



Universitat de València, 10-12 marzo de 2010

El laboratorio virtual PAG para la experimentación con modelos formales del lenguaje
The virtual laboratory PAG to experiment with formal language models

Ana Fernández-Pampillón Cesteros
Universidad Complutense de Madrid
apampi@filol.ucm.es

José Luis Sierra Rodríguez
Universidad Complutense de Madrid
jlsierra@fdi.ucm.es

María Matesanz del Barrio
Universidad Complutense de Madrid
mmatesanz@filol.ucm.es

Antonio Sarasa Cabezuelo
Universidad Complutense de Madrid
asarasa@fdi.ucm.es

Resumen

El laboratorio virtual PAG (Prototipado de Gramáticas de Atributos) surge de la necesidad, común en diferentes áreas de conocimiento, de ayudar a los estudiantes a diseñar y experimentar modelos formales del lenguaje, tanto artificial como natural. Ha sido creado, en el campus virtual de la Universidad Complutense de Madrid, por un equipo interdisciplinar de profesores de las Facultades de Informática y Filología, en el marco de los trabajos realizados en sucesivos PIMCD (Proyectos de Innovación y Mejora de la Calidad de la Docencia). Se utiliza, desde el curso 2006-07, en las asignaturas de Lingüística Computacional y de Procesadores de Lenguaje pertenecientes a las titulaciones de Lingüística e Ingeniería Informática, respectivamente.

El laboratorio tiene una aplicación didáctica e investigadora. Desde el punto de vista didáctico, el laboratorio ayuda a los estudiantes a: i) entender el valor que tiene el diseño de especificaciones formales para el desarrollo de procesadores del lenguaje de calidad; ii) aprender a formalizar y abstraer aspectos, fenómenos y procesos lingüísticos; iii) relacionar y aplicar todos los conocimientos adquiridos en sus titulaciones para resolver problemas sobre la estructura y funcionamiento de lenguajes; iv) mejorar las destrezas informáticas, en el caso de los estudiantes de lingüística; y iv) facilitar el conocimiento y trabajo interdisciplinar, lingüístico e informático. Esta última aplicación tiene particular relevancia ya que promueve la formación de futuros equipos mixtos de profesionales, capaces de cooperar para llevar a cabo tareas más complejas que requieran conocimientos de áreas diferentes. Desde el punto de vista investigador, el laboratorio aporta un entorno para: v) probar modelos formales del lenguaje sobre los que se realizarán futuros desarrollos, como por ejemplo, los modelos de gramáticas XML para el procesamiento

de lenguajes documentales (Sarasa et al. 2009); y vi) facilitar el trabajo de equipos interdisciplinarios en un espacio accesible para todos los miembros.

El laboratorio contiene i) la herramienta PAG, ii) un corpus de casos de estudio, iii) material didáctico, iv) espacios para el trabajo de los equipos, y v) un foro general (López Alonso, 2009). La experimentación de los modelos formales se realiza con la herramienta PAG (Sierra et al., 2006). Esta herramienta permite crear dinámicamente, a partir de especificaciones formales escritas con un formalismo basado en las gramáticas de atributos (Paaki, 1995), un analizador para dicha especificación. El usuario puede analizar ejemplos del lenguaje y comprobar los resultados del análisis de forma gráfica mediante un árbol de análisis, o varios si la gramática es ambigua, con los valores de los atributos (Sierra; Fernández-Pampillón, 2008). Los casos de estudio del laboratorio constituyen muestras que el usuario puede utilizar como base para crear las suyas propias o, simplemente, para entrenarse en el manejo de PAG. Las herramientas de grupos y de comunicación proporcionan el soporte necesario para crear redes de equipos de trabajo.

Este artículo presenta la experiencia didáctica e investigadora del laboratorio a través del corpus de casos de estudio creados por los estudiantes, y analiza las posibilidades de un laboratorio virtual de estas características para mejorar el aprendizaje experimental interdisciplinar con TICE.

Palabras clave: laboratorio virtual, modelos del lenguaje, aprender e investigar.

Abstract

The virtual laboratory PAG (Prototyping with Attribute Grammars) was created as part of the UCM Virtual Campus. It was born from the common necessity, in different areas of knowledge, to facilitate student work using spaces for the autonomous creation of and experimentation with formal language models (for both artificial and natural languages). The lab is a result of the collaboration between the Faculties of Computer Science and Philology, within the framework of the research performed in consecutive PIMCDs (Innovation and Improvement in Educational Quality Projects) carried out since the 2006-2007 academic year.

The PAG lab has a double purpose. On the one hand, from a didactic perspective, its aim is to assist students in the following ways: (1) to help them value, in the development of language processors, the importance of studying and designing formal quality specifications; (2) to help them correctly carry out the processes of formalisation and abstraction of linguistic aspects, phenomena or processes; (3) to assist them in using the knowledge acquired in their respective degree programs to solve problems related to language structure and performance; and (4) to provide interdisciplinary knowledge and work experience to our students in linguistics and computer science. We are highly interested in this last goal because this effort stimulates the creation of mixed working teams capable of cooperation in executing more complex tasks that require knowledge of different areas. On the other hand, from the point of view of research, the lab does the following: (5) facilitates experimentation on formal language models—for instance, XML grammar models; and (6) encourages interdisciplinary teamwork in an environment that is accessible to all its members. The lab is already being used in Computational Linguistics (in Linguistics degree) and Language Processors (in Computer Science degree); during the 2009-2010 school year, another course, Formal Languages and Automata, was scheduled to be offered.

In the lab, experimentation on formal models is performed using the PAG tool. This tool allows one to dynamically create an analyser according to written specifications with a formalism based on attribute grammars. The user can introduce instances and graphically test the results of the analysis through one or more analysis trees (depending on the ambiguity of the grammar) formed with the values of the attributes. Additionally, the laboratory has a case-study corpus made up of samples for the user to apply as a base to create his own specifications or, simply, to train using the PAG. Finally, the lab's communication and group tools provide the necessary support to create user networks.

In this article, we introduce our experience based on the corpora of case studies created by our students and demonstrate the advantages of sharing virtual environments for interdisciplinary and experimental learning as supported by ICT.

Key words: virtual laboratory, language models, learning and investigating.

Referencias bibliográficas / Bibliographical references:

- López Alonso, C., Fernández-Pampillón Cesteros, A., De Miguel García, E., Matesanz del Barrio, M. (2009) "E.Ling, un entorno b-learning para docencia, investigación e innovación". En *Las plataformas de aprendizaje: del mito a la realidad*. Madrid, Biblioteca Nueva, pp. 119-148.
- Paaki, J. (1995). "Attribute Grammar Paradigms – A High-Level Methodology in Language Implementation". *ACM Computer Surveys*, 27, 2, 196-255. DOI=<http://doi.acm.org/10.1145/210376.197409>
- Sierra, J.L.; Fernández-Valmayor. A. (2006). "A prolog Framework for the rapid prototyping of language processors with Attribute Grammars". *Electronic Notes in Theoretical Computer Science* 164(2), 19-36.
- Sierra, J. L.; Fernández-Pampillon, A.; Fernández-Valmayor, A. (2008). "An environment for supporting active learning in courses on language processing" .In *Proc. of the 13th Annual Conference on innovation and Technology in Computer Science Education*. ITiCSE08. ACM, NY.
- Sarasa, A.; Sierra. J.L.; Fernández-Valmayor, A. (2009) "Processing Learning Objects with Attribute Grammars". 8th IEEE Int. Conf. in Advanced Learning Technologies.