Semantic Data Integration for Virtual Production

Daniel Schilberg[1] **Tobias Meisen**[1] **Klaus Henning**[1]

[1]Center for Learning and Knowledge Management and Department of Information Management in Mechanical Engineering, RWTH Aachen University, Aachen, Germany

Abstract

The focus of this work is to promote the semantic data integration for virtual production systems. The phase of product-planning usually requires a huge range of numerical simulations that are used by different departments of a company. To increase productivity on the one hand and reduce complexity of simulation chains on the other hand, it is essential to be able to integrate the differing data of distributed numerical simulations by their semantics. The integration is realized by a simulation platform.

Keywords: Data Integration, Semantic Data Intergration, Simulation

1. Introduction

A manufacturing company in a high-wage country has to position itself in the dilemma between mass production with a very limited product range (scale) and the manufacturing of a large variety of products in small quantities (scope), as well as the dilemma between planning and value orientation. The vertices of these dilemmas shape the so-called polylemma of production technology (Fig. 1) [1]. In order to create a lasting competitive advantage over low-wage countries, it is necessary to develop methods to reduce or dissolve the polylemma.

The Cluster of Excellence "Integrative Production Technology for High-Wage Countries" at RWTH Aachen University is engaged in the reduction of the mentioned polylemma. One objective of the cluster is the development of a simulation platform for the interconnection of simulations that are used in product or production planning. The simulation platform builds up whole simulation chains according to a specific not predefined manufacturing process. The use of automatically interconnected simulations reduces the dilemma from planning orientation towards value orientation, whilst ruling out that there will be a reduction in the quality of planning. The challenge lies in combining dynamically monolithic simulations to represent an entire production process for virtual production [2].

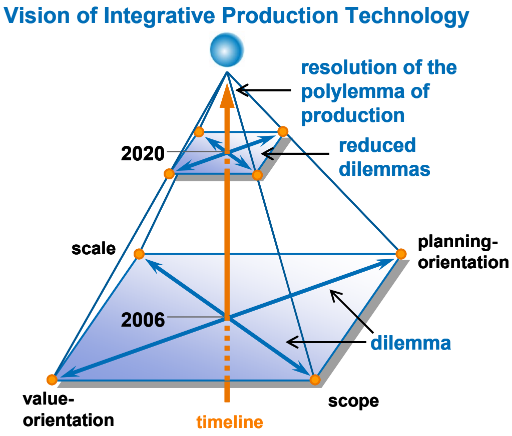

Fig. 1: Polylemma of production.

This paper shows the software architecture of the simulation platform. The simulation platform is used to integrate specific simulation input and output data in requested simulation input files within a simulation chain. The data integration connects simulations syntactically and semantically. Process simulations, such as welding or forming, are considered. The focus lies on the components needed for the semantic data integration. Then on the basis of the following use case the application of the platform is described. The paper closes with a summary of the idea of the simulation platform.

2. Use Case

The simulations used in the manufacturing process of a gear wheel are considered for the demonstration of the simulation platform. The gear wheel is manufactured in several working steps. Two levels of detail are examined in the manufacturing process of a gear wheel [3]. The first level of detail is linked to macro-structure processes like forming, heat treatment and welding. The second level of detail is linked to the examination of the micro-structure. The micro-structure data is homogenized on a transfer level to allow its usage on the macro-structure level.

Each simulation that is used in the manufacturing process must be docked to the simulation platform. That means that all used parameters of the five simulations must be represented by meaning and value. The meaning is stored in the simulation platform, while the values are provided by the simulations. By meaning the following is intended: unit, value range and context.

In the use case, the blank will be prepared with a heat treatment for the recast process. In order to use the best material data, a micro-structure analysis will be made after every step in the macrostructure process. With the results from the heat treatment the micro-structure data for a few representatives points of the blank will be determined. In order to use this information for the whole macro-structure, a homogenization tool is used to prepare the micro-structure data for the macro level. The results of the heat treatment and the micro-structure simulation can be used separately or together as input for the following recast process simulation. This procedure is repeated for the next heat treatment and the welding process [4].

The simulation platform gathers all parameters that are needed for a complete and automated run of the whole simulation chain through the user and the simulations. The user of the platform has to select the steps of the manufacturing process that should be simulated via the simulations. The user has to provide all data that is not generated by the simulations before the simulation chain can be started. A list of the missing data is provided by the simulation platform. The simulation platform will organize the order of the simulations to match the manufacturing process. Furthermore, it will provide the input data for each simulation.

Via the components of the simulation platform, the semantic data integration for simulation interconnection is realized. On the basis of this interconnection many manufacturing processes can be represented with the simulation platform simply by using the simulations in different orders. All parameters and the values generated by the simulations are available for every downstream simulation in the chain. Thus, in any further process step which will be simulated, the entire process history is recorded by the simulation platform; this ultimately leads to a better data set for all simulations.

3. Data Integration

The system architecture for the simulation platform consists of the platform itself and the simulations that are linked to it. Fig. 3 shows the system architecture of the platform.

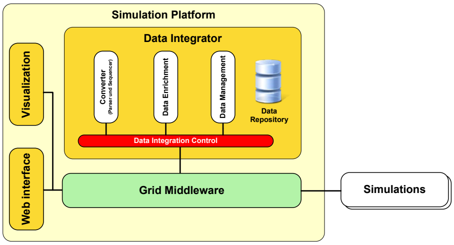

Fig. 3: Simulation Platform

The platform consists of the components Web Interface, Visualization, Middleware and the Data Integrator for data integration. The Web Interface is part of the Human Machine Interface which also includes the visualization. The user operates the platform via the Web. The user can choose the process steps that should be simulated and the order on which these will be implemented. Data that is not generated by the simulations or producible by the Data Integrator component out of the simulations results must be entered by the user via the Web Interface. The visualization presents the simulation results, either as individual data sets or in comparison with other data sets. For example, the results of the simulations can be displayed as single results or as interlinked results. For the interlinked representation all or a collection of results can be used. Through this representation, the user is able to visualize in parallel the effects of a change in the macro and microstructure at every time step of the manufacturing process. The middleware is used among other things for data exchange between the platform components and the platform, as well as the simulations. The components and the simulations are distributed through a network. The Data Integrator is required to create the syntactic and semantic interconnection between the simulations and to ensure that the chosen manufacturing process can be represented correctly by the simulation platform.

The components of the Data Integrator are used for the semantic data integration. The components of the Data Integrator are a Data Management, a Data Enrichment, a Converter, a Data Repository and a Data Integration Control. The Data Repository is used to store the input and output data of the simulation and process related information. As an interface for the communication between the Data Integrator and the other components of the platform, the Data Integrator uses a Data Integration Control component. All the incoming and out going data from the Data Integrator is administrated by this component. The Converter provides the simulation input files and converts the simulation output files into an internal data format.

The Data Enrichment (Fig. 4) consists of the following components: A Query Processor, a Reasoner, a Knowledge Base Parser, a Knowledge Base that is represented by an ontology, a Compiler and a Planner.

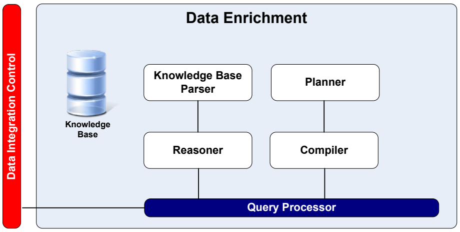

Fig. 4: Data Enrichment

The Data Enrichment integrates the simulation data with different semantics, allowing a comparison of the parameters used. All information concerning the parameters used by the simulations is stored in the Knowledge Base with information on value range, dependencies and context added. Via the Data Integration Control the Query Processor in cooperation with Data Management Control will check if a certain parameter is generated by a simulation, is introduced manually or if it can be generated and made accessible through the Data Enrichment. For example one simulation generates a value for the parameter amperage (I), another simulation generates a value for the parameter voltage (U), but the next simulation needs a value for the parameter resistance (R). The Data enrichment can than provide the information that amperage and voltage can be transformed to resistance. The needed formula U/I=R is stored in the Transformation Library of the Data Management.

The stored parameters are abstractly represented by an ontology [5]. The Knowledge Base Parser provides the Reasoner with information out of the Knowledge Base. To compute a certain query, the Query Processor has to gather all necessary information from the Knowledge Base by using the representation structure of the ontology. This is done by the Reasoner. After all needed information is available; the Planner is used to identify and order the next steps to gain a result for the query. The Compiler computes the results of the Planner. Then the Query Processor can provide the Data Integration Control with the information that the requested data can be generated.

Another example can be the identification of identical data this is represented by the simulations in different ways. In one simulation a material is represented by its chemical composition and in another simulation by its name (e.g. 20MnCr5 = Saarstahl). Without the necessary knowledge as to what this means, no direct conversion is possible. By using an ontology which represents this additional information, a transformation method can be chosen in order to provide the next simulation with this parameter [6].

Through cooperation of the components of the subsystem mentioned above, Data Enrichment enables the Data Integrator to connect simulations not only on the level of syntax but also on the level of semantics. The ontology, from which the importance of individual parameters is derived, makes semantic mappings possible. Parameters will be linked for comparison respectively their values, based on their meanings. The comparison performed by the Data Enrichment will run through the following steps. First it will be checked whether the meanings directly match. If this is the case, the parameter values will be directly circulated to the Converter. The second step is only used when the first does not generate a result. If the Data Enrichment does not find a direct match, it must be verified if the required data can be generated out of existing data via the use of the Transformation Library. If so, the parameter values will be generated. Should this not be possible, the missing data must be made available by the user via the Web Interface. The data that is provided by the user is also stored in the Data Base of the Data Management [7].

The component architecture of the subsystem Data Management is shown in Fig. 5. The Data Management consists of a Database, a Data Management Control, a Storage Manager, an Indexer, a Data Extractor and the Transformation Library.

Via the Data Integration Control, the core component of the Data Management, the Data Management Control, is executed. The task of the Data Management is to provide data and transformation methods to generate missing data. To be able to solve this task, the Data Management has to index the data and to store the indices related to the data and to the information. If certain data and information has to be provided, the Data Management uses the Data Extractor component to identify the needed data and information, and extract these from the Database via the indices. After that the transformation methods like U/I=R from the example are used to generate the values of the parameters. The needed values will be integrated by the Converter in the files, that are used by the simulation.

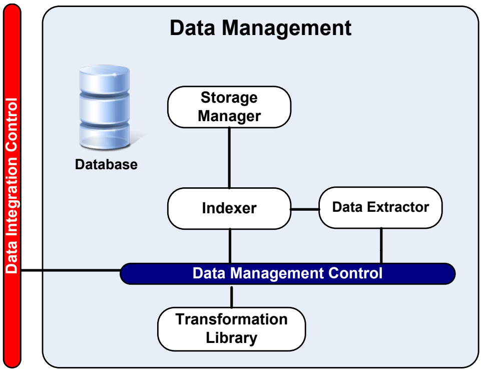

Fig. 5: Data Management

4. Conclusion

The need for semantic data integration derives from the need of reducing the planning effort of production processes and therefore the requirement of flexible simulation chains. Investigation in this field means increasing the insight and understanding of the scientific and technical processes. This paper has introduced a concept of semantic data integration. It also presents how it can be realized by using a Data Integrator. The approach shown of semantic data integration, has the advantage that all input and output data of the whole simulated process chain are correlated. This will prevent the loss of semantic information, as all data is linked to a certain step in the process of the simulation chain. This circumstance allows to identify the completeness of the requested data. If it is not complete, the system is able to localize the missing data and also specify where the missing data can be generated. The main research goal is to increase the understanding of how the semantic data integration can be successfully implemented as basis for the realization of virtual production systems. The development of tools and techniques to support the interoperability of simulations along the process chain supports the achievement of the main goal. The simulation platform is the first step in implementing a virtual production system that contains the product planning, the process planning, the process itself, the process chain and the factory design planning.

5. References

[1] Schuh, G., Klocke, F., Brecher, C. and Scmidt, R., Excellence in Production, Apprimus Verlag, Aachen Germany, (2007)

[2] Schilberg, D., Gramatke, A. and Henning, K., Semantic interconnection of distributed numerical simulations via SOA, *Proceedings of the World Congress on Engineering and Computer Science*, p.894-897, Newswood Limited, Hong Kong, (2008)

[3] Schmitz, G.J., Prahl, U., Toward a Virtual Platform for Materials Processing, JOM Vol. 61 No.5 p. 19-23, (2009)

[4] Schroeder, W., Martin, K. and Bill Lorensen, The Visualization Toolkit, Kitware, Inc., USA, (2004)

[5] Tudorache, T., Ontologies in Engineering. VDM Verlag, Saarbrücken Germany, (2008)

[6] Ciocoiu, M. and Nau, D., S., Ontology-Based Semantics, NIST, (1999)

[7] Lei, Y., An Instance Mapping Ontology for Semantic Web, *Proceedings of the 3rd International Conference on Knowledge Capture*, Banff Canada, (2005)