Tareas de análisis del contenido textual para la recuperación de información con realimentación*

Manuel J. Maña López[1], **L. Alfonso Ureña López**[2] y **Manuel de Buenaga Rodríguez**[3]

[1]Dpto. de Lenguajes y Sistemas Informáticos, Universidad de Vigo, mjlopez@uvigo.es
[2]Dpto. de Informática, Universidad de Jaén, laurena@ujaen.es
[3]Dpto. de Inteligencia Artificial, Universidad Europea de Madrid – CEES, buenaga@dinar.esi.uem.es

Resumen. La utilización de realimentación es una de las técnicas que proporciona mejoras más significativas en la efectividad del proceso de recuperación de información. Por otra parte, cada vez se utilizan en el proceso de recuperación de información, técnicas más avanzadas de análisis del contenido textual con vistas a mejorar la efectividad. En nuestro trabajo estudiamos los beneficios que proporciona la integración de mecanismos de análisis del contenido al utilizar la realimentación en el proceso de recuperación de información. Nos centramos en dos tareas de análisis: desambiguación de palabras y generación de resúmenes, presentando una sistemática para su utilización y experimentos asociados para la evaluación de las mejoras conseguidas.

1 *Introducción*

La cantidad de información a la que una persona puede tener acceso crece exponencialmente, gracias sobre todo a Internet, y aunque el tipo de esta información es cada vez más variado, la información textual hoy por hoy es la predominante. En [18], por ejemplo, se proporciona un dato representativo: aproximadamente el 90% del total de la información que maneja una empresa es texto. Podemos encontrar texto en documentos, manuales, informes, correos electrónicos, faxes y también en páginas web. Sólo para este último medio, hay estimaciones [2] de que la cantidad de texto disponible es de al menos 1.5 terabytes. En este escenario se entiende el creciente interés por los sistemas de acceso a la información y, en general, por las tareas de análisis automático del contenido textual.

En el ámbito de los sistemas de recuperación de información (RI), la constatación de la dificultad de formular consultas que se muestren efectivas recuperando información relevante ha suscitado un gran interés por las técnicas de modificación de consultas. La *realimentación* es una de estas técnicas y ha sido tradicionalmente una de las más utilizadas para mejorar la eficacia de los sistemas de RI. Esta técnica consiste en utilizar la opinión del usuario sobre la relevancia o no de algunos de los documentos recuperados para expandir la consulta inicial y repetir el proceso de recuperación.

La incorporación de variantes de la técnica de realimentación a buscadores de Internet como Excite[1] o AltaVista[2], y a bibliotecas digitales como la de ACM[3], refleja su popularidad. En todos estos casos se permite al usuario buscar documentos relacionados con uno recuperado previamente.

Numerosos experimentos sobre diferentes colecciones de resúmenes han demostrado la eficacia de la técnica tanto para el modelo de espacio vectorial [20] como para el probabilístico [8]. También se han realizado experimentos que aplican la realimentación a sistemas de recuperación basados en pasajes. En [13] se describe un sistema cuya principal característica es que en el proceso de realimentación sólo se emplean los pasajes que el usuario detecta como relevantes, en lugar de la totalidad del documento. Y en esta misma dirección apunta el trabajo de Allan [1], que realiza experimentos en los que los pasajes sustituyen a los documentos más largos en la realimentación.

* Este trabajo ha sido parcialmente financiado por:
- CICYT, proyecto TEL99-0335-C04-03.
- Ministerio de Industria y Energía, Iniciativa ATYCA, proyecto TS203/1999.

1 http://www.excite.com
2 http://www.altavista.com
3 http://www.acm.org/dl/

Las técnicas de análisis de contenido textual pueden mejorar el proceso de acceso a la información. En general, cuanto más completo sea el análisis y la comprensión de los textos que se procesan mejor será el acceso a la información. Así se muestra en trabajos centrados en tareas concretas como: categorización [10], generación de resúmenes [16], desambiguación del sentido de las palabras [22], *clustering* [11], segmentación [9] e integración de recursos léxicos [7]. En este sentido, pensamos que la realimentación también puede verse mejorada por la aplicación de algunas técnicas de análisis automático de contenido de documentos.

En concreto, en este trabajo nos centraremos en desambiguación del sentido de las palabras y generación de resúmenes de texto. En la primera tarea, usamos los documentos enjuiciados por el usuario en la fase de realimentación como contexto para desambiguar el sentido de los términos empleados en la consulta. Esta información se utiliza para expandir la consulta original con sinónimos obtenidos de la base de datos léxica WordNet [17]. En la segunda tarea, los resúmenes de los documentos originales se utilizan para sustituir a éstos en la realimentación.

El artículo está organizado como sigue. En el apartado siguiente nos ocuparemos de explicar con más detalle la técnica de realimentación. A continuación dedicamos sendos apartados a cada una de las tareas mencionadas, describiéndolas brevemente y explicando con detalle los experimentos y sus resultados. Acabamos con las conclusiones y el trabajo futuro.

2 La técnica de realimentación

Para el modelo del espacio vectorial, los experimentos de Salton y Buckley [20] mostraron que el método de realimentación *Ide dec-hi* [12] era el que, en conjunto, obtenía mejores resultados. Este método deriva el vector de la nueva consulta a partir de la consulta inicial, todos los documentos considerados relevantes y sólo el mejor clasificado de los documentos no relevantes. Más formalmente, sean Q_0 el vector de la consulta original, R_i el vector del documento relevante i y S el vector del documento no relevante mejor clasificado. En (1) se muestra como calcular el vector Q_1 correspondiente a la consulta expandida según este método.

$$Q_1 = Q_0 + \sum_{\substack{\text{todos los} \\ \text{relevantes}}} R_i - \frac{1}{2} S \qquad (1)$$

Los documentos relevantes e irrelevantes cuyos vectores se utilizan en (1) son aquellos documentos recuperados al utilizar la consulta inicial sobre los que el usuario emite una valoración. Sin embargo, cuando se evalúa la realimentación mediante colecciones de prueba (esto es, colecciones que incluyen un corpus de documentos, un conjunto de consultas y los juicios de relevancia correspondientes) no se necesita usuarios. Se supone que las valoraciones se realizan sobre los n primeros documentos recuperados. La relevancia de cada documento se obtiene de la propia lista de juicios de relevancia proporcionada por la colección de prueba. Para nuestros experimentos hemos fijado el valor de n a 15.

Para evaluar el incremento en efectividad producido por el método de realimentación no se puede, simplemente, comparar los resultados obtenidos por la consulta inicial y la consulta expandida. Parte de la mejora estaría provocada por el aumento de posiciones en la lista de resultados de los documentos utilizados en la realimentación y ya vistos por el usuario. Para evitar esta distorsión hemos utilizado el *método de la colección residual* [5] que supone que el usuario inspecciona los n primeros documentos de la lista de resultados para emitir su valoración. El resto de documentos, que recibe el nombre de colección residual, se utiliza para volver a medir la efectividad usando la consulta inicial y la expandida mediante realimentación.

Las medidas de efectividad que hemos elegido para mostrar los resultados son precisión y *recall* [21], clásicos en el ámbito de la recuperación de información. La *precisión* es la proporción de documentos recuperados que son relevantes. El *recall* es la proporción de documentos relevantes de la colección que se han recuperado. El incremento en la efectividad se mide comparando la media de los valores resultantes de la interpolación de la precisión en los 11 niveles estándar de *recall* (0.0, 0.1, 0.2, … 1.0).

Debe tenerse en cuenta que los valores de la precisión en los 11 niveles de *recall* suelen ser más bajos cuando se emplea el método de la colección residual que cuando se realiza una

Número de documentos	5000
Tamaño (Mb)	16
Número de consultas	50
Número de documento relevantes	385
Media 11-pt	0.2273

Tabla 1. Características del corpus usado en la evaluación de desambiguación

			Cons. expandidas	
Núm. términos	Corpus	Consultas	Todos los sinónimos	WSD
Máximo	6123	6	110	17
Mínimo	9	1	1	1
Media	383.74	2.88	26.96	5.16

Tabla 2. Número de términos significativos en el corpus y conjunto de consultas usados en la evaluación de desambiguación

evaluación estándar. La razón es que se excluyen los documentos relevantes mejor valorados. Esto no tiene gran importancia, ya que lo que realmente nos interesa medir es el incremento que se produce en la precisión debida a la realimentación.

3 Desambiguación del sentido de las palabras

3.1 Descripción de la tarea

La desambiguación no es un fin en sí misma, sino una tarea intermedia[4] muy necesaria para otras tareas del PLN. La tarea de desambiguación del significado de las palabras consiste en identificar el sentido de una palabra en un determinado contexto dentro de un conjunto de candidatos determinado.

En los últimos años se han propuesto varios enfoques de distinta naturaleza para WSD, que podemos clasificarlos de acuerdo con la fuente de conocimiento utilizada. Algunos enfoques se basan en la utilización de algún tipo de lexicón (o base de datos léxica) [30]. Otros, hacen uso de un corpus de texto no-anotados [19] o de corpus anotados semánticamente como colección de entrenamiento [3]. Finalmente, recientes trabajos proponen la combinación de varias fuentes de conocimiento, como bases de datos léxicas, corpus de texto, algunas heurísticas, colocaciones, etc. [29]. El enfoque empleado se encuadra dentro de esta clase, y se basa en la combinación de dos recursos lingüísticos, un corpus de texto y una base de datos léxica, como se describe en [26]. En el que se muestra la gran efectividad de dicho enfoque en la resolución de la ambigüedad léxica.

3.2 Uso de la desambiguación en RI

La expansión de las consultas ha resultado ser una técnica eficaz para mejorar la efectividad de los sistemas de RI. Esta mejora se puede conseguir, mediante realimentación y por medio de la inclusión de nuevos términos a la consulta original haciendo uso de recursos lingüísticos.

La expansión de consulta con WordNet ha mostrado ser potencialmente relevante para aumentar el *recall*, porque permite recuperar documentos relevantes que podrían no contener los términos de la consulta.

Proponemos una expansión de la consulta basada en la relación de sinonimia de WordNet, pero haciendo uso de la desambiguación. La expansión de la consulta con WordNet ha sido utilizada en otros trabajos de recuperación de información [27, 22] con resultados poco satisfactorios, debido fundamentalmente al tamaño de la consulta, que suele ser pequeño (3 términos de media), con lo cual hay una imposibilidad de resolver adecuadamente la ambigüedad.

Para solventar el problema de la escasa información contextual de las consultas, debido al tamaño de las mismas, le incorporamos la técnica descrita de realimentación a nuestro enfoque de desambiguación basado en la integración de recursos lingüísticos [26]. El proceso es como sigue, se realiza la consulta y

[4] Puede constituir un módulo para un sistema de procesamiento de lenguaje natural (PLN).

	Consulta original		Consulta expandida	Consulta expandida + WSD
	E. inicial	E. realiment.	E. realimentación	E. realimentación
Media 11-pt	0,1094	0,1853	0,1529	0.1896
% cambio		69.4	39,8	73,6
Media 3-pt	0,1106	0,1894	0,1530	0.1922
% cambio		71.2	38,3	73,8

Tabla 3. Precisión media y porcentaje de cambio en los 11 niveles estándar y 3 intermedios (0.2, 0.5, 0.8) de *recall*. Resultado de los experimentos de desambiguación

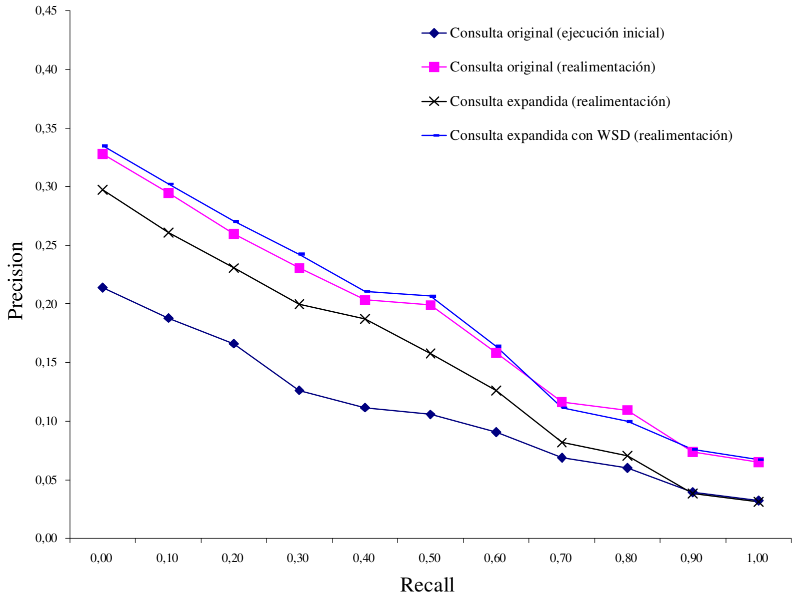

Figura 1. Efectividad para los diferentes tipos de expansión con feedback

se obtiene una relación de documentos ordenados por orden de relevancia. A continuación, se utiliza el juicio del usuario, es decir aquellos documentos relevantes, para realimentar la consulta original y así poder desambiguar y expandir los términos de la consulta original con los synsets adecuados de WordNet.

3.3 Entorno de evaluación

La mayor parte de los trabajos se centran en la evaluación directa de la desambiguación, es decir, en evaluar la efectividad en la asignación de los significados correctos a las palabras a desambiguar. Sin embargo, la desambiguación sirve fundamentalmente como ayuda a otras tareas. Por tanto, es deseable realizar una evaluación indirecta que mida la efectividad de la tarea a la que se aplica, dependiendo del método de desambiguación empleado [28, 25]. En este caso, aplicamos el método de desambiguación al proceso de RI.

Para la evaluación de desambiguación aplicada al proceso de RI, en este trabajo se utilizan 50 consultas elegidas aleatoriamente de la colección de prueba TREC-1[5] (Text REtrieval Conference) [23], ampliamente utilizada en la evaluación de sistemas de RI. Particularmente, hemos seleccionado 5.000 documentos, también aleatoriamente, del corpus Wall Street Journal (WSJ). De toda la información disponible en las consultas TREC (título, descripción, narrativa, conceptos, etc.) sólo utilizamos la relativa a la sección título, por considerarla la única representativa de una consulta típica de un usuario de un sistema de RI. Las tablas 1 y 2 muestran las características de esta colección de prueba.

[5] De entre las pertenecientes al intervalo 1-100 que tuvieran al menos un documento relevante en la subcolección creada.

Número de documentos	2379
Tamaño (Mb)	34
Número de consultas	50
Número de documentos relevantes	181
Media 11-pt – Ejecución inicial	0.0589
Media 11-pt – Ejecución realimentación	0.1105
% cambio	87.6

Tabla 4. Características del corpus usado en la evaluación de resúmenes

Núm. términos	Corpus	Consultas	Consultas expandidas
Máximo	8396	7	88
Mínimo	1250	1	1
Media	1655.78	3.44	23.66

Tabla 5. Número de términos significativos en el corpus y conjunto de consultas usados en la evaluación de resúmenes

Para la evaluación de nuestros experimentos hemos utilizado el sistema de recuperación de texto Smart [4] basado en el modelo del espacio vectorial y dotado de un módulo de evaluación de la eficacia sobre colecciones de prueba.

3.4 Resultados e interpretación

En la tabla 3 se muestra el porcentaje de cambio para el proceso de expansión. Destacando, en el caso de la expansión haciendo uso de la desambiguación (73,8%), el incremento mayor del 3,5% en 3 puntos de *recall*, con respecto a la consulta original (71.2%).

En la figura 1 podemos comparar las dos expansiones realizadas con la consulta original. En la cual se aprecia una ligera diferencia entre el proceso de recuperación haciendo uso de información de sinonimia con desambiguación (Consulta expandida con WSD), que sin hacer uso de ningún recurso lingüístico (Consulta original). Asimismo, la expansión de consulta para todos los sinónimos (Consulta expandida) no mejora los resultados de la consulta original los empeora, ya que se introducen muchos términos con diferentes significados. Sin embargo, haciendo uso de la desambiguación se obtiene una ligera mejora.

4 Generación automática de resúmenes de texto

4.1 Descripción de la tarea

La característica principal de la técnica de generación de resúmenes utilizada en nuestros experimentos es que emplea diversas heurísticas para valorar la importancia de cada frase, extrayendo las mejor puntuadas hasta llegar a un cierto número de ellas [6, 14, 15, 24]. Concretamente, el sistema utilizado se basa en el método propuesto en [15].

Las heurísticas que utiliza el sistema para valorar las frases son [15]: palabras clave, título, situación de la frase en el documento, palabras de la consulta y sinónimos de las palabras de la consulta (extraídos de WordNet).

4.2 Entorno de evaluación

Para la evaluación de la generación de resúmenes aplicada al proceso de RI con realimentación se han seleccionado todos los documentos del corpus WSJ, contenido en el volumen 2 de Tipster, con longitud superior o igual a 1250 palabras significativas (no pertenecientes a la lista de parada). Se eligieron también 50 consultas TREC, de entre las pertenecientes al intervalo 1-100, que tuvieran al menos un documento relevante en la subcolección creada. Las tablas 4 y 5 muestran las características de esta colección de prueba.

En los experimentos para esta tarea se utilizaron diferentes tipos de resúmenes:

- *Genéricos*. Se generaron utilizando las heurísticas: palabras clave, título y localización.
- *Adaptados a la consulta*. Resúmenes construidos usando las heurísticas anteriores y las palabras de la consulta.

Longitud		Colecciones de resúmenes			
		Genéricos	Adaptados a la consulta	AC expandida	Pseudopasajes
5%	Media 11-pt	0.1148	0.1124	0.1123	0.1025
	% cambio	94.9	90.8	90.7	74.0
10%	Media 11-pt	0.1146	0.1090	0.1090	0.0971
	% cambio	95.1	85.1	85.1	64.9
15%	Media 11-pt	0.1097	0.1098	0.1100	0.1082
	% cambio	86.2	86.4	86.8	83.7

Tabla 6. Precisión media y porcentaje de cambio en los 11 niveles estándar de *recall* Resultados de los experimentos de realimentación con resúmenes

- *Adaptados a la consulta expandida*. Se expanden las consultas con los sinónimos extraídos de WordNet y se aplican los mismos métodos de valoración que para los resúmenes adaptados.
- *Pseudopasajes*. Resúmenes generados utilizando sólo el método de palabras de la consulta. El objetivo es valorar la efectividad de la realimentación usando resúmenes que serían muy parecidos a pasajes.

Se experimenta con diferentes longitudes de resúmenes: 5, 10 y 15% (medida en número de frases del resumen respecto al documento completo).

4.3 Resultados e interpretación

Los resultados que presentamos en este apartado se obtuvieron ejecutando la consulta inicial sobre la colección original pero realimentando con los diferentes tipos de resúmenes, en vez de con el texto completo. Como consultas se utilizaron los títulos de los *topics* TREC seleccionados.

En la tabla 6 se muestra la media de la precisión en los 11 niveles de *recall* obtenida por cada tipo y longitud de resúmenes. También se presenta el porcentaje de cambio respecto a la ejecución inicial. En la tabla 4, se presentan los mismos datos obtenidos al realimentar con el texto completo.

Los mejores resultados se obtienen con los resúmenes genéricos y una longitud del 5%. A continuación, los resúmenes adaptados a las consultas y a la consulta expandida. Los peores resultados se obtienen cuando los resúmenes se construyen utilizando como única heurística las palabras de la consulta (pesudopasajes).

A partir de los resultados obtenidos puede concluirse que la realimentación con resúmenes permite, al menos, obtener tan buenos resultados, sino mejores en algunos casos, que usando el texto completo. Pero además, para el usuario leer un resumen de tan solo el 5% en lugar del texto completo para decidir sobre la relevancia de algunos documentos con los que luego realimentar la consulta representa una ventaja indudable. Sobre todo, con documentos largos como los pertenecientes al corpus con el que hemos experimentado.

Finalmente, la mejora de la precisión no es muy importante y es mayor para los resúmenes genéricos que para los adaptados al usuario. Pensamos que ambos efectos pueden deberse a la posibilidad de que el sistema de generación de resúmenes recoja información redundante. Quizás entre las frases puntuadas con valores más altos haya algunas con contenido similar. Es más, la posibilidad de que un documento contenga información redundante es mayor cuanto mayor es su longitud; y en nuestros experimentos hemos usado documentos muy largos.

5 *Conclusiones y futuros trabajos*

En este trabajo hemos estudiado la aportación de dos tareas de análisis del contenido textual en el entorno de los sistemas de recuperación de texto. En concreto la desambiguación del significado de las palabras y la generación automática de resúmenes.

Utilizamos la técnica de realimentación para ambas tareas. Por un lado, en la desambiguación hacemos uso de un enfoque basado en la combinación de recursos lingüísticos para mejorar la efectividad de los sistemas de recuperación de información. A pesar de la complejidad de la tarea, se ha mejorado la efectividad de la recuperación haciendo una expansión de los términos de la consulta, una vez desambiguada ésta empleando la realimentación, mediante información de sinonimia de WordNet. Los resultados demuestran que es preciso hacer uso de la desambiguación para sacar el máximo partido a la expansión de la consulta con WordNet. Y por otro, se han usado los resúmenes de los documentos en lugar del texto completo para realimentar el sistema. Los resultados demuestran que pueden sustituirlos sin pérdida de precisión (incluso, se obtienen ligeras mejoras en algunos casos). Además, aplicados a un sistema real, ofrecer un resumen en lugar de todo el texto puede ser un incentivo para usar la técnica de realimentación. Difícilmente un usuario utilizará esta técnica si para ello antes tiene que leer total o parcialmente documentos muy largos.

También hemos expuesto un método de evaluación sistemático de ambas tareas que nos permite comparar la efectividad haciendo uso de la realimentación, en el ámbito de los sistemas de recuperación de información.

Consideramos como principal línea de trabajo futuro el estudio de la sensibilidad de la RI a los errores de nuestro enfoque de desambiguación. Asimismo, estamos interesados en evaluar la aportación que pueden realizar ambas técnicas aplicadas al mismo tiempo a un sistema de RI que incorpore realimentación. El resumen del documento podría utilizarse para realimentar el sistema y como contexto suficiente para desambiguar el sentido de las palabras de la consulta. De esta manera podríamos expandir la consulta incorporando información proveniente de WordNet y de los resúmenes valorados por el usuario.

Referencias

[1] Allan, J. 1995. Relevance Feedback With Too Much Data. En *Proceedings of the 18th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieva*l, 337-343, Seattle, Washington.

[2] Baeza-Yates, R. y B. Ribeiro-Neto. 1999. *Modern Information Retrieval*. ACM Press Books, Nueva York. Página 371.

[3] Bruce, R. y W. Janyce. 1994. Word sense disambiguation using decomposable models. En *Proceedings of 33rd Annual Meeting of the Association for Computational Linguistics* (ACL'94).

[4] Buckley, C. 1985. *Implementation of the Smart Information Retrieval System*. Technical Report 85-686, Cornell University.

[5] Chang, Y.K., C. Cirillo y J. Razon. 1971. Evaluation of Feedback Retrieval Using Modified Freezing, Residual Collection, and Test and Control Groups. En G. Salton (ed.), *The SMART Retrieval System: Experiments in Automatic Document Processing*, 355-370. Prentice-Hall, Inc.

[6] Edmundson, H.P. 1969. New Methods in Automatic Abstracting. *Journal of the ACM*, 16(2):264-285.

[7] Gonzalo, J., F. Verdejo, I. Chugur, J. Cigarrán. 1998. Indexing with WordNet synsets can improve text retrieval. En *Proceedings of the COLING/ACL '98 Workshop on Usage of WordNet for NLP*, Montreal.

[8] Harman, D. 1992. Relevance Feedback Revisited. En *Proceedings of the 15th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval*, 1-10, Copenague.

[9] Hearst, M.A. 1994. *Context and Structure in Automated Full-Text Information Access.* PhD Thesis, Technical Report UCB/CSD-94/836, Computer Science Division, University of California at Berkeley.

[10] Hearst, M.A. 1994. Using Categories to Provide Context for Full-Text Retrieval Results. En *Proceedings of RIAO, Intelligent Multimedia Information Retrieval Systems and Management*.

[11] Hearst, M.A. y J.O. Pedersen. 1996. Reexamining the Cluster Hypothesis: Scatter/Gather on Retrieval Results. En *Proceedings of the 19th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval*, 76-84, Zurich.

[12] Ide, E. 1971. New Experiments in Relevance Feedback. En G. Salton (ed.), *The SMART Retrieval System: Experiments in Automatic Document Processing*, 337-354. Prentice-Hall, Inc.

[13] Knaus, D., E. Mittendorf, P. Schauble y P. Sheridan. 1995. Highlighting Relevant Passages for Users of the Intercative SPIDER Retrieval System. En *Proceedings of the Fourth Text REtrieval Conference (TREC-4)*, 233-238, Gaithersburg, Maryland.

[14] Luhn, H.P. 1958. The Automatic Creation of Literature Abstracts. *IBM Journal of Research and Development*, 2(2):159-165.

[15] Maña, M.J., M. Buenaga y J.M. Gómez. 1998. Diseño y Evaluación de un Generador de Resúmenes de Texto con Modelado de Usuario en un Entorno de Recuperación de Información. En *XIV Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural (SEPLN)*, 32-39, Alicante.

[16] Maña, M.J., M. Buenaga y J.M. Gómez. 1999. Using and Evaluating User Directed Summaries to Improve Information Access. En S. Abiteboul y A.M. Vercoustre (eds.), *Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries (ECDL'99)*, Lecture Notes in Computer Science, Vol. 1696, 198-214, Springer-Verlag.

[17] Miller, G. 1995. WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11).

[18] Oracle. 1997. *Managing Text with Oracle8TM ConText Cartridge*. An Oracle Technical White Paper.

[19] Pedersen, P. y R Bruce. 1997. Distinguishing word senses in untagged text. En *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*.

[20] Salton, G. y C. Buckley. 1990. Improving Retrieval Performance by Relevance Feedback. *Journal of the American Society for Information Science*, 41(4):288-297.

[21] Salton, G. y M.J. McGill. 1983. *Introduction to Modern Information Retrieval*. McGraw-Hill, New York.

[22] Sanderson, M. 1996. *Word sense disambiguation and information retrieval*. PhD thesis, Department of Computing Science, University of Glasgow.

[23] Sparck Jones, K. 1995. Reflections on TREC. *Information Processing and Management*, 31(3).

[24] Tombros, A. y M. Sanderson. 1998. *Advantages of Query Biased Summaries in Information Retrieval*. En *Proceedings of the 21st Annual International ACM/SIGIR Conference on Research and Development in Information Retrieva*l, 2-10, Melbourne.

[25] Ureña, L.A., M. Buenaga y J.M. Gómez. 2000. Integrating linguistic resources in TC through WSD. *Computers and the Humanities*. En prensa.

[26] Ureña, L.A., M. García, J.M. Gómez y A. Díaz. Integrando una base de datos léxica y una colección de entrenamiento para la desambiguación del sentido de las palabras. *Procesamiento del Lenguaje Natural*, (23), 1998.

[27] Voorhees, E.M. 1994. Query expansion using lexical-semantic relations. En *Proceedings of the 17th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieva*l, 61-69, Dublin.

[28] Wilks, Y. 1998. Is word sense disambiguation just one more NLP task? En *Proceedings of the SENSEVAL Conference*.

[29] Wilks, Y y M. Stevenson. 1998. Word sense disambiguation using optimised combinations of knowledge sources. En *Proceedings of COLING-ACL*.

[30] Xiaobin, L. y S. Spakowicz. 1995. WORDNET-based algorithm for word sense disambiguation. En *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)*, 1368-1374, Montreal.