Generation of Signed Directed Graphs Using Functional Models

Reinartz, Christopher Clarc; Kirchhübel, Denis; Ravn, Ole; Lind, Morten

Published in:
IFAC-PapersOnLine

Link to article, DOI:
10.1016/j.ifacol.2019.09.115

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

INTERNATIONAL FEDERATION
OF AUTOMATIC CONTROL

ICONS 2019

5th IFAC Conference on
Intelligent Control and Automation Sciences

21-23 August 2019, Belfast, Northern Ireland
http://www.qub.ac.uk/icons2019

PREPRINTS

ICONS 2019 Sponsors

IFAC TC Main-Sponsor

TC3.2 Computational Intelligence and Control

IFAC TC Co-Sponsors

TC1.1 Modelling, Identification and Signal Processing
TC1.2 Adaptive and Learning Systems
TC3.1 Computers and Control
TC3.3 Telematics: Control via Communication Networks
TC4.3 Robotics
TC4.5 Human Machine Systems
TC7.5 Intelligent Autonomous Vehicle

Sponsors:

We gratefully acknowledge our sponsors: Visit Belfast, Belfast City Council, Mathworks, Elsevier, and the Centre for Intelligent Autonomous Manufacturing Systems at Queen's University Belfast.

Generation of Signed Directed Graphs Using Functional Models ⋆

Christopher Reinartz * **Denis Kirchhübel** ** **Ole Ravn** *** **Morten Lind** ****

* *Department of Electrical Engineering Technical University of Denmark (e-mail: ccrein@elektro.dtu.dk)*
** *Department of Electrical Engineering Technical University of Denmark (e-mail: dekir@elektro.dtu.dk)*
*** *Department of Electrical Engineering Technical University of Denmark (e-mail: or@elektro.dtu.dk)*
**** *Department of Electrical Engineering Technical University of Denmark (e-mail: mli@elektro.dtu.dk)*

Abstract: Intelligent fault diagnosis systems can be a major aid to human operators charged with the high-level control of industrial plants. Such systems aim for high diagnostic accuracy while retaining the ability to produce results that can be interpreted by human experts on site. Signed directed graphs have been shown to be a viable method for plant-wide diagnosis that can incorporate both quantitative information about the process condition as well as qualitative information about the system topology and the functions of its components. Their range of application in industrial settings has been limited due to difficulties regarding the interpretation of results and consistent graph generation. This contribution addresses these issues by proposing an automated generation of signed directed graphs of industrial processes in the chemical, petroleum and nuclear industries using Multilevel Flow Modeling; a functional modeling method designed for operator support. The approach is demonstrated through a case study conducted on the Tennessee Eastman Process, showing that Multilevel Flow Modeling can facilitate a consistent modeling process for signed directed graphs. Finally, the resulting benefits regarding qualitative reasoning for plant-wide diagnosis are discussed.

Keywords: Human supervisory control, Fault diagnosis, Directed graphs, Intelligent knowledge-based systems, Decision support systems, Human-centered design.

1. INTRODUCTION

The operation of industrial plants in the chemical and petroleum industries poses a significant challenge to human operators on site. Especially the operation in abnormal plant states, which is characterized by e.g. the failure of one or multiple components, can be problematic. The production loss due to the operation in such abnormal plant states, which accounts for up to 18% of the total production loss (Crowl and Louvar, 2011), could be severely reduced if correct recovery strategies were executed by the operators in time. One of the main reasons for delayed or incorrect reactions by the operators during plant recoveries is that the information provided to them is not tailored to support quick, informed decision making. The alarm systems currently featured in industrial plants provide a descriptive overview of the plant state, featuring alarm signals, which indicate a deviation of the process from its nominal condition. Information about the potential relation of these alarms is not displayed (Rothenberg, 2009). It is imperative to provide the operators on site with a decision support system, which can generate comprehensive and contextual information about the current plant state to improve the overall performance of operations on industrial plants. This information should include probable root causes of existing disturbances and potential actions countering their effects. Due to the high degree of connectivity in most industrial plants, a local fault may propagate through large parts of the system. Most observed disturbances and alarms are in fact the result of such a propagation and do not necessarily give any indication about the actual fault. A system for fault-diagnosis in large-scale processes needs to capture the causal connections of the system to be able to reason about the origins of faults. Additionally, the diagnosis must be comprehensible for the human operators on site, if the software is intended for decision support (Yang et al., 2012).

Model-based (Venkatasubramanian et al., 2003a,b) and process history based methods (Venkatasubramanian et al., 2003c) for fault-diagnosis have been presented in the past. This paper will focus on Multilevel Flow Modeling (MFM) and Signed Directed Graphs (SDG). Both methods belong to the field of qualitative, model-based analysis and both possess the capability to capture causal connections within complex industrial processes. Research into improving the results obtained from SDG-based methods using quantitative approaches has been conducted with promising results

⋆ This work is supported by the Danish Hydrocarbon Research and Technology Centre.

(Maurya et al., 2007; Yang et al., 2012; Wan et al., 2013; Peng et al., 2014). A drawback of using SDGs is that both the model generation and the interpretation of obtained results is not straightforward. Multilevel Flow Modeling is capable of producing results that can be interpreted by process experts, but the combination of the method with quantitative modeling approaches, though possible (Kim and Seong, 2018; Hu et al., 2015; Larsson et al., 2004), has not been researched extensively yet. A mapping between SDG and MFM models offers the potential to make results obtained by one method accessible to the other and thereby extend the application range of both methods. The basic concepts of both methods are outlined in Sections 2 and 3, followed by an explanation of the method for the automated signed directed graph generation in Section 4. The method is tested on the Tennessee Eastman process in Section 5 and the results are discussed in Section 6.

2. REPRESENTING CAUSAL RELATIONS USING SIGNED DIRECTED GRAPHS

Signed directed graphs provide a generally applicable means of representing qualitative causal models (Venkatasubramanian et al., 2003a). A mathematical expression for SDG models $G = (V, E, \phi, \psi)$ is defined by Bondy and Murty (1976). The nodes (V) of a SDG represent system variables and the arcs (E) connecting the nodes represent the effect of these variables on each other. Each node has an assigned qualitative state $\psi : V \rightarrow \{+, 0, -\}$, which indicates whether the state of the variable represented by the node is higher, equal or lower than nominal. Each directed arc has either a positive or a negative sign $\phi : E \rightarrow \{+, -\}$, which is determined by the direction of effect between two variables. The direction of the arc is determined by the cause-effect relation of the connected nodes. Each arc points from the 'cause' node to the 'effect' node. Figure 1 displays the two basic connection types that can be expressed using signed directed graphs. States are propagated in SDGs by traversing the edges connecting the nodes. A state $\psi(v_i) =$ '+' will propagate to all nodes directly connected to v_i. Nodes connected via positive arcs will assume a high state (+) and nodes connected via negative arcs will assume a low state (-). The benefit of SDGs is that causal dependencies are explicitly expressed through the signed edges. Because of that, the state propagation is straight forward and easily traceable. SDG can furthermore be applied very broadly, since the modeling language is not derived from processes of any specific industry. A drawback of using SDG is that the direct representation of causalities does not enable a direct inference about the reason for the causality. This limits the usefulness of obtained results for further utilization by human experts.

Fig. 1. Basic positive (a) and negative (b) connection types between two nodes of a signed directed graph.

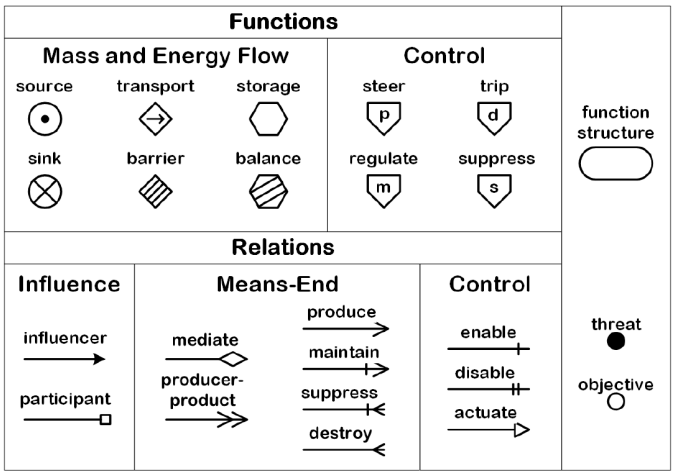


Fig. 2. Basic MFM Functions and Relations (Zhang, 2015).

3. MULTILEVEL FLOW MODELING

Multilevel Flow Modeling is a functional modeling method designed to model industrial processes such as nuclear and chemical plants. The benefit of MFM is that both the models and the reasoning results can be interpreted by process experts in the field, since the modeling language is based on concepts that they are already familiar with, such as mass and energy flows. MFM models are meant to capture the functionality and the causal relations between system parts rather than the topology on a component level. In Multilevel Flow Models, processes are divided into mass, energy, and control flow structures. These contain more detailed process representations, which are modelled using the basic function- and relation types displayed in Fig. 2. Each basic function type is assigned a qualitative state, which corresponds to commonly used terminology for alarm states used in industrial applications. The functions Storage, Transport, Source and Sink can assume the states high, normal and low. High and low states signify that the function variable is outside of its nominal range. The barrier function can assume a normal or a breach-state. The balance function has only a normal state and serves as a flow-distribution function. Lind (2011) provides an overview of the basic model components and modeling principles.

MFM is different from other qualitative modeling techniques like SDGs because causal relations are expressed implicitly. The causal relation between two MFM-functions is determined by the respective function types (Storage, Sink, Source, Balance, Transport), the connection types (influencer, participant) and the position of the functions towards each other, considering the direction of flow in the system (upstream, downstream). The direction of flow in the system is indicated by the arrow in the transport function symbol. Transport-type functions (transport, barrier) affect adjacent functions by default.

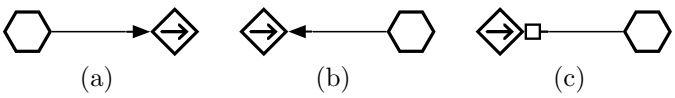


Fig. 3. Influencer (a,b) and participant (c) relations between "storage" and "transport" functions.

Table 1. Inferences for MFM-models in Fig. 3

Scenario	Initial state	Inference
Case "a"	transport: high	storage: low
	storage: high	transport: high
Case "b"	transport: high	storage: high
	storage: high	transport: low
Case "c"	transport: high	storage: high
	storage: high	transport: not affected

Non-transport functions (storage, sink, source, balance) affect adjacent functions, if they are connected via an influencer relation. Figure 3 shows three MFM models, which represent different explicit causal relations. The cases of a "high" state of the transport function and a "high"-state of the storage function are considered as examples. Table 1 summarizes the causal inference for the models in Fig. 3. In the case shown in Fig. 3a, a "high" state of the transport signifies an outflow of the storage, which is higher than expected. The result is a decrease and eventual "low" state of the level of the storage. The explicit cause-effect relation from the transport to the storage therefore has a negative sign. In the cases provided in Figures 3b and 3c, a "high" state of the transport function will lead to an eventual "high" state in the storage, since the storage will fill faster than expected due to the higher than nominal input. The explicit cause-effect relation from transport to storage has a positive sign in these cases. The case 3c is different from 3b, because it features a participant relation, which signifies that the state of the storage does not affect the transport function directly. Reasoning about causality in MFM requires the implementation of a fixed set of propagation-rules for each combination of MFM-functions (Zhang et al., 2013).

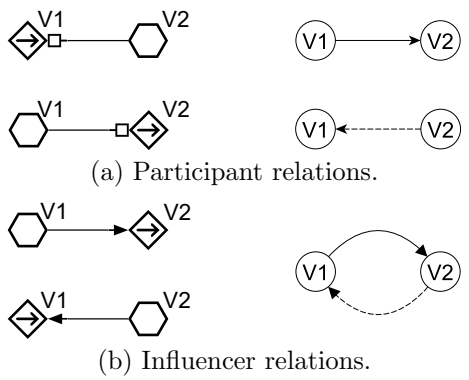
4. CONVERSION OF MFM MODELS TO SIGNED DIRECTED GRAPHS

The aim of the conversion is a signed directed graph that captures the explicit causal relations, which are implicitly expressed in the MFM model. Such a graph can be generated in three steps, if an MFM model is available. First, n SDG nodes are created, where n equals the amount of functions in the MFM model and each node corresponds to a specific function. The explicit causalities between the MFM-functions are then extracted by using the rules described by Zhang et al. (2013) in the second step. The signed edges of the SDG can be generated based on the information about explicit causalities that was extracted in step two, as shown in Fig. 4 for the most basic MFM models. Fig. 4a shows participant relations between transport and storage functions. It can be observed that the nodes representing the storage function in the SDG have no outgoing edges, since the participant relation implicitly states that the state of the storage does not have a direct effect on the state of the transport function. The lack of outgoing edges on the nodes representing the storages states the same relation explicitly. The same model with influencer instead of participant relations is shown in Fig. 4b. In this case both MFM models result in an identical signed directed graph representation, which is a direct result of the MFM-reasoning that is illustrated in Table 1, considering that the node "V1" corresponds to the storage function in the upper model and to the transport function in the lower model in Fig. 4b. It provides a good example for why Multilevel Flow Models are better suited for human interpretation than signed directed graphs. The MFM models in Fig. 4b represent two different physical processes, e.g. a tank being drained by an outflow (top) and a tank being filled by an inflow (bottom), which is not apparent in the signed directed graph representation. Because the signed directed graph does not capture and thus does not contain such implicit information about the process, it is not possible to generate a MFM model from a SDG directly. Reasoning results obtained from SDGs can, however, be transferred to MFM models if a mapping between the nodes of both models is established, which is always the case if the SDG is generated from an MFM model.

4.1 Reduction of Signed Directed Graphs

As mentioned above, a primary purpose of MFM is the representation of the process in a human-readable format. This necessitates that elements of the process, which are not monitored but still vital for the human operator's understanding, have to be considered in MFM models to ensure a comprehensible output. Signed directed graphs do not necessarily need to keep to this restriction. In most cases only process variables which are monitored or can be actuated are considered in SDG representations, since they can provide direct feedback to the reasoning system. Signed directed graphs generated using MFM will initially feature all variables that are captured in the MFM model, including those that are not monitored. The MFM representation does, however, include information about which MFM-functions are directly connected to process measurements using a "process variable" tag, thus defining a set of monitored nodes I. This information can be used to reduce generated signed directed graphs to exclusively include nodes representing monitored variables. The applied reduction scheme consisting of two main steps is described in Fig. 5 and illustrated by a simple example in Fig. 6.

Nodes, whose indegree $\eth^-$ or outdegree $\eth^+$ is equal to zero are removed in the first step, unless they represent a process variable. The remaining nodes that are not connected to process variables are recursively removed in



(a) Participant relations.

(b) Influencer relations.

Fig. 4. Representation of Signed Directed Graph equivalents of basic MFM relations.

Step 1:
  while $v_i \in V \setminus I;\ \eth^-(v_i) = 0 \vee \eth^+(v_i) = 0$ **do**
    $V \leftarrow V \setminus v_i$
  end while
Step 2:
  while $v_i \in V \setminus I$ **do**
    for $e^+ \in E;\ e^+ = (v_m\, v_i)$ **do**
      for $e^- \in E;\ e^- = (v_i\, v_n)$ **do**
        if $v_n \neq v_m$ **then**
          $e^* = (v_m\, v_n)$
          $E \leftarrow E \cup e^*$
        end if
      end for
    end for
    $V \leftarrow V \setminus v_i$
  end while

Fig. 5. Reduction algorithm used to obtain a graph featuring only monitored nodes. V, I, and E are defined as the sets of nodes, monitored nodes and edges, respectively.

the second step. A similar reduction scheme was previously described by Kramer and Palowitch (1987). The reduced graph represents the same causal relations as the original in regard to the measured process variables, but representations of internal connections between unmeasured nodes may be lost. This is acceptable if the target application uses the process data as its primary source of information.

5. CASE STUDY: GENERATING A SIGNED DIRECTED GRAPH OF THE TENNESSEE EASTMAN REACTOR

The Tennessee Eastman process is a well-known process from the chemical industries that was first introduced by Downs and Vogel (1993). It has since been used as a benchmark problem for studies in fault-diagnosis (Yin et al., 2012), state estimation (Ricker and Lee, 1995), control (Zerkaoui et al., 2010) and others. The five main components of the process are the reactor, condenser, vapour-liquid separator, stripper and the compressor. The open loop process presented by Downs and Vogel (1993) is unstable and needs to be shut down due to high reactor pressure, if left unchecked. Ma and Li (2017) adopt a control scheme presented by DAngelo et al. (2016) to test the performance of a fuzzy signed directed graph model of the reactor of the TE-process for fault diagnosis. The graph created by Ma and Li (2017) is used as a reference model for this case study, since it has been successfully applied for fault-diagnosis for the Tennessee Eastman reactor and serves as a good benchmark. The flowsheet and control scheme are displayed in Fig. 8.

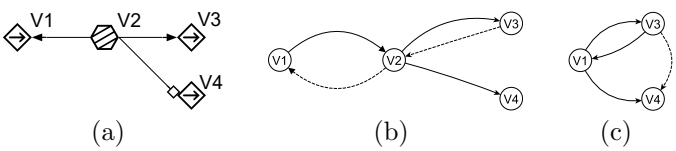


Fig. 6. Multilevel Flow Model (a), equivalent Signed Directed Graph representation (b) and SDG representation excluding "V2" (c) of a hypothetical process with branching causal relations.

The aim of this case study is to compare a SDG that was generated using a MFM model of the Tennessee Eastman reactor to a reference SDG presented by Ma and Li (2017) and reach conclusions about the merit of the method proposed in this article based on the consistency between the MFM model and the generated SDG on the one hand and the similarity between the generated and reference SDGs on the other hand.
The MFM model used as a basis for the graph generation is displayed in Fig. 9. It was designed from the knowledge represented in the flow sheet using the modeling strategy presented by Lind (2017). The graph shown in Fig. 7 is the direct result of the application of the principles presented in Sections 4 and 4.1. The consistency of MFM and generated SDG is tested by comparing the results of single fault propagation applied to both graphs. To this end, all scenarios resulting from a single fault input are generated using the propagation rules for MFM and SDG, respectively. It is then verified that nodes referring to the same process variable have the same state in all corresponding scenarios. This test was successfully run for all single fault scenarios (positive and negative deviation of each process variable) of the generated Tennessee Eastman Reactor model.
It is apparent that the generated (black edges) and reference (black and red edges) graphs are very similar. The automatically generated graph does not contain additional edges compared to the reference. The reference graph contains four edges and thus for causal relations which are not covered by the generated graph. It was expected that these relations would not be represented in the generated model, since they take the effect of the stripper and vapour-liquid separator into account, which have not been considered in the MFM model that focuses on the functionality of the reactor. It is noticeable that the generated graph does not feature any edges which imply "false" causal relations, which is important, since invalid causalities can lead to incorrect reasoning and thus impede the fault analysis.

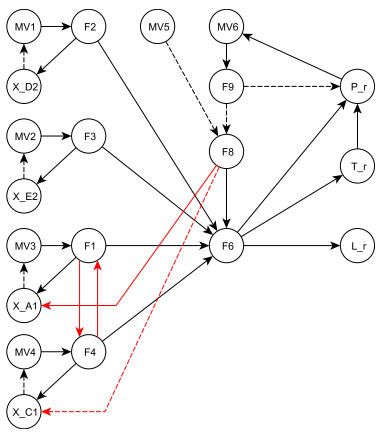


Fig. 7. SDG describing the causalities between measurable process variables that affect the TE-Reactor. Red edges: Causalities described by Ma and Li (2017) that do not appear in the MFM-generated SDG.

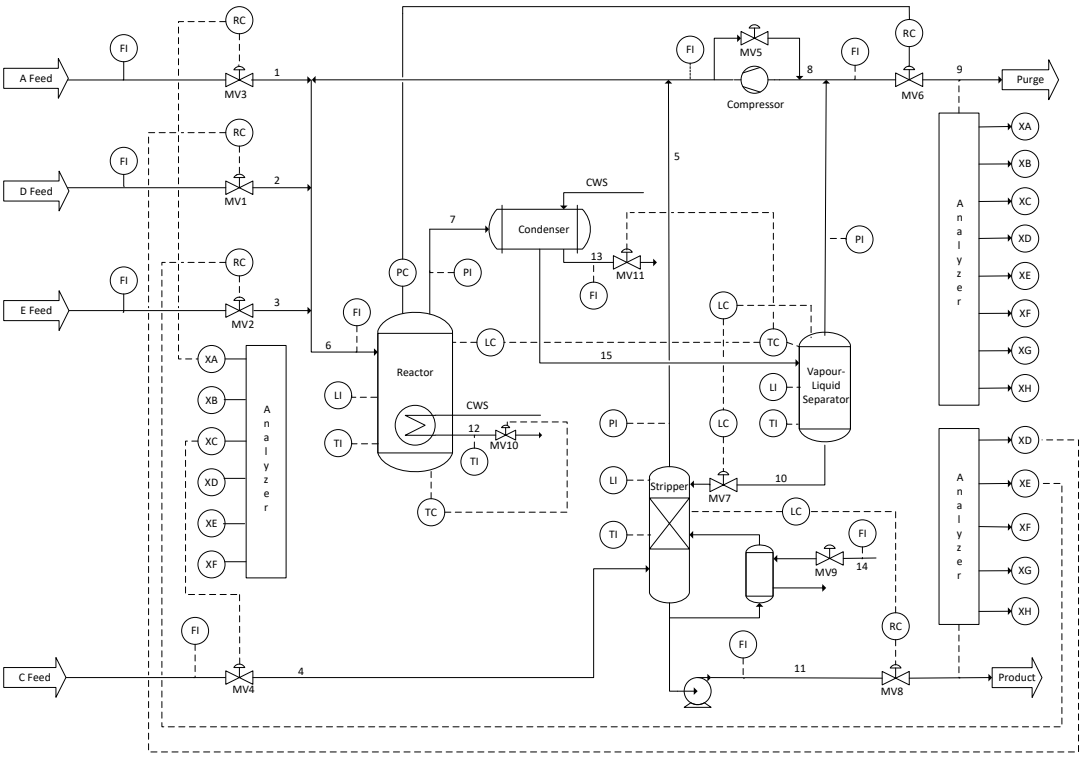


Fig. 8. Flowsheet of the Tennessee Eastman process containing the control structure used by Ma and Li (2017).

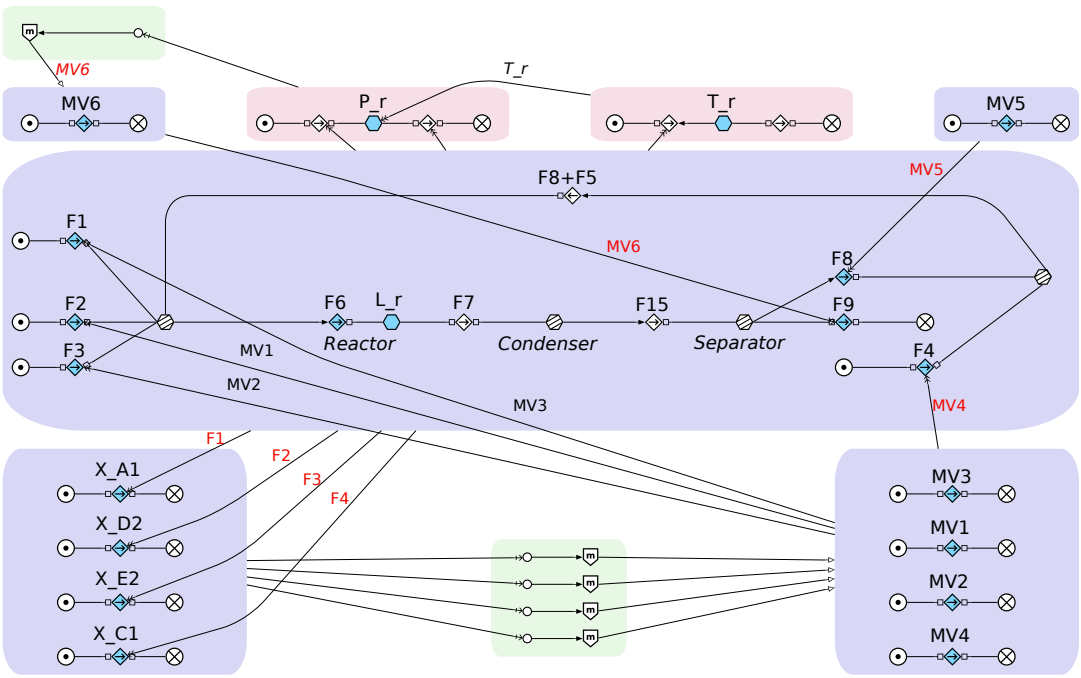


Fig. 9. Multilevel Flow Model of the Tennessee Eastman Reactor. Displayed are functional elements of the process which have either direct or indirect influence on the reaction. Measurable process variables are highlighted blue. Red print indicates the variables belonging to means-end relations. Process variable names are directly adopted from the used reference study conducted by Ma and Li (2017). The process is modelled using the MFM-modeling principles descibed by Lind (2017).

6. CONCLUSION

The conversion of functional process models to signed directed graphs is generally achieved by translating implicit causalities to explicit ones, as presented in Section 4. Such an approach is realizable for many functional modeling concepts, e.g. some of those presented by Erden et al. (2008), indicating that the presented method is in principle applicable for other approaches than Multilevel Flow Modeling as well.

Researchers focusing on signed directed graph based fault analysis profit from using functional models in two ways. They can use the existing modeling guidelines developed for the functional modeling concepts that are designed to identify causal structures in specific industrial applications. Following such guidelines ensures consistent model design and thereby consistent design of the signed directed graphs used for the analysis. In addition to that, the opportunity to express the diagnosis output in a functional modeling framework designed for the evaluation by process experts facilitates communication in the respective target industry.

A major benefit for researchers in the field of functional modeling is that the extensively researched field of fault diagnosis using signed directed graphs becomes directly accessible. This provides the opportunity to integrate such research into the development of functional reasoning systems with minimal effort.

The method presented in Section 4 facilitates the addressed consistent model generation for processes in the chemical, petroleum and nuclear industries. Initial testing on the Tennessee Eastman process has yielded promising results, including a automatically generated, well-formulated signed directed graph model of the Tennessee Eastman reactor which closely matches models presented in literature. Further case studies on systems from the chemical and petroleum industries as well as further research on the integration of signed directed graphs into the Multilevel Flow Modeling concept are planned for the near future.

REFERENCES

Bondy, J.A. and Murty, U.S.R. (1976). *Graph Theory with Applications*. Elsevier, New York.

Crowl, D.A. and Louvar, J.F. (2011). *Chemical Process Safety - Fundamentals with Applications*. Paul Boger, Prentice Hall, third edition.

DAngelo, M., Palhares, R.M., Camargos Filho, M.C., Maia, R.D., Mendes, J.B., and Ekel, P.Y. (2016). A New Fault Classification Approach Applied to Tennessee Eastman Benchmark Process. *Appl. Soft Comput.*,

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

Erden, M., Komoto, H., van Beek, T., D'Amelio, V., Echavarria, E., and Tomiyama, T. (2008). A review of function modeling: Approaches and applications. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*,

Hu, J., Zhang, L., Cai, Z., and Wang, Y. (2015). An intelligent fault diagnosis system for process plant using a functional HAZOP and DBN integrated methodology. *Engineering Applications of Artificial Intelligence*,

Kim, S.G. and Seong, P.H. (2018). Enhanced reasoning with multilevel flow modeling based on time-to-detect and time-to-effect concepts. *Nuclear Engineering and Technology*,

Kramer, M.A. and Palowitch, B.L. (1987). A rulebased approach to fault diagnosis using the signed directed graph. *AIChE Journal*,

Larsson, J.E., Ahnlund, J., Bergquist, T., Dahlstrand, F., Öhman, B., and Spaanenburg, L. (2004). Improving expressional power and validation for multilevel flow models. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*,

Lind, M. (2011). An introduction to multilevel flow modeling. *Nuclear Safety and Simulation*,

Lind, M. (2017). Knowledge Acquisition and Strategies for Multilevel Flow Modelling. In *International Symposium on Future I&C for Nuclear Power Plants (ISOFIC) 2017*, 1–8. Gyeongju, Korea.

Ma, X. and Li, D. (2017). A Hybrid Fault Diagnosis Method Based on Fuzzy Signed Directed Graph and Neighborhood Rough Set. In X. Ma and D. Li (eds.), *IEEE 6th Data Driven Control and Learning Systems Conference*, 253–258. Chongqing, China.

Maurya, M.R., Rengaswamy, R., and Venkatasubramanian, V. (2007). a Signed Directed Graph and Qualitative Trend Analysis-Based Framework for Incipient Fault. *Chemical Engineering Research and Design*,

Peng, D., Geng, Z., and Zhu, Q. (2014). A Multilogic Probabilistic Signed Directed Graph Fault Diagnosis Approach Based on Bayesian Inference. *Industrial & Engineering Chemistry Research*,

Ricker, N.L. and Lee, J.H. (1995). Nonlinear modeling and state estimation for the Tennessee Eastman challenge process. *Computers and Chemical Engineering*,

Rothenberg, D.H. (2009). *Alarm Management for Process Control: A Best-practice Guide for Design*. Momentum Press, New York.

Venkatasubramanian, V., Rengaswamy, R., and Kavuri, S.N. (2003a). A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies. *Computers & Chemical Engineering*,

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., and Yin, K. (2003b). A review of process fault detection and diagnosis: Part III: Process history based methods. *Computers & Chemical Engineering*,

Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S.N. (2003c). A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Computers & Chemical Engineering*,

Wan, Y., Yang, F., Lv, N., Xu, H., Ye, H., Li, W., Xu, P., Song, L., and Usadi, A.K. (2013). Statistical root cause analysis of novel faults based on digraph models. *Chemical Engineering Research and Design*,

Yang, F., Shah, S., and Xiao, D. (2012). Signed directed graph based modeling and its validation from process knowledge and process data. *International Journal of Applied Mathematics and Computer Science*,

Yin, S., Ding, S.X., Haghani, A., Hao, H., and Zhang, P. (2012). A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of Process Control*,

Zerkaoui, S., Druaux, F., Leclercq, E., and Lefebvre, D. (2010). Indirect neural control for plant-wide systems: Application to the Tennessee Eastman Challenge Process. *Computers and Chemical Engineering*,

Zhang, X. (2015). *Assessing Operational Situations*. Ph.D. thesis, Technical University of Denmark.

Zhang, X., Lind, M., and Ravn, O. (2013). Consequence Reasoning in Multilevel Flow Modelling. In *Analysis, Design, and Evaluation of Human-Machine Systems*, volume 12.