

Desambiguación Verbal Automática: un estudio sobre el rendimiento de la información semántica argumental

Verb Sense Disambiguation: a study about the performance of argumental semantic information

José Priego García

Universitat de Barcelona
Gran Via 585, 08007, Barcelona
jpriegga@gmail.com

Irene Castellón Masalles

Universitat de Barcelona
Gran Via 585, 08007, Barcelona
icastellon@ub.edu

Resumen: Una de las tareas fundamentales para la resolución de la ambigüedad en el ámbito del Procesamiento del Lenguaje Natural es la Desambiguación Semántica Automática; especialmente la tarea específica de Desambiguación Verbal Automática (DVA). En la presente investigación se lleva a cabo una tarea experimental con la finalidad de comprobar la viabilidad de una aproximación a la DVA basada en la información semántica de los argumentos verbales. Los buenos resultados obtenidos indicarían la necesidad de tener en cuenta este tipo de información en futuras propuestas de DVA.

Palabras clave: Semántica, procesamiento del lenguaje natural, desambiguación verbal automática, aprendizaje automático

Abstract: One of the key tasks for resolving the ambiguity in the field of Natural Language Processing is Word Sense Disambiguation; especially the specific task of Verb Sense Disambiguation (VSD). In the present study an experimental task is performed in order to test the feasibility of an approach to VSD based on semantic information about verbal arguments. The good results obtained indicate the need to take into account this information in future proposals for VSD.

Keywords: Semantics, natural language processing, verb sense disambiguation, machine learning

1 Introducción

Una de las características intrínsecas del lenguaje y que se hace patente de forma diaria en la comunicación mediante cualquier lengua es su ambigüedad. Todas las lenguas del mundo poseen multitud de palabras polisémicas que pueden hacer referencia a diferentes sentidos en función del contexto de uso, y resulta imprescindible saber identificar a cuál de ellos se está haciendo referencia en cada ocasión si se quiere comprender correctamente un mensaje. Por ejemplo, ante la oración “los participantes partieron la mañana del viernes” el hablante deberá identificar que el sentido activado por el verbo “partir” en esta ocasión es “abandonar un lugar” y no “fragmentar algo”. En el campo del Procesamiento del Lenguaje Natural (PLN) la ambigüedad supone un verdadero obstáculo para el correcto desempeño de cualquier aplicación y es por ello que desde sus inicios se ha intentado emular la capacidad de desambiguación humana

mediante el diseño y desarrollo de sistemas de Desambiguación Semántica Automática (DSA).

La investigación que se presenta a continuación se encuentra organizada del modo siguiente: en primer lugar, se realizará una breve descripción de la DSA, sus características y posibles aproximaciones, centrandose especial atención al caso concreto de la Desambiguación Verbal Automática (DVA). En segundo lugar, se propondrá y evaluará una aproximación a la DVA basada en información semántica acerca de los argumentos verbales.

2 Desambiguación Semántica Automática

La DSA podría describirse como una tarea de clasificación: un programa informático (clasificador) debe asignar automáticamente a una determinada palabra uno de sus sentidos posibles (clases) en un contexto de uso concreto (Navigli, 2009; Pal y Saha, 2015). En tanto que trata de emular una capacidad cognitiva humana,

la DSA es considerada una tarea propia del campo de la Inteligencia Artificial. Con el paso de los años, la necesidad de crear técnicas de desambiguación económicas y versátiles ha llevado al surgimiento de tres principales aproximaciones a la DSA que divergen en el tipo de fuente informativa empleada: métodos basados en conocimiento que emplean bases de datos estructuradas como tesauros y lexicones computacionales (Del Corro et al., 2014), métodos supervisados que combinan corpus anotados y sistemas de aprendizaje automático (Zhong y Tou, 2010) y, finalmente, métodos no-supervisados que extraen la información mediante cálculo estadístico a partir de corpus planos (Wang et al., 2013).

2.1 Desambiguación Verbal Automática

Del mismo modo que a la hora de llevar a cabo una tarea de DSA se puede escoger si el objeto a desambiguar serán todas las palabras del texto o únicamente una de ellas, también es posible centrar el objeto de desambiguación en una única categoría morfológica como, por ejemplo, el verbo. El actual interés por la desambiguación verbal responde a dos motivos principales: los escasos resultados conseguidos hasta la fecha en cuanto a la desambiguación de esta categoría y el papel central del verbo en la estructura oracional. Comenzando por el primero de ellos, en la actualidad los mejores resultados obtenidos en una tarea de DVA se hallan en torno a un 82% de verbos correctamente desambiguados (Del Corro et al., 2014). Se trata de unos resultados que se encuentran lejos de lo deseable en este tipo de tareas al dejar todavía un considerable margen de error y que, por lo tanto, invitan a seguir investigando con el fin de mejorarlos. Esta necesidad de mejora se corrobora si se tiene en cuenta, como se señalaba anteriormente, el papel relevante del verbo dentro de la estructura oracional como núcleo del predicado y administrador de los constituyentes de la oración mediante la selección de argumentos, así como la adjudicación de funciones sintácticas y papeles temáticos. Así pues, debido a su centralidad, una mejora en la desambiguación verbal podría suponer de forma colateral una mejora en aquellas tareas de DSA en las que se pretenda desambiguar todas las palabras de un texto y, por lo tanto, en el PLN en general.

3 Información semántica de los argumentos verbales en la DVA

El grueso de las investigaciones previas en el campo de la DSA de corte general acostumbra a emplear una serie de atributos concretos; mayoritariamente información acerca de la categoría morfológica y función sintáctica de las palabras próximas al término a desambiguar. Pese a funcionar de forma correcta en tareas generales, la aplicación de estos recursos a tareas específicas de DVA no parece que haya reportado hasta la fecha unos resultados tan buenos como se podría desear (Buscaldi et al., 2006). Ante esta situación resulta evidente la necesidad de explorar nuevas fuentes de información relacionadas con el verbo que contribuyan a mejorar el rendimiento de las tareas de DVA y, en definitiva, el campo de la DSA en su conjunto. Como se señalaba en la sección anterior, una característica intrínseca del verbo que quizás no ha recibido la atención merecida hasta tiempos recientes es la estrecha relación sintáctico-semántica que este establece con sus diferentes argumentos; una relación que puede condicionar la tipología de los argumentos en función del sentido activado por el verbo en un contexto de uso determinado. Puede suponerse, pues, que el rendimiento de las tareas de DVA podría aumentar si se contemplara en ellas información relativa a esta relación verbo-argumento; especialmente si, además de información morfológica y sintáctica, se contara también con información de tipo semántico. Esta necesidad de explotar la información argumental en el campo de la DVA ya ha sido señalada por recientes investigaciones que la incluyen de diversas formas en sus diseños experimentales; ya sea mediante el uso de preferencias de selección (Ye y Baldwin, 2006), el análisis estadístico y *clustering* de parejas verbo-argumento (Wagner et al., 2009) o mediante la inclusión de información semántica acerca de los argumentos verbales en un corpus anotado (Dang y Palmer, 2005; Dligach y Palmer, 2008).

Siguiendo la tendencia iniciada por estos estudios, en la presente investigación se llevará a cabo una tarea experimental mediante la cual se pretende evaluar el rendimiento de diferentes tipos de información semántica sobre los argumentos verbales en el proceso de DSA de una serie de verbos.

3.1 Fuentes de información empleadas

Debido a que el objetivo final de esta tarea es la evaluación del rendimiento de diversos tipos de información lingüística procedentes de diferentes fuentes, se ha considerado necesario adoptar una aproximación a la DVA de tipo supervisado. Como ya se ha apuntado anteriormente, la aproximación supervisada extrae la información de corpus etiquetados con información lingüística; en este caso, las características de los argumentos de cada verbo y el sentido activado por este en cada oración. De este modo, la aproximación supervisada otorga la posibilidad de realizar diferentes combinaciones informativas según se requiera en cada experimento.

La presente investigación toma como muestra un subconjunto del corpus Sensem (Alonso et al., 2007; Fernández y Vázquez, 2014). Dicho corpus ha sido anotado automáticamente en el nivel morfológico mediante Freeling (Padró, 2011) y de forma manual en los niveles sintáctico y semántico; incluyendo este último la desambiguación léxica de los sentidos verbales y argumentos nominales tomando WordNet como referencia (Castellón et al., 2012; Gonzalez-Agirre et al., 2012). El subconjunto empleado en esta investigación cuenta con un total de 1.033 oraciones correspondientes a 12 verbos que se hallan en un rango de entre dos y cinco sentidos (véase cabecera de las tablas de resultados); lo que supone entre 65 y 100 oraciones por cada uno de los verbos escogidos.

La selección de lemas verbales se ha realizado respetando tanto la frecuencia de aparición de los diferentes sentidos en el corpus como la necesidad de tener un mínimo de ejemplos por sentido a partir de los cuales entrenar eficientemente al clasificador. Así pues, ninguno de los sentidos verbales posee menos de un 6% de los ejemplos recogidos para su lema.

Al tratarse de un corpus anotado, las diversas oraciones que lo componen contienen información relativa a diferentes niveles de análisis lingüístico. A nivel léxico, cada una ha sido etiquetada con el lema correspondiente a su verbo principal; verbo que es, además, el candidato a desambiguar. Así mismo, cada oración cuenta con el sentido específico que toma dicho verbo en ese ejemplo concreto. Siguiendo la línea de investigaciones anteriores,

cada oración recoge también información acerca de la categoría morfológica y función sintáctica de cada uno de los argumentos verbales. En lo tocante al nivel semántico, el realmente relevante en esta investigación, los argumentos de cada oración han sido etiquetados según su papel temático y según su clasificación en tres ontologías de uso común en el ámbito del PLN: la Suggested Upper Merged Ontology (Pease et al., 2002), los Supersenses de WordNet (Fellbaum, 2005) y la Top Concept Ontology (Álvez et al., 2008).

4 Metodología

Como se ha señalado anteriormente, la DSA se puede definir básicamente como una tarea de clasificación. Siguiendo la aproximación supervisada, el clasificador toma los ejemplos que constituyen el corpus de entrenamiento, establece relaciones a partir de estos mediante un proceso de aprendizaje basado en un determinado algoritmo y, finalmente, realiza una clasificación de las instancias contenidas en un corpus de evaluación en el que estas carecen de un valor para el atributo que se pretende averiguar. En el caso concreto de esta investigación, el clasificador toma como ejemplo las oraciones contenidas en el corpus anteriormente descrito y, tras abstraer durante el aprendizaje aquellas características que los definen, otorga el sentido más probable, dentro de los posibles, para cada verbo de las diferentes oraciones existentes en el corpus de evaluación. Cada uno de los experimentos se ha realizado siguiendo un proceso de *10-fold cross validation*; de modo que el resultado de cada uno de estos se deriva de la media de diez pruebas realizadas con diferentes corpus de evaluación autogenerados a partir del corpus completo.

Los diferentes experimentos se han llevado a cabo mediante la plataforma de software para la minería de datos y el aprendizaje automático Weka (Witten et al., 2011), desarrollada por la Universidad de Waikato (Nueva Zelanda). Para ello se ha tenido que adaptar la información contenida en el corpus anotado a un formato específico capaz de ser procesado por esta aplicación. Tras esta adaptación, cada una de las oraciones ha sido transformada en una instancia en forma de vector que presenta un único valor categorial para cada uno de los atributos anteriormente descritos.

| | [2] Beneficiar | [2] Defender | [2] Facilitar | [3] Detener | [3] Interpretar | [3] Superar | [4] Hallar | [4] Partir | [4] Tratar | [5] Acabar | [5] Actuar | [5] Mover | Corpus |
|-----------------|----------------|--------------|---------------|-------------|-----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| <i>Baseline</i> | 61.8 | 64.6 | 65.0 | 52.3 | 48.3 | 41.9 | 36.1 | 41.7 | 76.8 | 40.6 | 60.0 | 32.3 | 52.1 |
| MS | 97.9 | 63.4 | 66.0 | 89.5 | 75.8 | 39.5 | 59.5 | 59.3 | 97.5 | 59.3 | 72.8 | 52.3 | 70.1 |
| MS + PTs | 96.9 | 95.1 | 76.6 | 97.6 | 76.9 | 98.7 | 89.3 | 95.6 | 97.5 | 73.6 | 88.5 | 60.0 | 87.7 |
| PTs | 96.9 | 95.1 | 80.5 | 96.5 | 79.1 | 98.7 | 89.3 | 95.6 | 97.5 | 80.2 | 84.2 | 67.6 | 89.2 |

Tabla 1: Resultados obtenidos a partir de *baseline*, información morfológico-sintáctica (MS) y papeles temáticos (PTs). Para cada verbo se indica entre [] su número de sentidos. Los mejores resultados de cada verbo se encuentran resaltados en negrita.

En el caso de esta investigación se ha considerado oportuno el uso de un algoritmo basado en el principio de *support vector machines* (SVM); concretamente, el algoritmo Sequential Minimal Optimization (Platt, 1998). Debido a sus características, un algoritmo basado en vectores de soporte se encuentra limitado a realizar clasificaciones binarias; es decir, solo puede realizar una clasificación entre dos valores distintos para un mismo atributo. De este modo, en la investigación presente, el algoritmo creará un único plano durante la clasificación de verbos con dos sentidos, pero deberá crear $(n \cdot (n-1))/2$ planos distintos cuando el verbo a desambiguar presente un mayor número de sentidos (n) con tal de realizar todas las combinaciones entre los valores existentes para, posteriormente, emitir un único resultado de clasificación a partir de estas. A pesar de esta particularidad, los algoritmos basados en SVM presentan mejores resultados que otros algoritmos típicamente empleados en DSA como pueden ser los de tipo probabilístico bayesianos como Naive Bayes o los basados en instancias vecinas como k-Nearest Neighbor (Escudero, 2006; Witten et al., 2011). En comparación con estos últimos, presentan, además, una clara ventaja en tanto que no son susceptibles a posibles desviaciones causadas por diferencias en el número de instancias contenidas en cada grupo al basarse únicamente en aquellas dos que actúan como vectores de soporte.

5 Resultados

A continuación, se realizará una descripción detallada de los resultados obtenidos en esta investigación. En primer lugar, se establecerán como punto de referencia los resultados obtenidos mediante el cálculo de la *baseline*, así como los resultados de una clasificación basada

únicamente en los atributos de categoría morfológica y función sintáctica. En segundo lugar, se llevará a cabo una serie de experimentos bajo diferentes configuraciones de los atributos semánticos anteriormente descritos. En todo momento se analizarán en primer lugar los resultados obtenidos en las pruebas realizadas de forma individual para cada verbo y en segundo lugar los resultados pertenecientes a aquella prueba realizada con el corpus completo.

5.1 Baseline e información morfológica y sintáctica

La *baseline* de cada verbo se ha obtenido siguiendo el proceso de clasificación *most frequent sense*, en el cual se asigna el valor predominante para el atributo “sentido” en todas las instancias pertenecientes a cada uno de los verbos presentes en el corpus. Los cálculos se han llevado a cabo mediante ZeroR, un algoritmo que permite el cálculo de la *baseline* de forma automática y que se halla implementado en Weka. En la Tabla 1 se puede observar que, como cabría esperar, tanto los resultados individuales (entre el 32,3% y el 76,8% de instancias correctamente clasificadas) como el de la clasificación realizada con el corpus completo (52,1%) son insuficientes para una tarea de este tipo.

En cuanto a las pruebas realizadas con información morfológica y sintáctica (MS), si bien destacan dos verbos que responden excepcionalmente bien a estas informaciones (97 puntos), las pruebas individuales arrojan unos resultados también pobres alrededor de los 60 puntos (con máximos de 89,5 y mínimos de 39,5). De modo similar, la prueba realizada con el corpus completo alcanza los 70,1 puntos. El hecho de que la desambiguación de determinados verbos como “beneficiar” y

| | [2] Beneficiar | [2] Defender | [2] Facilitar | [3] Detener | [3] Interpretar | [3] Superar | [4] Hallar | [4] Partir | [4] Tratar | [5] Acabar | [5] Actuar | [5] Mover | Corpus |
|------------|----------------|--------------|---------------|-------------|-----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Ont+MS+PTs | 96.9 | 89.0 | 85.4 | 93.0 | 83.5 | 98.7 | 90.4 | 94.5 | 97.5 | 75.8 | 87.1 | 70.7 | 88.2 |
| Ont+PTs | 90.7 | 90.2 | 88.3 | 94.1 | 83.5 | 97.5 | 89.3 | 93.4 | 97.5 | 78.0 | 80.0 | 69.2 | 88.6 |
| Ont+MS | 96.9 | 60.9 | 75.7 | 86.0 | 81.3 | 33.3 | 75.5 | 62.6 | 97.5 | 61.5 | 72.8 | 58.4 | 72.0 |
| Ont | 56.7 | 58.5 | 74.7 | 47.6 | 70.3 | 37.0 | 39.3 | 54.9 | 71.9 | 37.3 | 58.5 | 44.6 | 55.6 |

Tabla 2: Resultados obtenidos en la evaluación conjunta de las tres ontologías (Ont) y su combinación con información morfológico-sintáctica (MS) y papeles temáticos (PTs).

“tratar” logre unos resultados tan elevados únicamente con este tipo de información probablemente se deba a la existencia de unas fuertes preferencias de tipo sintáctico, según las cuales cada uno de los sentidos determina claramente la tipología a la que deben ajustarse sus posibles argumentos.

Pese a que la inclusión de información morfológica y sintáctica supone un aumento del rendimiento de estas tareas, lo cierto es que la levedad de este corrobora la necesidad de explorar nuevas fuentes de información como las que se evaluarán a continuación.

5.2 Papeles temáticos

En una nueva prueba, además de la información morfológica y sintáctica de cada argumento, se ha tenido en cuenta el atributo correspondiente al papel temático (PTs) representado por cada uno de los argumentos del verbo a desambiguar. Los resultados obtenidos parecen indicar que la inclusión de los papeles temáticos incrementa de forma considerable el porcentaje de verbos correctamente desambiguados; que en esta ocasión se halla entre los 60 y 98,7 puntos en las tareas individuales y alcanza los 87,7 puntos en la tarea conjunta.

Por otro lado, resultan especialmente interesantes los resultados obtenidos en una prueba basada únicamente en información sobre los papeles temáticos; prescindiendo, por lo tanto, de información morfológica y sintáctica. Como se recoge en la Tabla 1, los resultados obtenidos en esta ocasión muestran mejoras generales respecto a las tareas anteriores de entre 3 y 7 puntos en las tareas individuales y de 1,5 puntos al emplear el corpus completo.

5.3 Clasificación semántica

Mediante la siguiente batería de pruebas se tratará de comprobar hasta qué punto la

información semántica de los diferentes argumentos verbales, expresada según la clasificación de diversas ontologías (Ont), puede ayudar a mejorar los resultados obtenidos anteriormente o si bien puede incluso llegar a sustentar una aplicación de DVA por sí sola. Los resultados de estas se muestran en la Tabla 2.

En primer lugar, se ha podido comprobar que la adición simultánea de la información contenida en cada una de las tres ontologías antes presentadas (SUMO, Supersenses de WordNet y TCO) produce resultados heterogéneos si se aplica a la configuración anterior basada en información morfológico-sintáctica y papeles temáticos. El resultado de las tareas individuales tiende a mostrar mejoras de hasta 10 puntos respecto a la configuración anterior mientras que en la tarea con el corpus completo, en cambio, la mejora es de tan solo 0,5 puntos. Como podría esperarse si se tienen en cuenta las pruebas anteriores, si se realiza la misma prueba sustrayendo esta vez la información morfológica y sintáctica puede observarse una leve mejora en los resultados (0,4 puntos con el corpus completo). Por otro lado, si se mantiene esta y se prescinde de los papeles temáticos se produce un claro descenso de los resultados (16 puntos con el corpus completo). Cabe destacar que este descenso es especialmente notable en las pruebas individuales, que llegan a perder de 2 a 65 puntos respecto a la tarea inicial. Finalmente, si se trata de llevar a cabo una tarea de desambiguación basada únicamente en la combinación de estas tres ontologías, los resultados parecen indicar que se trata de algo inviable: en ninguno de los casos se logra mantener el resultado de pruebas anteriores y los resultados individuales sufren caídas de entre 11 y 61 puntos mientras que la prueba con el corpus completo cae en 33 puntos.

| | [2] Beneficiar | [2] Defender | [2] Facilitar | [3] Detener | [3] Interpretar | [3] Superar | [4] Hallar | [4] Partir | [4] Tratar | [5] Acabar | [5] Actuar | [5] Mover | Corpus |
|---------------|----------------|--------------|---------------|-------------|-----------------|-------------|-------------|-------------|------------|-------------|-------------|-------------|-------------|
| PTS+SUMO | 96.9 | 93.9 | 83.4 | 96.5 | 83.5 | 97.5 | 88.2 | 92.3 | 97.5 | 80.2 | 84.2 | 75.3 | 89.3 |
| PTs+Ss | 96.9 | 95.1 | 83.4 | 94.1 | 87.9 | 97.5 | 85.1 | 95.6 | 97.5 | 79.1 | 82.8 | 81.5 | 90.1 |
| PTs+TCO | 96.9 | 93.9 | 83.4 | 96.5 | 84.6 | 97.5 | 91.4 | 93.4 | 97.5 | 78.0 | 82.8 | 76.9 | 90.3 |
| PTs+TCO+Clust | 96.9 | 95.1 | 88.3 | 97.6 | 87.9 | 97.5 | 91.4 | 93.4 | 97.5 | 78.0 | 80.0 | 81.5 | 90.1 |

Tabla 3: Resultados obtenidos en la evaluación individual de las ontologías SUMO, Supersenses de Wordnet (Ss) y TCO junto a los papeles temáticos (PTs). En último lugar, se indican también los resultados obtenidos al incluir a esta última los *clusters* (Clust) basados en TCO.

La heterogeneidad de los resultados obtenidos en la serie de pruebas anterior, así como la diferente naturaleza de las tres ontologías empleadas, hace que sea prudente examinar el rendimiento de cada una de las ontologías de forma individual. Así pues, se ha realizado de nuevo la serie de pruebas anterior empleando esta vez las ontologías de forma aislada y no al unísono. Los resultados proporcionados por esta serie de pruebas han permitido establecer dos conclusiones principales. En primer lugar, se mantiene la gradación de rendimiento de las configuraciones empleadas: de nuevo, la configuración basada en papeles temáticos y una determinada ontología resulta la más adecuada; seguida de su variante que incluye información morfológico-sintáctica, a continuación por la que emplea información morfológica y sintáctica pero no papeles temáticos y, finalmente, por aquella que no emplea otra información que la recogida en la ontología. La segunda conclusión que se deriva de esta serie de pruebas es que, pese a que en las tareas individuales cada verbo parece recibir un mayor beneficio por parte de una determinada ontología, los resultados tanto de la prueba con el corpus completo como del conjunto de las tareas individuales parecen establecer una jerarquía de rendimiento entre las diferentes ontologías empleadas: TCO supera ligeramente a los Supersenses de WordNet en la mayoría de ocasiones y estos dos superan con claridad a SUMO. La diferencia de rendimiento entre las diferentes ontologías puede observarse en los resultados representados en la Tabla 3.

5.4 Aumentando el rendimiento de TCO

Los resultados obtenidos hasta el momento dejan patente que la inclusión de TCO produce un mayor rendimiento que los Supersenses de

WordNet y, especialmente, SUMO. Sin embargo, como se apuntaba anteriormente en la descripción de la metodología de esta investigación, el hecho de que el formato requerido por Weka admita un único valor para cada atributo limita en gran medida la utilidad de TCO si se tiene en cuenta su naturaleza jerárquica y componencial. De este modo, la representación compactada de los rasgos de TCO lleva al clasificador a entender como valores diametralmente distintos todos aquellos que se diferencien por un solo rasgo aun compartiendo parte o la totalidad de los demás.

Esta situación lleva a preguntarse si sería posible generalizar de algún modo la información sobre los argumentos verbales expresada según esta ontología, de forma que se pudiera establecer una relación entre los diferentes argumentos encontrados para cada sentido de un determinado verbo y, de este modo, caracterizar las preferencias de selección de cada sentido verbal. Para ello se ha optado por realizar la descomposición de la información aportada por TCO en sus categorías semánticas individuales y, posteriormente, llevar a cabo un proceso de *clustering* en tantos grupos como sentidos posee un determinado verbo. El proceso de *clustering* se ha llevado a cabo mediante el algoritmo Simple-K-Means implementado en Weka.

Al realizar una nueva prueba de clasificación donde además de papeles temáticos y TCO se han incluido dichos *clusters* (Clust), se ha observado que los resultados no solo se mantienen respecto a la prueba anterior, sino que además, como se muestra en la Tabla 3, pueden mejorar entre 1 y 5 puntos en las pruebas individuales de ciertos verbos. En la prueba realizada con el corpus completo, en cambio, se ha observado una pérdida de 0,2 puntos. Pese a

este ligero descenso, parece ser que la inclusión de información producto del *clustering* permite salvar, al menos en parte, las limitaciones impuestas por el formato establecido por el clasificador y ayuda a mejorar los resultados obtenidos hasta el momento.

5.5 Discusión de los resultados

Los resultados obtenidos en las pruebas anteriores permiten realizar una serie de consideraciones sobre la utilidad de las diferentes informaciones empleadas. En primer lugar, se ha podido comprobar que el uso exclusivo de información morfológica y sintáctica acerca de los argumentos verbales mejora los resultados proporcionados por la *baseline* pero, a pesar de ello, resulta insuficiente para llevar a cabo una tarea de desambiguación con resultados aceptables. En segundo lugar, se ha comprobado con claridad que la adición de información relativa a los papeles temáticos desempeñados por cada argumento incrementa de forma notable el rendimiento de estas tareas. Es más, de hecho, se ha podido observar que la información sobre papeles temáticos puede ofrecer mejores resultados si prescinde de información morfológica y sintáctica; probablemente debido a que dichos papeles temáticos ya suponen de forma implícita una representación de las características morfológicas y sintácticas de cada argumento verbal. En tercer lugar, se ha podido comprobar también cómo el uso de la información proporcionada por ciertas ontologías mejora la desambiguación de la mayoría de los verbos examinados. En todo caso, los mejores resultados se han obtenido con el apoyo de los papeles temáticos, prescindiendo de la información morfológica y sintáctica, y empleando únicamente una de las tres ontologías señaladas. En cuanto a las diferencias de rendimiento entre estas, se ha podido observar que SUMO se ve claramente superada por los Supersenses y TCO; siendo esta última la más fructífera de las tres. Esto probablemente se deba al hecho de que las 2.302 categorías que conforman SUMO la hacen una ontología demasiado específica; cosa que impediría establecer de forma adecuada generalizaciones a partir de diferentes argumentos relacionados con un mismo sentido verbal. Por otro lado, esto no parece suceder con recursos como los Supersenses o TCO, que con un menor número de categorías (25 y 64, respectivamente) reflejan un mayor grado de abstracción y, por lo tanto, sí

permitirían relacionar argumentos diversos. En relación con esta última ontología, cabe destacar que los resultados obtenidos a partir de la adición de información basada en el *clustering* de sus categorías parecen ser el indicador de una posible vía para aumentar su contribución en tareas de desambiguación.

6 Conclusión

En este estudio se ha propuesto una aproximación a la DVA basada en la información semántica de los argumentos verbales representada en forma de los papeles temáticos que interpretan y la clasificación de dichos argumentos según diversas ontologías. Los buenos resultados obtenidos (90% para el corpus completo empleando papeles temáticos y TCO) corroboran la necesidad de tener en cuenta estas informaciones en futuros sistemas de DVA. De cara a validar estos resultados, y como trabajo futuro, sería necesario realizar de nuevo los diferentes experimentos llevados a cabo en este estudio con un corpus de mayor tamaño; no solo en número de verbos sino también en número de sentidos y ejemplos por verbo. Así mismo, resultaría interesante realizar un contraste con otras fuentes de información como WordNet y, posteriormente, comparar el rendimiento de esta aproximación con el de otras propuestas vigentes.

Agradecimientos

Esta investigación se ha llevado a cabo gracias al proyecto ReTeLe (TIN 2015-68955-REDT).

Bibliografía

- Alonso, L., J.A. Capilla, I. Castellón, A. Fernández y G. Vázquez. 2007. The Sensem project: Syntactico-semantic annotation of sentences in Spanish. En *Recent Advances in Natural Language Processing IV. Selected papers from RANLP 2005*, páginas 89-98, John Benjamins Publishing Co, USA.
- Álvez, J., J. Atserias, J. Carrera, S. Climent, A. Oliver y G. Rigau. 2008. Consistent annotation of EuroWordNet with the Top Concept Ontology. En *Proceedings of the 4th Global WordNet Conference*, University of Szeged, Hungría.
- Buscaldi, D., P. Rosso, F. Pla, E. Segarra y E.S. Arnal. 2006. Verb sense disambiguation using support vector machines: Impact of

- WordNet-extracted features. En *Computational Linguistics and Intelligent Text Processing*, páginas 192-195, Springer, USA.
- Castellón, I., S. Climent, M. Coll-Florit, M. Lloberes y G. Rigau. 2012. Constitución de un corpus de semántica verbal del español: Metodología de anotación de núcleos argumentales. En *Revista de Lingüística Teórica y Aplicada*, número 50, páginas 13-38, Universidad de Concepción, Chile.
- Dang, H.T. y M. Palmer. 2005. The role of semantic roles in disambiguating verb senses. En *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, páginas 42-49, Association for Computational Linguistics, USA.
- Del Corro, L., R. Gemulla y G. Weikum. 2014. Werdy: Recognition and disambiguation of verbs and verb phrases with syntactic and semantic pruning. En *Conference on Empirical Methods in Natural Language Processing*, páginas 374-385, Association for Computational Linguistics, USA.
- Dligach, D. y M. Palmer. 2008. Novel semantic features for verb sense disambiguation. En *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers*, páginas 29-32, Association for Computational Linguistics, USA.
- Escudero, G. 2006. Machine Learning Techniques for Word Sense Disambiguation (tesis), Universitat Politècnica de Catalunya, España.
- Fellbaum, C. 2005. WordNet and Wordnets. En *Encyclopedia of Language and Linguistics*, 2ª edición, páginas 665-670, Elsevier, U.K.
- Fernández-Montraveta, A. y G. Vázquez. 2014. The SenSem Corpus: An annotated corpus for Spanish and Catalan with information about aspectuality, modality, polarity and factuality. En *Corpus Linguistics and Linguistic Theory*, volumen 10, páginas 273-288, De Gruyter, Alemania.
- Gonzalez-Agirre A., E. Laparra y G. Rigau. 2012. Multilingual Central Repository version 3.0. En *8th international conference on Language Resources and Evaluation*, páginas 2525-2529, Turquía.
- Navigli, R. 2009. Word sense disambiguation: A survey. En *ACM Computing Surveys*, volumen 41, ACM, USA.
- Padró, L. 2011. Analizadores Multilingües en Freeling. En *Linguamatica*, volumen 3, número 1, páginas 13-20, Portugal.
- Pal, A. y D. Saha. 2015. Word sense disambiguation: A survey. En *International Journal of Control Theory and Computer Modeling*, volumen 5, número 3, AIRCC, India.
- Pease, A., I. Niles y J. Li. 2002. The Suggested Upper Merged Ontology: A large ontology for the semantic web and its applications. En *Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web*, volumen 28, páginas 7-10, Association for the Advancement of Artificial Intelligence, USA.
- Platt, J. 1998. Fast training of support vector machines using Sequential Minimal Optimization. En *Advances in Kernel Methods - Support Vector Learning*, The MIT Press, USA.
- Wagner, W., H. Schmid y S.S. Im Walde. 2009. Verb sense disambiguation using a predicate-argument-clustering model. En *Proceedings of the CogSci Workshop on Distributional Semantics beyond Concrete Concepts*, páginas 23-28, Países Bajos.
- Wang, X., W. Zuo y Y. Wang. 2013. A novel approach to word sense disambiguation based on topical and semantic association. En *The Scientific World Journal*, Hindawi Publishing Co, UK.
- Witten, I., E. Frank y M.A. Hall. 2011. Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, USA.
- Ye, P. y T. Baldwin. 2006. Verb sense disambiguation using selectional preferences extracted with a state-of-the-art semantic role labeler. En *Proceedings of the Australasian Language Technology Workshop*, páginas 139-148, Australasian Language Technology Association, Australia.
- Zhong, Z. y H. Tou Ng. 2010. It Makes Sense: A wide-coverage word sense disambiguation system for free text. En *48th Annual Meeting of the Association for Computational Linguistics*, páginas 78-83, Association for Computational Linguistics, USA.