ONTOLOGY-DRIVEN INFORMATION INTEGRATION
Networked Organisation Configuration

Alexander Smirnov, Tatiana Levashova, Nikolay Shilov
St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS)
14 Line VO, 19178, St.Petersburg, Russia

Keywords: Networked organisation, decision support, ontology, information integration, semantic interoperability.

Abstract: Distributed networks of independent companies (networked organisations) are currently of high interest. This new organisational form provides for flexibility, tolerance, etc. that are necessary in the current market situation characterised by increasing competition and globalisation. Configuration of a networked organisation is a strategic task that requires intelligent decision support and integration of various tasks constituting the configuration problem. Achieving efficient integration of tasks is possible when it is done taking into account semantics. The paper proposes an approach to this problem based on ontology-driven knowledge integration. The knowledge in the approach is presented using formalism of object-oriented constraint networks. Such formalism simplifies problem formulation and interpretation since most of the tasks in the areas of configuration and management are constraint satisfaction tasks. The paper describes the developed approach and the ontological model that is the core of the approach. Application of the developed approach is demonstrated at two levels: (a) at the level of information integration within one company and (b) at the level of information integration across a networked organisation.

1 INTRODUCTION

Global changes in the economy worldwide have led to changes in priorities and strategies of market players. This has caused appearance of new network-driven organisational forms such as virtual enterprises, extended enterprises, supply chains, etc.

A networked organization (Laudon, et. al., 2000; Lipnack, et. al., 1994; Skyrme, 2003) is usually defined as an organization formed by geographically distributed independent partners on the basis of information technologies. Efficient creation of an effective configuration of the networked organisation can give its members a competitive advantage in getting an order. Hence, configuration of the networked organisation is a problem of the strategic level requiring intelligent decision support.

This problem has been addressed in numerous research efforts. However most of them solve particular tasks of the complex problem. A complex approach is required to provide for integration of the tasks to be solved. Among the tasks the following most important ones can be selected (they do not pretend to be a complete list):

1) order configuration (configuration of the product / service in accordance with existing constraints and customer preferences);

2) partner choice among existing companies – potential members of the networked organization;

3) resource allocation among the networked organisation members;

4) transportation network configuration (this logistics related task is required due to the distributed nature of the networked organisation);

5) configuration of technological resources of the networked organisation members.

During the process of the networked organisation configuration and management the above tasks have to be solved jointly. Hence, it is reasonable to speak about integration of them. Integration of tasks is more than data integration. It requires integration at the level of semantics or semantic interoperability. In other words, it requires knowledge integration.

Knowledge management has shown its efficient applicability in this area. It is a complex cooperative network-centric process to support multi-object and multi-disciplinary areas including modelling, design, knowledge representation and acquisition, decision

Smirnov A., Levashova T. and Shilov N. (2006).
ONTOLOGY-DRIVEN INFORMATION INTEGRATION - Networked Organisation Configuration.
In *Proceedings of the Eighth International Conference on Enterprise Information Systems - AIDSS*, pages 124-131
DOI: 10.5220/0002456901240131

support and supporting environment (Liu, et. al., 2004). A number of efforts have been done in the area of sharing information and processes between applications, people and companies. However knowledge sharing / exchange requires more than this. It requires information coordination and repository sharing with regard to semantics.

To address this, the paper proposes usage of the ideas of Knowledge Logistics (KL) that stands for acquisition, integration, and transfer of the right knowledge from distributed sources to right persons (decision makers) at the right time for the right business purpose in the right context (Smirnov, et. al., 2003a). KL with regard to individual customer requirements, available knowledge sources, and current situation analysis in an open information environment addresses problems of intelligent support of customer activities.

One of the main issues to resolve is interoperability. It can be defined as the ability of enterprise software and applications to interact. Interoperability is considered to be achieved if the interaction can, at least, take place at three levels: data, application and business enterprise (Interoperability Research for Networked Enterprises Applications and Software, 2004). Semantic interoperability assumes interaction at one more level, namely at the level of semantics. To provide for semantic interoperability KL uses ontologies as one of the most advanced approaches to knowledge mark-up and description. Ontologies establish a joint terminology between members of a community of interest (Semantic Web, 2005). This makes it possible to provide for semantic interoperability between various tasks of configuration and management.

The paper is structured as follows. Approach description is presented in sec. 2. Sec. 3 outlines the principles of the central integrated ontological model creation. The case study is given in sec. 3. Some results are summarised in conclusions.

2 APPROACH

The KL problem in the presented approach is considered as a configuration of a network including end-users, knowledge resources, and a set of tools and methods for knowledge processing located in the network-centric environment. Such a network of loosely coupled sources is referred to as a knowledge source network or “KSNet” (detailed description of the approach can be found in Smirnov, et. al., 2003b), and the approach is called KSNet-approach. The approach is built upon constraint satisfaction / propagation technology for problem solving since application of constraint networks allows simplifying the formulation and interpretation of real-world problems that are usually presented as constraint satisfaction problems in such areas as management, engineering, etc. (e.g., Baumgaertel, 2000).

Selected tasks are integrated using a single integrated ontological model of the networked organisation. In other words it can be stated that these tasks are formulated using the semantics provided by the ontological model. Figure 1 represents the integration of the tasks with their input and output parameters. It also shows methods for task solving used for the implementation of the approach. As a notation the triad "input data – method – output data" is used. Customer order is considered as a driver for the entire system and the output is a feasible configuration of the networked organisation.

The ontology-driven architecture proposed deals with ontologies of different types. The ontologies are represented by means of a common notation and a common vocabulary supported by an ontology library. The common representation enables performance of operations on ontology integrations as alignment and merging, and operations on context integrations. Main components of the ontology library are domain, tasks & methods, and application ontologies. All the ontologies are interrelated so that an application ontology is a specialization both of domain and tasks & methods ontologies.

The classification of knowledge according to the abstraction and types (Neches, et. al., 1991) distinguishes *universal*, *shared*, *specific*, and *individual* knowledge abstraction levels. In the knowledge sharing model of the system “KSNet” (figure 2) the universal level is considered as the common knowledge representation paradigm. The universal level is based on the formalism of object-oriented constraint networks represented by means of a knowledge representation language. The abstractions provided at this level are shared by the ontologies stored in the ontology library. Both shared abstraction level and specific abstraction level are considered sharable and reusable since ontologies of these levels share common representation paradigm and common vocabulary.

The level of knowledge representation provides with a common notation for knowledge description and enables compatibility of different formats (e.g., KIF, OWL, etc.). Knowledge sharing level

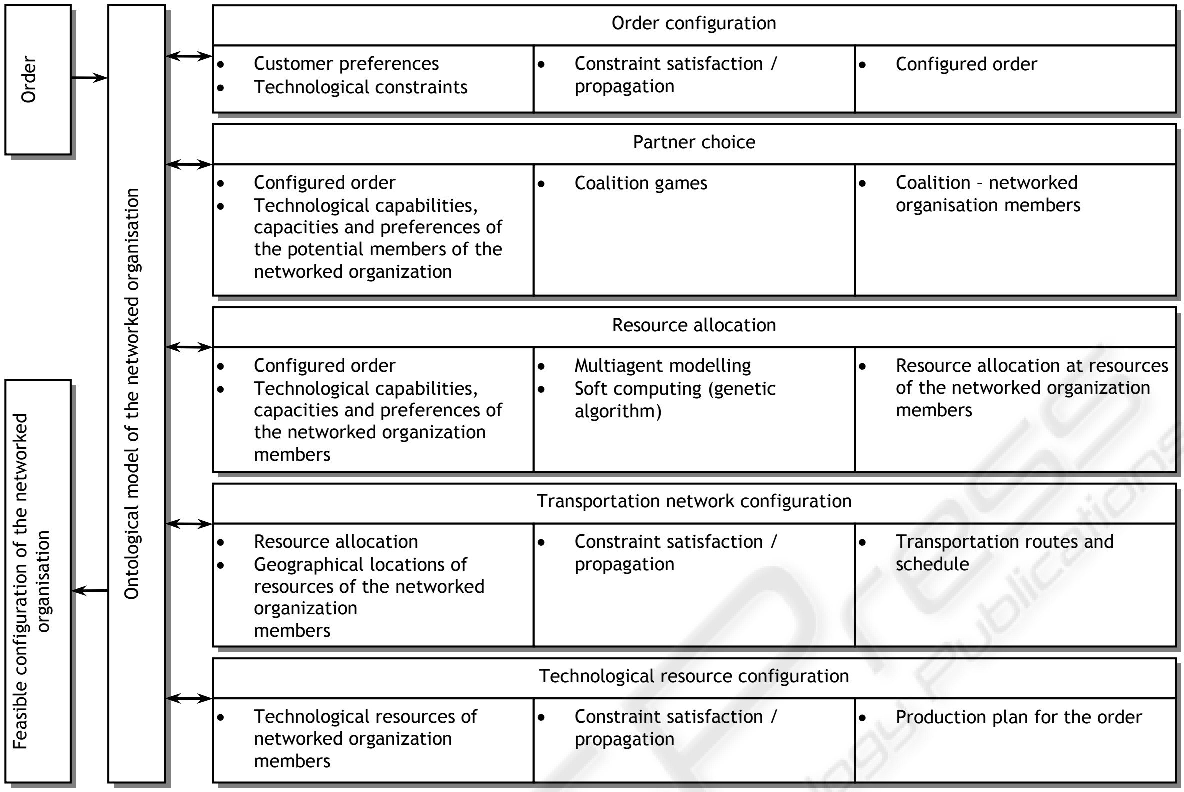

Figure 1: Tasks solved during configuration of the networked organisation.

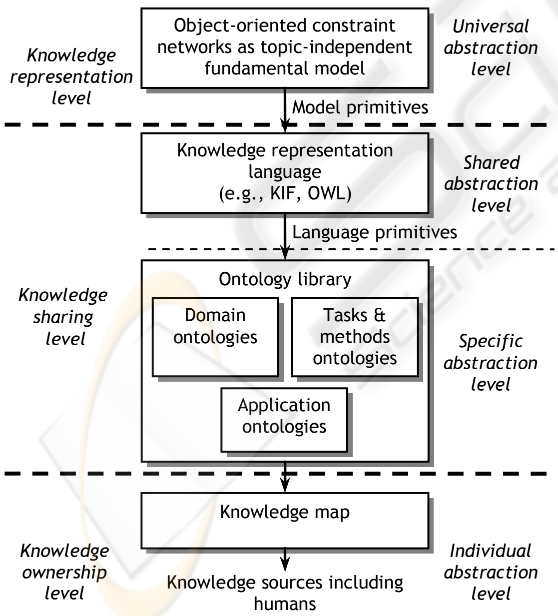

Figure 2: Ontology-driven knowledge sharing.

focuses on ontological knowledge common for a particular area. Knowledge represented by this level suits well for sharing and reuse, since, on the one hand, this level does not concentrate on any specific properties, on the other hand, knowledge of this level is not a universal abstraction rarely taken into account when the case considers practical knowledge sharing and reuse. The knowledge ownership level increases scalability of the system regarding to number of knowledge sources that can be attached to the system and users that can be served.

3 ONTOLOGICAL MODEL

As a general model of ontology representation in the system "KSNet" implementing the approach, object-oriented constraint network paradigm (Smirnov, et. al., 2003a) is used. This model defines the common ontology notation used in the system. In accordance with this representation the ontological model is defined as follows: $M = (O, Q, D, C)$. This formalism includes a set of classes O and attributes Q, Cartesian product of which is a set of variables. Each variable may have values from a certain domain $D(1), \ldots, D(n), \ldots$. The model also includes constraints of six types: $C_1, \ldots, C_6 \subset C$, defining which values the variables may take simultaneously, and relationships between classes. To solve a constraint satisfaction task means to assign values to each variable so that all constraints hold. Class "*Thing*" is used as a parent class for all classes of the ontological model, i.e., any class of the ontological model is a direct or indirect child ("is-a" relationship) of the class "*Thing*".

The following types of the constraints have been defined:

$C^I = \{c^I\}$, $c^I = (o, q)$, $o \in O$, $q \in Q$ – accessory of attributes to classes;

$C^{II} = \{c^{II}\}$, $c^{II} = (o, q, d)$, $o \in O$, $q \in Q$, $d \in D$ – accessory of domains to attributes;

$C^{III} = \{c^{III}\}$, $c^{III} = (\{o\}, \text{True} \vee \text{False})$, $|\{o\}| \geq 2$, $o \in O$ – classes compatibility (compatibility structural constraints);

$C^{IV} = \{c^{IV}\}$, $c^{IV} = \langle o', o'', \text{type}\rangle$, $o' \in O$, $o'' \in O$, $o' \neq o''$ – hierarchical relationships (hierarchical structural constraints) "is a" defining class taxonomy (*type*=0), and "has part" / "part of" defining class hierarchy (*type*=1). The most abstract class is "*Thing*".

$C^V = \{c^V\}$, $c^V = (\{o\})$, $|\{o\}| \geq 2$, $o \in O$ – associative relationships ("one-level" structural constraints);

$C^{VI} = \{c^{VI}\}$, $c^{VI} = f(\{o\}, \{q\}) \rightarrow \text{True} \vee \text{False}$, $|\{o\}| \geq 0$, $|\{q\}| \geq 0$, $q \in Q$ – functional constraints referring to the names of classes and attributes.

$|c|$ – is a number of parameters included into a constraint (constraint cardinality).

Below, some example constraints are given:

- the attribute *costs* (q_1) belongs to the class *order* (o_1): $c^I_1 = (o_1, q_1)$;
- the attribute *costs* (q_1) belonging to the class *order* (o_1) may take positive values: $c^{II}_1 = (o_1, q_1, R^+)$;
- instances of the class *standard operation* (o_2) can be compatible with instances of the class *resource* (o_3): $c^{III}_1 = (\{o_2, o_3\}, True)$;
- an instance of the class *order related operation* (o_4) can be a part of an instance of the class *order* (o_1): $c^{IV}_1 = \langle o_4, o_1, 1\rangle$;
- the order related operation (o_4) is an operation (o_5): $c^{IV}_1 = \langle o_4, o_5, 0\rangle$;
- an instance of the class *order related operation* (o_4) can be connected to an instance of the class *resource* (o_3): $c^V_1 = (o_2, o_3)$;
- the value of the attribute *cost* (q_1) of an instance of the class *order* (o_1) depends on the values of the attribute *cost* (q_1) of instances of the class *order related operation* (o_4) connected to that instance of the class *order* and on the number of such instances: $c^{VI}_1 = f(\{o_4\}, \{(o_4, q_1), (o_1, q_1)\})$.

A graphical interpretation of this model of the networked organization at macro level is presented in figure 3. The model contains one taxonomy (Thing, second level classes, third level classes) and two hierarchies (networked organization – networked organization member – Resource, and order – order related operation).

4 CASE STUDY

This section demonstrates application of the above approach in two areas: knowledge sharing within one company and knowledge sharing across a networked organisation.

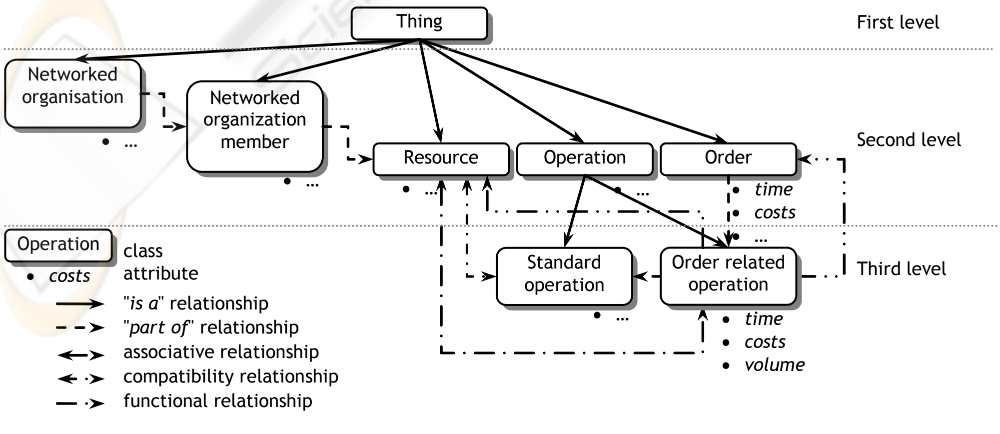

Figure 3: Ontological model of a networked organisation at macro level.

4.1 Knowledge Sharing within One Company

In this case study it was necessary to access information about products and solutions stored in various sources (documents, databases, rule bases and Web sites) for an industrial company that has more than 300.000 customers in 176 countries supported by more than 50 companies worldwide with more than 250 branch offices and authorised agencies in further 36 countries (Hinselmann et. al., 2004). Among the major tasks that had to be solved the following should be outlined:

1) keep existing facilities of the applications and avoid doubling of data;

2) extend opportunities of fast provision of information about the company's products by new features (like free text search, feature prioritisation and other);

3) provide multilingual interface;

4) implement local and Web versions of the software;

5) index existing documents against information stored in the databases.

To adopt the developed approach to the company's requirements the following tasks were solved:

1) knowledge sources were selected and interfaces for accessing them were developed;

2) an ontological model, which is a part of the company's ontology, based on available structured data was built and extended by user-defined elements and synonyms;

3) special methods to convert documents into machine readable formats were developed

4) an interface to other corporate databases was developed;

5) documents were indexed against the ontological model vocabulary and the knowledge map was created;

6) methods for calculation of the results relevance, fuzzy string comparison, and document ranking were developed.

The ontological model, which is a basis for corporate knowledge description, was built using structured information from databases and rule bases (figure 4). For this purpose a number of software modules were created that then were used for automatic creation of the ontology. It made possible to access all available information as if it was stored in a single knowledge base.

As a result a system was built that based on user (customer) requests activated appropriate knowledge sources (Web sites, documents, databases and rule bases) and provided access to them. The system was successfully tested within the company.

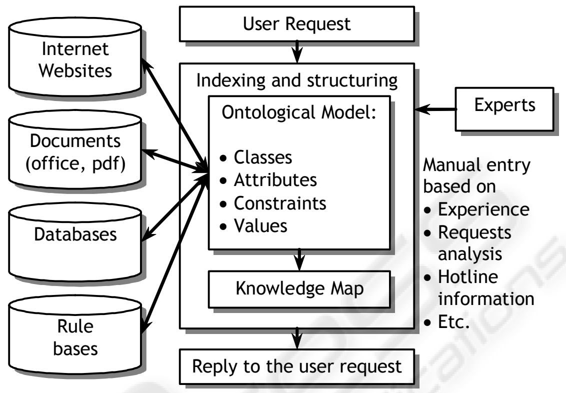

Figure 4: Relation between application ontology and knowledge sources.

4.2 Knowledge Sharing Across a Networked Organisation

The FP6 project "Intelligent Logistics for Innovative Product Technologies" (ILIPT) is devoted to development of new methods and technologies to facilitate the implementation of a new manufacturing paradigm (Stone, et. al., 2005). This new paradigm, "the 5-day car" will approach the building of 'cars to order' in a reduced time scale. ILIPT project will address the conceptual and practical aspects of delivering cars to customers only within several days after placing the order, the automotive industry's exciting and radical new business model (ILIPT, 2005). One of the tasks of the ILIPT project is development of a common knowledge management platform to support interoperability within the "5-day car" production network. This will make it possible to accumulate, share, reuse and process knowledge across the "5-day car" production network that in turn can significantly help in increasing the supply chain effectiveness and in decreasing the lead time.

This case study demonstrates creation of the ontological model for a networked organisation (a virtual production network). It was built using several source ontologies and setting relationships between their elements. Figure 5 represents establishment of relationships between task &

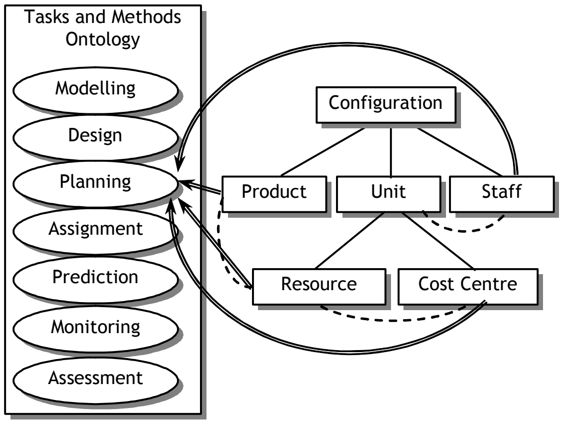

Figure 5: Task & Method and Management ontologies (a fragment).

methods ontology and domain ontology "Configuration".

Summing up definitions of networked organisation configuration (e.g., Cooper et. al., 1997, Simchi-Levi, et. al., 2000) it can be defined as configuration of flows of products and services, finances and information between different stages from a supplier to a consumer / customer and managing operational activities of procurement and material releasing, transportation, manufacturing, warehousing and distribution, inventory control and management, demand and supply planning, order processing, production planning and scheduling, and customer service across the networked organisation.

The resulting networked organisation domain ontology is given in figure 6. The figure presents the class hierarchy for the classes of the taxonomy level following the root. Management concepts are constructed to cover various stages, functions, decisions, and flows.

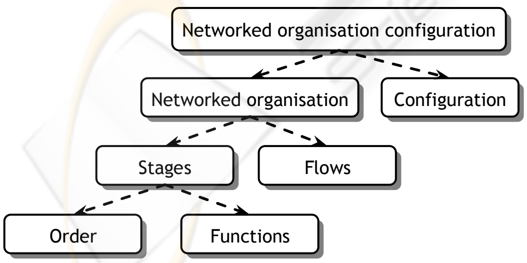

Figure 6: Networked organisation configuration domain ontology: top-level classes view.

Networked organisation activities include *flow* of information, materials, and finances between different stages from suppliers to customer (figure 7). *Information flow* includes capacity, promotion plans, delivery schedules, sales, orders, inventory, quality; *material flow* contains row materials, intermediate products, finished goods, material returns, repairs, servicing, recycling, disposal; *finances flow* is made up of credits, consignment, payments (Chopra, et. al., 2001). Detailed specializations for products and services can be found in various product ontologies and classification systems (e.g., UNSD, 2004, UNSPSC, 2001, NAICS, 2002) and mapped onto the presented classification level of the material flow.

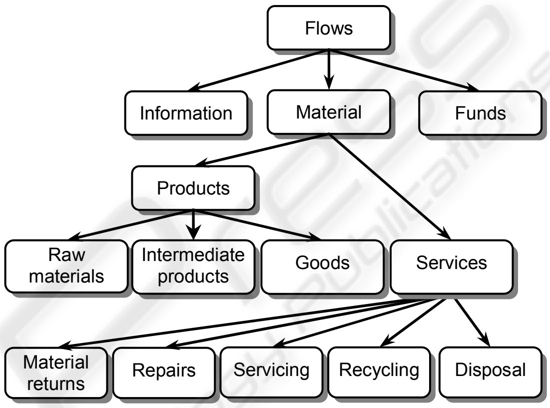

Figure 7: Networked organisation flows: taxonomy.

Virtual production network as a networked organisation is a mechanism to integrate production functions taking place at the separate stages. Most of the *functions* (figure 8) happen within various stages, some of them cross the boundaries among several stages (Chopra, et. al., 2001). The functions operate on the networked organisation flows.

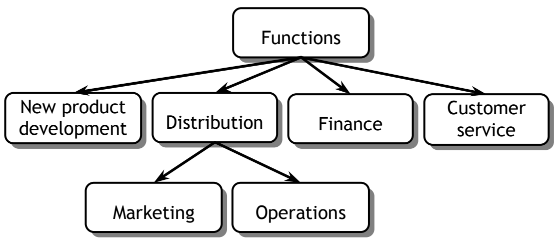

Figure 8: Networked organisation functions: taxonomy.

Part of the built ontological model presented in figure 9 focuses on the partner choice task. Since the problem considered is very complex a part of the ontology is given.

As a characteristic influencing networked organisation performance, cost is considered. In fact many cost items make up the total costs of the product required by the customer and the

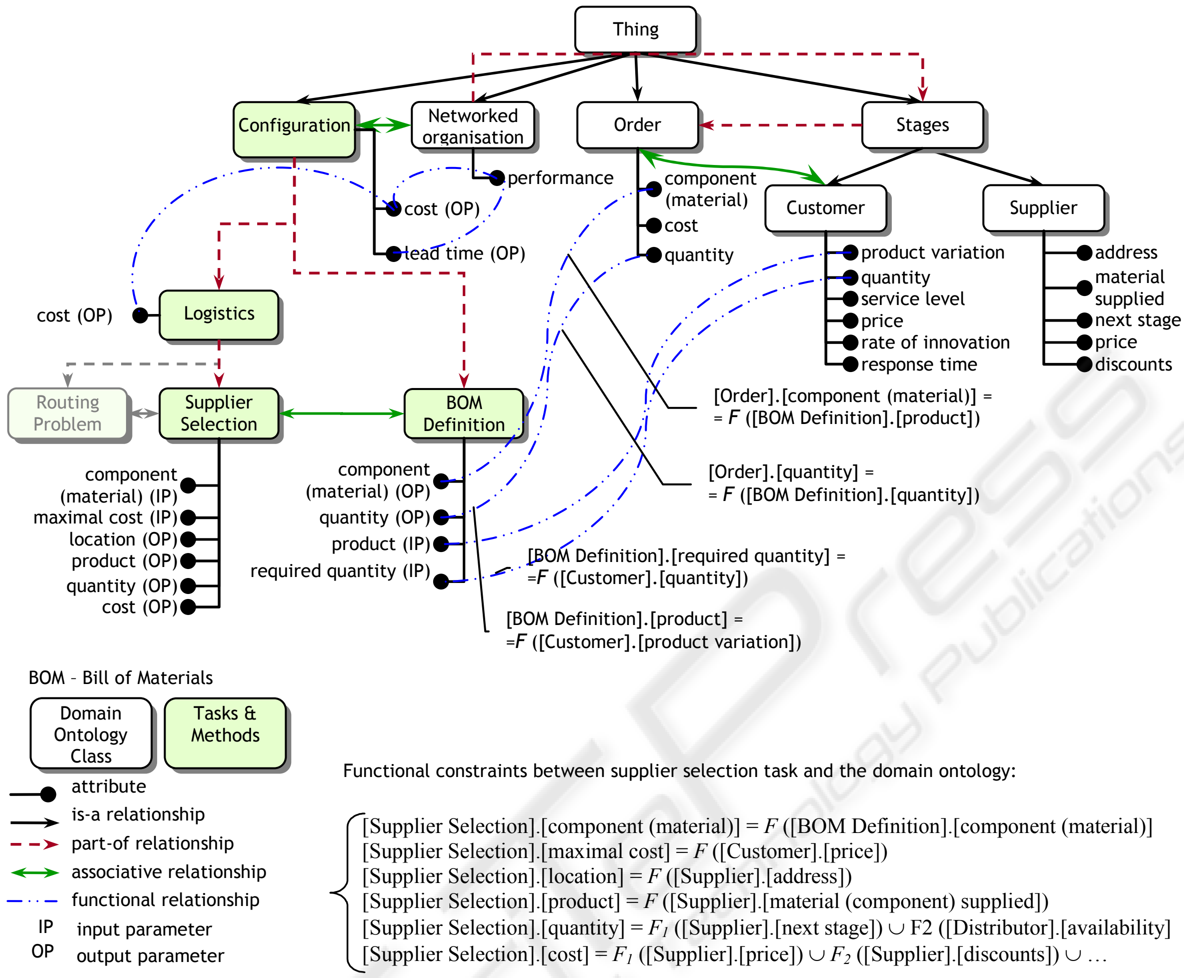

Figure 9: Application ontology for networked organisation management (a fragment).

production, among them manufacturing costs, shipping costs, and other are. This means that the complete ontological model includes all domain ontology classes that have an influence on costs. To simplify illustration interrelations between the domain ontology and the set of tasks are given by the example of the task of forming order for bill of materials (BOM). This task defines a set of materials and components that compose the product ordered by the customer.

The supplier selection task follows BOM definition task and has a set defined as input parameters. The task also takes into account maximal cost of the product that the customer is ready to pay, if any. Within the limits of the considered example the supplier selection task and the domain ontology are interrelated by the set of functional constraints shown in the bottom of figure 9.

Analogously, the supply chain performance depends on supply chain configuration cost combined with other influencing items:

[Supply Chain].[performance] =

F_1 ([Supply Chain Configuration].[cost]) $\cup$ F_2 …

5 CONCLUSIONS

The paper presents an approach to semantic information integration for intelligent decision support in networked organisations. Usage of ontological knowledge description made it possible to provide for common terminology and notation what, in turn, enabled integration of different tasks, constituting a common complex problem.

The approach has been tested in production related projects described in the section 4. One of them was implemented for an industrial company and was oriented to providing an access for users to different sources containing information and knowledge about company's products and services. The aim of other project was configuration of a BTO (build-to-order) production network consisting of several manufacturing facilities (suppliers).

Among the limitation of the approach the complexity of the common ontological model creation can be mentioned. However, the advantage of the ontological model is that it is a conceptual model of a high abstraction. Hence it can be defined for most general concepts and detailed concepts can be described only in the tasks.

ACKNOWLEDGEMENTS

The paper is due to the research carried out as a part of Integrated Project FP6-IST-NMP 507592-2 "Intelligent Logistics for Innovative Product Technologies" sponsored by European Commission, projects supported by the Russian Academy of Sciences # 16.2.35 of the research program "Mathematical Modelling and Intelligent Systems" and # 1.9 of the research program "Fundamental Basics of Information Technologies and Computer Systems", and project funded by grant # 05 01 00151 of the Russian Foundation for Basic Research.

REFERENCES

Baumgaertel, H., 2000. Distributed constraint processing for production logistics, *IEEE Intelligent Systems*, Vol. 15, No. 1, pp. 40-48.

Chopra, S., Dougan, D., Taylor, G., 2001. B2B e-Commerce Opportunities, *Supply Chain Management Review*, May / June, pp. 50-58.

Chopra, S., Meindl, P., 2001. Supply Chain Management, *Strategy, Planning, and Operation*, Prentice Hall.

Cooper, M.C., Lambert, D.M., Pagh, J.D., 1997. Supply Chain Management: More Than a New Name for Logistics, *The International Journal of Logistics Management*, Vol. 8, No. 1, pp. 1-14.

Hinselmann, T., Smirnov, A., Pashkin, M., Chilov, N., Krizhanovsky, A. Implementation of Customer Service Management System for Corporate Knowledge Utilization. *Proceedings of the 5th International Conference on Practical Aspects of Knowledge Management (PAKM 2004)*, Vienna, Austria, December, 2004, LNAI 3336, pp. 475-486.

Intelligent Logistics for Innovative Product Technologies. Web-site. URL: http://www.ilipt.org.

Interoperability Research for Networked Enterprises Applications and Software, 2004. URL: http://www.interop-noe.org.

Laudon, K. C., Laudon, J. P. 2000. *Management Information Systems: Organisation and Technology in the Networked Enterprise*. Prentice Hall International, New York, 6th ed., p. 588.\

Lipnack, J., Stamps, J. 1994. *The Age of the Network: Operating Principles for the 21st century*, John Wiley & Sons, p. 264.

Liu, M., Zhong, P., Meng, X., Liu, D., Cheng, H., 2004. Knowledge-based integrated process management in lifecycle of product development. Concurent Engineering: The Worldwide Engineering Grid, proceedings of the 11th ISPE International Conference on Concurrent Engineering (Sobolewsky, M., Cha, J., eds.), Tsighua University Press, Springer, pp. 181-186.

Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T. and Swartout, W., 1991. Enabling technology for knowledge sharing, *AI Magazine*, Vol. 12, No. 3, pp. 36-56.

Semantic Web, 2005. URL: http://www.semanticweb.org.

Simchi-Levi, D., Kaminsky, P, Simchi-Levi, E., 2000. Designing and Managing the Supply Chain, *Concepts, Strategies, and Case Studies*, McGraw-Hill, Boston.

Skyrme, D. J. 2003. The Networked Organization. URL: http://www.skyrme.com/insights/1netorg.htm.

Smirnov, A., Pashkin, M., Chilov, N., Levashova, T, Haritatos, F., 2003a. Knowledge source network configuration approach to knowledge logistics, *International Journal of General Systems*, Vol. 32, No. 3, pp. 251-269.

Smirnov, A., Pashkin, M., Chilov, N., Levashova, T, 2003b. Agent-based support of mass customization for corporate knowledge management, *Engineering Applications of Artificial Intelligence*, Vol. 16, No. 4, pp. 349-364.

Stone, G., Miemczyk, J., Esser, R., 2005. Making Build to Order a Reality: The 5-Day Car Initiative, *Strengthening Competitiveness through Production Networks. A prospective from European ICT research projects in the field of "Enterprise Networking"*, pp. 26-37.

The North American Industry Classification System (NAICS), United States, 2002. URL: http://www.census.gov/epcd/www/naics.html.

The UNSPSC Code (Universal Standard Products and Services Classification Code), DAML Ontology Library, Stanford University, 2001. URL: http://www.ksl.stanford.edu/projects/¬DAML/-UNSPSC.daml.

UN Classifications Registry, United Nations Statistics Division (UNSD), 2004 February, http://unstats.un.org/unsd/cr/registry/default.asp.