Towards Adaptive Data Integration in Autonomous Cooperating Logistics Processes

K. A. Hribernik, C. Kramer, C. Hans and K.-D. Thoben

BIBA - Bremer Institut für Produktion und Logistik GmbH, Hochschulring 20, 28359 Bremen, Germany
E-Mail: {hri, kra, han, tho}@biba.uni-bremen.de

Summary: The characteristics of the IT landscape underlying autonomous cooperating logistics processes pose a number of challenges towards data integration. The heterogeneity of the data sources, their highly distributed nature, along with the limited availability of dynamic sources make the application of traditional approaches problematic. A combination of semantic data integration with the principles of service-oriented architecture has proven an adequate approach to meet these challenges. However, whilst the approach succeeds at providing a scalable, robust, flexible, platform-independent and uniform data integration approach, its limitations with respect to its adaptability especially towards new data sources need to be addressed. Consequently, this paper investigates a number of starting points for research towards an adaptive, (semi-)automatic data integration for use in autonomous cooperating logistics processes.

Keywords: Distributed Control, Artificial Intelligence, Information, Integration.

1. Introduction

Today's challenges in logistics are characterised by the goods structure, logistics and structural effects [1]. Autonomous cooperating logistics processes [2] aim to meet these challenges by introducing autonomy and self-organisation into control, information processing and decision-making in logistics [3]. *"Autonomous control"* refers to *"...processes of decentralised decision-making in heterarchical structures. It presumes interacting elements in non-deterministic systems, which possess the capability and possibility to render decisions independently."*[4] The rationale behind applying autonomous control to logistics processes is the expectation of an increase in their robustness, flexibility, adaptability and reactivity in responding to changing business environments, requirements and conflicting objectives [2].

A prominent characteristic of autonomous cooperating logistics processes is the decentralisation of decision-making responsibilities in contrast to traditional, hierarchical process control. Logistics objects are entrusted with acting in their own "best interest" within the bounds of their operational, tactical or strategic autonomies [5]. Logistics objects in this context can be material (e.g. component parts, machines and conveyors) or immaterial items (e.g. a production order) within a networked logistics system. They are capable of interacting with other logistics objects in the system [4]. According to the criteria for autonomous cooperating processes defined by Windt and Böse [4], the less central the data storage and processing of the underlying logistics IT systems supporting the processes are, the higher the level of autonomous control can be [6].

Regardless of autonomous control, the logistics IT landscape is highly complex, distributed and heterogeneous. Significant effort has been directed towards achieving integration between systems within supply chains by "bridging the technological islands" through specific ICT solutions [7]. However, most of these solutions quickly become obsolete due to the continuous development of the "technological islands" as well as the highly dynamic partnerships within today's enterprise networks. Instead of developing solutions for one-to-one relationships, a general solution needs to be found which allows uniform access to all relevant logistics data while accepting the diversity of existing systems and standards [8].

The situation is exacerbated in autonomous cooperating logistics processes. Depending on the application, relevant data may be stored in heterogeneous logistics systems, such as Warehouse Management Systems, Enterprise Resource Planning Systems or disposition systems. Simultaneously, data from RFID and item-level tracking and tracing systems may need to be taken into account. Dynamic data generated by logistics objects needs to integrated, for example sensor networks monitoring the temperature of a refrigerated container. Decision-making instances such as software agents representing individual logistics objects need to be able to access data relevant to their decision making processes, regardless which "technological island" that data may be located on.

The characteristics of the underlying IT landscape pose a number of challenges towards data integration. The heterogeneity of the data sources, their highly distributed nature, along with the limited availability of dynamic sources make the application of traditional approaches problematic. A combination of semantic data integration with the principles of service-oriented architecture has proven an adequate approach and is described in the following section [6], [8], [9], [10], [11]. However, whilst it succeeds at providing a scalable, robust, flexible, platform-independent and uniform data integration approach, its limitations with respect to adaptability especially towards new data sources need to be addressed. Consequently, Section 3 investigates a number of starting points for research towards adaptive data integration in autonomous cooperating logistics processes.

2. Semantic Data Integration in Autonomous Control

This section describes the concept of service-oriented, semantic data integration for autonomous cooperating logistics processes. Central to the concept is an ontology-based mediator which is capable of composing queries to any combination of relevant logistics data sources [12], [13], [14]. This is achieved by semantic mediation. Each data source is fully described syntactically and semantically by an ontology, which is internally mapped onto the others by the mediator.

Wrapper components handle the transformation to and from the relevant data sources in a rule-based or algorithmic fashion, depending on the characteristics of that data source. The wrappers query data from the data sources and transform it transparently to the core mediator component. This allows for a complete abstraction from the data sources. Heterogeneity conflicts are solved either by the mediator component itself or by the respective wrapper, depending on the type of conflict. The Web Ontology Language (OWL-DL) [15] is used to specify the ontologies which describe the individual data exchange formats. It is adequately expressive to cover the semantic description of both the standard exchange formats used in transport logistics and the overarching concepts of autonomous logistics processes. The query language SPARQL, developed for querying ontologies [16], defines the query interface of the system. To support bidirectional queries, it was extended by the SPARQL/Update [17] language. A service layer is defined upon the SPARQL/Update interface. This has the benefit of facilitating a process-oriented, model-driven approach to the definition of logical views upon mediated data. Finally, to facilitate the direct integration of dynamic data sources such as RFID, sensor networks and other systems integrated into physical logistics objects, an abstraction layer is defined to provide a reliable interface to the data sources, regardless of physical accessibility at any time. It is responsible for buffering, filtering and routing data to and from the respective data sources.

3. Towards Adaptive Data Integration in Autonomous Cooperating Logistics Processes

In order to contribute to an increase in flexibility, dynamism, adaptability and distribution of the IT environments underlying autonomous controlled logistics processes, the semantic data integration approach described needs to be extended. A demand-oriented, adaptive, and ad-hoc approach to the integration of new data sources – ideally without increased development effort – would be advantageous. The data integration approach described in Section 2 offers a number of starting points for research into an adaptive, (semi-)automatic data integration of autonomous cooperating logistics processes:

1. Semantic descriptions of data sources
2. Definition of transformation mechanisms
3. Wrapper configuration and deployment
4. Definition and deployment of logical views

With regards to the first point, in the approach presented above, data sources need to be semantically described manually with the ontology language OWL-DL. Modelling tools can be used to make the process easier. The effort required to describe new data sources however remains considerable. Promising approaches towards automating the process can be found in the field of ontology learning. Ontology learning approaches can be differentiated by the type of data source to be described [18]. For autonomous cooperating logistics processes, semi-structured schemata [19], [20], [21], [22], relational schemata [23], [24], [25] and knowledge databases [26], [27] are most relevant. Examples of used approaches of ontology learning from these categories are Naive Bayesian Learning, Prediction Combination or Meta Learning. These approaches are suitable for semi-automated description of the respective categories of static data sources. For dynamic data sources, such as sensor networks and embedded systems, different approaches need to be investigated. For example, the context of the data source to be integrated may be taken into consideration [28]. Process and workflow oriented learning methods may also prove advantageous [29]. Point 2, the automation of the definition of transformation mechanisms, can be approached from a number of different angles. Specific methods exist for the transformation of Internet data sources [30]. Methods of schema integration can be used as points of departure for relational data sources [31], [32]. Furthermore, algorithmic methods of ontology mapping which support automated ontology modelling should be investigated [33]. Promising approaches can also be found in the algorithmic determination of semantic equivalence [34]. Relevant examples in literature mainly refer to the area of geospatial data [35], [36]. Looking at the automation of the definition of transformation mechanisms and generation of wrappers (points 3 and 4), a number of promising approaches can be found in literature. From a software engineering point of view, the deployment of wrappers into the mediator can be facilitated using approaches like UPnP (Universal Plug and Play), OSGi, PMI/QMI (PROMISE/Quantum Messaging Interface) or programming interfaces like the Reflections API (Java). Finally, mechanisms of service orchestration and composition can be applied to automate the definition and deployment of logical views. This can be facilitated on the basis of semantic descriptions [37].

4. Conclusions and Outlook

To summarise, different methods of ontology learning promise to be applicable to the problem of automating the semantic description of different types of data sources. Schema integration and algorithmic methods of ontology mapping can be applied to automate the definition of transformation mechanisms. With regards to the automation of wrapper configuration and deployment as well as the definition and deployment of logical views, a number of promising software engineering methods exist. Future work will investigate each of the methods identified in more detail. The applicability of each method will be evaluated. Finally, a draft solution concept will be proposed for an adaptive service-oriented, semantic data integration approach in autonomous cooperating logistics processes based on these methods.

Acknowledgements

The authors gratefully acknowledge the contribution of the following projects' consortia and funding organisations to the research presented in this paper: CRC 637 Autonomous Cooperating Logistics Processes – A Paradigm Shift and its Limitations Subproject C2 Data Integration (DFG) and LinkedDesign - Linked Knowledge in Manufacturing, Engineering and Design for Next-Generation Production (EU FP7 NMP).

References

[1] Aberle, G., 2003, Transportwirtschaft: einzelwirtschaftliche und gesamtwirtschaftliche Grundlagen, 4th edn. Oldenbourg, Munich.

[2] Freitag, M., Herzog, O., Scholz-Reiter, B., 2004, Selbststeuerung logistischer Prozesse - Ein Paradigmenwechsel und seine Grenzen, Industrie Management, 20/1:23-27.

[3] Ehnert, I., Arndt, L., Mueller-Christ, G., 2006, A sustainable management framework for dilemma and boundaries in autonomous cooperating transport logistics processes, Int. J. of Environ and Sustainable Dev, 5:355-371.

[4] Böse, F., Windt, K., 2007, Catalogue of Criteria for Autonomous Control in Logistics, Understanding Autonomous Cooperation and Control in Logistics - The Impact on Management, Information and Communication and Material Flow:57-72.
[5] Timm, I.J., 2006, Strategic Management of Autonomous Software Systems: Overview Article, Technical Report 35. University of Bremen, Centre for Computing and Communication Technologies TZI.
[6] Hribernik, K., Kramer, C., Hans, C., Thoben, K.-D., 2010, A Semantic Mediator for Data Integration in Autonomous Logistics Processes, Enterprise Interoperability IV. Making the Internet of the Future for the Future of Enterprise:157-167.
[7] Hannus, M., 1996, Islands of automation in construction, Construction on the information highway, CIB Publication, University of Ljubljana.
[8] Hans, C., Hribernik, K.A., Thoben, K.-D., 2008, An Approach for the Integration of Data within Complex Logistics Systems, Dynamics in Logistics, 1st International Conference LDIC 2007 Proceedings:381-389.
[9] Hribernik, K., Pille, C., Jeken, O., Thoben, K.-D., Windt, K., Busse, M., 2010, Autonomous control of intelligent products in beginning of life processes, Proceedings of the 7th International Conference on Product Lifecycle Management.
[10] Hribernik, K., Warden, T., Thoben, K.-D.; Herzog, O., 2010, An Internet of Things for Transport Logistics - An Approach to Connecting the Information and Material Flows in Autonomous Cooperating Logistics Processes, Proceedings of the 12th International MITIP Conference on Information Technology & Innovation Processes of the Enterprises:54-67.
[11] Hribernik, K., Kramer, C., Hans, C., Thoben, K.-D., 2011, Impacts of Data Integration Approaches on the Limitations of Autonomous Cooperating Logistics Processes, Autonomous Cooperation and Control in Logistics:247-270.
[12] Ullman, J.D., 1997, Information integration using logical views, Proceedings of the 6th International Conference on Database Theory:19-40.
[13] Wache, H., Vögele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., Hübner, S., 2001, Ontology-Based Integration of Information - A Survey of Existing Approaches, IJCAI 2001 Workshop on Ontologies and Information Sharing:108-117.
[14] Wache, H., 2003, Semantische Mediation für heterogene Informationsquellen, Akademische Verlagsgesellschaft:1-308.
[15] Smith, M.K., Welty, C., McGuinness, D.L., 2004, OWL Web Ontology Language Guide, http://www.w3.org/TR/owl-guide/, Accessed 15 July 2009.
[16] Prud'hommeaux, E., Seaborne, A., 2008, SPARQL Query Language for RDF, http://www.w3.org/TR/rdf-sparql-query/, Accessed 05 March 2012.
[17] Seaborne, A., Manjunath, G., Bizer, C., Breslin, J., Das, S., Harris, S., et al., 2008, SPARQL Update. A language for updating RDF graphs, http://www.w3.org/Submission/SPARQL-Update/, Accessed 05 March 2012.
[18] Gómez-Pérez, A., Manzano-Mach, D., 2003, A Survey of Ontology Learning Methods and Techniques, Technical Report, Universidad Politécnica de Madrid.
[19] Deitel, A., Faron, C., Dieng, R., 2001, Learning ontologies from RDF annotations, Proceedings of the IJCAI Workshop in Ontology Learning, Seattle:3-13.
[20] Doan, A. H., Domingos, P., Levy, A., 2000, Learning Source Descriptions for Data Integration, Proceedings of the Third International Workshop on the Web and Databases, Dallas:81-86.
[21] Murata, M., Lee, D, Mani. M., 2005, Taxonomy of XML Schema Languages using Formal Language Theory, ACM Transactions on Internet Technology (TOIT), 5/4:660-704.
[22] Reynaud, C., 2001, Semantic integration of XML heterogeneous data sources, Database Engineering & Applications, 2001 International Symposium on:199-208.
[23] Johannesson, P., 1994, A Method for Transforming Relational Schemas into Conceptual Schemas, 10th International Conference on Data Engineering, Houston:115-122.
[24] Kashyap, V., 1999, Design and Creation of Ontologies for Environmental Information Retrieval, Proceedings of the 12th Workshop on Knowledge Acquisition, Modeling and Management:3-21.
[25] Rubin, D.L., Hewett, M., Oliver, D.E., Klein, T.E., Altman, R.B., 2002, Automatic data acquisition into ontologies from pharmacogenetics relational data sources using declarative object definitions and XML, Proceedings of the Pacific Symposium on Biology: 88-99.
[26] Modica, G., Gal, A., Jamil, H. M., 2001, The Use of Machine-Generated Ontologies in Dynamic Information Seeking, Proceedings of the 6th International Conference on Cooperative Information Systems:433-447.
[27] Suryanto, H, Compton, P., 2001, Discovery of Ontologies from Knowledge Bases, Proceedings of the 1st International Conference on Knowledge Capture:171-178.
[28] Porzel, R., Zorn, H.P., Loos, B., Malaka, R., 2006, Towards a Separation of Pragmatic Knowledge and Contextual Information, Proceedings of ECAI 06 Workshop on Contexts and Ontologies, Riva del Garda, Italy.
[29] Pathak, J., Caragea, D., Honavar, V., 2004, Ontology-extended component-based workflows: a framework for constructing complex workflows from semantically heterogeneous software components, Proceedings of the VLDB-04 Second International Workshop on Semantic Web and Databases (SWDB 2004):41-56.
[30] Ashish, N., Knoblock, C.A., 1997, Semi-automatic wrapper generation for internet information sources, Proceedings of the Second IFCIS International Conference on Cooperative Information Systems:160-169.
[31] Batini, C., Lenzerini, M., Navathe, S. B., 1986, A comparative analysis of methodologies for database schema integration, ACM Computing Surveys, 18/4:323-364.
[32] Sheth, A. P., Gala, S. K., Navathe, S. B., 1993, On automatic reasoning for schema integration, Interntational J. of Intelligent and Cooperative Information Systems, 2/1:23-50.
[33] Kalfoglou, Y., Schorlemmer, M., 2003, Ontology mapping: The state of the art, Cambridge Univ. Press, UK:1-31.
[34] Rodríguez, M. A., Egenhofer, M. J., 2003, Determining semantic similarity among entity classes from different ontologies, IEEE Transactions on Knowledge and Data Engineering, 15/2:442-456.
[35] Cruz, I. F., Sunna, W., Chaudhry, A., 2004, Semi-automatic ontology alignment for geospatial data integration, Lecture Notes in Computer Science, 3234/2004:51-66.
[36] Cruz, I. F., Xiao, H., 2005, The role of ontologies in data integration, J. of Engineering Intelligent Systems, 13/4:245-252.
[37] Sirin, E., Hendler, J., Parsia, B., 2003, Semi-automatic composition of web services using semantic descriptions, Proceedings of the 1st Workshop on Web Services: Modeling, Architecture and Infrastructure (WSMAI-2003):17-24.