

Identificación automática de la función sujeto en oraciones con verbos de cognición

Ana Paula González Muñoz¹, Marta Lourdes Bauguer Díaz-Romañach², Gretel Gutierrez Fuentes³

¹Facultad de Matemática y Computación, Universidad de La Habana

²Facultad de Matemática y Computación, Universidad de La Habana

³Facultad de Artes y Letras, Universidad de La Habana

Resumen

La identificación automática de la función sujeto en las oraciones y su relación posicional con respecto a un verbo de cognición es una tarea que involucra el área de Procesamiento del Lenguaje Natural (PLN). Aunque se acepta en general que el español sigue, en cuanto al orden de palabras, el esquema SVO (sujeto-verbo-objeto directo), con verbos de cognición como *creer*, puede variar no solo por la posición del verbo, sino también del sujeto. La presencia o ausencia del sujeto, así como la posición pre- o posverbal de esa función sintáctica puede tener repercusiones en la interpretación que se haga sobre el grado de certeza y la responsabilidad con que el hablante presenta sus juicios. Esto resulta de gran interés para los lingüistas por ser las Cartas a la Dirección, del periódico Granma, textos que propician la expresión de la opinión de quienes escriben. Se propone un etiquetador que reconozca el sujeto (su manifestación sintáctica y posición) con estos verbos para permitir extender el análisis de estas relaciones en grandes volúmenes de textos.

Trabajos similares encontrados en la literatura apuntan a dividir las oraciones en diferentes roles semánticos que respondan a preguntas como ¿qué ocurrió?, ¿quién lo hizo?, ¿cómo lo hizo?, ¿dónde ocurrió?, ¿cuándo pasó?, entre otras. En un primer enfoque se divide una oración en varios roles semánticos donde los actantes juegan un tipo de rol semántico y son clasificados como el agente; que es el encargado de realizar la acción y adicionalmente puede corresponder con el sujeto de la oración (Reyes, Montes, González, y Pinto, 2013).

Otro enfoque aborda la identificación de roles semánticos mediante la simplificación de oraciones y el uso de características sintácticas. Se asignan estos roles a los argumentos de un verbo objetivo en una oración para identificar y etiquetar los roles semánticos, o argumentos, de ese verbo (Gildea y Jurafsky, 2002).

Siguiendo un enfoque diferente se propone una metodología basada en técnicas de aprendizaje automático más recientes para detectar directamente la función sujeto de una oración y su posición con respecto al verbo de cognición.

Palabras clave: Procesamiento del Lenguaje Natural, función sujeto, verbos de cognición, aprendizaje automático, sintaxis, roles semánticos.

Referencias

- Gildea, D., y Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational Linguistics*, 28(3), 245–288.
- Reyes, J. A., Montes, A., González, J. G., y Pinto, D. E. (2013). Clasificación de roles semánticos usando características sintácticas, semánticas y contextuales. *Computación y Sistemas*, 17(2), 263–272.