowledge (usually incomplete) about a defined by sample values of diagnostic signals. In comparison to reasoning only on the basis of signatures symptom sequence allows increased fault isolability in the occurrence, or even the same one, but with a different delay, allows fault isolation, which is impossible when analyzing only the values of diagnostic signals.

The knowledge of the sequence of symptoms is utilized, as well as using expert knowledge to determine diagnoses. An alternative or complementary method of object modeling, taking into account fault influence, expert knowledge, and learning.

The primary method for increasing fault isolability is obviously increasing the number of the measured signals and the generation of additional residuals as a consequence. Not always the extension of a set of the measured signals is possible or economically justified.

Another possibility is designing secondary residuals. This offers great possibilities when the internal form of residuals is known and when models with fault influence are available. There are also some possibilities for the generation of secondary residuals when the only computational form is known; however, they are rather limited. An alternative or complementary method of increasing isolability can be utilization of polyvalent or continuous residual evaluation instead of binary one. Exchanging the BDM with the FIS and utilizing trivalent computational form is known; however, they are rather limited. An alternative or complementary method of increasing isolability can be utilization of polyvalent or continuous residual evaluation instead of binary one. Exchanging the BDM with the FIS and utilizing trivalent computational form is known; however, they are rather limited.

Another possibility is designing secondary residuals. This offers great possibilities when the internal form of residuals is known and when models with fault influence are available. There are also some possibilities for the generation of secondary residuals when the only computational form is known; however, they are rather limited. An alternative or complementary method of increasing isolability can be utilization of polyvalent or continuous residual evaluation instead of binary one.

Exchanging the BDM with the FIS and utilizing trivalent computational form is known; however, they are rather limited. An alternative or complementary method of increasing isolability can be utilization of polyvalent or continuous residual evaluation instead of binary one.

7. Summary

The paper presented formal conditions of fault isolability and unisolability on the basis of the binary diagnostic matrix, the knowledge of the sequence of emerging symptoms and the FIS information. Those relations between faults and symptoms are designed based on object modeling, taking into account fault influence, expert knowledge, and learning.

The primary method for increasing fault isolability is obviously increasing the number of the measured signals and the generation of additional residuals as a consequence. Not always the extension of a set of the measured signals is possible or economically justified.

Another possibility is designing secondary residuals. This offers great possibilities when the internal form of residuals is known and when models with fault influence are available. There are also some possibilities for the generation of secondary residuals when the only computational form is known; however, they are rather limited. An alternative or complementary method of increasing isolability can be utilization of polyvalent or continuous residual evaluation instead of binary one. Exchanging the BDM with the FIS and utilizing trivalent computational form is known; however, they are rather limited. An alternative or complementary method of increasing isolability can be utilization of polyvalent or continuous residual evaluation instead of binary one.

This approach can be used for object modeling including fault influence, as well as using expert knowledge to determine the FIS.

An additional possibility for increased isolability results from including the knowledge on symptom sequence emergence. A different order of symptom occurrence, or even the same one, but with a different delay, allows fault isolation, which is impossible when analyzing only the values of diagnostic signals. The knowledge of the sequence of symptoms is complementary to the knowledge of the signatures of faults contained in the BDM or the FIS, and the knowledge of the directions of faults in the residual space. The conclusion is that the methods of reasoning about faults based on different forms of notation of the faults–symptoms relation are worth using. They should be integrated into diagnostic systems.

Acknowledgment

This work was partially supported by the National Science Center in Poland under the project DEC-2011/01/B/ST7/06183 and by the Warsaw University of Technology, Faculty of Mechatronics Dean’s Grant 504/02141.

References


Fault isolability with different forms of the faults–symptoms relation


Kościelny, J.M. (1999). Application of fuzzy logic for fault isolation in a three-tank system, 14th IFAC World Congress, Beijing, China, pp. 73–78.


Jan Maciej Kościelny received his M.Sc. (in industrial automation), Ph.D. (in machine construction and operation) and D.Sc. (in automatic control and robotics) degrees as well as the professorial title from the Warsaw University of Technology in 1973, 1980, 1991 and 2003, respectively. His research interests include the diagnostics of industrial processes and fault tolerant systems. He is a member of the Committee on Automation and Robotics of the Polish Academy of Sciences as well as the Technical Committee: TC 6.4. on Fault Detection, Supervision & Safety of Technical Processes SAFEPROCESS of the IFAC, the editorial board of the International Journal of Applied Mathematics and Computer Science, and the Programme Committee of the journal Measurement Automation and Monitoring.
Michał Syfert received the M.Sc. degree in automatic control and robotics as well as the Ph.D. degree in machine construction and exploitation from the Warsaw University of Technology (WUT), Warsaw, Poland, in 1997 and 2004, respectively. Since 2005, he has been with the Institute of Automatic Control and Robotics of that university. His professional experience is mainly related to the development and applications of diagnostic systems for industrial processes. The main fields of his interests include applications of theory and artificial intelligence methods for advanced monitoring and diagnostics of complex industrial installations.

Kornel Rostek was born in ˙Zyrardów, Poland, in 1988. He received an M.Sc. degree in automatic control and robotics from the Warsaw University of Technology (WUT), Warsaw, Poland, in 2012. Since 2012, he has been with the Institute of Automatic Control and Robotics of that university. The main fields of his interests include development of methods for advanced diagnostics of industrial installations and measures of isolability.

Anna Sztyber was born in Warsaw, Poland, in 1986. She received an M.Sc. degree in automatic control and robotics in 2010, a Ph.D. degree in 2015, and an M.Sc. degree in a mathematics in 2013 from the Warsaw University of Technology. She is currently with the Institute of Automation Control and Robotics of that university as an assistant professor. Her research interests include applications of qualitative models in fault diagnosis, industrial data compression, and applications of remote sensing in environmental monitoring.