

Project Year

2011-2012

Project Title

Natural Language Processing Java-Based e-Library

Project Team

Sanjeev Khudanpur, Whiting School of Engineering, Center for Language & Speech, Faculty
Xiaoxu Kang, Whiting School of Engineering, Biomedical Engineering Major, Fellow

Audience

Students enrolled in Natural Language Processing courses.

Pedagogical Issue

Undergraduates are burdened with many courses and much of the time they are taught techniques and theories separately, so they are often at a loss as to how to use a tool to solve real world problems. We observe this in courses involving the theories and applications of natural language processing and translation.

Solution

It is time for a more integrated educational approach. A Java-based e-library will allow students to solve real problems in natural language processing and translation with a freely available general-purpose programming language (Java) they are all familiar with. Abstract theories and algorithms will be turned into real implementations that students can test.

Technologies Used

Java, HTML, Adobe, PowerPoint

Project Abstract

Undergraduates are burdened with many courses and much of the time they are taught techniques and theories separately, so they are often at a loss as to how to use a tool to solve real world problems.

The proposed project is designed following the spirit of integrated educations. A Java-based e-library will show students how to solve real problems, in natural language processing and translation in our case, with a freely available general purpose programming language (Java) they are all familiar with. Abstract theories and algorithms will be turned into real implementations that students can try and test. With this resource, the learning curve for understanding theories will be much less steep. At the same time, students are more likely to be motivated to learn advanced programming techniques, and the principles they acquire can be carried into the learning of other programming languages later in their studies.

The project will be composed of three major natural language processing and translation modules namely, a semantics processing unit, a grammar induction unit, and vocabulary library construction. Corresponding user manuals will be written. On completion, the project will be made available online for students to use and provide feedback for further improvement and development.

The completion of the project will benefit several departments that work with natural language processing such as the Department of Computer Science, the Department of Biomedical Engineering, and the Department of Cognitive Science. Courses such as *Introduction to Java Programming* and *Natural Language Processing* will receive direct benefit.

Our proposed integrated educational model is unique in the sense it can boost students' performance in both theoretical and technical subject learning at the same time. The final e-library will be an invaluable learning and research tool that will be accessible online by any student who is interested in Java programming learning and natural language processing and translation research.