Interoperability between Distributed Systems and Web-Services Composition

Christophe Nicolle
Université de Bourgogne, France

INTRODUCTION

An information system is a multi-axis system characterized by a “data” axis, a “behavioral” axis, and a “communication” axis. The data axis corresponds to the structural and schematic technologies used to store data into the system. The behavioral axis represents management and production processes carried out by the system and corresponding technologies. The processes can interact with the data to extract, generate, and store data. The communication axis relates to the network used to exchange data and activate processes between geographically distant users or machines. Nowadays, technologies required for interoperability are extended to deal with the semantic aspect of the information systems. The aim of the semantic axis is to take into account new aspects of the sharing of the data and the processes, such as the understanding of the data and the processes, the access security, and owner rights (OWL Services Coalition, 2006).

Information system interoperation has emerged as a central design issue in Web-based information systems to allow data and service sharing among heterogeneous systems. Data heterogeneity stemming from the diversity of data formats or models used to represent and store information in the Web is a major obstacle to information systems interoperability. These data models range from the structured data models (network, relational, OO) found in traditional databases to flat files and emerging Web oriented semi-structured models. Information system interoperability aims at supporting the amalgamation autonomous heterogeneous systems to create integrated virtual environments or architectures in which information from multiple disparate sources can be accessed in a transparent and efficient manner. As an example of such integrated virtual systems, consider an airline reservation system based on the integration of a group of airlines reservation and ticket sale information systems. The specific airline systems provide various types of fares and services.
special discount trips. That can be searched and compared to respond to user queries for finding the best available prices for specified flights.

**BACKGROUND**

Database interoperability issues have been extensively studied in the past. Several approaches, including database translation, distributed systems, federations, language based multidatabase, ontology, and mediation, have been proposed to bridge the semantic gaps among heterogeneous information systems.

The *database translation* approach is a point-to-point solution based on direct data mappings between pairs of information systems. The mappings are used to resolve data discrepancies among the systems (Yan & Ling, 1992). The database translation approach is most appropriate for small-scale information processing environments with a reduced number of participants. The number of translators grows with the square of the number of components in the integrated system. For example, consider two information systems, IS1 and IS2 in the earlier travel agency example. The corresponding translators must be placed between the information systems as shown in Figures 2 and 3. Information in IS1 is represented by vertical lines, while the information in IS2 is shown as horizontal lines.

In the *standardization* approach, the information sources use the same model or standard for data representation and communication. The standard model can be a comprehensive metamodel capable of integrating the requirements of the models of the different components (Atzeni & Torlone, 1997). The use of a standard metamodel reduces the number of translators (this number grows linearly with the number of components) to resolve semantic differences. However, the construction of a comprehensive metamodel is difficult; the manipulation of high-level languages is complex; and there are no unified database interfaces. In our example, the travel agencies must define a common model to export their data. A centralized information system can be built to replace the original information systems (IS1, IS2). The global centralized schema is a combination of the entire data (horizontal and vertical lines) contained in IS1 and IS2.

*Federated systems* consist of a set of heterogeneous databases in which federation users can access and manipulate data transparently without knowledge of the data location (Sheth & Larson, 1990). Each federation database includes a federated schema that incorporates the data exported by one or more remote information systems. There are two types of federations. A tightly coupled federation is based on a global federated schema that combines all participant schemas. The federated schema is constructed and maintained by the federation administrator. A loosely coupled federation includes one or more federated schemas that are created by users or the local database administrator. The federated schema incorporates a subset of the schema available in the federation. This approach becomes complex when the number of translators required becomes large. In our example, the existing information systems are completely operational for local users. Only the shared data