Structure Graph of Production: A basic concept for process data integration and analysis

Wan Li*, Michael Winter*, Christian Roos, Tobias Kleinert***

** Chair of Information and Automation Systems for Process and Material Technology, RWTH Aachen University 52064 Aachen, Germany (e-mail: {w.li, m.winter, kleinert}@plt.rwth-aachen.de)*
*** International Partners in Glass Research (IPGR) e.V., Krantzstr. 7, 52070 Aachen (e-mail: Christian.roos@ipgr.com)*

Keywords: interdisciplinary modelling, integrated system model, causal graph, knowledge representation

1. INTRODUCTION

With the development of Industry 4.0 technologies, numerous manufacturers are establishing production digitization as their corporate strategy to maintain and increase competitiveness. More and more data on production facilities and processes are collected and stored to answer recurring questions e.g., causes of machine downtimes or product quality defects, so that measures for correcting the failure can be intervened as early as possible.

For such cause-effect analysis, influencing factors have to be gathered, evaluated, and visually represented. Potential influencing factors can be found using priori knowledge about the production plant, e.g., the plant structure presented in a CAE (computer aided engineering) model or the project configuration in a control application program. Such prior knowledge is considered as context information to the production data of the plant, however, there is usually no direct link created to the production data. Therefore, influencing factors and their belonging production data have to be gathered and interrelated with laborious manual efforts, and the established relationships are often not reusable for new tasks of causal analysis. In addition, heterogeneity in production data regarding, e.g., data structure and sampling rate, complicates the integration of production data needed for analysis. Furthermore, the influencing factors found have to be represented in an easy human-readable way. At last, saving the context information, the influencing factors, and their relationships in a reusable way is challenging as well.

To address the above described challenges in managing and representing of production data for data analysis purposes, manufacturers need a sensible and practical concept. In this paper, we present a novel and generalized concept, the *Structure Graph of Production (SGP).*

2. CONCEPT OF THE SGP

The SGP is a graph model based on the *phase model of production (PMP)* (Polke 1992). It includes property clusters representing components of the considered production system, and relationships between the properties as well as the PMP elements (Fig. 1).

A production system in the current consideration is composed of production processes, machines, raw materials, produced (intermediate) products, and field devices (including actuators, measuring devices), which are abstractly modelled with clusters of properties of these components (Fig. 1).

Nodes and relationships are the most basic model elements that build the property clusters and the PMP. Fig. 1. shows the generalized SGP class with nodes and relationships, which are specialized as shown in Fig. 4 and Fig. 5, respectively.

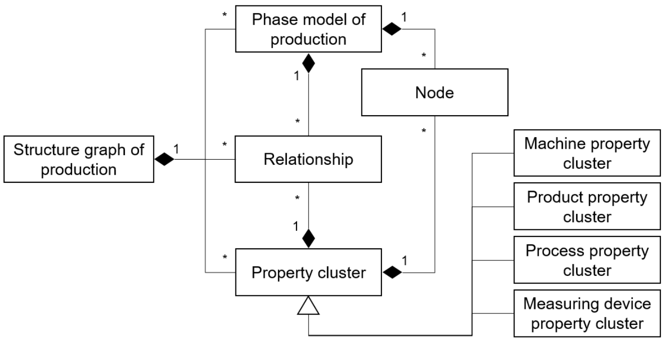

Fig. 1. General SGP class model

In the following, the concept of SGP is explained with a demonstrative example of a heating system of our laboratory plant. Its piping and instrumentation diagram (P&ID) is shown in Fig. 2. Pump N13 feeds the outflow from an upstream tank B1 into a heat exchanger and feeds it further to the downstream processes through valve Y16. A temperature sensor (T15) and a flow meter (F17) measures the process and product properties, respectively. The pumping power, heating power and valve opening are controlled automatically.

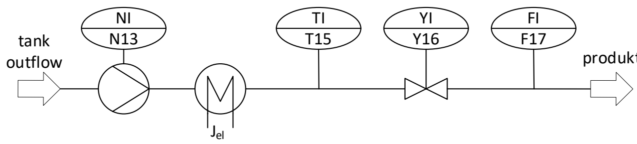

Fig. 2. P&ID of a part of a laboratory plant

With the given information, the SGP is created as shown in Fig. 3. The processes and products are modelled as PMP represented by the blue nodes and edges. The respective properties of the heating system components are modelled as nodes and clusters on the side of the PMP elements, which are pipelines with actuators and sensors as well as water.

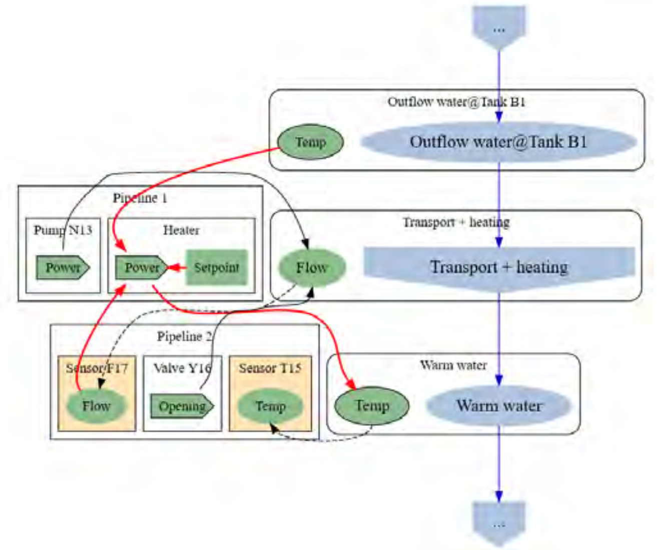

Fig. 3. SGP model of the heating process shown in Fig. 2

Each node (Fig. 4) in SGP has a uniquely identifiable *ID* that can be represented by a textual label. The class node has two subclasses: structural node and property node. A *structural node* can either be a product node (e.g., “warm water”) or a process node (e.g., “transport + heating”), representing process steps or products in the PMP. A *Property node* (e.g., the green nodes) describes either a certain aspect of a machine, a product, a process or a measuring device. Each property node has two further attributes: (1) *Associated cluster* to which the node belongs to; (2) *Associated data variable* with basic metadata (e.g., name, data type and location) which are referencing the data stored in a database. The nodes are displayed in corresponding colors and shapes, depending on controllability, control mode and measurability.

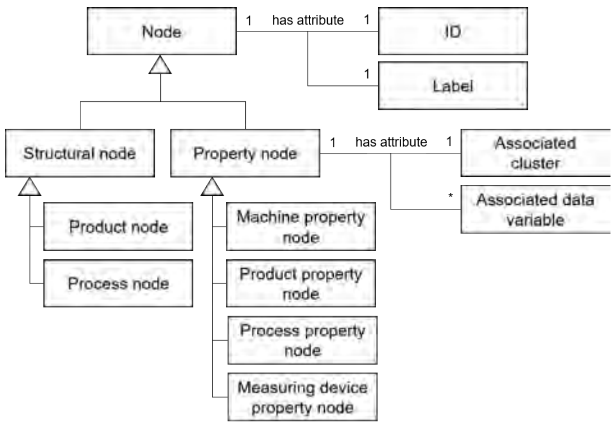

Fig. 4. Node class model

A *relationship* (Fig. 5) is modelled as an *edge* in the graph model, which is represented by a directed line pointing from a source node to a target node. Each relationship is uniquely identifiable in an SGP by its *ID* or by a node tuple (source node, target node). A relationship has two subclasses: structural or influencing relationship. A *structural relationship* represents the direction of the *material flow* in the PMP, or the *data flow* between the property nodes of a measuring device cluster and a machine, product or process cluster, respectively. An *influencing relationship* (causal or correlational) indicates that the source node may have an influence on the target node with a certain probability and intensity, which can be determined by using methods of either data analysis, experimental observation, or expert knowledge (see Section 4). When an influencing relationship is valid, its attribute *isValidated* is set TRUE.

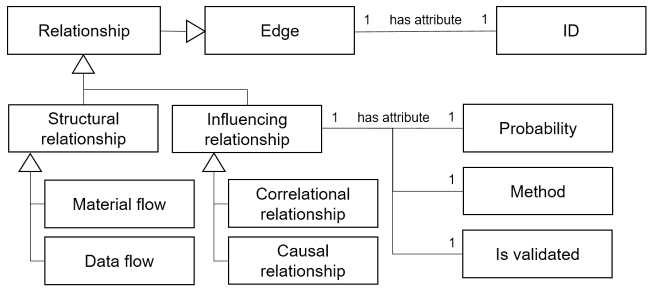

Fig. 5. Relationship class model

6. DISCUSSION

Causal questions can be answered with a combination of process model and data analysis (Pearl 2018). In the concept of SGP, prior knowledge about the production plant is modelled as PMP and clusters of the production system properties. The PMP containing structural relationships and nodes forms the fundamental structure of the SGP, which allows categorization and association of property clusters. The PMP model implies also time constraints for inferencing causal and influencing relationships, because the structural nodes in PMP are arranged regarding the execution sequence of the production processes. In addition, measurements stored in databases are associated with properties of process steps, products and machines using the attribute *associated data variable*. Thus, semantics and metadata of measurements are available, such that the measured data can be extracted, integrated and prepared in an efficient way. Influencing factors of a certain property can be gathered using existing relationships and topological distances between the property nodes. Furthermore, PMP allows the property nodes to be presented in a structured way, so that human can easily interact with the SGP and carry out causal analysis. With the above advantages, the SGP provides a basis for a straight through data integration and analysis. In future work, we will specify the modelling systematics of SGP, and develop SGP towards automated data integration and analysis. Moreover, aspects of changeability, scalability and user-friendly visualization will also be considered.

REFERENCES

Pearl, J. (2018). *The book of why: the new science of cause and effect*. Chap. 10. Basic Books, New York, U.S.

Polke, M. (1992). Phasenmodell der Produktion als nicht-mathematische Präsentationsform für Automatisierungs-aufgaben. *Entwurf komplexer Automatisierungs-systeme*, Braunschweig, Germany