

Un Sistema Basado en Conocimiento para el Reconocimiento de Implicación Textual *

Ó. Ferrández, R. M. Terol, R. Muñoz, P. Martínez-Barco y M. Palomar

Dept. de Lenguajes y Sistemas Informáticos (Universidad de Alicante)

Carretera San Vicente s/n 03690 Alicante España

{ofe,rafamt,rafael,patricio,mpalomar}@dlsi.ua.es

Resumen: Este artículo presenta un sistema para el reconocimiento del fenómeno de implicación textual basado en conocimiento. El sistema propuesto está compuesto por dos componentes: el primero de ellos deriva las formas lógicas asociadas a un par de fragmentos de texto (denominados texto e hipótesis), y una vez obtenidas, el segundo componente aplica varios enfoques con el objetivo de obtener un factor de similitud semántica entre ellas y determinar si se produce implicación textual entre los textos. Todos los enfoques están basados en la base de datos léxica WordNet y las relaciones entre conceptos que ésta establece. El sistema ha sido evaluado con los corpus y la metodología de evaluación de la competición *PASCAL Second Recognising Textual Entailment* consiguiendo un 60% de precisión media.

Palabras clave: Implicación textual, formas lógicas, similitud semántica, WordNet

Abstract: This paper covers the recognition of textual entailment by means of an approach based on knowledge. Our approach consists of two stages. The first stage infers the logic forms from two fragments of text. These logic forms are obtained by analysing the dependency relations between words. And the second stage carries out several methods in order to achieve a score that determines the semantic similarity between the derived logic forms. Depending on this score the system establishes the existence of an entailment relation. All the methods use the WordNet lexical database as a knowledge source and obtain a semantic similarity score by means of WordNet relations. Our approach has been evaluated within the PASCAL Second RTE Challenge and achieved 60% average precision.

Keywords: Textual Entailment, logic forms, semantic similarity, WordNet

1. Introducción

En el lenguaje natural un fenómeno frecuente es la variabilidad de expresión semántica, mediante la cual un mismo significado puede ser expresado por o inferido de diferentes textos. La resolución de este fenómeno ayudaría a mejorar numerosas aplicaciones de Procesamiento del Lenguaje Natural (PLN) e implicaría un gran avance dentro de este área.

La tarea de implicación textual (del inglés, *Textual Entailment*) ha sido recientemente definida como una solución común para el modelado de la variabilidad del lenguaje (Dagan y Glickman, 2004). Una implicación textual estricta define una relación existente entre dos expresiones en lenguaje natural, de manera que el significado de

una de ellas puede deducirse a partir de la otra. El fragmento de texto que mantiene la implicación es definido como T , mientras que la hipótesis deducida por T se define como H . Un ejemplo de una relación de implicación textual sería el siguiente:

T : *But only one major oil company was shamed by the 623-page report: Texaco, part of Chevron, the US's second largest energy group.*

H : *Chevron owns Texaco.*

en el ejemplo H puede ser deducida por T , ya que en T se establece que *Texaco* es parte de *Chevron* y ésta es una afirmación similar a la que se dice en H (*Chevron* es dueño de *Texaco*).

El par de expresiones (T-H) pueden mostrar diferentes hechos, pero para que se produzca la implicación textual tanto el texto como la hipótesis deben ser expresiones con

* Esta investigación ha sido parcialmente financiada bajo los proyectos CICYT número TIC2003-07158-C04-01

sentido, coherentes y del significado del texto debe ser deducida la hipótesis.

La implicación textual es un fenómeno que se da en muchas aplicaciones PLN, y resultaría muy beneficioso disponer de un sistema que lo detectara y mejorara la eficiencia de dichas aplicaciones. En el caso de Búsquedas de Respuestas (QA) una misma respuesta puede ser expresada de diferentes maneras sintáctico-semánticas y un módulo de implicación textual ayudaría a detectar cuales de las posibles respuestas pueden deducir la respuesta esperada. Del mismo modo, en Extracción de Información (IE) descubrir las diferentes formas de expresar un mismo concepto también resultaría provechoso, por último en sistemas de generación automática de Resúmenes (SUM) identificar información redundante entre las frases con mayor contenido de información también mejoraría esta tarea. Por lo tanto, disponer de un sistema que ayude a la resolución de la implicación textual resulta una tarea motivadora por muchos aspectos.

En este artículo se presenta un sistema basado en conocimiento para el reconocimiento de la implicación textual estricta. La implicación textual estricta abarca tanto transformaciones sintácticas como semánticas, las cuales pueden producirse del texto a la hipótesis, con la única restricción de que la hipótesis debe ser deducida a partir del significado del texto. El sistema que se presenta, resuelve la implicación textual mediante la derivación de las formas lógicas de los textos y encontrando relaciones semánticas entre sus predicados utilizando WordNet.

El artículo está organizado en las siguientes secciones. La sección 2 introduce un breve estado de la cuestión de la tarea. La sección 3 describimos nuestro sistema en detalle. La evaluación y discusión de los resultados se muestra en la sección 4, y las conclusiones y trabajos futuros en la sección 5

2. Estado de la cuestión

En la literatura podemos encontrar dos grandes grupos o aproximaciones: los enfoques que utilizan técnicas basadas en conocimiento, los cuales normalmente hacen uso de recursos lingüísticos, y los enfoques que se basan en aprendizaje automático y métodos estadísticos para detectar las relaciones de implicación.

La implicación textual es una tarea nove-

dosa y que ha tenido un fuerte interés en la comunidad científica. Un claro ejemplo es la organización de varios *Workshops* como *PASCAL Challenge Workshops*¹ (Dagan, Glickman, y Magnini, 2005) y *ACL Workshop on Empirical Modeling of Semantic Equivalence and Entailment*². En los siguientes apartados comentamos algunas de las aproximaciones que fueron expuestas en estos *Workshops*.

2.1. Basadas en aprendizaje automático y métodos estadísticos

En las aproximaciones basadas en métodos estadísticos, destacar el trabajo desarrollado por (Glickman y Dagan, 2005), propone un entorno probabilístico que formalice la noción de implicación textual. En este trabajo sus autores describen un modelo basado en las probabilidades de coocurrencia.

En (Bos y Markert, 2005), un árbol de decisión es entrenado usando características obtenidas a partir de diferentes métodos de PLN como *word overlap*, *CCG-parser*, lógica de primer orden, *Vampire* y *Paradox* (un generador de autómatas finitos).

Otro trabajo interesante es el desarrollado por (Zanzotto, Pazienza, y Pennacchiotti, 2005). En este trabajo, los autores se centran en encontrar la implicación textual entre los verbos, ya que el significado de la frase viene normalmente gobernado por el verbo. Investigan prototipos de formas textuales que describan relaciones de implicación, a las que ellos llaman *textual entailment patterns*. Estos patrones son analizados para detectar expresiones que denotan implicación textual. El análisis de los patrones lo realizan sobre grandes colecciones de textos y aplicando medidas estadísticas relevantes para la tarea.

2.2. Basados en técnicas de conocimiento

En esta sección mostramos varias de las aproximaciones que se basan en el uso de recursos lingüísticos. Por ejemplo, (Akhmatova, 2005) describe un sistema basado en un análisis semántico dirigido por la sintaxis. Hace uso de la noción de proposición atómica como principal elemento para reconocer la implicación textual. Una proposición atómica

¹<http://www.pascal-network.org/Challenges/RTE/> y <http://www.pascal-network.org/Challenges/RTE2>

²<http://acl.ldc.upenn.edu/W/W05/>

ca es una mínima exposición declarativa (o pequeña idea). Establece que, si para todas las proposiciones atómicas en la hipótesis hay una en el texto por la que podría ser implicada entonces se produce la implicación textual, en el caso de que no se encuentren las proposiciones atómicas correspondientes no se producirá la implicación.

(Herrera, Peñas, y Verdejo, 2005) presentan un sistema que convierte el texto y la hipótesis en árboles de dependencias. Estos árboles son comparados mediante un algoritmo simple de emparejamiento centrado en analizar todas las ramas empezando desde cualquier hoja del árbol de la hipótesis y mostrar el emparejamiento con cualquiera de las ramas de los árboles asociados al texto. Un nodo de la rama del árbol de la hipótesis casa con un nodo del árbol del texto si hay una implicación léxica entre ambos. Los autores definen esta implicación léxica considerando las relaciones que proporciona WordNet (sinonimia, antonimia, hiperonimia e implicación).

Por último, destacar también el trabajo realizado por (Kouylekov y Magnini, 2005). En este trabajo se utiliza un algoritmo sobre la distancia de edición aplicado a los árboles de dependencias derivados del texto y de la hipótesis. Si existe una secuencia de transformaciones aplicables sobre el árbol del texto, por las cuales se pueda obtener la hipótesis con un coste inferior a un umbral específico, se produce implicación textual.

Aunque hay múltiples aproximaciones que utilizan métodos estadísticos y de aprendizaje automático, consiguiendo muy buenos resultados, la tendencia principal se orienta hacia la resolución de la implicación textual mediante recursos lingüísticos. Esto es debido a que el reconocimiento de esta tarea es dependiente de una comprensión semántica detallada del texto y, además el uso de recursos como WordNet parece que son apropiados para dicha tarea.

3. Arquitectura del sistema

Nuestro sistema detecta la implicación textual estricta entre dos fragmentos de textos aplicando un proceso basado en dos componentes. El primer componente aplica una serie de técnicas basadas en el procesamiento del lenguaje natural que derivan la forma lógica del texto y de la hipótesis. Mientras que, el segundo componente proporciona un factor característico de la similitud semántica

existente entre los predicados de las formas lógicas. En base al valor que adquiera este factor, el sistema toma la decisión referente a si se produce o no la implicación entre ambas formas lógicas y, en consecuencia, la implicación textual entre el texto y la hipótesis.

La Figura 1 muestra la arquitectura de nuestro sistema. En ella se pueden distinguir los dos componentes que pasamos a describir con detalle a continuación.

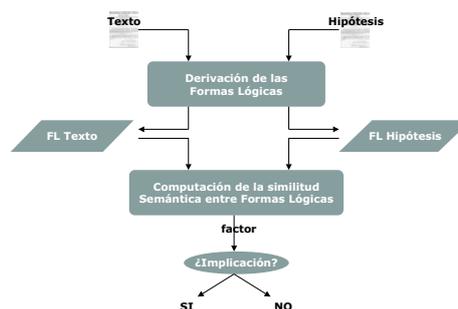


Figura 1: Arquitectura del sistema

3.1. Derivación de las formas lógicas

Nuestro sistema hace uso de las formas lógicas asociadas al texto y a la hipótesis con el objetivo de aplicar un procesamiento de lo más natural posible. El formato de una forma lógica es similar al formato empleado en el recurso léxico “*Logic Form Transformation of eXtended WordNet*” (LTF) (Harabagiu, Miller, y Moldovan, 1999). La forma lógica de una frase es derivada a través de la aplicación de reglas basadas en PLN a las relaciones de dependencia existentes entre las palabras de la frase. Así, la primera tarea necesaria para derivar la forma lógica consiste en obtener las relaciones de dependencia entre las palabras. Con este propósito utilizamos las dependencias entre palabras dadas por el recurso léxico “*MINIPAR*” (Lin, 2005). Una vez calculadas las relaciones de dependencias entre las palabras, la siguiente tarea consiste en derivar automáticamente la forma lógica a través del análisis de estas relaciones de dependencia. De este modo, podemos concluir que la derivación de la forma lógica es un proceso composicional que comienza en las hojas del árbol de dependencias, continua por las ramificaciones y finaliza en la raíz.

A continuación aplicamos este complejo proceso de derivación de la forma lógica a

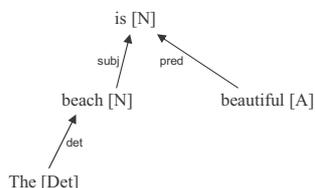


Figura 2: Árbol de dependencias de la frase

la frase “*The beach is beautiful*”. La primera tarea consiste en encontrar las relaciones de dependencia entre las palabras de la frase. La figura 2 muestra estas relaciones de dependencia. La segunda tarea consiste en aplicar las reglas a las hojas del árbol de dependencias derivando así un predicado de la forma lógica asociado a cada hoja. Esta tarea composicional continúa por las ramificaciones del árbol hasta llegar a su raíz.

Una vez finalizada esta segunda tarea, la forma lógica asociada a la frase queda instanciada como “*beach:NN(x2) be:VB(e1, x2, x3) Atributo:IN(e1, x1) beautiful:JJ(x1)*”. Cabe destacar que el verbo de la frase es intransitivo. Este hecho implica que en la forma lógica aparezca el predicado “*Atributo*” que relaciona dicho verbo y el adjetivo “*beautiful*”.

3.2. Similitud Semántica entre las Formas Lógicas

El objetivo de este componente es calcular el factor que le permita al sistema decidir si entre las formas lógicas asociadas al texto y a la hipótesis se produce una implicación. Para ello, es necesario un análisis de las relaciones semánticas de los predicados de ambas formas lógicas. Lo primero que hace este componente es comprobar la relación semántica existente entre los predicados asociados a los verbos de ambas formas lógicas. Aunque en esencia la idea es similar a la expuesta por (Zanzotto, Pazienza, y Pennacchiotti, 2005) que consiste en darle una mayor importancia semántica al verbo de la frase, ellos en su trabajo establecen a partir de un corpus precondiciones o propiedades que debe satisfacer un verbo y dependiendo de si el verbo implicado las satisface o no se producirá la implicación. En nuestro caso la relación semántica entre los predicados verbales se establece explorando las relaciones de WordNet mediante alguno de los enfoques que desarrollamos y que serán explicados más adelante. Si se

produce alguna relación semántica entre los predicados verbales, se procede a analizar las relaciones semánticas existentes entre los predicados asociados a las palabras dependientes de dichos verbos. Este conjunto de análisis parciales proporciona factores semánticos que son sumados y normalizados obteniendo así el factor característico de la similitud semántica. El pseudo-código del algoritmo es el siguiente:

```

pesoSem = 0
Tvb = obtenerVerbos(T)
Hvb = obtenerVerbos(H)
for i = 0 ... size(Tvb) do
    for j = 0 ... size(Hvb) do
        if semSim(Tvb(i),Hvb(j)) ≠ 0 then
            pesoSem += semSim(Tvb(i),Hvb(j))
            Telem = obtenerElem(Tvb(i))
            Helem = obtenerElem(Hvb(j))
            pesoSem += semSim(Telem,Helem)
        end if
    end for
end for
if pesoSem > umbral then
    return CIERTO
else
    return FALSO
end if
    
```

Hemos adoptado tres enfoques diferentes con el objetivo de obtener la similitud semántica existente entre los predicados de ambas formas lógicas ($semSim(x,y)$). Todos nuestros enfoques se basan en la jerarquía de WordNet y en sus relaciones. El primero de ellos se caracteriza por aplicar un estudio de relaciones semánticas complejo entre los predicados de ambas formas lógicas. Por otro lado, realizamos otro enfoque que trate este análisis semántico de manera mucho más relajada y se centre únicamente en las relaciones que consideramos adecuadas a la tarea de implicación textual. Por último, nuestro tercer enfoque hace uso de la medida de similitud definida por Lin (Lin, 1998), que aunque no se trata de una medida específica para la implicación textual se adapta bastante bien a la tarea. En las siguientes subsecciones se describen en detalle cada uno de los enfoques.

Antes de describir estos enfoques, destacar que nuestro sistema no emplea ningún módulo de desambiguación de sentidos a la hora de obtener las relaciones existentes entre dos predicados. Para ello, sólo hemos considerado

la primera mitad de los sentidos de una palabra en WordNet. Tomamos esta decisión, debido a que los sentidos están organizados por frecuencia de uso y consideramos que con obtener la primera mitad de ellos conseguiremos un detalle semántico de la palabra suficiente para esta tarea. Por otra parte, el umbral que decide si se produce implicación textual, ha sido calculado empíricamente a partir del corpus de desarrollo. La Figura 3 de la siguiente sección muestra y detalla este proceso.

3.2.1. Análisis Semántico Complejo utilizando relaciones de WordNet

En la base de datos léxica WordNet (Miller, 1990), un synset es un conjunto de conceptos que representan el mismo significado. Se define un concepto como el uso de una palabra en un determinado contexto (sentido). Así, este enfoque trata de determinar si dos conceptos distintos están relacionados entre sí a través de la composición de diferentes relaciones de WordNet, concretamente las relaciones de hiperonimia, hiponimia, implicación, sinonimia, meronimia y holonimia. La longitud del camino que relaciona los dos conceptos no debe ser nunca superior a 4, dado que a longitudes mayores la relación semántica que puedan tener dos conceptos es tan débil que introduciría más ruido que beneficio. Cada una de estas relaciones de WordNet tiene asociado un peso: 0.8 para la relación de hiperonimia, 0.7 para las relaciones de hiponimia e implicación, 0.9 para las relaciones de sinonimia, y 0.5 para las relaciones de meronimia y holonimia. De este modo, el peso del camino existente entre dos conceptos se calcula como el producto de los pesos asociados a las relaciones que conectan dichos conceptos. Esta técnica deriva en su esencia de la técnica empleada por el algoritmo *Spread-Weights* (Moldovan y Novischi, 2002), aunque ambas divergen en el tratamiento entre conceptos.

A pesar de que estas relaciones de WordNet han sido aplicadas al fenómeno de implicación textual, también ha sido usado en la resolución de otros problemas del PLN como la Búsqueda de Respuestas (Terol, Martínez-Barco, y Palomar, 2005). Además, hemos considerado que la aplicación de estas relaciones de WordNet es muy interesante a la hora de comparar su impacto en la tarea de implicación textual. De este modo, podemos estudiar que relaciones de WordNet son más

adecuadas para dicha tarea.

3.2.2. Análisis Semántico Superficial utilizando relaciones de WordNet

En este caso, determinamos si dos conceptos están relacionados a través de la composición de las relaciones de WordNet que hemos considerados más específicas para la tarea de implicación textual. Hiponimia, implicación y sinonimia son las relaciones de WordNet que hemos manejado en este enfoque. Consideramos que estas tres relaciones son las más adecuadas para el fenómeno de implicación textual, y queremos comprobar que con el único uso de estas tres relaciones, el reconocimiento de la implicación textual mejora con respecto al enfoque anterior.

Al igual que en el anterior enfoque, la longitud del camino que relaciona los dos conceptos diferentes debe ser inferior o igual a 4 synsets, y los pesos asociados a cada una de las relaciones de WordNet son los mismos (0.7 para la hiponimia y la implicación, y 0.9 para la sinonimia). Como en el enfoque anterior, el peso de este camino se calcula como el producto de los pesos asociados a las relaciones que conectan los synsets intermedios.

3.2.3. Medida de similitud de Lin

En este caso, la similitud entre los predicados de las formas lógicas ha sido obtenida usando la medida de similitud de Lin (Lin, 1998) tal y como es implementada en la aplicación WordNet::Similarity³ (Pedersen, Patwardhan, y Michelizzi, 2004). WordNet::Similarity es una herramienta de código abierto desarrollada por la Universidad de Minnesota. Proporciona tres medidas de relación y seis de similitud entre un par de conceptos. Todas estas medidas se basan en la jerarquía de la base de datos léxica WordNet.

Las medidas de similitud analizan las relaciones *is-a* de WordNet para obtener la similitud semántica existente entre dos conceptos. Podemos dividirlos en dos grupos: por un lado las que se basan en los caminos de conexión entre conceptos (*path-based*), y por otro lado agruparíamos aquellas que utilizan el contenido de información (*information content-based*) para obtener la similitud.

Para nuestros experimentos, hemos escogido la medida de similitud de Lin. Esta medida escala el contenido de información del

³<http://www.d.umn.edu/~tpederse/similarity.html>

LCS (Least Common Subsumer⁴) con el contenido de información de cada concepto individualmente.

4. Evaluación del sistema

Con el objetivo de evaluar nuestro sistema sobre la tarea de implicación textual, hemos utilizado los corpus proporcionados por *PASCAL Second Recognising Textual Entailment Challenge*. Los organizadores de esta competición proveen a los participantes con un corpus para el desarrollo y adaptación de los sistemas, y un corpus para evaluar la precisión de los sistemas en la tarea de reconocimiento de implicación textual. Los corpus poseen ambos 800 pares texto-hipótesis manualmente anotados, los cuales deberán ser predecidos por los sistemas como relaciones de implicación si del significado del texto se deduce la hipótesis. De éstos el 50% posee una relación de implicación correcta, mientras que el resto son ejemplos incorrectos. A su vez, los pares texto-hipótesis corresponden a fallos y aciertos extraídos de diferentes aplicaciones reales de Extracción de Información (IE), Recuperación de Información (IR), Búsquedas de Respuestas (QA) y técnicas de Resúmenes (SUM).

Para evaluar la precisión de los sistemas se establecen dos medidas: el porcentaje de aciertos con respecto a la anotación manual (*accuracy*) y la precisión media (*Average Precision*). Ésta última evalúa la capacidad de los sistemas en determinar las implicaciones correctas de acuerdo a su coeficiente de confianza, en orden decreciente desde el par texto-hipótesis que posee un coeficiente mayor hasta el que posee el menor coeficiente. Formalmente:

$$PrecMedia = \frac{1}{R} \left(\sum_{i=1}^n E(i) \frac{\#correctos_hasta_par_i}{i} \right)$$

donde n es la cantidad de pares en el corpus, R es la cantidad de pares positivos, $E(i)$ es 1 si el par es positivo y 0 si no lo es, e i recorre todos los pares ordenados según el coeficiente de confianza.

Consideramos apropiado evaluar los diferentes enfoques de nuestro sistema sobre el corpus proporcionado por *PASCAL Second Recognising Textual Entailment Challenge*. Todos nuestros enfoques se basan en derivar

⁴LCS es el primer concepto antecesor común a dos conceptos en la jerarquía

las formas lógicas del texto y la hipótesis, y a partir de sus predicados obtener la similitud semántica entre ellos.

Los enfoques que utilizan las relaciones de WordNet para determinar la similitud entre dos conceptos, los denominaremos *WNsuperficial* y *WNcomplejo*. Donde *WNsuperficial* hace uso de tres relaciones de WordNet (sinonimia, hiponimia y entailment, ver sección 3.2.2) que se consideran más relevantes para la tarea de implicación textual, mientras que *WNdetallado* utiliza seis relaciones (hiperonimia, hiponimia, implicación, sinonimia, meronimia y holonimia, ver sección 3.2.1) para obtener el peso de similitud entre dos conceptos. El último de nuestros enfoques es el que utiliza la medida de similitud de Lin, al que denominaremos *WNLin*.

En el Cuadro 1, se muestran los resultados obtenidos para cada uno de los enfoques mediante el *script* de evaluación que se proporciona en la competición *PASCAL Second Recognising Textual Entailment*.

El umbral que determina si la hipótesis es inducida por el texto, ha sido obtenido empíricamente. Varios experimentos fueron realizados sobre el corpus de desarrollo, los cuales establecieron un umbral de decisión para cada uno de nuestros enfoques. La figura 3 muestra el incremento empírico del umbral para cada enfoque, en relación con la precisión que se obtiene con cada incremento.

El umbral que mejores resultados obtuvo sobre el corpus de desarrollo, fue de un valor de 0.24 para los enfoques de *Lin* y *WNdetallado*, mientras que para el enfoque que sólo utilizaba tres relaciones de WordNet (*WNsuperficial*) el mejor umbral empírico fue de 0.25. Aunque los umbrales son bastante parecidos y sufren un comportamiento similar conforme se van aumentando, el tratamiento de diferentes relaciones de WordNet por parte de nuestros enfoques ocasiona esta ligera diferencia entre ellos.

Los resultados del Cuadro 1 muestran que en todos los casos, tanto para el corpus de desarrollo como para el test, el enfoque que utiliza la medida de similitud de Lin obtiene mejores resultados que el resto de enfoques. Sin embargo, cabe destacar la mejora que produce el enfoque que hace uso únicamente de las tres relaciones de WordNet que consideramos más relevantes para la tarea con respecto al enfoque que utiliza seis relaciones de WordNet. Esta mejora, aunque pequeña, la consi-

			overall	IE	IR	QA	SUM
desarrollo	WNlin	Accuracy	0.5462	0.5421	0.5440	0.5722	0.5260
	WNcomplejo	Accuracy	0.5273	0.5510	0.5345	0.4677	0.5686
	WNsuperficial	Accuracy	0.5375	0.5026	0.5357	0.5641	0.5474
test	WNlin	Accuracy	0.5563	0.4950	0.5800	0.6100	0.5400
		Average Precision	0.6089	0.5722	0.6159	0.6431	0.6215
	WNcomplejo	Accuracy	0.5475	0.4750	0.5850	0.6150	0.5150
		Average Precision	0.5743	0.5853	0.6113	0.5768	0.5589
	WNsuperficial	Accuracy	0.5513	0.5150	0.5350	0.5950	0.5600
		Average Precision	0.6027	0.5689	0.5891	0.6385	0.6105

Cuadro 1: Resultados obtenidos con el script de evaluación PASCAL RTE2

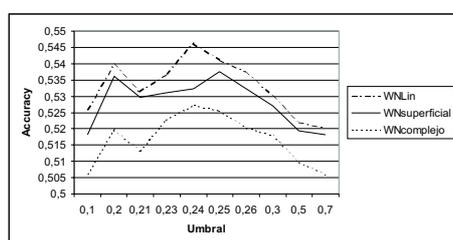


Figura 3: Ajuste de los umbrales sobre el corpus de desarrollo

deramos importante ya que demuestra nuestra hipótesis sobre qué relaciones de WordNet son más adecuadas para la implicación textual y, además el tratamiento de un menor número de relaciones ocasiona en el sistema un coste de procesamiento reducido.

La razón por la cual nuestro enfoque que hace un uso más complejo de las relaciones de WordNet obtiene resultados más bajos, se debe a que intenta establecer una comparación semántica objetiva entre las formas lógicas más que una relación de implicación textual.

No obstante, nuestro sistema en todos sus enfoques se equivoca cuando encuentra un factor alto de comparación semántica entre los predicados de las formas lógicas, pero con un significado final diferente. Por ejemplo:

T: Jose Reyes scored the winner for Arsenal as they ended a three-game league losing streak with a victory over battling Charlton.

H: Jose Reyes scored the winner against Arsenal.

nuestro sistema determina que se produce una relación de implicación textual debido al elevado factor de similitud entre los predicados, y sin embargo no es así.

La razón de que se produzca este fenómeno se debe a que, aunque los verbos y sus

predicados dependientes son muy similares semánticamente, en la hipótesis la palabra *against* causa un significado diferente que no es detectado por el sistema. Con el fin de solucionar estos casos y como trabajo futuro, será necesario un mayor procesamiento sintáctico para detectar aquellas palabras y/o negaciones que impliquen cambios de significado de las frases.

5. Conclusión y trabajo futuro

En este artículo se presenta un sistema para la resolución de la implicación textual estricta, este tipo de implicación tiene en cuenta todo tipo de transformaciones tanto sintácticas como semánticas, teniendo como única restricción que la hipótesis sea inferida a partir del texto. El sistema deriva las formas lógicas asociadas, y establece un factor de similitud semántica entre ellas para detectar implicación textual. Se implementan tres enfoques para obtener dicho factor. De entre todos, es el enfoque que hace uso de la medida de similitud de Lin el que consigue mejores resultados. Sin embargo, los tres detectan mejor la implicación textual cuando el par de textos pertenecen a aplicaciones de Búsquedas de Respuestas (ver Cuadro 1).

Otra de las aportaciones de este artículo es la evaluación del impacto de los diferentes tipos de relaciones de WordNet sobre la tarea de implicación textual. Tras evaluar nuestros experimentos concluimos que el uso de un grupo reducido de relaciones de WordNet (en concreto sinonimia, hiponimia e implicación) consigue una mejora con respecto al tratamiento de un mayor número de relaciones. Además, el tratamiento de un menor número de relaciones también reduce el tiempo de procesamiento del sistema. Sin embargo, evaluar nuestro enfoque que hace uso masivo de las relaciones de WordNet sobre la tarea de implicación textual, nos pareció muy

interesante ya que ha sido usado para otras aplicaciones como Búsquedas de Respuestas (Terol, Martínez-Barco, y Palomar, 2005).

Como trabajos futuros nos concentraremos en dos aspectos. El primero de ellos, consiste en mejorar nuestro sistema analizando con mayor detalle el árbol sintáctico y así, detectar los posibles cambios de significado de las frases. Por otro lado, también nos atrae la idea de estudiar el impacto que produciría la adición de otras herramientas de PLN sobre la tarea de implicación textual. Por ejemplo, incorporar un reconocedor de entidades podría ayudarnos a detectar la implicación textual entre dos fragmentos de texto.

Bibliografía

- Akhmatova, Elena. 2005. Textual entailment resolution via atomic propositions. En *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*, páginas 61–64, Southampton, UK.
- Bos, Johan y Katja Markert. 2005. Combining shallow and deep nlp methods for recognizing textual entailment. En *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*, páginas 65–68, Southampton, UK, April.
- Dagan, Ido y Oren Glickman. 2004. Probabilistic textual entailment: Generic applied modeling of language variability. En *PASCAL Workshop on Learning Methods for Text Understanding and Mining*, páginas 26–29, Grenoble, France, January.
- Dagan, Ido, Oren Glickman, y Bernardo Magnini. 2005. The pascal recognising textual entailment challenge. En *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*, páginas 1–8, Southampton, UK, April.
- Glickman, Oren y Ido Dagan. 2005. A probabilistic setting and lexical cooccurrence model for textual entailment. En *Proceedings of the ACL Workshop on Empirical Modeling of Semantic Equivalence and Entailment*, páginas 43–48, Ann Arbor, Michigan, June. Association for Computational Linguistics.
- Harabagiu, S., G.A. Miller, y D.I. Moldovan. 1999. Wordnet 2 - a morphologically and semantically enhanced resource. En *Proceedings of ACL-SIGLEX99: Standardizing Lexical Resources*, páginas 1–8, Maryland, June.
- Herrera, Jesús, Anselmo Peñas, y Felisa Verdejo. 2005. Textual entailment recognition based on dependency analysis and wordnet. En *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*, páginas 21–24, Southampton, UK, April.
- Kouylekov, Milen y Bernardo Magnini. 2005. Recognizing textual entailment with tree edit distance algorithms. En *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*, páginas 17–20, Southampton, UK, April.
- Lin, D. 2005. Dependency-based evaluation of minipar. En *Workshop on the Evaluation of Parsing Systems*, páginas 17–20, Southampton, UK, April.
- Lin, Dekang. 1998. An information-theoretic definition of similarity. En *ICML '98: Proceedings of the Fifteenth International Conference on Machine Learning*, páginas 296–304, San Francisco, CA, USA.
- Miller, G.A. 1990. Wordnet: An on-line lexical database. En *International Journal of Lexicography* 3, 4, páginas 235–312.
- Moldovan, Dan y Adrian Novischi. 2002. Lexical chains for question answering. En *Proceedings of the 19th international conference on Computational linguistics - Volume 1*, páginas 1–7, Taipei, Taiwan.
- Pedersen, Ted, Siddharth Patwardhan, y Jason Michelizzi. 2004. Wordnet::similarity - measuring the relatedness of concepts. En *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*, San Jose, CA, July.
- Terol, Rafael M., Patricio Martínez-Barco, y Manuel Palomar. 2005. Applying logic forms to biomedical Q-A. En *Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications*, Istanbul, Turkey, June.
- Zanzotto, Fabio Massimo, Maria Teresa Pazienza, y Marco Pennacchiotti. 2005. Discovering entailment relations using “textual entailment patterns”. En *Proceedings of the ACL Workshop on Empirical Modeling of Semantic Equivalence and Entailment*, páginas 37–42, Ann Arbor, Michigan, June.