A Study of Diagnosability in Dynamic Systems: Integral and Derivative Causality

Hamed Khorasgani[1] and **Gautam Biswas** [2]

[1]Vanderbilt University and Institute for Software Integrated Systems, Nashville, TN, USA
email: hamed.g.khorasgani@vanderbilt.edu
[2] Vanderbilt University and Institute for Software Integrated Systems, Nashville, TN, USA
email: gautam.biswas@vanderbilt.edu

ABSTRACT

This paper attempts a systematic analysis of diagnosability when comparing derivative versus integral causal forms of dynamic system models. We formally define the notions of structural detectability and structural isolability, and use these definitions in a case study where we compare these factors for a three-tank system model. We prove that integral and derivative causality can lead to different isolability properties and by using derivate and integral causality forms of the models concurrently we can maximize fault isolability of the system.

1 Introduction

The control systems-based FDI community and AI-based DX community have developed a number of methodologies for fault detection and isolation in dynamic systems using structural models [1-6]. Structural methods for diagnosis have a number of advantages. These methods are based on the interconnectedness of system components and variables, and numerical values of the parameters do not affect the design of the diagnostic process. Therefore, structural methods are more general and easier to implement, and they can be easily applied to different systems and domains. Also, many common complexities in other methods like existence of answer or singularity of the matrices in the design process do not directly affect the structural methods.

The DX community has developed several structural approaches for diagnosis, e.g., the Temporal Causal Graph (TCG) approach which derives Qualitative Fault Signatures (QFS) [2] and the Possible Conflicts (PC) approach that exploits local redundancy in system measurements [3]. In parallel, the FDI community has developed ARR schemes (e.g., [4]) based on system equations. One of these is the Diagnostic Bond Graph approach that derives ARRs in an organized manner from system bond graph models [1]. Since system equations can be reformulated in many different ways to exploit the analytic redundancy relations for fault detection and isolation, researchers have often restricted the form of the equations used to integral or derivative causality forms. The integral causality form of the models has advantages that they are more robust to measurement noise, but they have the disadvantage that the initial system state has to be known for residual generation. The derivative causality models do not require the initial state, but computing derivatives in noisy environments is hard, and may lead to false alarms, and incorrect diagnoses.

The diagnosis community has often debated if the method of representing the model of a dynamic system affects the inferred diagnosability of the system. On the surface this may not appear to be the case, because irrespective of how the constraints are represented, the collective set of equations defines the dynamic behavior of the system. In this paper, we undertake a preliminary investigation of this topic, by comparing the diagnosability of system models represented in integral and derivative causality – the two most common forms of representation for dynamic system models.

Recently, Frisk, et al. [7] presented through an example that using different forms of causality in the system equations can lead to different isolability properties. The results of this paper is a primary motivation for this paper, i.e., study in more detail how different forms of causality can lead to different isolability properties and how we can utilize the information provided by integral and derivative causality models concurrently to maximize isolability in the system. In this paper, we make an attempt to make these results more systematic.

The rest of this paper is organized as follows. In section 2, we formally present the notions of integral and derivative causality forms of system behavior models. The three tank system is used as a case study to demonstrate the various concepts and the results derived in this paper. In section 3, we first formally define the notions of detectability and isolability in dynamic systems. Then Possible Conflicts (PC), a structural diagnosis approach developed by Pulido and Alonso [3] is utilized to demonstrate the diagnosability results on the three tank system using integral and derivative causality. In previous work [10], we have demonstrated the equivalence between the PC and ARR approaches to diagnosis.

Section 5 discusses an approach where the derivative and integral causality forms of the dynamic system models are integrated to achieve the highest levels of diagnosability in a system, for a given set of sensors. Finally, Section 6 discusses the advantages and drawbacks of using the two forms of causality concurrently, and concludes by laying out a brief plan for future work.

2 Test Case: A Three Tank System

To study fault detectability and isolability in dynamic systems a simple three tank system model shown in Figure 1 is considered as a running example for this paper.

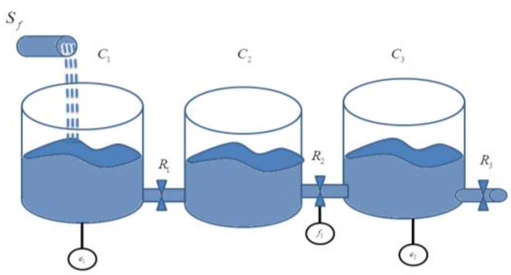

Figure 1: Three tank system configuration.

The three tank system consists of a flow input source (S_f), three tanks (C_1, C_2 and C_3) and three valves (R_1, R_2 and R_3). Also three sensors measure the pressure of the first tank (e_1), the volumetric flow rate of the second valve (f_5) and the pressure of the last tank (e_3). Six possible faults in the system components (R_1, R_2, R_3, C_1, C_2 and C_3) are considered in this study.

This section first reiviews Bond Graph modeling language and its relationship with structral fault diagnostability and then uses bond graph model to derive system equations in integral and derivative causalities.

2.1 Bond Graph Modeling Methodoloy

Bond Graphs are topological, energy based, modeling methodology for physical processes [8]. Nodes in the bond graphs represent elements of the dynamic system and directional links or bonds show energy path and its positive direction between the elements. Effort and flow are the energy variables in the bond graph language. They represent different variables in different domains. For instance in the hydraulic domain effort represents pressure and flow is volumetric flow rate. System elements in the bond graphs are modeled as sources (source of effort and source of flow), energy storage elements (capacities and inertias) and dissipative elements (resistors). Two ideal junctions (0− and 1−junctions) are also defined in the bond graph modeling language.

Series (1−)junctions are common flow junctions and parallel (0−)junctions are common effort junctions. The bond graph model of the three tank system in Figure 1 is shown in Figure 2. There is an effort and a flow variable associated with each bond (half arrow) in the bond graph. The product of the effort with flow variable represents the rate of energy flow between connected components. In Figure 2 the bonds connected to a common effort junction (0–junction) only one effort variable and for the bonds connected to a common flow junction (1–junction) only one flow variable is independent, and determines the value of that variable on all incident bonds.

Mosterman and Biswas [2] define the set of hypothesized faults candidates for a system as the parameters of components of its bond graph model. In [5] bond graphs are used to derive ARRs and [9] suggests a method to derive PCs from the bond graph model of the system. Based on these similarities we utilize bond graph models as a common framework for comparing diagnosability of different structural methods [10]. In bond graphs, causality is represented by a vertical bar at the side of link, and this defines the side of the effort receiver. For example, in Figure 2 R_1 represents a resistor with input effort, and, consequently, the flow value is the output defined by the equation associated with the resistor, $f_3 = \frac{e_2}{R_1}$. There are four kinds of causalities in the bond graphs. (1)-Fixed causality :for sources and nonlinear elements which can accept one form of causal model. (2) Constrained causalities for junctions: 0–junctions can accept effort only from one of the connected links and similarly, 1–junctions can accept flow only from one of the connected links. (3) Preferred causality: for energy storage elements. If we prefer integral causality, we assign the causality in a way that capacities receive flow and inertias receive effort. For derivative preferred causality, capacities receive effort and inertias receive flow. (4) Arbitrary causality for dissipative elements: Assigning a specific causality form to the bond graph represents the computational order that the model should be solved. For example, in figure 2 the assigned causality to R_1 informs the simulator to compute e_2, and then use e_2 to compute f_3. In the next two subsections the equations of the three tank system model are derived in integral and derivative causality forms.

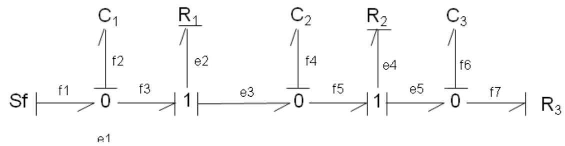

Figure 2 Bond graph of 3 tank system in integral causality.

2.2 Deriving System Equations: Integral versus Derivative Causality in Dynamic Systems

To derive the equations of the system in integral causality form we assign integral causality to the bond graph bonds as shown in Figure 2. Then for each element and junction we write the associted equation based on the assigned cauasality, as shown below:

$$
\begin{aligned}
&f_2 = f_1 - f_3 && e_5 = \frac{1}{C_3}\int f_6 \\
&e_1 = \frac{1}{C_1}\int f_2 && f_7 = \frac{e_5}{R_3} \\
&e_2 = e_1 - e_3 && e_4 = e_3 - e_5 \\
&f_3 = \frac{e_2}{R_1} && f_5 = \frac{e_4}{R_2} \\
&f_4 = f_3 - f_5 && f_6 = f_5 - f_7. \\
&e_3 = \frac{1}{C_2}\int f_4 &&
\end{aligned}
\tag{1}
$$

For modeling dynamic systems usually integral causality is prefered for several reasons. The most important one is that derivative operator is not causal, i.e., to calculate the derivation of a variable at time t its value at the next sample time is needed.

$$\frac{de(t)}{dt} = \lim_{\Delta t \to 0} \frac{e(t+\Delta t) - e(t)}{\Delta t}. \tag{2}$$

However, the simulator can wait for one sample time and calculate the derivative. So it is possible to simulate three tank systems in derivative causality as well. To derive the equations of the system in derivative causality we assign derivative causality to the bond graph model of the system. Figure 3 shows the bond graph model of the three tank system in derivative causality. Set of equations of the elements and junctions of the bond graph of three tank system in derivative causality, form the set of system equations in derivative causality as represented in (3).

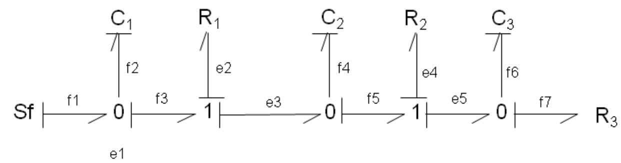

Figure 3 Bond graph of three tank system in derivative causality.

$$
\begin{aligned}
&f_3 = f_1 - f_2 && e_3 = e_4 + e_5 \\
&f_2 = C_1 \frac{de_1}{dt} && e_4 = R_2 f_5 \\
&e_1 = e_2 + e_3 && f_7 = f_5 - f_6 \\
&e_2 = R_1 f_3 && f_6 = C_3 \frac{de_5}{dt} \\
&f_5 = f_3 - f_4 && e_5 = R_3 f_7. \\
&f_4 = C_2 \frac{de_3}{dt} &&
\end{aligned}
\tag{3}
$$

To complete the set of system equations input and outputs variables are expressed as:

$$u = f_1 \tag{4}$$

$$y_1 = e_1, \quad y_2 = f_5, \quad y_3 = e_5. \tag{5}$$

Combinations of integral and derivative causality, i.e., mixed causality is not considered in this paper. In the next section we will see how integral and derivative causalities can lead to different fault isolation properties.

3 Fault Detection and Isolation: Possible Conflicts Approach

Possible conflicts (PCs) are localized methods for structural diagnosis. Given a set of system equations, PCs capture all the minimal subsets of constrains which produce inconsistencies when faults occur. Any constraint in the system which contains fault components and has sufficient analytical redundancy to capture faults is a possible conflict. In the first part of this section, detectable and isolable faults are defined. Then possible conflicts method is used in integral and derivative causality forms to isolate the faulty components in the three tank system.

3.1 Detectability and Isolability in Dynamic Systems

In this section, we start with relevant definitions.

Definition 1: Given the set of component faults $F = \{f_1, ..., f_n\}$ and the set of possible conflicts

$P=\{p_1,...,p_m\}$, the fault connection matrix (FCM) is a $m*n$ matrix and its elements are defined as:

$$FCM(i,j)=\begin{cases}0 & if\ p_j\ does\ not\ include\ f_i\ component\\ 1 & if\ p_j\ includes\ f_i\ component.\end{cases}$$

Using definiton 1 we can define detectable and isolatable faults [10].

Definition 2: A fault parameter f_i is structurally detectable, if there exists a non-zero entry in the FCM for at least one possible conflict.

Definition 3: Fault parameter f_i and f_j are structurally isolable from each other if they have different signatures in the FCM.

Using these definitions the isolability of different component faults for integral and derivative causality is analyzed in the following subsections.

3.2 Diagnosability of System Faults using PC in Integral Causality

Fault detection and isolation starts with the system dynamic model. In this subsection, the set of system equation in integral causality (1) is applied to derive Possible Conflicts. Consider $\frac{1}{D}$ as the integral operator we have:

$$e_2=e_1-e_3=e_1-\frac{f_4}{C_2D}=e_1-\frac{\frac{e_2}{R_1}-f_5}{C_2D} \quad (6)$$

$$\Rightarrow e_2=\frac{R_1C_2D}{(R_1C_2D+1)}(e_1+\frac{f_5}{C_2D})$$

and $$e_1=\frac{1}{DC_1}f_2=\frac{1}{DC_1}(f_1-f_3)=\frac{1}{DC_1}(f_1-\frac{e_2}{R_1}). \quad (7)$$

By substituting (6) in (7) we get:

$$e_1=\frac{1}{DC_1}(f_1-\frac{R_1C_2D}{(R_1C_2D+1)R_1}(e_1+\frac{f_5}{C_2D})). \quad (8)$$

Since e_1, f_1 and f_5 are all known, equation (8) is a possible conflict. From (1) one can also conclude that:

$$e_3=\frac{f_4}{DC_2}=\frac{f_3-f_5}{DC_2}=\frac{\frac{e_2}{R_1}-f_5}{DC_2}=\frac{e_2-R_1f_5}{DC_2R_1}=\frac{e_1-e_3-R_1f_5}{DC_2R_1}$$

$$\Rightarrow e_3=\frac{e_1-R_1f_5}{(1+DC_2R_1)}, \quad (9)$$

and:

$$f_5=\frac{e_4}{R_2}=\frac{e_3-e_5}{R_2}. \quad (10)$$

By substituting (9) in (10) one can get:

$$f_5=\frac{e_1-R_1f_5-e_5(1+DC_2R_1)}{R_2(1+DC_2R_1)}. \quad (11)$$

Since e_1, e_5 and f_5 are all known (11) is also a possible conflict. For e_5 we can say:

$$e_5=\frac{f}{DC_{3\ 6}}=\frac{f_5-f_7}{DC_3}=\frac{f_5-\frac{e_5}{R_3}}{DC_3}=\frac{R_3f_5-e_5}{DC_3R_3}. \quad (12)$$

Here also e_5 and f_5 are both known so equation (12) is also a conflict. Based on (8), (11) and (12) we can derive FCM as:

Table 1: FCM in Integral Causality

	R_1	R_2	R_3	C_1	C_2	C_3
e_1	*1*	*0*	*0*	*1*	*1*	*0*
f_5	*1*	*1*	*0*	*0*	*1*	*0*
e_5	*0*	*0*	*1*	*0*	*0*	*1*

Using FCM, one can derive the maximum isolability matrix as:

Table 2: Maximum Isolability Matrix in Integral Causality

	R_1	R_2	R_3	C_1	C_2	C_3
R_1	×				×	
R_2		×				
R_3			×			×
C_1				×		
C_2	×				×	
C_3			×			×

Table 2 shows R_1 is not isolatable from C_2 and R_3 is not isolatable from C_3 but all the other faults are isolatable from each other.

3.3 Diagnosability of System Faults using PC in Derivative Causality

To perform fault detection and isolation in derivative causality we start with the set of equations in derivative causality (3) and consider the same set of inputs (4) and sensors (5) as the integral causality model. Representing the derivation operator as D and doing some algebraic manipulations we have:

$$e_1=e_2+e_3=R_1f_3+e_4+e_5=R_1(f_1-C_1De_1)+R_2f_5+e_5. \quad (13)$$

Since e_1, f_5 and e_5 are all known so equation (13) is a possible conflict. We also have:

$$f_5 = f_3 - f_4 = f_1 - f_2 - C_2 De_3 = f_1 - C_1 De_1 - C_2 D(e_4 + e_5)$$
$$= f_1 - C_1 De_1 - C_2 D(R_2 f_5 + e_5), \quad (14)$$

where e_1, f_1, f_5 and e_5 are all known variables and (14) is also a possible conflict. Finally, we can say:

$$e_5 = R_3 f_7 = R_3(f_5 - f_6) = R_3(f_5 - C_3 De_5). \quad (15)$$

It is clear that (15) is a possible conflict as well. Table 3 shows the FCM for the three-tank system in derivative causality form. The maximum isolability matrix for derivative causality is presented in Table 4.

Table 3: FCM in Derivative Causality

	R_1	R_2	R_3	C_1	C_2	C_3
e_1	*1*	*1*	*0*	*1*	*0*	*0*
f_5	*0*	*1*	*0*	*1*	*1*	*0*
e_5	*0*	*0*	*1*	*0*	*0*	*1*

Table 4: Maximum Isolability Matrix in Derivative Causality

	R_1	R_2	R_3	C_1	C_2	C_3
R_1	×					
R_2		×		×		
R_3			×			×
C_1		×		×		
C_2					×	
C_3			×			×

In this case, R_1 is isolatable from C_2 but R_2 is not isolatable from C_1 and R_3 and C_3 are still non isolatable. The example is quite interesting because it shows changing the causality can change the isolability of the faults. A method to isolate maximum possible faults from each other is suggested in the next section.

4 Concurrent Causality

In this section, we use the definitions presented in section 3.1 and the FCM derived in sections 3.2 and 3.3 to show if we consider derivative and integral causality simultaneously, we can improve the isolability of the faults. Then we use this theoretical discussion to suggest fault isolation using concurrent causality.

4.1 Fault Isolability in Dynamic Systems

Isolability information provided in integral and derivative causalities is discussed in the following theorem.

Theorem: Structural isolability information provided by considering derivative and integral causalities together (concurrent causality) provides equal or more isolability than by considering just derivative or just integral causality.

Proof: To prove this theorem we first prove that the isolability information provided in derivative and integral causality are not the same. To prove this part we simply use contradiction. Assume that isolability information provided by derivative and integral causality are the same. If we can provide one example where integral and derivative causalities lead to different solutions, the first part is proved. Consider the three-tank system in figure 1 with the set of measurements provided in (5). One can see from Table 2 that in integral causality faults in R_1 and C_2 are not isolatable from each other. But Table 4 shows these faults are isolable in derivative causality. Also from Table 4 it can be seen that faults in R_2 and C_1 are not isolable from each other in derivative causality but Table 2 shows they are isolable in integral causality. So it is proved that derivative and integral causality provide different isolability information and the information provided by each of them is not a subset of the other. Therefore, by using both of them we can get more or in some possible cases equal information of using one of them.

4.2 Fault Diagnostic in the Three-tank System using Concurrent Causality

Based on the theoretical discussions in the previous subsection we expect that by using concurrent causality we achieve to the maximum isolability in our case study. Figure 4 shows fault diagnostic structure using concurrent causality.

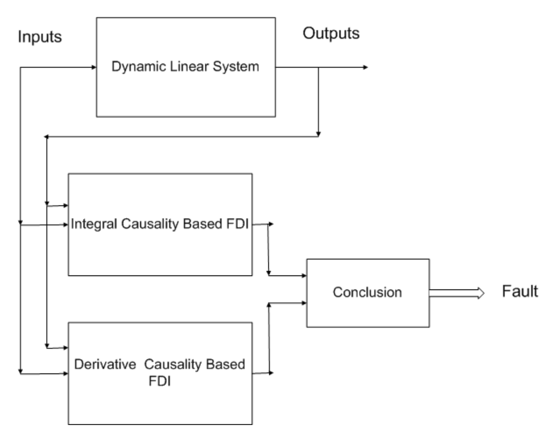

Figure 4 Concurrent causalities.

We know that non of the integral and derivate causality does not provide maximum isolability information. So to isolate maximum faults concurrent causality is applied. To this end equations (8), (11), (12) and equations (13), (14), (15) are all used to derive possible conflicts. In this case FCM is:

Table 8: FCM in Concurrent Causality

	R_1	R_2	R_3	C_1	C_2	C_3
$e_1(i)$	1	0	0	1	1	0
$f_5(i)$	1	1	0	0	1	0
$e_5(i)$	0	0	1	0	0	1
$e_1(d)$	1	1	0	1	0	0
$f_5(d)$	0	1	0	1	1	0
$e_5(d)$	0	0	1	0	0	1

where index i represents integral and index d represents derivative causality. The maximum isolability matrix in this case would be:

Table 10: Maximum Isolability Matrix in Concurrent Causality

	R_1	R_2	R_3	C_1	C_2	C_3
R_1	×					
R_2		×				
R_3			×			×
C_1				×		
C_2					×	
C_3			×			×

It can be seen from Table 10 that the faults C_1 and R_2 or R_1 and C_2 are structurally isolatable in this method.

5 Conclusions

In this paper, we showed that by considering integral and derivative causality simultaneously we can isolate some faults which were not isolatable in just integral or derivative case. However, the important point is that each of the integral and derivative causalities has some limitations. For example, derivative causality is not proper for noisy environments because possibly we have to get derivation from some measurements and in noisy environments it can lead to huge errors and false alarms. On the other hand, integral causality may lead to unstable residuals. Also considering both of the causalities needs twice computations. So in the cases that we do not have these limitations concurrent causality could be considered as an interesting tool to isolate faults as much as possible. In future work, we will study mixed causality and discus which mixed causality can provide maximum fault isolation in the system.

REFERENCES

[1] Samantaray, A. K., & Bouamama, B. O. (2008). Model-based process supervision: a bond graph approach. Springer.

[2] Mosterman, P. J., & Biswas, G. (1999). Diagnosis of continuous valued systems in transient operating regions. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *29*(6), 554-565.

[3] Pulido, B., & González, C. A. (2004). Possible conflicts: a compilation technique for consistency-based diagnosis. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *34*(5), 2192-2206.

[4] Staroswiecki, M., & Comtet-Varga, G. (2001). Analytical redundancy relations for fault detection and isolation in algebraic dynamic systems. *Automatica*,*37*(5), 687-699.

[5] Samantaray, A. K., Medjaher, K., Ould Bouamama, B., Staroswiecki, M., & Dauphin-Tanguy, G. (2006). Diagnostic bond graphs for online fault detection and isolation. *Simulation Modelling Practice and Theory*, *14*(3), 237-262.

[6] Krysander, M. (2006). Design and analysis of diagnosis systems using structural methods (Doctoral dissertation, Linköping).

[7] Frisk, E., Bregon, A., Aslund, J., Krysander, M., Pulido, B., & Biswas, G. (2012). Diagnosability analysis considering causal interpretations for differential constraints. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *42*(5), 1216-1229.

[8] Karnopp, D. C., Margolis, D. L., & Rosenberg, R. C. (2012). System dynamics: modeling, simulation, and control of mechatronic systems. Wiley.

[9] Bregon, A., Biswas, G., & Pulido, B. (2012). A decomposition method for nonlinear parameter estimation in TRANSCEND. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 42(3), 751-763.

[10] Bregon. A., Biswas, G, Pulido, B, Alonso-Gonzalez, C & Khorasgani, H, (2013) "A Common Framework for Compilation Techniques Applied to Diagnosis of Linear Dynamic Systems" Manuscript submitted for publication.