

Crónica

Curso sobre Tecnologías de la lengua (segunda edición)

Language technology Program (2nd edition)

Toni Badia, Gemma Boleda y Martí Quixal

El tratamiento del habla, sus aplicaciones y usos sociales fueron los temas centrales en la segunda edición del curso dedicado a las tecnologías de la lengua. Un salto cualitativo que permite el fomento del estudio de las tecnologías aplicadas al lenguaje y la exposición al público de los trabajos más destacados en este ámbito tecnológico.

A diferencia de la primera edición del curso, que había tenido un contenido claramente general con la pretensión de ofrecer una panorámica amplia de las tecnologías de la lengua (véase *QUARK* setiembre-diciembre 2000, núm. 19), el presente año se ha centrado en uno de los aspectos más relevantes de las tecnologías del lenguaje: al tratamiento del habla. El objetivo de este curso era que los ponentes se adentrasen de forma más detallada en los contenidos específicos de la materia. Las conferencias de este curso, impulsado y patrocinado por el Ayuntamiento de Barcelona (a través de su Concejalía Ciudad del Conocimiento) y el Forum Universal de las Culturas Barcelona 2004, se impartieron en el Consorcio de la Universidad Menéndez Pelayo (UIMP) - Centro Ernest Lluch. La dirección estuvo a cargo de Marsal Gavaldà, de la Carnegie Mellon University en Pittsburg (Estados Unidos), y Toni Badia, de la Universitat Pompeu Fabra de Barcelona, se encargó de la secretaría técnica. Con esta edición se inicia una serie de tres cursos, una cada verano hasta el 2004, sobre uno de los aspectos centrales en las tecnologías del lenguaje.

Los contenidos del curso se agruparon en dos grandes bloques: el primero con información sobre los procesos técnicos de la descodificación del habla y, el segundo, centrado en las aplicaciones y usos sociales de los productos de tratamiento del habla. En las lecciones sobre los procesos técnicos intervinieron Climent Nadeu, de la Universitat Politècnica de Catalunya, y Monika Woszczyna, de la Carnegie Mellon University. Los profesores Alex Waibel, de la Universität Karlsruhe y de la Carnegie Mellon University, y Xuedong Huang de Microsoft Speech.Net centraron sus exposiciones en las aplicaciones de los sistemas de tratamiento del habla.

El profesor Climent Nadeu tituló su conferencia «Producción, percepción y codificación del habla». Tras una breve introducción sobre la producción y percepción del habla en los humanos, realizó un análisis en profundidad sobre los temas de la representación del habla y del procesamiento digital de la señal. En este sentido, trató del análisis espectral de la voz, del modelaje y parametrización de la señal, de la codificación específica de la voz, de la extracción de las características necesarias para el reconocimiento del habla y de las técnicas usadas para el reconocimiento robusto de la voz frente a los problemas provocados por el entorno acústico.

«Descodificación de la señal acústica» fue el tema central de la exposición de la profesora Monika Woszczyna y se ocupó de las técnicas más habituales de tratamiento de la señal acústica. Así, explicó cómo se construyen modelos acústicos, especialmente mediante las cadenas de Markov, y modelos del lenguaje. En este último campo expuso los modelos estocásticos, conocidos como *N-gramas*, y las gramáticas más habituales. Finalmente, discutió los algoritmos de búsqueda más usados en este ámbito, en especial el *Dynamic time warping* y el algoritmo de Viterbi.

Por otro lado, Xuedong Huang disertó sobre «Nuevas aplicaciones del tratamiento de la voz». En su intervención se centró tanto en la arquitectura como en los estándares para las aplicaciones de voz. Presentó tres modelos según el lugar de uso: la oficina (el uso típico de la estación de trabajo habitual), el hogar (que debería incorporar tecnología para trabajar sin manos, puesto que no nos podemos imaginar a alguien en la cocina o el comedor arrastrando todo el día el teclado) y entornos

móviles (con ordenadores portátiles o teléfonos móviles). Remarcó el papel del tratamiento de la voz en cada una de ellas y el diseño, con especial énfasis en las interfaces de usuario. Al final de su intervención presentó el *MiPad*, un ejemplo de herramienta que combina la voz y el teclado para llevar a cabo las tareas que ordena el usuario.

El profesor Alex Waibel en su exposición titulada «Sistemas completos» habló sobre las aplicaciones desde el punto de vista del uso global, es decir, cómo puede utilizarlos el usuario. Explicó el estado de las tecnologías del reconocimiento del habla, indicó los factores que afectan a su desarrollo y presentó datos sobre el rendimiento de los sistemas actuales. Asimismo repasó los distintos retos que presentan cinco aplicaciones distintas, en función de sus características: el dictado (persona-máquina), el sistema de pedido de vídeo (datos de persona-máquina), el diálogo (persona-máquina, interactivo), la traducción de voz (persona-máquina-persona), y los resúmenes de reuniones (persona-persona). Y, por último, analizó el problema del reaprovechamiento de los resultados obtenidos en un campo o lengua determinados para aplicarlos en otra situación o aplicación.

En el ámbito de las aplicaciones tuvo un interés particular la mesa redonda «Implicaciones sociales del tratamiento informático del habla» en la que participaron los profesores Santiago Aguilera, de la Universidad Politécnica de Madrid; Xuedong Huang, de Microsoft Speech.Net, y Joaquim Llisterri, de la Universidad Autónoma de Barcelona. En el transcurso de la sesión se plantearon, entre otros temas, el soporte tecnológico para personas con discapacidades, los recursos tecnológicos en los entornos educativos, los retos sociales de las nuevas tecnologías y el tratamiento informático del habla.

Entre los asistentes a este curso, la mayoría estudiantes de doctorado o jóvenes investigadores de las universidades catalanas, se contó también con la presencia de alumnos del resto de España, así como de otros países. Un aspecto destacado fue la asistencia de un numeroso grupo de personas que trabajan en el ámbito empresarial, lo que indica que, finalmente, este tipo de formación empieza a ser interesante para compañías del sector de servicios con departamentos informáticos.

En conclusión, podemos afirmar que la edición de este año ha supuesto un salto cualitativo respecto al curso del año pasado. Entre otros motivos por la especificidad del tema, el aumento de la participación de las empresas y también el incremento en el número de asistentes a las jornadas. Así, el objetivo de fomentar el estudio de las nuevas tecnologías aplicadas al lenguaje y dar a conocer a la sociedad el trabajo que se realiza en este ámbito tecnológico se ha cumplido satisfactoriamente gracias a este curso.