

AMPLÍA LA REPRESENTACIÓN COMPUTACIONAL DE LAS LENGUAS

Crean un modelo computacional inteligente de la gramática descriptiva de la lengua española

El primer modelo computacional inteligente de la gramática descriptiva de la lengua española, aplicable a gramáticas descriptivas de otras lenguas, ha sido desarrollado por investigadores de la Facultad de Informática de la Universidad Politécnica de Madrid. Han aplicado metodologías de elicitación del conocimiento propias de la Ingeniería del Conocimiento (rama de la Inteligencia Artificial) a la gramática descriptiva de la lengua española, que ha sido utilizada como fuente de conocimiento.

UPM

16/7/2008 09:23 CEST

Investigadores del [Grupo de Validación y Aplicaciones Industriales](#) (VAI) de la Facultad de Informática de la Universidad Politécnica de Madrid (FIUPM), han desarrollado un modelo computacional inteligente de la gramática descriptiva de la lengua española, abriendo así nuevas posibilidades a la representación computacional de las lenguas y a sus aplicaciones de procesamiento de lenguaje natural.

La lingüística computacional se ha basado fundamentalmente en teorías lingüísticas para construir modelos de representación de las lenguas para aplicaciones computacionales. Las teorías lingüísticas son modelos formales (es decir, expresables matemáticamente) que necesitan entre 5 y 10 años para desarrollar un modelo en una lengua determinada y con una cobertura del 55%. Es decir, la cobertura es muy limitada y el coste es enorme, lo cual dificulta que lenguas diferentes del inglés o de los idiomas

más dominantes tengan aplicaciones útiles.

Para superar este obstáculo, los investigadores [Carolina Gallardo](#) y [Jesús Cardeñosa](#) han estudiado la posibilidad de sustituir las teorías lingüísticas por las gramáticas descriptivas, que aunque no son formales, representan el uso real de una lengua.

A pesar de su “aspecto” poco formal, las gramáticas descriptivas contienen una gran cantidad de conocimiento lingüístico, explican estos investigadores. La gran ventaja de estas gramáticas descriptivas es que existen para todas las lenguas, son de bajo coste y no requieren la presencia de expertos lingüistas.

La experimentación se ha llevado a cabo a partir de la Gramática Descriptiva de la Lengua Española (GDLE), de la Real Academia de la Lengua Española, creando un modelo computacional que será aplicable a gramáticas descriptivas de otras lenguas. La novedad radica en que se han aplicado metodologías de elicitación del conocimiento propias de la Ingeniería del Conocimiento (rama de la Inteligencia Artificial) a la GDLE, que ha sido utilizada como fuente de conocimiento.

Este modelo ha sido probado por medio de una aplicación mediante una “arquitectura” de Pizarra, que es uno de los diseños más genuinos de la Inteligencia Artificial para aplicaciones distribuidas y de cierta complejidad. Ha sido probado con numerosos casos y sus resultados son esperanzadores.

El modelo servirá para abordar la construcción de aplicaciones de procesamiento de lenguaje natural desde el análisis hasta la generación de lenguaje, y aplicable a cualquier lengua donde las necesidades de desarrollar un módulo de lenguaje natural requieran una cierta rapidez y una cobertura razonable.

Un avance de estos trabajos ha sido publicado en las actas del congreso [IKE'08](#), (The 2008 International Conference on Information and Knowledge Engineering), conferencia que tiene lugar del 14 al 17 de Julio en Las Vegas (USA).

Derechos: **Creative Commons**

Puedes copiar, difundir y transformar los contenidos de SINC. [Lee las condiciones de nuestra licencia](#)