Some theoretical results on sensor placement for diagnosis based on fault isolability specifications

Mattias Krysander and Erik Frisk
Dept. of Electrical Engineering, Linköping University
SE-581 83 Linköping, Sweden
{matkr,frisk}@isy.liu.se

February 1, 2007

This report presents the theoretical results of the work in [3]. The report is not self-contained and should be considered to be complementary to the paper.

1 Theoretical Background

The objective of this section is to introduce notation and theoretical tools regarding structural models and bipartite graphs that will be used in proofs of results in coming section.

The structural representation of a set of equations M with unknown variables X is a bipartite graph with variables and equations as node sets. There is an edge in the graph between a node representing an equation $e \in M$ and node representing an unknown variable $x \in X$ if the variable x is contained in e. For notational convenience, we will denote the node representing an equation e or a variable x simply by the equation name e and the variable name x respectively. A bipartite graph can be described by a biadjacency matrix where the rows and columns correspond to the node sets and the position (i, j) is one if there is an edge between node i and j, and a zero otherwise.

1.1 Dulmage-Mendelsohn decomposition

In the analysis of these graphs, we will frequently use the Dulmage-Mendelsohn decomposition [2] which is illustrated in Figure 1. The decomposition defines a partition $(M_0, M_1, \ldots, M_n, M_\infty)$ of the set of equations M, a similar partition of the set of unknowns X, and a partial order on the sets M_i. If the rows and columns are rearranged according to this order, the biadjacency matrix has the form shown in Figure 1. There are zero entries in the white parts of the matrix and there might be ones in the gray-shaded parts. Three main parts of M can be identified in the partition, M_0 is called the structurally underdetermined part,

$\cup_{i=1}^{n} M_i$ is the structurally just-determined part, and M_∞ is the structurally overdetermined part.

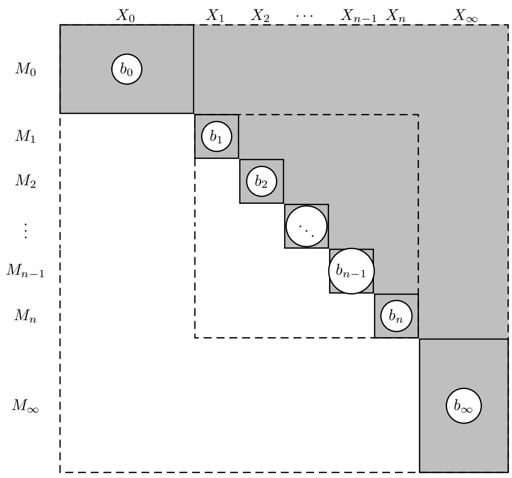

Figure 1: Dulmage-Mendelsohn decomposition

Next we will describe how these partitions can be defined and to do this we need to introduce a couple of definitions. Let $|A|$ denote the cardinality of the set A. Given a bipartite graph with node sets M and X, let the variables in $E \subseteq M$ be denoted by $\mathrm{var}(E)$ and the surplus of the equation set E be defined by

$$\varphi(E) = |E| - |\mathrm{var}(E)|$$

Given a model M, there is a family of subsets of M with the maximum surplus:

$$\mathcal{L} = \{E \subseteq M | \varphi(E) \geq \varphi(E'), \forall E' \subseteq M\} \tag{1}$$

Let $E_0 \supset E_1 \supset \cdots \supset E_{n-1} \supset E_n$ be any maximal descending chain of $\mathcal{L}$, then the partition of M is defined as $M_0 = M \setminus E_0$, $M_i = E_{i-1} \setminus E_i$ for $i = 1, \ldots, n$, and $M_\infty = E_n$, see [5]. The partition of X is defined as

$$X_i = \mathrm{var}(M_i) \setminus \mathrm{var}(E_i) \tag{2}$$

for $i \in \{0, 1, \ldots, n\}$ and $X_\infty = \mathrm{var}(M_\infty)$.

The partial order $\leq$ can be defined on the sets M_i by

$$M_i \leq M_j \quad \Leftrightarrow \quad (M_j \subseteq E \in \mathcal{L} \Rightarrow M_i \subseteq E) \tag{3}$$

In the figure, each pair (M_i, X_i) is related to a block which is denoted by b_i. Since there is a one-to-one correspondence between the sets M_i and the blocks b_i, we will also partially order the blocks b_i in the same way:

$$b_i \leq b_j \quad \Leftrightarrow \quad M_i \leq M_j \tag{4}$$

1.2 Structural formulation of fault diagnosis

In this section, we will give structural characterizations of fault diagnosis properties. By doing this, the sensor placement problem can be formulated as a graph theoretical problem.

A fault f is detectable if there exists an observation that is consistent with fault mode f and inconsistent with the no-fault mode. This means that a detectable fault can violate a monitorable equation in the model describing fault free behavior. An equation is, in the generic case, monitorable in an equation set M if it is contained in the structurally overdetermined part of M [1]. If the structurally overdetermined part of a set of equations M is denoted by M^+, then the structural characterization of detectability can be defined as follows.

Definition 1. *A fault f is structurally detectable in a model M if $e_f \in M^+$.*

Detection is a special case of isolation, i.e. a fault is detectable if the fault is isolable from the no-fault mode. By noting this similarity, it holds that a fault f_i, isolable from f_j, can violate a monitorable equation in the model describing the behavior of the process having a fault f_j. The equations valid with a fault f_j is $M \setminus \{e_{f_j}\}$ and the monitorable part of these equations is, in the generic case, equal to $(M \setminus \{e_{f_j}\})^+$. This motivates the following structural characterization of isolability.

Definition 2. *A fault f_i is structurally isolable from f_j in a model M if*

$$e_{f_i} \in (M \setminus \{e_{f_j}\})^+ \tag{5}$$

2 Theorems and proofs

The lemmas and theorems formulated in [3] are in this section proved.

Lemma 1. *Let M be an exactly determined set of equations, b_i a strongly connected component in M with equations M_i, and $e \notin M$ an equation corresponding to measuring any variable in b_i. Then*

$$(M \cup \{e\})^+ = \{e\} \cup (\cup_{M_j \leq M_i} M_j) \tag{6}$$

Proof. The proper overdetermined part $(M \cup \{e\})^+$ is defined by the minimal subset of $M \cup \{e\}$ with maximum surplus. The maximum surplus of all subsets of M is 0. By adding one equation e, we know that the maximum surplus of any subset of $M \cup \{e\}$ is at most 1. Since $\mathrm{var}(\{e\}) \subseteq \mathrm{var}(M)$, it follows that $\varphi(M \cup \{e\}) = 1$. Hence the minimal set with surplus 1 is the proper overdetermined part of $M \cup \{e\}$. Any such set contains e since all other sets have surplus less or equal to 0. This means that the sought set can be written as $E \cup \{e\}$ where $E \subseteq M$. Since the surplus of $E \cup \{e\}$ is one and the surplus of E can be at most 0, it follows that the surplus of E is 0. Let $\mathcal{L}$ be a sub-lattice of the subset-lattice of M defined similar to the set defined in (1). This means that $E \in \mathcal{L}$. Furthermore $\varphi(E \cup \{e\}) = 1$ only if $\mathrm{var}(\{e\}) \subseteq \mathrm{var}(E)$. This implies that $M_i \subseteq E$. The minimal set E in $\mathcal{L}$ such that $M_i \subseteq E$ is according to (3) $E = \cup_{M_j \leq M_i} M_j$ and this completes the proof. $\square$

Theorem 1. *Let M be an exactly determined set of equations, F the corresponding set of faults, $P \subseteq X$ the set of possible sensor locations, and M_S the equations corresponding to adding a set of sensors S. Then maximal detectability of faults F in $M \cup M_S$ are obtained if and only if S has a non-empty intersection with $D([f])$ for all $[f] \in F_m$ where F_m is the set of maximal fault classes among the fault classes with $D([f]) \neq \emptyset$.*

Proof. First, note that faults in fault classes with $D([f]) = \emptyset$ can not be made detectable with any of the available sensor locations. Therefore, let F_m be, among the fault classes with $D([f]) \neq \emptyset$, the set of maximal elements with respect to the partial order. Then maximal fault detectability is obtained if and only if the fault classes in F_m are detectable. This follows from Lemma 1 and Definition 1 which states that if a sensor is added such that a fault in a higher ordered fault class is detected, detectability for the lower ordered fault classes is also obtained.

Furthermore, Lemma 1 also states that a fault f in $\mathcal{F}$ becomes detectable if and only if we measure at least one unknown variable in blocks that are greater or equal than the block that includes the fault equation, i.e. if we measure a variable in $D([f])$. A sensor addition that makes all faults in F detectable must thus have a non-empty intersection with $D([f])$ for all $[f] \in F_m$. $\square$

Theorem 2. *Let M be a set of equations with no structurally underdetermined part, F a set of structurally detectable faults in M, $P \subseteq X$ the set of possible sensor locations, and M_S the equations added by adding the sensor set S. For an arbitrary fault f_j, assume that M^0 is the just-determined part of $M \setminus \{e_{f_j}\}$, F^0 is the faults contained in M^0, and $\mathcal{D} = \texttt{Detectability}(M^0, F^0, P)$. Then the maximum possible number of faults $f_i \in F \setminus \{f_j\}$ are structurally isolable from f_j in $M \cup M_S$ if and only if S have a non-empty intersection with all sets in $\mathcal{D}$.*

Proof. Given a sensor set S, a fault f_i is structurally isolable from f_j in the model $M \cup M_S$ if

$$e_{f_i} \in ((M \setminus \{e_{f_j}\}) \cup M_S)^+ \tag{7}$$

according to Definition 2. This is equivalent to say that f_i is structurally detectable in $(M \setminus \{e_{f_j}\}) \cup M_S$. Since all faults are structurally detectable, it follows that $e_{f_j} \in M^+$. This implies that the underdetermined part of $M \setminus \{e_{f_j}\}$ is empty. The faults in the structurally overdetermined part of $M \setminus \{e_{f_j}\}$ are according to Definition 1 structurally detectable. From Theorem 1, maximal detectability of faults F^0 in the structurally just-determined part M^0 of $M \setminus \{e_{f_j}\}$ is obtained if and only if S has a non-empty intersection with all detectability sets contained in $\mathcal{D} = \texttt{Detectability}(M^0, F^0, P)$. $\square$

A key property in the determination of structural isolability is the set $(M \setminus \{e_{f_j}\})^+$ which is determined by the result of the combined operation of removing an equation and then computing the overdetermined part. The resulting set of the combined operation has been studied in [4] and can be characterized as follows. There exists a partition $(M_1, M_2, \ldots, M_p)$ of the overdetermined part M^+ such that for any equation $e \in M_k$, it holds that

$$(M \setminus \{e\})^+ = M^+ \setminus M_k \tag{8}$$

Theorem 3. *Given a model M, let f_i and f_j be two structurally detectable faults in M. The fault f_i is structurally isolable from f_j if and only if e_{f_i} and e_{f_j} belong to different sets in the partition defined in* (8)*.*

Proof. The fault f_i is structurally isolable from f_j if and only if (5) holds according to Definition 2. By using (8), (5) can be expressed as

$$e_{f_i} \in M^+ \setminus M_k \tag{9}$$

where M_k is the set in the partition such that $e_{f_j} \in M_k$. Since f_i is structurally detectable, i.e. $e_{f_i} \in M^+$, it follow that (9) is equivalent to $e_{f_i} \notin M_k$ and this completes the proof. □

Theorem 4. *Let M be a model with no underdetermined part and let $x \in var(M)$ be measured with a sensor described by an equation $e \notin M$. Then, a sensor fault violating e will be structurally detectable in $M \cup \{e\}$.*

Proof. The sensor fault is structurally detectable if $e \in (M \cup \{e\})^+$. Since there is no underdetermined part in M, it follows that $\varphi(M)$ is equal to the maximal surplus for any set contained in M. The maximal surplus of any set in $M \cup \{e\}$ is $\varphi(M)+1$. Any set with surplus $\varphi(M)+1$ have to include e, and especially the minimal set of the maximal surplus $\varphi(M)+1$. This implies that $e \in (M \cup \{e\})^+$ which was to be proved. □

Theorem 5. *Let M be a model with no underdetermined part and F a set of structurally detectable faults in M. Furthermore, let M_S be an equation set describing additional sensors and F_S the associated set of sensor faults. Then for any sensor fault $f \in F_S$ and for any fault $f' \in (F \cup F_S) \setminus \{f\}$, it holds that f is isolable from f' and f' is isolable from f in $M \cup M_S$.*

Proof. The faults in F are detectable by condition and the faults in F_S are detectable according to Theorem 4. Since both f' and f are structurally detectable it is sufficient to show that f' is structurally isolable from f in $M \cup M_S$ according to Theorem 3.

First, assume that $f' \in F$. All faults in F are structurally detectable and it follows that f' is structurally detectable, i.e.

$$e_{f'} \in M^+ \tag{10}$$

From the fact that $M \subseteq (M \cup M_S) \setminus \{e_f\}$, it follows that $M^+ \subseteq ((M \cup M_S) \setminus \{e_f\})^+$. This and (10) imply that $e_{f'} \in ((M \cup M_S) \setminus \{e_f\})^+$, i.e. f' is structurally isolable from f according to Definition 2.

Finally, assume that $f' \in F_S \setminus \{f\}$. From Theorem 4, we get that f' is structurally detectable in $M \cup \{e_{f'}\}$, i.e.

$$e_{f'} \in (M \cup \{e_{f'}\})^+ \tag{11}$$

From the fact that $M \cup \{e_{f'}\} \subseteq (M \cup M_S) \setminus \{e_f\}$, it follows that $(M \cup \{e_{f'}\})^+ \subseteq ((M \cup M_S) \setminus \{e_f\})^+$. This and (11) imply that $e_{f'} \in ((M \cup M_S) \setminus \{e_f\})^+$, i.e. f' is structurally isolable from f according to Definition 2 and this completes the proof. $\square$

References

[1] M. Blanke, M. Kinneart, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer-Verlag, 2003.

[2] A. L. Dulmage and N. S. Mendelsohn. Coverings of bipartite graphs. *Canadian Journal of Mathematics*, 10:517–534, 1958.

[3] E. Frisk and M. Krysander. Sensor placement for maximum fault isolability. In *Proceedings DX-2007*, Nashville,USA, submitted to DX-2007.

[4] Mattias Krysander, Jan Åslund, and Mattias Nyberg. An efficient algorithm for finding minimal over-constrained sub-systems for model-based diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, Accepted for publication.

[5] K. Murota. *Matrices and Matroids for System Analysis*. Springer-Verlag, 2000. ISBN 3-540-66024-0.