USO DEL SISTEMA UMCC_DLSI PARA DETERMINAR LA SIMILITUD TEXTUAL SEMÁNTICA EN SEMEVAL-2013 USANDO UNA ÓPTICA MULTIDIMENSIONAL CON ENFOQUE LÉXICO-SEMÁNTICO.

Ing. Alexander Chávez López[1], Ing. Héctor Dávila Díaz[1], Ing. Armando Collazo Amable[1]

1. Universidad de Matanzas "Camilo Cienfuegos", Vía Blanca Km.3, Matanzas, Cuba.

CD de Monografías 2013

Resumen.

En los últimos años se ha visto mucho interés en el crecimiento de la investigación para determinar cuándo oraciones o frases "significan lo mismo". El presente artículo describe las especificaciones y resultados del sistema UMCC_DLSI que participó en la tarea Similitud Textual Semántica (STS) de SemEval-2013. El sistema supervisado usa tipos diferentes de atributos léxicos y semánticos para entrenar con un clasificador de *Bagging* que genere un modelo para decidir un resultado correcto. Relacionado a los atributos diferentes se puede resaltar el recurso ISR-WN usado para extraer las relaciones semánticas entre las palabras y el uso de algoritmos diferentes que establecen las similitudes semánticas y léxicas. Para establecer qué atributos son los más apropiados se participó en tres corridas. La mejor corrida alcanzó la posición 44 en la clasificación jerárquica oficial, obteniendo un coeficiente de la correlación general de 0.61.

***Palabras claves**: Similitud Textual Semántica; Procesamiento del Lenguaje Natural; Aprendizaje Automático.*

Introducción

SemEval-2013 (Agirre, et al., 2013) nuevamente presenta la tarea de Similitud Textual Semántica (STS). En STS, los sistemas participantes deben examinar el grado de equivalencia semántica entre dos frases. La meta de esta tarea es crear un armazón unificado para la evaluación de módulos de similitud textuales semánticos y caracterizar su impacto en las aplicaciones de Procesamiento del Lenguaje Natural (PLN).

STS se relaciona a Implicación Textual (*Textual Entailment, TE,* por sus siglas en inglés) y Paráfrasis. La diferencia principal es que STS asume un grado de equivalencia bidireccional entre el par de frases.

En caso de TE, la equivalencia es direccional (por ejemplo un estudiante es una persona, pero una persona necesariamente no es un estudiante). Además, STS difiere de TE y Paráfrasis, pues en lugar de ser una decisión de si o no (*yes/no*) binaria, STS es un grado de similitud (por ejemplo un estudiante es más similar a una persona que un perro a una persona).

El grado se similitud bidireccional planteado en las tareas de SemEval-2013 es útil para tareas de PLN como la Traducción automática (MT), Extracción de Información (IE), Sistemas de Pregunta y Respuesta (QA), y Resumen Automático. Podrían agregarse varias tareas semánticas como los módulos en el armazón de STS, como la Desambiguación e Inducción, Substitución Léxica, el Rol de Etiquetado Semántico, el descubrimiento de Multi-palabras, detección de fecha y hora en textos, Detección de Entidades, entre otros.[1]

---

[1] http://www.cs.york.ac.uk/semeval-2012/task6/

Esta edición de SemEval-2013 permanece con los mismos acercamientos de clasificación de la primera versión en 2012. Los resultados de los sistemas diferentes se compararon con las puntuaciones de referencia proporcionadas por el *Gold Standing* de SemEval-2013 con el rango de cinco a cero según los siguientes criterios[2]: (5) Las dos frases son equivalentes, significan lo mismo. (4) Las dos frases son principalmente equivalentes, pero difieren en pocos detalles que son insignificantes. (3) Las dos frases son aproximadamente equivalentes, pero alguna diferencia de información importante. (2) Las dos frases no son equivalentes, pero comparten algunos detalles. (1) Las dos frases no son equivalentes, pero están en el mismo tema. (0) Las dos frases están en temas diferentes.

Existen innumerables literaturas dedicadas a medir la similitud entre frases. Quizás el más reciente escenario lo constituye la competición de SemEval-2012, específicamente su tarea 6: *A Pilot on Semantic Textual Similarity* (Aguirre y Cerd, 2012). En SemEval-2012 se usaron diferentes herramientas y recursos como la lista de *stop words*, corpus multilingües, diccionarios, siglas, y tablas de paráfrasis, pero WordNet fue el recurso más usado, seguido de los corpus monolingües y Wikipedia. (Aguirre y Cerd, 2012).

Según (Aguirre y Cerd, 2012) se usaron las herramientas de PLN de diversas formas. Entre las herramientas usadas se destacan las destinadas a la lematización y el etiquetado sintáctico (*POS- Tagging*). En menor escala fueron utilizadas las destinadas a la desambiguación, etiquetado de roles semánticos y reconocimiento de expresiones de fecha y hora. Además, fueron grandemente usados métodos basados en conocimiento y métodos distribuidos. En (Aguirre y Cerd, 2012) se dice que se usaron alineamientos y/o traducción automática estadística, substitución léxica, implicación textual y software de evaluación de traducción automática en menor grado. Puede notarse que se usó ampliamente el aprendizaje automático para generar modelos capaces de evaluar resultados de similitud textual.

Uno de los sistemas mejor ubicados en SemEval-2012 (Šarić, Glavaš et al., 2012) tiende a usar la mayoría de los recursos mencionados y herramientas. En este se predicen las evaluaciones humanas de similitud de la frase usando un modelo de regresión de soporte vectorial con múltiples atributos midiendo el solapamiento de palabras y la similitud de la sintaxis de las frases. También computan la similitud entre frases usando la alineación semántica de lemas. Primero, computan la similitud de las palabras entre todos los pares de lemas de la primera a la segunda frase usando el método basado en conocimiento o la similitud semántica basada en corpus. Nombraron este método Alineamiento Goloso de Lemas Solapados.

Desarrollo

Como se muestra en la Figura 1, las tres corridas de UMCC_DLSI empiezan con el pre-procesamiento del conjunto de entrenamiento de SemEval-2013. Cada par de frases es tokenizada, lematizada y etiquetada sintácticamente usando la herramienta Freeling 2.2 (Atserias, Casas et al., 2006). Después, varios métodos y algoritmos son aplicados para

[2] http://www.cs.york.ac.uk/semeval-2012/task6/data/uploads/datasets/train-readme.txt

extraer todos los atributos para el Sistema de Aprendizaje de Máquina (MLS) o Sistema de Aprendizaje Automático. Cada corrida usa un grupo particular de atributos.

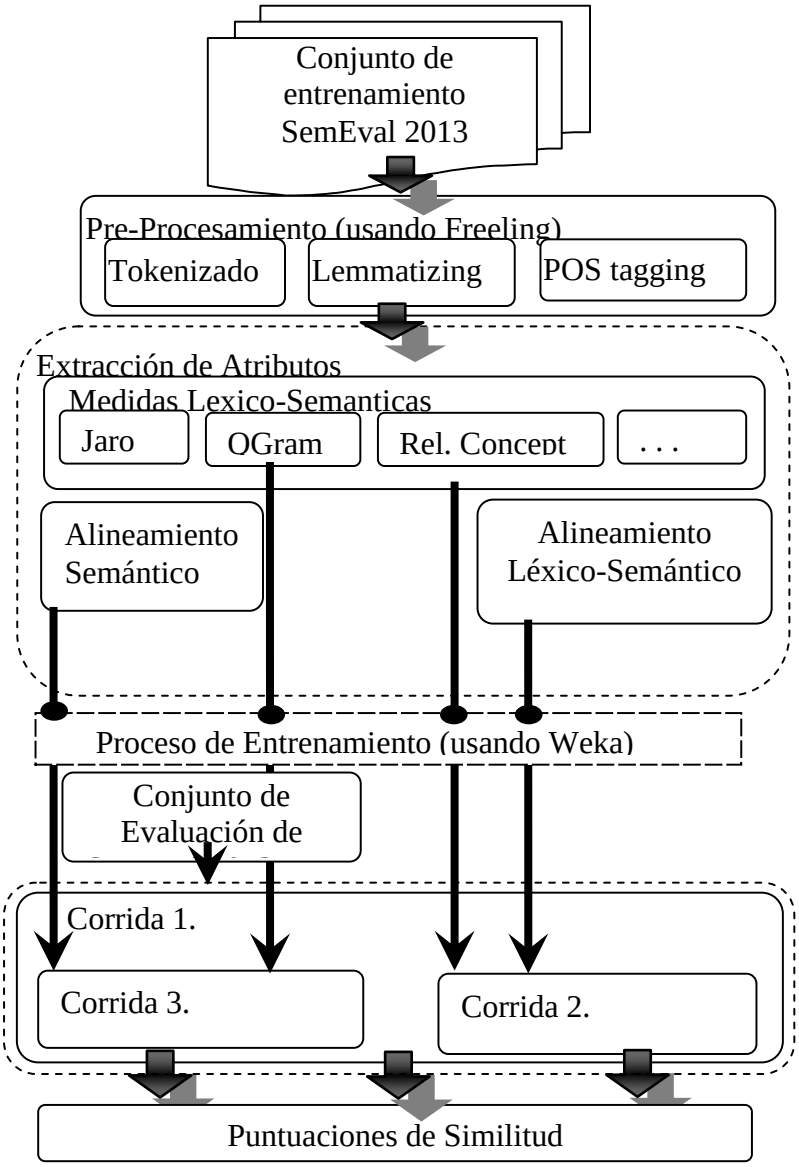

Figura 1: Arquitectura del Sistema.

La Corrida 1, nombrada *MultiSemLex*, es la corrida principal. Tiene en cuenta todos los atributos extraidos y entrena un modelo con un clasificador de *Bagging*, usando *REPTree*. El corpus de entrenamiento fue proporcionado por la competición de SemEval-2013, en concreto por la tarea de Similitud Textual Semántica.

La Corrida 2, nombrada *MultiLex* y la Corrida 3, nombrada *MultiSem*, use el mismo clasificador, pero incluyendo atributos diferentes.

Muchas veces cuando dos frases son muy similares están lexicalmente en un grado alto de solapamiento una de la otra. Inspirado en este hecho, se desarrollaron varios algoritmos que miden el nivel de solapamiento para calcular la cantidad de emparejamiento de palabras en un par de frases. En el sistema, se usan medidas de similitud léxicas y semánticas. Se extrajeron otros atributos de un alineamiento léxico-semántico de las frases y se usó una variante de alineación semántica.

Fueron usadas por el sistema un conjunto de medidas de similitud textual y distancias léxicas como: *Needleman-Wunch, Smith-Waterman, Smith-Waterman-Gotoh, Smith-Waterman-Gotoh-Windowed-Affine, Jaro, Jaro-Winkler, Chapman-Length-Deviation, Chapman-Mean-Length, QGram-Distance, Block-Distance, Cosine Similarity, Dice Similarity, Euclidean Distance, Jaccard Similarity, Matching Coefficient, Monge-Elkan y Overlap-Coefficient.* Estos algoritmos se han obtenido de una API (*Application Program Interface*) de la biblioteca de *SimMetrics* v1.5 para .NET 2.0. *Copyright (c) 2006 by Chris Parkinson*, disponible en http://sourceforge.net/projects/simmetrics/. Se obtuvieron 17 atributos para el Sistema de Aprendizaje Automático de estas medidas de similitud.

Usando la distancia de Levenshtein (LED) (Levenshtein, 1965), se computaron también dos algoritmos diferentes para obtener el alineamiento de las frases. En el primero, se consideró un valor del alineamiento como LED entre dos frases. Contrariamente a (Tatu et al., 2006), no se quitaron las puntuaciones ni las *Stop Words* de las frases, ni se consideraron costos diferentes para el funcionamiento de la transformación, se usaron todas las operaciones (la eliminación, inserción y substitución). La segunda variante se nombró Distancia de Levenshtein Doble (DLED). Para este algoritmo, se usó LED para medir la distancia entre las frases, pero para comparar las palabras, se usa nuevamente LED (Fernández et al., 2009; Fernández et al., 2012).

Otra distancia usada es una extensión de LED nombrada Distancia Extendida (EDx) (Fernández et al., 2009; Fernández et al., 2012). Este algoritmo es una extensión del algoritmo de Levenshtein donde las penalizaciones son aplicadas considerando el tipo de transformación (la inserción, eliminación, substitución, o ninguna operación) y la posición donde fue llevada a cabo, junto con el carácter involucrado en la operación. Además de las matrices de costo usadas por el algoritmo de Levenshtein, EDx obtiene también la Sub-secuencia Común Máxima (LCS) (Hirschberg, 1977) y otros atributos útiles para determinar la similitud entre cadenas de caracteres en una sola iteración. Merece señalar que las penalizaciones hechas en EDx la hacen buena candidata para el acercamiento del sistema de STS.

Otro atributo obtenido por el sistema es un valor que se corresponde con la suma de las distancias más pequeñas (usando *QGram-Distance*), entre las palabras o los lemas de la frase uno, con cada palabra de la frase dos. Como parte de los atributos extraídos por el sistema, también se destaca el valor de la suma de las distancias más pequeñas (usando *Levenshtein*) entre los tallos (*stems*), pedazos cortos y gruesos (*chunks*) y entidades de ambas frases.

Alineamiento Léxico-Semántico

Otro algoritmo que se creó es el Alineamiento Léxico-Semántico. En este se intentó alinear las frases por sus lemas. Si los lemas coinciden se buscan las coincidencias entre el *parts-of-speech*[3] (POS-tagging), y entonces la frase es realineada usando ambos criterios. Si las palabras no comparten el mismo POS, no se alinearán. En este punto, se tuvo en cuenta sólo una alineación léxica. De ahora en adelante, se aplicará una variante semántica. Después de todo el proceso, las palabras no-alineadas se analizarán teniendo en cuenta las relaciones de WorldNet (la sinonimia, el hiponimia, la hiperonimia, el similar a, el grupo verbal, implicación y relación de causa); y un conjunto de equivalencias como las abreviaciones de meses, países, capitales, días y monedas. En caso de la relación hiperonimia e hiponimia, las palabras van a ser alineadas si hay una palabra en la primera frase que está en la misma relación (hiperonimia o hiponimia) con otra de la segunda frase. Para las relaciones causa e implicación, las palabras se alinearán si hay una palabra en la primera frase que causa o implica a otra de la segunda frase. Todos los otros tipos de relaciones se llevarán a cabo de la manera bidireccional, es decir, hay una alineación si una palabra de la primera frase es un sinónimo de otra de la segunda frase o viceversa.

Finalmente, se obtuvo un valor al cual se le llamó la relación de alineamiento. Este valor es calculado como FAV = NAW / NWSP. Donde FAV es el valor final del alineamiento, NAW es el número de palabras alineadas, y NWSP es el número de palabras de la frase más corta. El valor de FAV también es otro atributo para el sistema. Otros atributos extraídos son la cantidad de palabras alineadas y la cantidad de palabras no alineadas. El centro de la alineación se lleva a cabo de formas diferentes que se obtienen de varios atributos. Cada forma puede compararse por:

- El etiquetado sintáctico o *parts-of-speech* (POS-tagging).
- La morfología y el etiquetado sintáctico.
- El lema y el etiquetado sintáctico.
- La morfología, el etiquetado sintáctico y las relaciones de WordNet (la sinonimia, el hiponimia, etc).
- El lema, el etiquetado sintáctico y las relaciones de WordNet (la sinonimia, el hiponimia, etc.).

Alineamiento Semántico

Este método de alineación depende de calcular la similitud semántica entre las frases basándose en un análisis de las relaciones, en ISR-WN, de las palabras que componen las frases. Primero, las dos frases se pre-procesan con Freeling y las palabras son clasificadas según su POS, creando grupos diferentes. La distancia entre dos palabras será la distancia, basado en WordNet, del sentido más probable de cada palabra en el par, al contrario del sistema en SemEval 2012. En SemEval-2012 se asumió el sentido seleccionado después de aplicar un Algoritmo Hungariano Doble (Kuhn, 1955), para más detalles por favor refiérase a (Fernández, et al., 2012). La distancia se computa según la formulación siguiente:

[3] Sustantivos, verbos, adjetivos, adverbios, preposiciones, conjunciones, pronombres, determinantes, modificadores, etc.

$$d(x,y) = \sum_{i=0}^{i=m} w * r(L[i], L[i+1]); \quad (1)$$

Donde *L* es la colección de sentidos (*synsets*) que corresponde al camino mínimo entre los nodos *x* y *y*, m es la longitud de *L* sustrayendo uno, *r* es una función que busca la relación conectando los nodos *x* y *y*, *w* es un peso asociado a la relación buscada por *r* (vea Tabla 1).

Relación	Weight
Hiponimia, Hiperonimia	2
Miembro_Holónimo, Miembro_Merónimo, Causa, Implicación	5
Similar_A	10
Antónimo	200
Otras relaciones diferentes a sinonimia	60

Tabla 1: Peso de las relaciones de WordNet.

La Tabla 1 muestras los pesos asociados a las relaciones de WordNet entre dos sentidos. Véase el ejemplo siguiente:

Se toma el par 99 de corpus MSRvid (del conjunto de entrenamiento de SemEval-2013) con una pequeña transformación en el orden de hacer una buena explicación del método implementado.

Par original:
A: *A polar bear is running towards a group of walruses*.
B: *A polar bear is chasing a group of walruses*.
Par transformado:
A1: *A polar bear runs towards a group of cats*.
B1: *A wale chases a group of dogs*.

Después, usando la formulación mostrada previamente, se crea una matriz con las distancias entre todos los grupos de ambas frases (vea Tabla 2).

GRUPOS	polar	bear	runs	towards	group	cats
wale	Dist:=3	Dist:=2	Dist:=3	Dist:=5		Dist:=2
chases	Dist:=4	Dist:=3	Dist:=2	Dist:=4		Dist:=3
group					Dist:=0	
dogs	Dist:=3	Dist:=1	Dist:=4	Dist:=4		Dist:=1

Tabla 2: Distancias entre grupos.

Usando el Algoritmo Hungariano (Kuhn, 1955) para la Asignación de Costo Mínimo, cada grupo de la primera frase se verifica con cada elemento de la segunda frase, y el resto es marcado como palabras que no se alinearon.

En el ejemplo anterior las palabras "*toward*" y "*polar*" no se alinearon, entonces el número de palabras no alineadas es dos. Hay sólo un alineamiento perfecto: group-group (el alineamiento con el costo=0). La longitud de la frase más corta es cuatro. La Tabla 3 muestras los resultados de este análisis.

Número de coincidencias exactas	Distancia Total del Alineamiento Óptimo	Número de palabras no alineadas
1	5	2

Tabla 3: Atributos extraidos al analizar el par.

Este proceso tiene que ser repetido para los sustantivos (ver Tabla 4), verbos, adjetivos, adverbios, preposiciones, conjunciones, pronombres, determinantes, modificadores, dígitos y fechas. La Tabla 4 muestra rasgos extraídos del análisis de sustantivos.

GRUPOS	bear	group	cats
wale	Dist := 2		Dist := 2
group		Dist := 0	
dogs	Dist := 1		Dist := 1

Tabla 4: Distancia del grupo de sustantivos.

Número de coincidencias exactas	Distancia Total del Alineamiento Óptimo	Número de palabras no alineadas
1	3	0

Tabla 5: Atributos extraidos del análisis de sustantivos.

Se extraen varios atributos del par de frases (ver Tabla 3 y Tabla 5). Tres atributos que consideran sólo verbos, sólo sustantivos, sólo adjetivos, sólo adverbios, sólo preposiciones, sólo conjunciones, sólo pronombres, sólo determinantes, sólo modificadores, sólo modificadores, y sólo fechas. Estos atributos son:

- Número de coincidencias exactas.
- Distancia Total del Alineamiento Óptimo.
- Número de palabras no alineadas.

Para el grupo de adjetivos se agregó un atributo que indica la distancia entre los sustantivos que modifican.

Para los verbos, se buscaron los sustantivos que lo preceden, y los sustantivos que están luego del verbo, y se definieron dos grupos. Se calcula la distancia para alinear cada grupo con cada par de verbos alineados. Los verbos tienen otro atributo que especifica si todos los verbos están en el mismo tiempo verbal. Con los adverbios, se busca el verbo que se modifica por él, y se calcula la distancia con su similar alineado. Para la determinación de todas estas funciones de las palabras en las frases fue usada la herramienta Freeling.

Como resultado, se obtienen 42 atributos de este método de alineamiento. Es importante resaltar que cada una de las palabras de la primera frase se trata de alinear con cada una de las palabras de la segunda frase (ver Tabla 4).

Descripción de todos los atributos generados por el Sistema.

Del proceso de alineamiento, se extraen atributos diferentes que ayudan a dar un resultado bueno del Sistema de Aprendizaje Automático. La Tabla 6 muestra el grupo de atributos con el apoyo léxico y semántico, basado en la relación de WordNet. Cada uno se nombró

CPA_FCG, CPNA_FCG, SIM_FCG, CPA_LCG, CPNA_LCG, SIM_LCG, CPA_FCGR, CPNA_FCGR, SIM_FCGR, CPA_LCGR, CPNA_LCGR, SIM_LCGR

Tabla 6: Grupo de atributos semánticos.

con un prefijo, un guión y un sufijo. La Tabla 7 describe el significado de cada prefijo, y la Tabla 8 muestras el significado de los sufijos.

CPA	Número de palabras alineadas.
CPNA	Número de palabras no alineadas.
SIM	Similaridad.

Tabla 7: Significado de los prefijos.

Atributo	Palabras comparadas por…
FCG	Morfología y POS
LCG	Lema y POS
FCGR	Morfología, POS y relaciones de WordNet.
LCGR	Lema, POS y relaciones de WordNet.

Tabla 8: Sufijos para describir cada tipo de alineamiento.

LevForma	Distancia de Levenshtein entre las dos frases comparando las palabras por la morfología.
LevLema	Lo mismo de arriba, pero comparando las palabras por sus lemas.
LevDoble	Igual, pero comparando las palabras de Nuevo por Levenshtein y haciendolas alinear si la distancia entre las palabras es ≤ 2.
EDx	Distancia Extendida.
NormLevF, NormLevL	Normalizaciones de LevForma y LevLema.

Tabla 9: Atributos del alineamiento léxico.

NWunch, SWaterman, SWGotoh, SWGAffine, Jaro, JaroW, CLDeviation, CMLength, QGramD, BlockD, CosineS, DiceS, EuclideanD, JaccardS, MaCoef, MongeElkan, OverlapCoef.

Tabla 10: Medidas léxicas de la librería *SimMetrics*.

AxAQGD_L	Todos contra todos aplicando QGramD y comparando por el lemas de las palabras.
AxAQGD_F	Lo mismo de arriba, pero aplicando QGramD y comparando por la

	morfología.
AxAQGD_LF	Igual, no sólo comparando por el lema sino también por la morfología.
AxALev_LF	Todos contra todos aplicando Levenhstein y comparando por la morfología y lemas.
AxA_Stems	Igual, pero aplicando Levenhstein y comparando por los *stems* de las palabras.

Tabla 11: Alineamiento todos contra todos.

Descripción de la fase de entrenamiento.

Para el proceso de entrenamiento, se usó un sistema de aprendizaje automático supervisado, incluyendo todo el conjunto de entrenamiento de SemEval-2013 como corpus de entrenamiento para el sistema.

Se usó una validación cruzada con paquetes de 10 y un clasificador de Bagging con un REPTree en su interior (experimentalmente seleccionado).

Como se puede ver en la Tabla 12 los atributos correspondientes a la Prueba 1 (solo atributos léxicos) obtuvieron un 0.7534 de correlación. Por otro lado, los atributos de la Prueba 2 (los atributos léxicos con apoyo semántico) obtuvieron 0.7549 de correlación y todos los atributos 0.7987, demostrándose la necesidad de atacar el problema de la similitud desde un punto de vista bidimensional.

Atributos	Correlación en los datos de entrenamiento de SemEval-2013		
	Prueba 1	Prueba 2	Prueba 3
Tabla 6		0.7549	0.7987
Tabla 9			
Tabla 10	0.7534		
Tabla 11			

Tabla 12: Influencia de los atributos. Las celdas grises significan que no se tuvieron en cuenta.

Resultados y discusión.

La tarea de Similitud Textual semántica de SemEval-2013 ofreció dos medidas oficiales para evaluar los sistemas[4]: *Mean* - el valor de la evaluación principal, *Rank* - el *ranking* de la subida ordenada por el resultado de *Mean*.

Los conjuntos de datos de evaluación a los que fueron sometidos los sistemas en SemEval-2013 se describen a continuación:

Headlines:	Los titulares de las noticias de varias fuentes de noticias de *European Media Monitor* usando alimentación por RSS.
OnWN:	mapeado de recursos léxicos *OnWN*. Las frases son las definiciones de sentidos de *WordNet* y *OntoNotes*.

[4] http://ixa2.si.ehu.es/sts/index.php?option=com_content&view=article&id=53&Itemid=61

FNWN:	Las frases son las definiciones de sentidos de *WordNet* y *FrameNet*.
SMT:	El conjunto de datos *SMT* viene de *DARPA GALE HTER* e *HyTER*. Una frase es una salida de MT y la otra es una traducción de la referencia dónde la referencia se genera basado en la corrección humana.

Usando estas medidas, la segunda corrida (Corrida 2) obtuvo los mejores resultados. Como se puede ver en la Tabla 14, la corrida léxica ha obtenido el mejor resultado. Esto demuestra que atacando el problema combinando múltiples medidas de similitud léxicas produce buenos resultados en concordancia a corpus de prueba específicos.

Para explicar la Tabla 14 se presentan las siguientes descripciones: Titulo encima de las filas significa: 1 - *Headlines*, 2 - *OnWN*, 3 - *FNWN*, 4 - *SMT* y 5 - *mean*.

Corrida	**1**	**R**	**2**	**R**	**3**	**R**	**4**	**R**	**5**	**R**
1	0.5841	60	0.4847	54	0.2917	52	0.2855	66	0.4352	58
2	0.6168	55	0.5557	39	0.3045	50	0.3407	28	0.4833	44
3	0.3846	85	0.1342	88	-0.0065	85	0.2736	72	0.2523	87

Tabla 14: Resultados oficiales de los conjuntos de prueba de SemEval-2013. *Ranking* (R).

La Corrida 1 es la corrida principal que contiene la unión de todos los atributos (los atributos léxicos y semánticos). La Tabla 14 muestras los resultados de todas las corridas para los diferentes corpus de la fase de prueba. Como se puede ver, la Corrida 1 no obtuvo los mejores resultados entre todas las corridas. Por otra parte, la Corrida 3 usa un análisis más semántico que la Corrida 2. La Corrida 3 debe mejorar los resultados en el corpus de *FNWN*, porque este corpus se extrae de *FrameNet* (Panadero, et al., 1998) (una red semántica). *FNWN* proporciona más volumen semántico que léxico. La Corrida 3 obtuvo un coeficiente de correlación de 0.8137 para todo el corpus de entrenamiento de SemEval-2013, mientras que la Corrida 2 y la Corrida 1 obtuvieron 0.7976 y 0.8345 respectivamente con el mismo clasificador (*Bagging* usando *REPTree*, y validación cruzada con 10 paquetes). Estos resultados presentan una contradicción entre la prueba y evaluación del entrenamiento; probablemente como consecuencia de algunos obstáculos presentes en los corpus de prueba, por ejemplo:

- En el corpus de *Headlines* hay gran cantidad de entidades, siglas y gentilicios que no se tienen en cuenta en el sistema presentada.

- El corpus *FNWN* no presenta equilibrio en cuanto a la longitud de las frases.

- En el corpus de prueba *OnWN*, se piensa que algunas evaluaciones no son adecuadas en correspondencia con el corpus de entrenamiento. Por ejemplo, en la línea 7, el resultado propuesto era 0.6, sin embargo ambas frases son semánticamente similares. Las frases son:

    - the act of lifting something

    - the act of climbing something

Se piensa que 0.6 no es una evaluación correcta para este ejemplo. El resultado del sistema, para este caso particular, fue 4.794 en la Corrida 3 y 3.814 en la Corrida 2, finalmente 3.695 en la Corrida 1.

Conclusiones y trabajos futuros.

Este artículo ha introducido un nuevo armazón para reconocer Similitud Textual Semántica que depende de la extracción de varios atributos que pueden inferirse de una interpretación convencional de un texto. Se han dirigido tres corridas diferentes, estas corridas sólo difieren en el tipo de atributos usados. Se puede ver en la Tabla 14 que todas las corridas obtuvieron resultados alentadores. La mejor corrida se situó en la posición 44 de 89 corridas de la clasificación jerárquica de SemEval-2013. La Tabla 12 y la Tabla 14 muestran las posiciones alcanzadas para las tres corridas y la clasificación jerárquica según el resto de los equipos. Se ha dirigido la extracción de los atributos semánticos en un contexto multidimensional que usa el recurso ISR-WN, el que permitió navegar por varios recursos semánticos (*WordNet, WordNet Domains, WordNet Affect, SUMO, SentiWorNet y Semantic Classes*).

Finalmente, se puede decir que el sistema tuvo un desempeño y resultado satisfactorios. En el trabajo actual, se muestra que este acercamiento puede usarse para clasificar varios ejemplos correctamente del STS de SemEval-2013. Comparado con la mejor corrida de la clasificación jerárquica (*UMBC_EBIQUITY - ParingWords*) (ver Tabla 15) la corrida principal tiene los resultados muy cerrados en los *headlines* (1), y *SMT* (4) del conjunto de datos de prueba.

Corrida	1	2	3	4	5	6
(Mejor ubicada de la competencia)	0.7642	0.7529	0.5818	0.3804	0.6181	1
(Nuestra) RUN 2	0.6168	0.5557	0.3045	0.3407	0.4833	44

Tabla 15: La mejor corrida de SemEval-2013.

Como trabajo futuro se planea enriquecer el método de alineamiento semántico con *Extended WordNet* (Moldovan y Rus, 2001), con esta mejora se pretende poder aumentar los resultados obtenidos en corpus con las características de *OnWN*.

Bibliografía.

AGUIRRE, E., D. CER, et al. (2013). *SEM 2013 Shared Task: Semantic Textual Similarity including a Pilot on Typed-Similarity. *SEM 2013: The Second Joint Conference on Lexical and Computational Semantics, Association for Computational Linguistics.

AGUIRRE, E. and D. CERD (2012). SemEval 2012 Task 6: A Pilot on Semantic Textual Similarity. First Join Conference on Lexical and Computational Semantic (*SEM), Montréal, Canada, Association for Computational Linguistics.

ATSERIAS, J., B. CASAS, et al. (2006). FreeLing 1.3: Syntactic and semantic services in an open source NLP library. Proceedings of LREC'06, Genoa, Italy.

FERNANDEZ, A., Y. GUTIERREZ, et al. (2012). UMCC_DLSI: Multidimensional Lexical-Semantic Textual Similarity. {*SEM 2012}: The First Joint Conference on Lexical and Computational Semantics -- Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation {(SemEval 2012)}, Montreal, Canada, Association for Computational Linguistics.

FERNANDEZ, A. C., D. B. JOSVAL, et al. (2009). Un algoritmo para la extracción de características lexicográficas en la comparación de palabras. IV Convención Científica Internacional CIUM, Matanzas, Cuba.

GUTIERREZ, Y., A. FERNANDEZ, et al. (2010a). Integration of semantic resources based on WordNet. XXVI Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural, Universidad Politécnica de Valencia, Valencia, SEPLN 2010.

GUTIERREZ, Y., A. FERNANDEZ, et al. (2010b). UMCC-DLSI: Integrative resource for disambiguation task. Proceedings of the 5th International Workshop on Semantic Evaluation, Uppsala, Sweden, Association for Computational Linguistics.

GUTIERREZ, Y., A. FERNANDEZ, et al. (2011). Enriching the Integration of Semantic Resources based on WordNet. Procesamiento del Lenguaje Natural 47: 249-257.

HIRSCHBERG, D. S. (1977). Algorithms for the longest common subsequence problem. J. ACM 24: 664–675.

KUHN, H. W. (1955). The Hungarian Method for the assignment problem. Naval Research Logistics Quarterly 2: 83–97.

LEVENSHTIN, V. I. (1965). Binary codes capable of correcting spurious insertions and deletions of ones. Problems of information Transmission.

MILLER, G. A., R. BECKWITH, et al. (1990a). Five papers on WordNet. Princenton University, Cognositive Science Laboratory.

MILLER, G. A., R. BECKWITH, et al. (1990b). Introduction to WordNet: An On-line Lexical Database. International Journal of Lexicography, 3(4):235-244.

MOLDOVAN, D. I. and V. RUS (2001). Explaining Answers with Extended WordNet. ACL.

ŠARIC, F., G. GLAVAS, et al. (2012). TakeLab: Systems for Measuring Semantic Text Similarity. Montréal, Canada, First Join Conference on Lexical and Computational Semantic (*SEM), pages 385-393. Association for Computational Linguistics.

TATU, M., B. ILES, et al. (2006). COGEX at the Second Recognizing Textual Entailment Challenge. Proceedings of the Second PASCAL Recognizing Textual Entailment Challenge Workshop, Venice, Italy.

WINKLER, W. (1999). The state of record linkage and current research problems. Technical Report, Statistical Research Division, U.S, Census Bureau.