An Improved Natural Language Interface to a Complex Anatomical Knowledge Base

Gregory M. Distelhorst, Kevin P. Hinshaw, PhD,
Cornelius Rosse, MD, DSc, and James F. Brinkley, MD, PhD
Structural Informatics Group, Departments of Biological Structure and Medical Education and Biomedical Informatics,
University of Washington, Seattle, USA

Abstract

We are developing a natural language interface to a large knowledge base of anatomy. We describe a new version of this interface, which is more efficient, user-friendly, and flexible than the first version.

The Foundational Model of Anatomy (FMA) is a large, complex knowledge base representing structural anatomical information (http://sig.biostr.washington.edu/projects/fm). To facilitate human queries to the FMA, we are developing a natural language interface that should be applicable to other large knowledge bases [1]. This program, called GAPP, takes English questions as input and attempts to answer them by generating corresponding queries to the FMA. An input question is parsed, expressed in the StruQL database query language, and submitted to a query server to the FMA , which returns the results of the query in an XML document (http://sig.biostr.washington.edu/projects/gapp/).

Our latest version, GAPP2, contains a number of improvements that bring it closer to an evaluation by end-users. To improve *efficiency* we now run the program as an online server. We also keep a cache of anatomical entities in memory, thereby avoiding a costly search through the FMA's huge namespace. In GAPP2 most queries take less than 1 second to answer, as opposed to 5-10 seconds in the earlier version

Figure 1 – A failed query

To make GAPP more *user-friendly*, we have added a diagnostic message that shows users how GAPP interpreted the question, and if it failed, which parts of the question it was unable to understand (Fig. 1).

Figure 2 – Natural language output

We have also added an option for "natural language" output that reformats the XML results into a set of English sentences. For example, "What innervates the liver?" generates the answer, "The Hepatic nerve plexus is the nerve supply of the Liver." A similar example is shown in Fig. 2. These results include hyperlinks to entries in the Foundational Model Explorer, an online browser for the FMA.

Finally, to accommodate a *greater range of question formats*, we updated the syntactic parser to MINIPAR, a more complex and customizable parser than our original version. Fig. 2 shows a new type of question that only GAPP2 can answer.

These improvements have resulted in greater acceptability to local users, and should allow us to evaluate GAPP on a wider scale.

1.Distelhorst G, Srivastava V, Rosse C, Brinkley JF. A prototype natural language interface to a large complex knowledge base, the Foundational Model of Anatomy. In: AMIA Fall Symposium. In Press.; 2003.