Chapter IV

The Weather Tool: An Agent-Based Approach to Information Integration

Gunjan Kalra, Quantum Leap Innovations, USA

Abstract

This chapter discusses the process of providing information in its most accurate, complete form to its users and the difficulties faced by the users of the current information systems. The chapter describes the impact of prevalent technologies such as the multi-agent systems and the Semantic Web in the area of information supply via an example implementation and a model use case. The chapter offers a potentially more efficient and robust approach to information integration and supply process. The chapter intends to highlight the complexities inherent in the process of information supply and the role of emerging information technologies in solving these challenges.
Introduction

The progress in distributed computing has played a significant role in facilitating replacement of large, monolithic databases with smaller and conceptually self-contained databases. This has resulted in efficient yet “ever-expanding” clusters of heterogeneous sources of data (data sources) such as the traditional databases, Web sites, e-mails, and operational applications. The need to view and analyze data from various data sources unified under one application context has become more prevalent and has gained prominence among data analysts, both in the information technology industry and academia. For example, large-scale enterprises have a growing demand for enterprise integration products, while data mining researchers need a unified view of data from different sources to select and test their algorithms. The task of providing integrated data management within one umbrella is particularly challenging. In this chapter, we describe the need and evolution of the Weather Tool system, a weather data provisioning and integration system accessible via a standard application programming interface (API). The chapter begins with an introduction to the evolution of sources of data and to the problem of weather data integration. Some of the previous and existing data integration technologies and products are then summarized. The chapter will also address some of the new technologies that the data/information integration systems benefit greatly from, including the agent technology and the Semantic Web. We then provide an overview of the Weather Tool architecture (and an example application) that enables integration of several data sources to provide a unified data environment. The architecture discussion describes how data processing services are provisioned in a dynamically-changing environment. The advantages and pitfalls of such a system are elaborated, followed by a discussion of potential improvements in the architecture. The chapter briefly discusses the core ideas of Quantum Leap Innovation, Inc. (QLI) intelligent data management (IDM) framework, the generalized, enhanced version of the Weather Tool effort. This framework is aimed at performing beyond the Weather Tool’s single repository approach to data integration by incorporating several strategies for on-demand, data access and analysis remotely, thereby avoiding creation of redundant intermediary repositories. Finally, the chapter highlights the need, evolution, and importance of data/information integration efforts.

Evolution from Large Databases to Large Data Environment

Historically, due to high set-up and maintenance costs for databases, extremely large volumes of data were stored in one database; this strategy was considered ef-