Proceedings of the Symposium on Intelligent Software Systems for the New Infostructure, 16th International Conference on Systems Research, Informatics and Cybernetics (InterSymp-2004), Baden-Baden, Germany, July 2004

Model-based Data Management for Mediation Services for Intelligent Software Agents

Andreas Tolk, Ph.D.
Virginia Modeling Analysis and Simulation Center
Old Dominion University
Norfolk, VA 23529, United States
+1 (757) 686 – 6203
atolk@odu.edu

Johnny J. Garcia
General Dynamics/ AIS
7025 Harbour View #101
Suffolk, VA 23435, United States
+1 (757) 673 – 2411
john.garcia@gd-ais.com

Keywords

Service-oriented Architectures, Extensible Markup Language (XML), Command & Control Information Exchange Data Model (C2IEDM), Mediation Services, Global Information Grid (GIG)

Abstract

Mediation services can be generally defined as a mechanism to map interchange formats (map them to what?), thus increasing the ability for disparate systems to exchange information through common methods. However, when intelligent software agents use these meditation services, syntactical translations of formats are not sufficient.

The semantic context has to be captured and interchanged as well; a common ontology is needed as the basis for the mediation service. While in the commercial world several recent publications are looking at possible automated solutions, in complex environments (is the commercial world not complex?), data engineering is necessary in order to support semantically meaningful mediation layers. Model-based data management uses a common reference model to map data models to data sources to support intelligent software agents for their internal decision processes

This paper defines the phases of data engineering, shows potential conflicts and how they can be solved, and gives an example from the military application domain by showing how the Command and Control Information Exchange Data Model (C2IEDM) developed by the North Atlantic Treaty Organization (NATO) can be used as a common reference model for military applications.

Overview

In order to support operations with rapidly changing requirements, service oriented architectures are being developed in lieu of the often too inflexible traditional solutions. As an alternative to having a system fulfilling a set of predefined requirements, services fulfilling requirements are identified, composed, and orchestrated to achieve the current users' needs in an ongoing operation. Grid Computing, System-of-Systems engineering approaches, and the Global Information Grid are examples of this trend. Intelligent agents are required to identify the service requirements in a given situation, find applicable services, compose these services to support the operation, orchestrate their execution, and evaluate the result in a way meaningful to the user.

One of the most urgent requirements in this circle is to ensure meaningful interoperability of the information exchanged between the services. This is a real challenge during the design and implementation of the service since the designer does not know with which other services this service will communicate. This can only be determined during runtime. Although in the commercial world, some applications try to do such mappings automatically, their domains are relatively simple, such as address mapping problems (Su et al. 2001). For more complex problem sets, explicit mapping of data must be done. A common way must be used to capture the format (syntax), meaning (semantics), and the use (pragmatics) of data in order to avoid ambiguity and structural variances when composing the services. This is the domain of model based data management. The results can be used to establish the required mediation layers necessary for the intelligent agents to support the process effectively.

Service Oriented Architectures

Traditional information technology (IT) followed a waterfall model starting with a set of user requirements, which led thru several stages to the system definition, system design, and system implementation. The reality of required distributed computing and the necessity for combining information resources using very heterogeneous IT infrastructures – different hardware, middleware, languages, etc. – cannot be met by such traditional efforts. Starting with the ideas of net-centric operations and setting up a system of systems, the commercial world, as well as the military world, is moving from system components delivering the operationally required functionality, towards service oriented architectures. Within the commercial world, distributed computing environments operate as a uniform service, which looks after resource management and security independently of individual technology choices. Grid computing is a means of network computing that harnesses the unused processing cycles of numerous computers to solve intensive problems that are often too large for a single computer to handle. In other words, grid computing enables the virtualization of distributed computing and data resources such as processing, network bandwidth, and storage capacity, to create a single system image which grants users and applications seamless access to vast IT capabilities. Just as an Internet user views a unified instance of content via the Web, a grid user essentially sees a single, large virtual computer. In order to access the functionality, services are defined based on common open standards and bridge the gap between the heterogeneous worlds of different languages, middleware solutions, and hardware. The authors perceive web services to be currently the strongest candidate for a technical solution to instantiate a service-oriented architecture.

This trend can be observed in the military world. Following the ideas of net-centric warfare (Alberts and Hayes, 2003), future military operations will be characterized by the seamless sharing of information and other resources. The technical backbone enabling this vision will be the Global Information Grid (DoD, 2002), which will be implemented using the Internet Protocol version 6 as the technical baseline. It will establish a service-oriented architecture of military services, from command and control to modeling and simulation, supporting the soldiers in all relevant military operations.

The real potential of service oriented architectures lies in the possibility to compose services and to orchestrate their execution, thus enabling new functionality compositions to fulfill the current often changing user requests within an ongoing operation. Information must be exchangeable between all composed services in a meaningful manner, i.e., not simply exchanging bits and bytes, but ensuring the interpretation of data in a consistent way leading to the same information, knowledge, and ultimately awareness within the services and their users; each service has to know what data is located where, the meaning of data and its context, and into what format the data has to be transformed to be used in respective services composed into a distributed application within the overall system. To generate the answers to these questions is the objective of data administration, data management, data alignment, and data transformation. These can be defined as the building blocks of a new role in the interoperability process: Data Engineering (Tolk, 2004). The composing terms are defined as follows:

Data Administration is the process of managing the information exchange needs that exist between the services, including the documentation of the source, the format, context of validity, and fidelity and credibility of the data. Data Administration therefore is part of the overall information management process for the service architecture. Data Management is planning, organizing and managing of data by defining and using rules, methods, tools and respective resources to identify, clarify, define and standardize the meaning of data as of their relations. Data Alignment ensures that the data to be exchanged exist in the participating systems as an information entity or that the necessary information can be derived from the data available, e.g., using the means of aggregation or disaggregation. Finally, Data Transformation is the technical process of aggregation and/or disaggregation of the information entities of the embedding systems to match the information exchange requirements including the adjustment of the data formats as needed.

Model based data management uses a reference model to capture the planning, organizing and managing of data. Instead of mapping point-to-point, the information exchange requirements of a service are mapped to a common information exchange reference model, which can be seen as the ontology of the application domain.

C2IEDM

In 1978, NATO's Long-Term Defense Plan (LTDP) Task Force on C2 recommended that an analysis be undertaken to determine if the future tactical Automatic Data Processing (ADP) requirements of the Nations (including that of interoperability) could be obtained at a significantly reduced cost when compared with previous approaches. In early 1980, the then Deputy Supreme Allied Commander Europe initiated a study to investigate the possibilities of implementing the Task Force's recommendations. This resulted in the establishment of the Army Tactical Command and Control Information System (ATCCIS) Permanent Working Group (APWG) to deal with the challenge of the future C4I systems of NATO. The ATCCIS approach was designed to be an overall concept for the future command and control systems of the participating nations. One constraint was that each nation could still build independent systems. To meet this requirement, ATCCIS defined a common kernel to facilitate common understanding of shared information. In 1999, ATCCIS became a NATO standard with the new name Land Command and Control Information Exchange Data Model (LC2IEDM). In parallel to this, the project managers of the Army Command and Control Information Systems (C2IS) of Canada, France, Germany, Italy, the United Kingdom, and the United States of America established the Multilateral Interoperability Program (MIP) in April 1998. By 2002, the activities of LC2IEDM and MIP were very close, expertise was shared, and specifications and technology were almost common. The merger of ATCCIS and MIP was a natural and positive step. LC2IEDM became the data model of MIP. Finally, in 2003 the name was changed to Command and Control Information Exchange Data Model (C2IEDM).

There are two application domains for the C2IEDM within NATO: Data Management and Information Exchange. The NATO Data Administration Group used the C2IEDM to map all information exchange requirements between the national command and control systems to it in order to ensure semantic (What do the data mean?) and pragmatic (What are the data used for?) interoperability between the systems. The MIP data managers will continue this task after the merger between MIP and C2IEDM is finished. MIP also uses the C2IEDM to exchange data between national command and control systems in order to foster sharing information and gain a common understanding on what is happening on the battlefield. To this end, the national systems establish data translation layers mapping their internal data presentation to the data elements of C2IEDM for information exchange with the other systems.

A technical view on the data model goes far beyond the scope of this paper, as C2IEDM comprises data elements describing a common vocabulary consisting of 176 information categories that include over 1500 content elements. In order to cope with these needs, C2IEDM is divided into a Generic Hub comprising the core of the data identified for exchange across multiple functional areas. It lays down a common approach to describe the information to be exchanged and is not limited to a special level of command, force category, etc. In general, C2IEDM describes all objects of interest on the battlefield, e.g., organizations, persons, equipment, facilities, geographic features, weather phenomena, and military control measures such as boundaries. In addition to this, special functional areas are defined extending the Generic Hub under national responsibility to cope with information exchange needs of national concern. A tutorial on C2IEDM is given in (Loaiza, 2004). The complete data model documentation and additional information is available on the Internet (MIP, 2004).

C2IEDM for Intelligent Agents

The use of C2IEDM for intelligent agents to collaborate and support within the Global Information Grid is already presented in (Pohl, 2004). The authors share this view and are motivating this as follows: The application of the Extensible Mark-up Language (XML) enabled a new level of interoperability for heterogeneous IT systems. However, although XML enables separation of data definition and data content, it doesn't ensure that data exchanged is interpreted correctly by the receiving system. A common reference model defining the XML tag sets and the structure is needed to ensure meaningful interoperability. For military operations, the C2IEDM has the potential to become such a reference model. Intelligent agents can use both approaches. C2IEDM structured in XML schemas can become the "language" spoken by the agents and used to communicate between agents and services. As the operational use is also part of the C2IEDM agreements, even pragmatic interoperability can be reached; the C2IEDM becomes an ontological layer for the GIG. The ultimate use is that the combination of services, agents, and the common ontology enables a quantum leap in command and control quality, as described in (Alberts and Hayes, 2003): When data are put into context, the result is information; applied information in form of procedural access leads to knowledge; and knowledge applied to analyses leads finally to awareness. When the text based messages where replaced by a common operating picture, we stepped up from data to information. The introduction of modeling and simulation services introduces procedural knowledge. If intelligent software agents can now use these services and map them to observations of the real world, they can support the analyses of military experts.

Hence, the command of control systems incorporating agents, services, and a common ontology can even support awareness, which traditionally is seen in the cognitive domain only.

Summary

Service-oriented architectures are evolving rapidly. In order to make better use of the services delivering the necessary functionality, intelligent agents are required. The agents need to make sense of the services functionality, their use of data, and their behavior. To this end, effective (metadata has not been mentioned up to the point, why is it mentioned in the summary???) meta-data and meta-model management is necessary. One of the most challenging tasks in this context is the management of information exchange requirements of services in a way that these services can be composed and orchestrated with other services during runtime delivering the functionality as specified by user during the ongoing operation. In the military environment, the Global Information Grid is the technical backbone. The Command and Control Information Exchange Data Model (C2IEDM) is a matured approach for a military ontology in the domain of command and control. C2IEDM has been proven to be flexible enough to cope with all information exchange requirements of services. Technically, the definition of mediation layers to make this information available to intelligent agents is feasible. XML can be used as a syntax layer, the C2IEDM definitions can be used for ensuring semantic interoperability, and the C2IEDM structures and views – which have been agreed to by military operators of the developing nations – can insure that the pragmatic view, i.e. how the data is used – is aligned as well.

References

Alberts, D. and Hayes, R. (2203); Power to the Edge – Command, Control, in the Information Age; Command and Control Research Program Publications

Department of Defense (2002); DoD Directive 8100.1: Global Information Grid (GIG) Overarching Policy; The Pentagon, Washington, D.C.

Loaiza, F. (2004); C2IEDM Tutorial; Proceedings of the 2004 C2IEDM Workshop, Fairfax, VA

MIP (2004); Multilateral Interoperability Programme website; http://www.mip-site.org [last visited June 2004]

Parent, C. and Spaccapietra, S. (1998); Issues and approaches of database integration; Communications of the ACM; Volume 41 (5), pp. 166 - 178

Pohl, J. (2004); C2IEDM as a Component of the GIG Architecture; Proceedings of the 2004 C2IEDM Workshop, Fairfax, VA

Spaccapietra, S., Parent, C. and Dupont, Y. (1992); Model Independent Assertions for Integration of Heterogeneous Schemas; Very Large Database (VLDB) Journal, Vol. 1 (1), pp. 81-126

Su, H., Kuno, H. and Rundensteiner, E. (2001); Automating the transformation of XML documents; Proceedings of the third international workshop on Web information and data management; Session: Web Information Integration, pp. 68-75

Tolk, A. (2004); Common Data Administration, Data Management, and Data Alignment as a Necessary Requirement for Coupling C4ISR Systems and M&S Systems; Journal on Information & Security, Vol. 12