Semi-automatic Database Design for Neuroscience Experiment Management Systems

Hao Li, BSc, James F. Brinkley, MD, PhD, John Gennari, PhD

Structural Informatics Group,
Departments of Medical Education and Biomedical Informatics, and Biological Structure,
University of Washington, Seattle, WA 98195

Abstract

Neuroinformatics provides tools for neuroscience researchers to study brain function. In order to handle experiment paradigms that change frequently, we are developing a semi-automatic database design tool that will enable an experiment management system (EMS) to manage data with flexibility while retaining the efficiency of a relational database.

Keywords:

Neuroinfomatics, Protégé, databases, information system.

Electronic experiment management systems (EMS), which often consist of a backend database and a web-based browser front end, enable the management of large quantities of neuroscience experiment data. In addition to accommodating heterogeneous data, the task of designing an EMS is challenging due to frequent experiment paradigm shifts[1]. The process of design and development of an EMS often cannot catch up with the rate at which experiment paradigms change. This is the motivation for the University of Washington Human Brain Project effort to develop a generalized EMS design tool.

Figure 1. EMS System Design Flow

The driving application for the project is a single-cell recording (SCR) database developed for a neuroscience research group, which studies the correlation of human brain anatomy with language identification and memory ability. In these experiments, which are conducted during open brain surgeries, single neuron activity in cortical language areas is recorded and analyzed later for correlation with patient verbal responses.

The EMS design tool allows neuroinformaticists to specify the data model using Protégé[2], a knowledge base design tool developed at Stanford. Protégé provides a graphical user interface for quick manipulation of complex data object relationships (Fig. 1.1). The use of Protégé separates the design process from EMS implementation. Given the data model the system then semi-automatically generates the database schema (Fig. 1.2), which is stored in a PostgreSQL object relational database (Fig 1.3). It is important that PostgreSQL is an object-relational system, which allows easier transformation from the Protégé frame-based system. The resulting PostgreSQL schema captures the data model described in Protégé. Finally, a web based user interface is semi-automatically generated in JSP and Java Servlets (Figure 1.4). The user interface allows neuroscientists to manage and access data in the database via a webpage. Thus, Figure 1 shows the process for which we rapidly generate both databases and data entry clients as models of experimental data change.

In the near future, we will evaluate the prototype EMS for single-cell-recording. We will then evaluate the EMS design tool for its extensibility and portability to other neuroscience experiment research projects. As the tool becomes more robust and automatic we expect that data modeling using Protégé, followed by semi-automatic schema and user interface evolution, will allow quick and easy redesign to accommodate experiment paradigms that change frequently. Such flexibility will enable neuroscience and other basic science researchers to manage and share data with greater efficiency.

References

1. G. M. Shepherd, J. S. Minsky, M. D. Healy, et. al. The Human Brain Project: neuroinformatics tools for integrating, searching and modeling multidisciplinary neuroscience data. *Trends in Neurosciences*, 1998, 21:*11*:460-468.

2. J. Gennari, M. A. Musen, R. W. Fergerson, et. al. The Evolution of Protégé: An Environment for Knowledge-Based System Development. *International Journal of Hum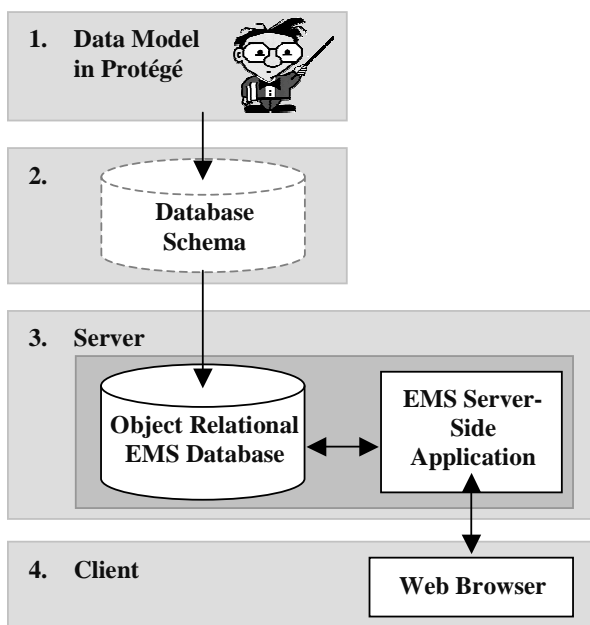an-Computer Studies*, vol 58 (1): 89-123, 2003.