

# Un algoritmo para la comprensión automática del habla sobre grafos de palabras\*

## *An algorithm for automatic speech understanding over word graphs*

Marcos Calvo, Jon Ander Gómez, Emilio Sanchis, Lluís-F. Hurtado

Departament de Sistemes Informàtics i Computació

Universitat Politècnica de València

Camí de Vera s/n - 46022 València

{mcalvo, jon, esanchis, lhurtado}@dsic.upv.es

**Resumen:** En este trabajo se propone un algoritmo para la comprensión automática del habla que toma como entrada un grafo de palabras. Este grafo es procesado en primer lugar mediante un algoritmo de programación dinámica, obteniendo como resultado un segundo grafo enriquecido con información semántica. El cálculo del mejor camino sobre este segundo grafo permite obtener la secuencia de conceptos más verosímil de acuerdo con la evidencia acústica reflejada en el grafo de palabras. También como resultado de la decodificación semántica se obtiene la secuencia de palabras asociada a dicha secuencia de conceptos, así como la segmentación semántica de la secuencia de palabras.

**Palabras clave:** Comprensión automática del habla, grafos de palabras, programación dinámica, algoritmos.

**Abstract:** In this work we propose an algorithm for automatic speech understanding that takes a word graph as its input. First, this word graph is processed by means of a dynamic programming algorithm which gives as a result a second graph that includes semantic information. Computing the best path over this second graph allows us to obtain the most likely concept sequence, given the acoustic evidence reflected on the input word graph. As a result of the semantic decoding, the word sequence attached to the concept sequence as well as its semantic segmentation are also obtained.

**Keywords:** Spoken language understanding, word graphs, dynamic programming, algorithms.

## 1. Introducción

La comprensión automática del habla es el proceso por el que, dada una pronunciación emitida por un locutor, se extrae una interpretación semántica de la información contenida en ésta basada en un conjunto de conceptos (o etiquetas semánticas) definido a priori. El ámbito donde los sistemas de comprensión del habla tienen mayor aplicación práctica es el de los sistemas de diálogo hablado, en los cuales es crucial que el sistema pueda extraer la información asociada a la pronunciación de entrada (la “comprenda”) para devolver una respuesta coherente con lo que ha dicho el usuario. En la estructu-

ra típica de un sistema de diálogo hablado, el módulo de comprensión toma como entrada la salida del módulo de reconocimiento de voz (ASR) y pasa su salida al módulo gestor del diálogo, el cual utilizará esta información para decidir la siguiente acción a realizar. Por tanto, teniendo en cuenta que al módulo de comprensión le llega una estructura de datos en la que está representada la información que ha extraído el ASR, la comprensión del habla consiste de dos subtareas, que son:

1. La identificación de la secuencia de conceptos y su asignación a secuencias de palabras.
2. La extracción de la información relativa a estos conceptos codificada en las secuencias de palabras asignadas y la construcción de estructuras de datos que representen dicha información.

\* Este trabajo ha sido subvencionado por el Ministerio de Economía y Competitividad en el marco del proyecto TIN2011-28169-C05-01 y por el Vicerrectorat d'Investigació, Desenvolupament i Innovació de la Universitat Politècnica de València con proyecto 20110897.