

# Método basado en Marcas de Especificidad para WSD

Andrés Montoyo

Grupo de Procesamiento del Lenguaje y Sistemas de Información  
Departamento de Lenguajes y Sistemas Informáticos. Universidad de Alicante.

Teléfono: 965903772 Ext. 2725 Fax: 965909326

E-mail: montoyo@dlsi.ua.es

## Resumen

Este artículo presenta un método que resuelve la ambigüedad léxica de nombres en textos escritos en inglés. El método depende en gran medida de la taxonomía de nombres (en particular, Hipernimia/Hiponimia) que utiliza WordNet [8], y la noción de Marcas de Especificidad definida en este artículo para desambiguar automáticamente el sentido de las palabras. El método es completamente automático por lo que no necesita procesos de entrenamiento, codificación manual léxica de las entradas ni etiquetado de los nombres del texto.

La evaluación del método se ha realizado sobre el corpus Semantic Concordance (Semcor) [9] y sobre la enciclopedia electrónica "Microsoft Encarta 98 Enciclopedia Deluxe", obteniendo unos porcentajes de sentidos correctos de 65,8 % y de 65,6 % respectivamente. Estos porcentajes demuestran que el método obtiene resultados muy similares sobre dos corpus con diferentes dominios. Por lo tanto el método propuesto se puede aplicar satisfactoriamente a cualquier corpus.

## 1. Introducción

Cualquier sistema de Procesamiento del Lenguaje Natural (PLN) necesita utilizar abundante conocimiento sobre las estructuras del lenguaje, siendo este conocimiento de tipo morfológico, sintáctico, semántico y pragmático. El conocimiento morfológico nos proporciona información de cómo se construyen las palabras, el sintáctico de cómo combinar las palabras para formar oraciones, el semántico qué significan las palabras y cómo contribuye el significado de las mismas al significado completo de la oración, y por último, el pragmático de cómo el contexto afecta a la interpretación de las oraciones.

Todas las formas anteriores de conocimiento lingüístico tienen un problema asociado, la

ambigüedad. Por lo tanto, la resolución de este tipo de problema es uno de los objetivos principales de cualquier sistema de PLN. Se distinguen diversos tipos de ambigüedades: estructural, léxica, de ámbito de cuantificación, de función contextual y referencial.

En el presente artículo nos centramos en la resolución de la ambigüedad léxica, la cual aparece cuando las palabras presentan diferentes significados. A esta tarea se le conoce como Desambiguación del sentido de las palabras (Word Sense Disambiguation, WSD). Esta es una "tarea intermedia" [19] que sirve de ayuda cuando necesitamos conocer el sentido de las palabras en algunas aplicaciones del PLN, como en Traducción Automática (TA), Recuperación de la Información (RI), Clasificación de Textos, Análisis del Discurso, Extracción de Información (EI), etc.

De forma más genérica, la WSD consiste en la asociación de una palabra dada en un texto con una definición o significado, el cual la distingue de otros significados atribuibles a esa palabra. La asociación de palabras a los sentidos, se cumple dependiendo de dos recursos de información (contexto<sup>1</sup> y recursos de conocimiento externos<sup>2</sup>).

Para WSD existen diferentes métodos de trabajo como puede verse en [4], sin embargo, el presente artículo se centra en el método que se basa en el emparejamiento del contexto de la palabra a ser desambiguada con cualquier información de un recurso de conocimiento léxico externo, conociéndose como desambiguación del sentido de las palabras basada en el conocimiento (*WSD knowledge-driven*).

---

<sup>1</sup> El **contexto** de la palabra a ser desambiguada es considerado como un conjunto de palabras que acompañan a la palabra a desambiguar, junto con las relaciones sintácticas, categorías semánticas, etc. En este artículo el contexto se compone de oración en oración.

<sup>2</sup> Los **recursos de conocimiento externos** son los recursos léxicos, enciclopédicos, recursos de conocimiento léxico (WordNet) desarrollados manualmente que proporcionan datos valiosos para asociar palabras con sentidos, etc.

Particularmente para este artículo, se ha utilizado WordNet como recurso léxico externo porque combina las características de los diccionarios y los tesauros, además de incluir relaciones semánticas (hiperonimia, hiponimia, meronimia, etc) entre palabras. WordNet incluye definiciones para sentidos individuales de palabras (como un diccionario) y define grupos de palabras sinónimas, mediante "Synsets", representando un simple concepto léxico y organizándolo en una jerarquía conceptual (como un tesauro).

Muchas investigaciones, sobre WSD basado en el conocimiento, han sido realizadas durante los últimos años. Lesk en [5], propone un método para descifrar el sentido de una palabra en un contexto, contando el número de coincidencias que aparecen entre el contexto y la definición del diccionario. Cowie *et al.* en [2] describe un método para resolver la ambigüedad léxica de textos usando la definición dada en Longman's Dictionary of Contemporary English (LDOCE) obteniendo unos resultados del 47 % en cuanto a distinguir los sentidos y un 72 % para homógrafos. Yarowsky en [20] deriva clases de palabras a partir de palabras en categorías comunes del Roget's International Thesaurus. Wilks *et al.* en [18] utiliza co-ocurrencia de datos, extraídos del LDOCE, para construir vectores de contexto y de sentidos asociados a las palabras. Voorhees en [17] define la construcción denominada *hood* utilizando los hipónimos para nombres en WordNet. Sussna en [16] define una métrica basada en la distancia semántica entre los términos de un texto, la cual consistía en asignar pesos a los enlaces de WordNet según los tipos de relación (sinónimos, hiperónimos, etc) y en contar el número de arcos del mismo tipo que salen del nodo y la profundidad del arco en total. Resnik en [10] define una métrica basada en la similaridad semántica para las palabras en la jerarquía WordNet. Aguirre *et al.* en [1] describe un algoritmo no supervisado usando la Distancia Conceptual para desambiguar nombres en Semcor. Rigau *et al.* en [13] combina un conjunto de algoritmos no supervisados para desambiguar el sentido de las palabras en un corpus no etiquetado. Hale *et al.* en [3] presenta los resultados obtenidos de la combinación de Roget's International Thesaurus y la taxonomía de WordNet usando la similaridad semántica como medida. Stetina *et al.* en [15] introduce un método para WSD, basado en un corpus de entrenamiento etiquetado sintácticamente y semánticamente. Este método explota la información del contexto

de la oración y sus relaciones semánticas. Resnik en [12] presenta una medida para la similaridad semántica in una taxonomía IS-A, y la aplica en un algoritmo para resolver la ambigüedades sintácticas y semánticas. Mihalcea *et al.* en [7] expone un método para desambiguar nombres, verbos, adverbios y adjetivos de un texto, referenciando el sentido proporcionado por WordNet.

En el presente artículo se presenta un método, que resuelve la ambigüedad léxica de nombres, basándose en el conocimiento que nos proporciona la taxonomía semántica de nombres para el inglés, WordNet. En la siguiente sección se explica de forma intuitiva y detallada el método propuesto.

## **2. Método con Marcas de Especificidad**

El método que se presenta consiste básicamente en desambiguar automáticamente el sentido de las palabras que aparecen dentro del contexto de una oración. Normalmente, las palabras que aparecen en un mismo contexto tienen sus sentidos muy relacionados entre sí. Por tal motivo, y para resolver el problema interesa tener un recurso que tenga las palabras y los conceptos organizados alrededor de clases (jerarquías), de tal forma que describan todas sus características semánticas. Así, si dos palabras pertenecen a una misma clase quiere decir que sus sentidos están fuertemente relacionados. El método propuesto fue diseñado para que obtuviera las ventajas de lo comentado anteriormente, por eso se usarán las relaciones jerárquicas (hiperonimia/ hiponimia) que proporciona WordNet.

En primer lugar se explicará intuitivamente la noción de Marca de Especificidad para posteriormente aplicarla en la desambiguación. Cuanta más información común comparten dos conceptos, más relacionados estarán, y la información común que comparten esos dos conceptos es indicada por el concepto padre de ambos en la jerarquía, al cual se llamará de ahora en adelante Marca de Especificidad (ME). Estas ME se obtendrán a partir de clases semánticas de la jerarquía de WordNet.

El método recorrerá todos los subárboles de la jerarquía semántica de WordNet para el contexto de entrada y para cada ME calculará cuantas palabras del contexto de entrada se agrupan alrededor de ella. Aquella ME que agrupe el máximo número de palabras del contexto, será elegida como el sentido de la palabra. Con otras

palabras, el método busca aquella ME que tenga mayor densidad de palabras debajo de su subárbol, queriendo decir que sus sentidos están fuertemente relacionados. Este método se aplica sobre las jerarquías de WordNet de la siguiente manera.

La entrada al método WSD será el grupo de palabras  $W=\{W_1, W_2, \dots, W_n\}$  que se obtienen de la oración y forman el contexto. Cada palabra  $W_i$  se busca en WordNet y se obtienen los sentidos asociados a ellas  $S_i=\{S_{i1}, S_{i2}, \dots, S_{in}\}$  y para cada sentido  $S_{ij}$  se obtendrá todos los synsets hiperónimos en su taxonomía IS-A. Inicialmente, se busca el concepto común a todos los sentidos de las palabras que forman el contexto de entrada. A este concepto se le denomina Marca de Especificidad Inicial (MEI). Si esta MEI no resuelve la ambigüedad de las palabras, se va descendiendo nivel a nivel a través de la jerarquía WordNet asignando nuevas ME. Para cada ME anterior, se calculará el número de conceptos que forman parte del contexto y que están contenidos en la subjerarquía. Aquella ME que en su subjerarquía, tenga el mayor número de palabras del contexto será la elegida, asignando el sentido que nos devuelve WordNet a cada una de estas palabras que forman parte de la ME seleccionada.

En la figura 2, se puede apreciar gráficamente como la palabra  $W_1$  tiene cuatro sentidos diferentes y varias palabras de contexto. La MEI no resuelve la ambigüedad léxica, ya que la palabra  $W_1$  aparece en tres subjerarquías con diferentes sentidos. Sin embargo, la ME con el símbolo (\*) contiene el mayor número de palabras del contexto (tres) y, por lo tanto, será elegida para resolver el sentido  $S_2$  de la palabra  $W_1$ . Las palabras  $W_2$  y  $W_3$  también son desambiguadas eligiendo el sentido  $S_1$  para ambas. Para la palabra  $W_4$ , que no ha sido desambiguada satisfactoriamente, se le aplicarán las heurísticas especificadas a continuación y que están previstas para estos casos.

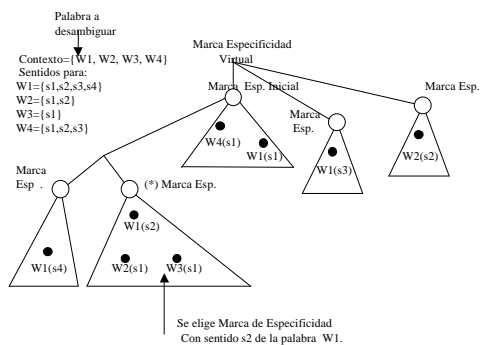


Figura 2: "Marcas de Especificidad" (ME).

Después de evaluar el método se observó que se podía mejorar, obteniendo unos niveles de desambiguación bastante mejores. Para ello se definieron tres heurísticas : de hiperónimo, de definición y de Marca de Especificidad Común. Estas tres heurísticas se definen a continuación:

- **Heurística de hiperónimo:** Esta heurística se utiliza para resolver la ambigüedad de aquellas palabras, que formando parte del contexto, no son directamente relacionadas por WordNet (plant and leaf, etc). Pero, sin embargo, a veces si que aparece alguna de las palabras del contexto como miembro de un synset de alguna relación de hiperonimia para algún sentido de la palabra a ser desambiguada (leaf#1 → plant organ). Esta heurística actúa de la siguiente manera para desambiguar una palabra dada. Todas las otras palabras que forman el contexto son chequeadas en cada uno de los synsets obtenidos de la relación de hiperonimia a partir de cada uno de los sentidos de la palabra dada. Si se encuentra algún synset hiperónimo conteniendo palabras del contexto, se asigna un peso en relación a la profundidad de la subjerarquía. Y el sentido con mayor peso será el elegido como correcto. En caso de tener varios sentidos con el mismo peso, o el peso es 0 para todos se pasará a aplicar la siguiente heurística. Ejemplo:

Palabras del contexto: *plant, tree, leaf, perennial*  
 Palabras no desambiguadas: *leaf*.  
 Sentidos finales: *leaf#1, leaf#2, leaf#3*.

Para leaf#1

=> *entity, something*  
 => *object, physical object*  
 => *natural object*  
 => **plant** part  
 => **plant** organ  
 => *leaf#1, leafage*

Como puede observarse, en los hiperónimos del sentido leaf#1 aparece **plant** que es otra de las palabras pertenecientes al conjunto de entrada o contexto. El peso para este sentido sería el siguiente: **peso** = peso + (Nº nivel / Nº niveles total) = (4/6) + ( 5/6) = **1,5**.

- **Heurística de definición:** WordNet incluye definiciones (glosas) para cada sentido asociado a una palabra. Estas definiciones son útiles porque tienen asociadas un micro-contexto para cada sentido. Por tal motivo, esta heurística actúa de la siguiente manera para desambiguar una palabra dada. Todas las

palabras que forman parte del contexto son chequeadas para buscar coincidencias en las definiciones asociadas a cada sentido que nos suministra WordNet. Cada vez que coinciden las palabras en la definición de un sentido, se incrementa su peso en una unidad. El sentido que finalmente tiene el mayor peso es elegido. En caso de tener varios sentidos con el mismo peso, o el peso es 0 para todos se pasará a aplicar la siguiente heurística. Ejemplo:

Palabras del contexto: *person, sister, musician*.  
 Palabras no desambiguadas: *sister, musician*.  
 Sentidos finales: *sister#1, sister#2, sister#3 sister#4*.

Para *sister#1* → peso = 2

1. *sister, sis* -- (a female **person** who has the same parents as another **person**; "my sister married a **musician**")

Para *sister#3* → peso = 1

3. *sister* -- (a female **person** who is a fellow member (of a sorority or labor union or other group); "none of her sisters would betray her")

- Heurística de Marca de Especificidad Común:** Con esta heurística se resuelve el problema de la granularidad fina (*year, month*). Esta heurística actúa de la siguiente manera para desambiguar una palabra dada. La primera ME que sea común a todos los sentidos resultantes de la heurística anterior es elegida, ya que proporciona el concepto común más informativo a tales sentidos. Mediante esta heurística se intenta resolver el problema de la granularidad fina que posee WordNet, ya que en la mayoría de los casos, los sentidos resultantes de las palabras se diferencian en un pequeño matiz y debido a que el contexto es muy general no se consigue dar con el sentido más acertado. Ejemplo:

Palabras del contexto: *year, month*.  
 Palabras no desambiguadas: *year*.  
 Sentidos finales: *year#1, year#2, year#3*.

Para *year#1*:

```
=> abstraction
=> measure, quantity
=> time period, period
=> year#1, twelvemonth
```

Para *year#2*:

```
=> abstraction
=> measure, quantity
=> time period, period
=> year#2
```

Como puede observarse en este ejemplo, debido a la granularidad tan fina que tiene la versión de WordNet 1.6 y que *month* no especifica nada sobre alguno de los sentidos de *year*, lo que más se puede afinar es diciendo que el sentido para *year* es el de **time period**.

### 3. Evaluación y Discusión.

#### 3.1 Test de prueba

La evaluación del método con Marcas de Especificidad se ha realizado sobre textos del corpus Semantic Concordance (Semcor) y de la enciclopedia electrónica *Microsoft Encarta 98 Encyclopedia Deluxe*. Los textos a desambiguar han sido escogidos al azar, sumando en total 100 frases y 619 nombres para el Semcor y 100 frases y 697 nombres para la enciclopedia. La evaluación del método se ha realizado sin y con heurísticas, para poder demostrar el porcentaje de mejora que aportan estas heurísticas.

#### 3.2 Resultados

Las siguientes tablas muestran los porcentajes obtenidos cuando se aplica el método base sólo o junto con las tres heurísticas. Estos resultados demuestran que cuando el método se aplica con heurísticas se obtienen unos porcentajes de mejora considerables.

Para el Semcor

%	Bien	Mal	Sin Desambiguar
Método Base	52'5%	28%	19'5%
1ª heurística	53%	29'7%	17'3%
2ª heurística	53'5%	30'2%	16'3%
3ª heurística	65'8%	32'5%	1'7%

Para Encarta

%	Bien	Mal	Sin Desambiguar
Método Base	55'2%	26'7%	18'1%
1ª heurística	56'4%	27'2%	16'4%
2ª heurística	58'3%	27'5%	14'2%
3ª heurística	65'6%	33'3%	1'1%

Como comentario final y considerando los datos anteriores, hay que resaltar que los porcentajes de las palabras que quedan sin desambiguar, entre 1,1% y 1,7%, se deben a que no se obtiene ninguna ME común que relacione a las palabras a ser desambiguadas.

### 3.3 Comparación con otros trabajos

Según Resnik *et al.* en [11], los métodos de WSD son muy difíciles de comparar unos contra otros, debido a la gran cantidad de diferencias que se encuentran en las palabras elegidas para desambiguar, además de los diferentes tipos de métodos empleados (i.e. métodos basados en el conocimiento, en el corpus o en estadística).

Resnik en [10] aplica su método a una tarea diferente, que consiste en desambiguar automáticamente el sentido de las palabras (nombres) utilizando la agrupación de los nombres. La entrada para este tipo de evaluación proviene de las categorías de Roget's Thesaurus. Como se puede deducir es muy difícil compararlo con nuestro método.

Yarowsky en [21] propone un método que obtiene un porcentaje del 91,4 % de sentidos correctos. Sin embargo, su método está orientado a un número limitado de palabras y a dos sentidos únicamente. Además, este método requiere de un corpus de entrenamiento, por lo que es muy difícil que nuestro método se pueda comparar con este.

Mihalcea *et al.* en [7] propone un método que utiliza dos recursos de información diferentes. El primero busca en Internet con preguntas previamente formateadas. Los sentidos son clasificados por el número de veces que la palabra aparece. El siguiente paso es utilizar WordNet para medir la densidad semántica de un par de palabras. Aunque este método tiene unos resultados del 85,6% para nombres, es tremendamente ineficiente. Esto es debido a que requiere un filtrado previo a causa del problema de combinatoria que acarrea. Nuestro método resuelve este problema, porque es completamente automático y no requiere ningún proceso de filtrado ni preguntas manuales en un buscador de Internet.

Los resultados publicados por Stetina *et al.* en [15] no se pueden comparar con los nuestros por varias razones. Utiliza un método supervisado para desambiguar el sentido de las palabras con una corpus de entrenamiento semánticamente etiquetado. Por lo tanto, requiere de un proceso de entrenamiento para ver como utiliza la limitación del dominio.

El método propuesto en este artículo se puede comparar con los de Agirre *et al.* en [1] y Yarowsky en [20], ya que estos son los métodos más similares al propuesto. El trabajo en [20] se tuvo que adaptar para poder trabajar con WordNet

y esto se hizo en [1]. Los resultados<sup>3</sup> de la comparación se muestran en la siguiente tabla:

%	"Coverage"	"Precision"	"Recall"
Specification Marks	98.5 %	66.6 %	65.7 %
Agirre and Rigau	86.2 %	71.2 %	61.4 %
Yarowsky	100 %	64.0 %	64.0 %

Una comparación más amplia y detallada se podría haber realizado entre los métodos de arriba, pero no ha podido ser posible debido a que las comparaciones se han realizado sobre diferentes versiones de WordNet y de oraciones de entrada. En este artículo se ha utilizado la versión 1.6 de WordNet, lo cual implica una mayor granularidad y un mayor número de sentidos para cada palabra que en la versión 1.4.

### 3.4 Discusión

Los resultados obtenidos demuestran que el método propuesto tiene un mejor recall que todos los demás métodos comparados. También tiene una mejor precisión que el método de Yarowsky y una mejor coverage que el método de Aguirre-Rigau. Por lo que el método propuesto mejora considerablemente bastantes aspectos de los otros métodos comparados.

## 4. Conclusiones y trabajos futuros

El método propuesto en este artículo para desambiguar el sentido de las palabras tiene la ventaja de no necesitar procesos de entrenamiento, ni codificación manual de las entradas, ni etiquetado manual. Por lo tanto, a partir de un texto de cualquier dominio se obtendrán de forma automática los sentidos de las palabras cuyas categorías léxicas sean nombres. Las salidas de este método serán las palabras con el sentido correspondiente de WordNet.

El inconveniente encontrado en los experimentos realizados con el Semcor y con la enciclopedia electrónica Microsoft Encarta 98 Encyclopedia Deluxe es que el método depende en gran medida de las relaciones semánticas (Hiperonimia/Hiponimia) y la organización jerárquica conceptual que utiliza WordNet. Hay palabras que semánticamente deberían estar relacionadas, sin embargo WordNet no refleja esta

<sup>3</sup> "Coverage" es obtenida del resultado entre el número total de sentidos contestados y el número total de sentidos. "Precision" se define como el resultado entre los sentidos desambiguados correctamente y el número total de sentidos contestados. "Recall" se define como el resultado entre los sentidos desambiguados correctamente y el número total de sentidos.

relación. Por ejemplo los nombres en inglés "plant" y "leaf" pertenecen al contexto de "flora" y WordNet únicamente incluye a "plant" en esta categoría mientras que a "leaf" le asigna la de "object". Por tal motivo se introdujeron las heurísticas del Hiperónimo y de la Definición ya que en la gran mayoría de casos solucionan este problema.

WordNet no es un recurso perfecto para desambiguar el sentido de las palabras, ya que tiene el problema de la granularidad fina para la distinción de los sentidos [4]. Esto crea muchas dificultades a la hora de desambiguar el sentido de las palabras automáticamente, debido a que hay que hacer elecciones en cuanto al significado, que a veces es difícil hasta manualmente. Incluso varios autores como [14] han dicho que las divisiones de un sentido en los diccionarios son demasiado finas para el propósito de los trabajos de Procesamiento del Lenguaje Natural. Para solucionar este problema se introdujo la heurística de la Marca de Especificidad Común ya que proporciona el concepto común mas informativo a tales sentidos.

Como trabajo futuro se pretende modificar el método para utilizar más y mejor las relaciones semánticas de WordNet, además de añadir más categorías léxicas a la hora de desambiguar, como los verbos y los adjetivos. Esto hará que se tenga más información de contexto y mejor relacionado. A consecuencia de lo dicho anteriormente y según McRoy en [6], también se pretende combinar otros recursos léxicos para obtener más información.

Pero quizá el cambio más importante sea utilizar relaciones sintácticas combinadas con diferentes técnicas y recursos para producir todos juntos una mejor desambiguación del sentido de las palabras.

## 5. Reconocimientos

Quiero agradecer a M. Palomar sus comentarios y revisiones en el presente trabajo, así como a German Rigau y a Eneko Aguirre por sus consejos para mejorarlo.

## 6. Referencias

- [1] Agirre E. and Rigau G. (1996) *Word Sense Disambiguation using Conceptual Density*. Proc. 16<sup>th</sup> International Conference on COLING. Copenhagen.
- [2] Cowie J., Guthrie J. and Guthrie L. (1992) *Lexical disambiguation using simulated annealing*. Proc. DARPA Workshop on Speech and Natural Language. 238-242. New York.
- [3] Hale, Michael L. Mc. A comparison of WordNet and Roget's taxonomy for measuring semantic similarity.
- [4] Ide N. and Véronis J. (1998) *Introduction to the Special Issue on Word Sense Disambiguation: The State of the Art*. Computational Linguistics. 24 (1), 1-40.
- [5] Lesk, M. (1986) *Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone*. Proc. 1986 SIGDOC Conference, ACM 24-26, New York.
- [6] McRoy S. (1992) *Using Multiple Knowledge Sources for Word Sense Discrimination*. Computational Linguistics 18 (1).
- [7] Mihalcea R. and Moldovan D. (1999) *A Method for word sense disambiguation of unrestricted text*. Proc. 37th Annual Meeting of the ACL 152-158, Maryland, Usa.
- [8] Miller G. A., Beckwith R., Fellbaum C., Gross D., and Miller K. J. (1990) *WordNet: An on-line lexical database*. International Journal of Lexicography, 3(4): 235-244.
- [9] Miller G., Leacock C., Rande T. and Bunker R. (1993) *A Semantic Concordance*. Proc. 3<sup>rd</sup> DARPA Workshop on Human Language Technology, 303-308, Plainsboro, New Jersey.
- [10] Resnik P. (1995) *Disambiguating noun groupings with respect to WordNet senses*. Proc. Third Workshop on Very Large Corpora. 54-68. Cambridge, MA.
- [11] Resnik P. and Yarowsky D. (1997) *A perspective on word sense disambiguation methods and their evaluation*. Proc. ACL Siglex Workshop on Tagging Text with Lexical Semantics, why, what and how?, Washington DC.
- [12] Resnik P. (1999) *Semantic similarity in a taxonomy: an information-based measure and its application to problems of ambiguity in natural language*. In Journal of Artificial Intelligence Research 11. 95-130.
- [13] Rigau G., Atserias J. and Agirre E. (1997) *Combining Unsupervised Lexical Knowledge Methods for Word Sense Disambiguation*. Proc. 35th Annual Meeting of the ACL, 48-55, Madrid, Spain.
- [14] Slator B. and Wilks Y. (1987) *Towards semantic structures from dictionary entries*. Proc. 2<sup>nd</sup> Annual Rocky Mountain Conference on Artificial Intelligence, 85-96. Boulder, CO.
- [15] Stetina J., Kurohashi S. and Nagao M. (1998) *General word sense disambiguation method based on full sentencial context*. In Usage of WordNet in Natural Language Processing. COLING-ACL Workshop, Montreal, Canada.
- [16] Sussna M. (1993) *Word sense disambiguation for free-text indexing using a massive semantic network*.

- Proc. Second International CIKM, 67-74, Arlington, VA.
- [17] Voorhees E. (1993) *Using WordNet to disambiguation word senses for text retrieval*. Proc. 16<sup>th</sup> Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. 171-180, Pittsburgh, PA.
- [18] Wilks Y., Fass D., Guo C., McDonal J., Plate T. and Slator B. (1993) *Providing Machine Tractable Dictionary Tools*. In *Semantics and the lexicon* (Pustejovsky J. Ed.) 341-401.
- [19] Wilks Y. And Stevenson M. (1996) *The grammar of sense: Is word sense tagging much more than part-of-speech tagging?* Technical Report CS-96-05, University of Sheffield, UK.
- [20] Yarowsky D. (1992) *Word Sense disambiguation using statistical models of Roget's categories trained on large corpora*. Proc. 14<sup>th</sup> COLING, 454-460, Nantes, France.
- [21] Yarowsky, D. (1995) *Unsupervised word Sense disambiguation rivaling supervised methods*. Proc. 32nd Annual Meeting of the ACL.