

Nueva Propuesta de Desambiguación de Sentidos de Palabras para nombres en un sistema de Búsqueda de Respuestas*

Sergio Ferrández; Sandra Roger; Antonio Ferrández; Antonia Aguilar; Pilar López-Moreno

Universidad de Alicante

sferrandez@dlsi.ua.es; sroger@dlsi.ua.es; antonio@dlsi.ua.es; aaguilar7671@cv.gva.es; P.Lopez@ua.es

Resumen: Este artículo describe el impacto de un algoritmo de Desambiguación de Sentidos de Palabras (WSD) para nombres en AliQAn, el sistema de Question Answering con el cual hemos participado en el CLEF-2005. Al aplicar el WSD tradicional, el rendimiento se decrementa en un 4.7% en el Mean Reciprocal Rank (MRR). Para resolver este problema, proponemos dos aplicaciones de WSD: (1) elegir un grupo de synsets en vez de uno como hace el WSD tradicional; (2) desambiguar las palabras inexistentes en EuroWordNet (EWN). Con nuestra propuesta de WSD el MRR se incrementa en un 6.3% en relación con el baseline sin WSD. Además, nuestra propuesta de WSD incrementa el MRR en un 11% respecto al WSD tradicional. Por último, la implementación de nuestro enfoque de WSD es computacionalmente eficiente gracias a un preprocesamiento de EWN.

Palabras clave: Búsqueda de Respuestas, Desambiguación de sentidos, CLEF.

Resumen: This paper studies the influence of a Word Sense Disambiguation algorithm for nouns on AliQAn, the Question Answering with which we have participated in the CLEF-2005. Applying the traditional WSD decreases the performance in 4.7% on the Mean Reciprocal Rank (MRR). To solve this problem, we propose two different uses of WSD: (1) to choose a set of synsets instead of the traditional use of WSD, in which only one synset is chosen; (2) to disambiguate the words not present in EuroWordNet (EWN). Using our proposal of WSD the MRR increases a 6.3% with regard to the baseline without WSD. Furthermore, our proposal of WSD increases the MRR with regard to the traditional use of WSD in an 11%. Finally, the implementation of our approach of WSD is computationally efficient by means of a preprocessing of EWN.

Keywords: Question Answering, Word Sense Disambiguation, CLEF.

1 Introducción

En este artículo analizamos las ventajas de un algoritmo de Desambiguación de Sentidos de Palabras (WSD) para nombres en AliQAn (Roger et al., 2005), un sistema de Question Answering (QA), con el cual hemos participado en la competición CLEF-2005¹.

El objetivo de QA consiste en identificar la respuesta a preguntas dentro de un gran

conjunto de documentos. QA no es una tarea simple de Information Retrieval (IR), QA intenta ir más allá y devolver una determinada contestación en respuesta a una pregunta arbitraria. Para los usuarios, es interesante encontrar información precisa, gracias al incremento de la información disponible. Los sistemas QA son capaces de responder a cuestiones formuladas por los usuarios en lenguaje natural.

Los enfoques actuales hacia QA están basados principalmente en herramientas de Procesamiento de Lenguaje Natural (NLP) o en el aprendizaje automático. Existen diferentes y posibles implementaciones para sistemas QA. Generalmente, la mayoría de ellas se basan en herramientas NLP (Acebo et

* Esta investigación ha sido parcialmente financiada por el Gobierno Español bajo el proyecto CICyT número TIC2003-07158-C04-01 y por el Gobierno Valenciano bajo el proyecto número GV04B-268.

¹ <http://www.clef-campaign.org/>

al., 1994), (Ageno et al., 2004), (Herrera, Peñas y Verdejo, 2004), (Méndez-Díaz, Vilares-Ferro y Cabrero-Souto, 2004), (Pérez-Coutiño et al., 2004), como POS-taggers, analizadores sintácticos, bases de conocimiento consistentes en diccionarios, bases de datos léxico-semánticas, ontologías y muchos otros. No obstante, otros sistemas usan técnicas de aprendizaje automático con modelos estadísticos (Pablo et al., 2004), tales como Hidden Markov Models o Maximum Entropy.

El sistema AliQAn ha sido desarrollado durante los dos últimos años en el Departamento de Procesamiento del Lenguaje Natural y Sistemas Informáticos en la Universidad de Alicante. Se basa en el uso de herramientas NLP, destacándose el uso de WSD para mejorar la precisión del sistema.

El algoritmo WSD se usa en las fases de indexación y búsqueda. En el primer caso, este algoritmo permite la desambiguación de las palabras del corpus y en el segundo, resuelve las ambigüedades en las palabras de preguntas.

WSD tiene varios problemas críticos. El tiempo de ejecución de los algoritmos WSD hace difícil su uso en corpus enormes, como requieren los sistemas QA. Por otro lado, la baja precisión de los algoritmos WSD hace que esta técnica no sea apropiada para aplicarse en sistemas QA. Estas dos razones no permiten la obtención de resultados interesantes aplicando WSD en sistemas QA de tiempo real. Para resolver estos problemas, proponemos una herramienta de WSD que reduce el tiempo de ejecución en un 98.9%, gracias a un preproceso de EuroWordNet (EWN) y mejora la precisión por medio de: (1) selección de un conjunto de synsets por palabras (en vez de uno solo como se hace tradicionalmente); (2) desambiguar palabras que no están presentes en EWN.

El resto del artículo se organiza como sigue; la sección dos describe los antecedentes de QA en relación con WSD; la sección tres revisa brevemente el sistema AliQAn; la sección cuatro explica nuestra propuesta de algoritmo WSD en el sistema AliQAn; la sección cinco muestra los resultados de evaluación y finalmente, la sección seis expone nuestras conclusiones y trabajos futuros.

2 Antecedentes de QA respecto a WSD

La mayoría de los actuales sistemas QA monolingüales (Gómez et al., 2005), (Laurent, Negre y Séguéla, 2005), (Neumann y Sacaleanu, 2005), (Sutcliffe, Mulcahy y Gabbay, 2005) no aplican algoritmos de WSD. Hoy en día, el uso de algoritmos WSD en QA e IR producen, normalmente, un decremento en la precisión total y un incremento en el tiempo de ejecución.

Las técnicas WSD se han aplicado solamente en los sistemas IR (Gonzalo et al., 1998), (Vossen et al., 2006), (Stokoe, Oakes, y Tait, 2003). En el primer sistema analizado (Gonzalo et al., 1998), la indexación con los synsets de wordnet mejora los resultados del sistema IR en un 29% (desde 30% hasta 60%) pero tiene la desventaja de que se lleva a cabo manualmente.

El proyecto MEANING (Vossen et al., 2006), ha desarrollado herramientas para la adquisición automática de conocimiento léxico que ayudará WSD. El conocimiento léxico obtenido se almacena en el Multilingual Central Repository (Atserias et al., 2004), el cual se basa en el diseño de una base de datos EWN. Esta implementación se basa en el uso de WSD con dominios (Magnini et al., 2002). El problema de usar esta técnica es que los dominios tienen que crearse en una fase previa. La precisión observada en un sistema de IR es de 65.9% lo cual significa un incremento del 2% (Tabla 1, Sistema A) con WSD.

Finalmente, el último sistema de IR (Stokoe, Oakes, y Tait, 2003), utiliza una combinación de técnicas de alta precisión y estadísticas de frecuencia de sentidos en un intento de reducir el impacto de desambiguaciones erróneas. El incremento de la precisión en los resultados finales es 1,73% sobre 62.1 (Tabla 1, Sistema B).

	Sistema A	Sistema B
Sin WSD	65.9%	62.1%
Mejora WSD	2%	1.173%

Tabla 1: Resultados de diferentes sistemas IR con las mejoras una vez que se aplica WSD

3 El sistema AliQAn

AliQAn es un sistema QA de dominio abierto monolingual basado en el uso intensivo de

herramientas NLP. AliQAn ha participado en la competición española QA CLEF-2005 (Roger et al., 2005), en la cual quedó tercero. La arquitectura de AliQAn se divide principalmente en dos fases: fase de indexación y fase de búsqueda. En ambas fases, se aplica el mismo proceso NLP, en el corpus de documentos y de preguntas: PoS-tagger, análisis parcial y WSD.

Con el fin de realizar el análisis sintáctico, se usa el sistema SUPAR (Ferrández, Palomar y Moreno, 1999), el cual incluye el resultado del etiquetador MACO (Acebo et al., 1994). Por otro lado, se aplica WSD usando EWN.

AliQAn identifica las distintas estructuras gramaticales de una frase, llamadas bloques sintácticos (SB). Esto se realiza usando el resultado de SUPAR, el cual realiza un análisis sintáctico parcial. Estos bloques son frases verbales (VP), frases nominales (NP) o frases preposicionales (PP).

Por ejemplo en la frase: “Kim Il Sung died at the age of 80”, la lista obtenida de SB es: [NP, kim*il*sung] [VP, to die] [PP, at age [PP, of: 80]].

La fase de indexación, la primera fase de AliQAn, consiste en preparar los datos donde el sistema intenta encontrar la respuesta a las preguntas. Se llevan a cabo dos tipos de indexación: indexación de IR e indexación QA.

En la segunda fase, la búsqueda, se realizan estas tres tareas secuencialmente: análisis de la pregunta, selección de pasajes relevantes y extracción de la respuesta.

En la primera tarea el sistema detecta el tipo y caso de pregunta. Para el tipo de pregunta se considera el WordNet Based-Types y EWN Top-Concepts, el caso de la pregunta determina el conjunto de modelos sintácticos a usar en la extracción de la respuesta.

Además, se lleva a cabo la selección de los términos de preguntas (palabras clave). Estas palabras claves se usan en la siguiente tarea, la selección de pasajes relevantes, realizada por el sistema IR-n (Llopis y Vicedo, 2001). IR-n devuelve una lista de los pasajes donde el sistema aplica la extracción de respuesta, en la última tarea de AliQAn, se intenta extraer la respuesta correcta usando los modelos sintácticos.

A continuación, un ejemplo (pregunta 114, del CLEF 2003) de resolución de una pregunta, donde el sistema elige la solución correcta.

- **Pregunta:** A qué primer ministro abrió la Fiscalía de Milán un sumario por corrupción?
- **Tipo:** persona
- **Caso:** 3
- **Lista de SB:**
 - **NP1:** ([NP, primer*ministro])
 - **VP:** ([VP, abrir])
 - **NP2:** ([NP, fiscalia [PP, de: milan]])([NP, sumario [PP, por: corrupción]])
- **Texto donde la solución correcta es:** “[...] la **Fiscalía de Milán** abrió, hoy martes, un **sumario** al **primer ministro**, *Silvio Berlusconi*, por un supuesto delito de **corrupción** [...]”
- **Texto donde la solución incorrecta es:** “[...] **primer ministro** y líder socialista, Bettino Craxi, al que el pasado 21 de septiembre Paraggio **abrió** un **sumario** relacionado con el proyecto Limen por supuestos delitos de **corrupción** [...]”
- **Respuesta:** Silvio Berlusconi

4 El algoritmo WSD

El algoritmo WSD utilizado en este artículo es el propuesto por Agirre y Rigau (1995). Este algoritmo WSD se basa en el uso de la distancia conceptual con el fin de intentar suministrar una base para determinar la proximidad de sentidos entre palabras, tomando como referencia la estructura jerárquica de EWN. La distancia conceptual se obtiene por medio de la fórmula de Densidad Conceptual (CD). Dado un concepto c , CD se calcula usando las siguientes ecuaciones:

$$CD(c, m) = \frac{\sum_{i=0}^{m-1} nhyp^i}{descendientes_c} \quad (1)$$

$$descendientes_c = \sum_{i=0}^{h-1} nhyp^i \quad (2)$$

donde $nhyp$ es el número de hipónimos, m es el número de marcas de sentidos y h es la altura de la subjerarquía.

A diferencia del algoritmo propuesto por Agirre y Rigau (1995), nuestra implementación no selecciona un synset sino un conjunto de los synsets con mayor probabilidad de ser correctos. Esto se debe a que el algoritmo intenta descartar solo los synsets que son completamente erróneos para mejorar la precisión del sistema. Además, los nombres propios que no están en EWN se desambiguan también con respecto los siguientes synsets:

05369359 (persona) 07451540 (objeto)
08229827 01218276 (lugar).

El comportamiento de nuestra implementación se muestra en el siguiente ejemplo:

- En la frase
 - *El presidente de Guinea, Obiang, sugirió hoy, viernes, que su Gobierno podría rechazar la ayuda internacional*
- Synsets para la palabra “*presidente*” ordenado por (Agirre y Rigau, 1995)
 - 09400170 006140480 11176111 01090427 08956043
- Synset correcto de la palabra “*presidente*” en esta frase
 - 09400170
- Conjunto de synsets seleccionados de la palabra “*presidente*” usando nuestra implementación
 - 09400170 006140480

En este ejemplo, la palabra “*Obiang*” que no está en EWN es desambiguada también. En este caso, la herramienta de Agirre y Rigau (1995) devuelve el synset correcto: 05369359 (persona).

Nuestra propuesta permite mejorar la precisión de nuestro sistema QA como la sección de evaluación mostrará, porque normalmente los algoritmos WSD tienen baja precisión (alrededor del 66%). Por lo que es bastante probable descartar el synset correcto, lo cual es peor que mantener todos los synsets de la palabra. De este modo, nuestra propuesta solo descarta los synsets que son completamente erróneos, los cuales no se utilizan cuando se comparan los términos de la pregunta con los términos del documento. Por ejemplo, en la pregunta “¿Cuántos muertos al año causan las minas antipersona en el mundo?”, la frase del documento que habla sobre “minas de carbón” no será procesada. Además, nuestra propuesta supera la desambiguación que usa dominios EWN (Vossen et al., 2006) porque no requiere agrupación alguna de EWN, lo cual podría llevar a errores en el sistema.

El siguiente ejemplo muestra la mejora obtenida cuando se usa nuestra implementación de WSD.

- Para la pregunta:
 - *¿Quién es el director de la CIA?*
- Sin usar desambiguación, AliQAn devuelve:
 - Servicio Central de Información
 - Servicio Central de Inteligencia
 - Servicio Central de Información de EEUU
- Usando desambiguación, AliQAn devuelve:

- William Colby
- Robert Gates
- James Woolsey

El algoritmo WSD de Agirre y Rigau (1995) presenta una precisión del 66%. Se ha evaluado nuestra implementación de este algoritmo en el corpus EFE (detallado en la sección de evaluación), donde la precisión es de un 60% cuando se selecciona un solo synset. Cuando se selecciona el 40% de los synsets de una palabra, la precisión es del 76%. Considerando la precisión de nuestra propuesta para nombres que no están en EWN, la precisión es de un 65%. Finalmente, el tiempo de ejecución del algoritmo se ha reducido almacenando toda la información necesaria mostrada en las ecuaciones (1) y (2), como el número de descendientes, para cada synset en EWN. De esta forma, el tiempo de ejecución se ha reducido en un 98.9% (de 1400 segundos a 15 segundos en la desambiguación de 9 ficheros).

5 Evaluación

5.1 Conjunto de Datos

Los experimentos descritos en esta sección se han llevado a cabo usando el sistema AliQAn con preguntas y corpus del CLEF 2003. El conjunto indexado contiene aproximadamente 125.738 documentos (un total de 509 Megabytes) y un total de 200 preguntas.

5.2 Medidas de Evaluación

Para evaluar el sistema necesitamos una medida que valore los resultados generales del sistema. Mean Reciprocal Rank (MRR) es la medida que se usa en la campaña de CLEF 2003 para evaluar los sistemas.

$$MRR = \left(\sum_{i=1}^Q \frac{1}{far(i)} \right) / Q$$

donde Q es el número de preguntas (200 en nuestro caso) y $far(i)$ indica la posición de la primera respuesta correcta. El valor de $1/far(i)$ será 0 si el sistema no encuentra la respuesta.

5.3 Análisis de los Resultados

En las secciones anteriores se ha descrito la arquitectura y comportamiento de nuestro sistema. Ahora presentaremos los resultados

obtenidos y el estudio sobre la realización del sistema cuando se usa WSD en el corpus y preguntas.

La comparación será realizada usando el sistema AliQAn con tres algoritmos, los cuales tienen diferentes niveles de WSD. Estos algoritmos se muestran en la Tabla 2.

1. **Baseline sin WSD:** (primera columna de la Tabla 2). El baseline no tiene el algoritmo WSD, es decir, se usan todos los synsets por palabra. Este sistema base tiene un MRR de 44.5%.

2. **Algoritmo 1-Sense WSD:** (segunda columna de la Tabla 2). En este caso, el algoritmo WSD usado es el propuesto por Agirre y Rigau (1995). Este algoritmo elige un synset para desambiguar una palabra y el MRR obtenido es 42.4%.

3. **Nuestra propuesta de algoritmo WSD:** (tercera columna de la Tabla 2). En este caso se selecciona el conjunto de synsets con mayor probabilidad. Los synsets seleccionados son un 40% del total de synsets. El sistema tiene un MRR del 47.3%.

	1 Baseline sin WSD	2 1-Sense WSD	3 Nuestra propuesta de WSD
MRR	44.5	42.4 (-4.75%)	47.3 (+6.3%)
% de Respuesta Primera Correcta (FCA)	39	37.5 (-3.85%)	42.5 (+8.97%)
Mejora FCA (200 preguntas)		3 preguntas +11% de mejora en el MRR respecto a 1-Sense WSD	7 preguntas

Tabla 2: Resultados usando diferentes aplicaciones de WSD con 200 preguntas del CLEF 2003

Por lo tanto, nuestro WSD propuesto ha incrementado el MRR en un 11% comparado con el algoritmo 1-sense WSD y en un 6.3% con respecto al baseline (ver Figura 1), incrementa el número de preguntas contestadas en primer lugar en 7.

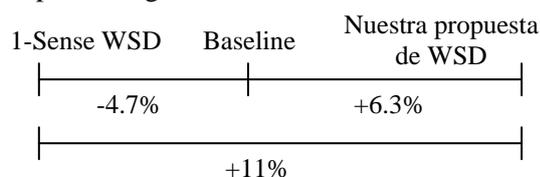


Figura 1. Relación respecto a los porcentajes del sistema WSD presentados en la Tabla 2

En comparación con el 42.4 del algoritmo 1-sense WSD, el porcentaje de incremento en el MRR es 8.3% cuando se seleccionan un 10% de los synsets. Un 9.2% cuando se selecciona el 20%. El 10.4% cuando se selecciona el 30%. Por último, un 11% cuando se selecciona el 40%.

Hemos llevado a cabo un análisis del número de nombres desambiguados en 200 preguntas (159,778) donde 46,194 nombres tienen un solo synset en WordNet, y el resto de

nombres (113,584) tienen una media de 3.9 synsets en WordNet. Nuestra propuesta de WSD selecciona una media de 1.9 synsets (en caso de empate entre dos synsets, se seleccionan ambos). Además, se han desambiguado 44,247 nombres que no están en WordNet.

Por otra parte, el tiempo empleado normalmente en el procesamiento del corpus de WSD es demasiado largo, en el peor caso NP completo. Muchos investigadores se han encontrado este problema cuando aplican sus enfoques de WSD, por esta razón solamente aplican WSD a preguntas. Con nuestro algoritmo WSD, hemos conseguido reducir considerablemente el tiempo de ejecución. Al comienzo, se necesitaban 1400 segundos para procesar 9 ficheros, ahora se realiza en 15 segundos, lo cual supone un decremento del 98.9%.

6 Conclusiones y trabajos futuros

En este artículo proponemos un algoritmo que intenta mejorar WSD para nombres en un sistema QA. Este algoritmo se basa en un algoritmo propuesto por Agirre et Rigau (1995). La diferencia es que mientras éste

considera un synset para desambiguar una palabra, nuestra propuesta selecciona los synsets más relevantes y añade la desambiguación de los nombres propios que no están incluidos en EWN 1.6. Con el fin de evaluar la viabilidad de nuestro algoritmo, se han llevado a cabo una serie de comparaciones. Los resultados confirman la viabilidad de nuestro algoritmo, mostrando una mejora de hasta un 11% sobre el algoritmo WSD tradicional y un 6.3% sobre el baseline. Además, hemos reducido en un 98.9% el coste computacional del proceso WSD, por medio de un preprocesamiento de EWN. Es importante enfatizar que el algoritmo propuesto no requiere una agrupación previa (como dominios) para el proceso de desambiguación, el cual puede implicar algún error.

La contribución de nuestro artículo al área de investigación de WSD se basa en que el algoritmo 1-sense WSD tradicional no mejora el QA, como se afirma en la Tabla 2 y Figura 1 (-4.7% en el MRR). Esto es debido a su baja precisión. Nuestra propuesta permite seleccionar un porcentaje de synsets en vez de solo uno como hace el algoritmo 1-sense WSD. De esta forma, se mejora el MRR de un algoritmo 1-sense WSD en un 11% y mejora el MRR de un sistema QA sin WSD en un 6.3%. Además, nuestra propuesta supera la desventaja tradicional del WSD que es el coste computacional, lo cual hace demasiado difícil su aplicación en un corpus bastante grande. Nuestro enfoque reduce el tiempo de ejecución del algoritmo WSD en un 98.9%.

Los resultados son prometedores. Por lo tanto, esperamos analizar los resultados en el CLEF 2004 y CLEF 2005. En el futuro, vamos a desarrollar otro algoritmo WSD para probar que nuestra propuesta se valida independientemente del algoritmo WSD.

Bibliografía

Acebo, S., A. Ageno, S. Climent, J. Farreres, L. Padró, R. Placer, H. Rodríguez, M. Taulé, y J. Turno. 1994. MACO: Morphological Analyzer Corpus-Oriented. *ES-PRIT BRA-7315 Aquilex II, Working Paper 31*.

Agno, A., D. Ferrés, E. González, S. Kankan, H. Rodríguez, M. Surdeanu y J. Turmo. 2004. TALP-QA System for Spanish at CLEF-2004. *In Workshop of Cross-*

Language Evaluation Forum (CLEF), páginas 425-434.

Agirre, E., y G. Rigau. 1995. A proposal for word sense disambiguation using conceptual distance. *1st Intl. Conf. on recent Advances in NLP*. Bulgaria.

Atserias, J., L. Villarejo, G. Rigau, E. Agirre, J. Carrol, B. Magnini, y P. Vossen. 2004. The meaning multilingual central repository. *In Proceeding of the Second Internacional WordNet Conference-GWC.*, páginas 23-30.

de Pablo, C., J.L. Martínez-Fernández, P. Martínez, J. Villena, A.M. García-Serrano, J.M. Goñi y J.C. González. 2004. miraQA: Inicial experiments in Question Answering, *In Workshop of Cross-Language Evaluation Forum (CLEF)*, páginas 371-376.

Ferrández, A., M. Palomar y L. Moreno. 1999. An Empirical Approach to Spanish Anaphora Resolution. *Machine Translation. Special Issue on Anaphora Resolution In Machine Translation*, 14(3/4):191-216, Diciembre.

Gómez, J.M., E. Bisbal, D. Buscaldi, P. Rosso y E. Sanchis. 2005. Monolingul and Cross-Language QA using a QA-oriented Pasaje Retrieval System. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, Septiembre.

Gonzalo, J., F. Verdejo, I. Chugur y J. Cigarrán. 1998. Indexing with wordnet synsets can improve text retrieval. *In the Proceedings of the ACL/COLING Workshop on Usage of WordNet for Natural Language Processing*, páginas 38-44.

Herrera, J., A. Peñas y F. Verdejo. 2004. Question Answering Pilot Task at CLEF 2004. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, páginas 445-452.

Laurent, D., S. Negre y P. Séguéla. 2005. Cross Lingual Question Answering using QRISTAL for CLEF 2005. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, Septiembre.

Llopis, F. y J.L. Vicedo. 2001. IR-n, a passage retrieval system. *In Workshop of Cross-Language Evaluation Forum (CLEF)*.

- Magnini, B., C. Strapparava, G. Pezzulo y A. Gliozzo. 2002. The role of domains information in word sense disambiguation. *Nat. Lang. Eng.*, 8(4):359-373.
- Méndez-Díaz, E., J. Vilares-Ferro y D. Cabrero-Souto. 2004. COLE at CLEF 2004: Rapid prototyping of a QA system for Spanish. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, páginas 413-384.
- Neumann, G. y B. Sacaleanu. 2005. DFKT's LT-lab at the CLEF 2005 Multiple Language Question Answering Track. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, Septiembre.
- Pérez-Coutiño, M., T. Solorio, M. Montes y Gómez, A. López-López y L. Villaseñor-Pineda. 2004. The use of Lexical Context in Question Answering for Spanish. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, páginas 377-384.
- Roger, S., S. Ferrández, A. Ferrández, J. Peral, F. Llopis, A. Aguilar y D. Tomás. 2005. AliQAn, Spanish QA System at CLEF-2005. *In Workshop of Cross-Language Evaluation Forum (CLEF)*.
- Stokoe, C., M.P. Oakes, y J. Tait. 2003. Word sense disambiguation in information retrieval revisited. *In SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, páginas 159-166, New York, NY, USA., ACM Press.
- Sutcliffe, R. F. E., M. Mulcahy e I. Gabbay. 2005. Cross-Language French-English Question Answering Using the DLT system at CLEF 2005. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, Septiembre.
- Vossen, P., G. Rigau, I. Alegria, E. Agirre, D. Farwell y M. Fuentes, y. 2006. Meaning-ful results for information retrieval in the meaning Project. *Third International WordNet Conference*.