

Sentiment Classification Using Semantic Features Extracted from WordNet-based Resources

Yoan Gutiérrez

Department of Informatics
University of Matanzas, Cuba.
{yoan.gutierrez}@umcc.cu

Sonia Vázquez and Andrés Montoyo
Department of Software and Computing
Systems

University of Alicante, Spain.
{svazquez, montoyo}@dlsi.ua.es

Abstract

In this paper, we concentrate on the 3 of the tracks proposed in the NTCIR 8 MOAT, concerning the classification of sentences according to their opinionatedness, relevance and polarity. We propose a method for the detection of opinions, relevance, and polarity classification, based on ISR-WN (a resource for the multidimensional analysis with Relevant Semantic Trees of sentences using different WordNet-based information sources). Based on the results obtained, we can conclude that the resource and methods we propose are appropriate for the task, reaching the level of state-of-the-art approaches.

1 Introduction

In recent years, textual information has become one of the most important sources of knowledge to extract useful and heterogeneous data. Texts can provide from factual information such as descriptions, lists of characteristics or instructions to opinionated information such as reviews, emotions or feelings. This heterogeneity has motivated that dealing with the identification and extraction of opinions and sentiments in texts require special attention. In fact, the development of different tools to help government information analysts, companies, political parties, economists, etc to automatically get feelings from news and forums is a challenging task (Wiebe et al., 2005). Many researchers such as Balahur et al., (2010), Hatzivassiloglou et al.(2000), Kim and Hovy (2006), Wiebe et al. (2005) and many others have been working in this way and related areas.

Moreover, in the course of years we find a long tradition on developing Question Answering (QA) systems. However, in recent years, researchers have concentrated on the development of Opinion Questions Answering (OQA) systems (Balahur et al., 2010). This new task has to deal with different problems such as Sentiment Analysis where documents must be classified according to sentiments and subjectivity features. Therefore, a new kind of evaluation that takes into account this new issue is needed.

One of the competitions that establishes the benchmark for opinion question answering systems, in a monolingual and cross-lingual setting, is the NTCIR Multilingual Opinion Analysis Task (MOAT)¹. In this competition, researchers work hard to achieve better results on Opinion Analysis, introducing different techniques.

In this paper, we only concentrate on three tracks proposed in the NTCIR 8 MOAT, concerning to the classification of sentences according to their opinionatedness, relevance and polarity. We propose a method for the detection of opinions, relevance and polarity classification, based on ISR-WN which is a resource for the multidimensional analysis with Relevant Semantic Trees of sentences using different WordNet-based information sources.

2 Related works

Related to Opinion Analysis task we can find many points of view. Some researchers say that adjectives combined with semantic characteristics provide vital information to the performance of Opinion Analysis (Hatzivassiloglou et al., 2000). Others like Zubaryeva and Savoy (2010) assume

¹<http://research.nii.ac.jp/ntcir/ntcir-ws8/meeting/>

that the extraction of relevant terms on the documents could define their polarity, designing a method capable of selecting terms that clearly belong to one type of polarity. Another research based on features extraction was conducted by Lai et al. (2010), they developed a trained system on Japanese Opinionated Sentence Identification. And Balahur and Montoyo (2009) proposed a method to extract, classify and summarize opinions on products from web reviews. It was based on the prior building of product characteristics taxonomy and on the semantic relatedness given by the Normalized Google Distance (Cilibrasi and Vitányi, 2007) and SVM learning. As we can see, the usage of features extraction is a suitable mode to work on Opinion Analysis task. Apart from that other authors have used semantic resources, for example, Kim and Hovy (2006, 2005) used semantic resources to get an approach on Holder Detection and Opinion Extraction tasks.

In general, using semantic resources is one of the most applied procedures over different tasks such as Document Indexing, Document Classification, Word Sense Disambiguation, etc. In Natural Language Processing (NLP), one of the most used resources for WSD and other tasks is WordNet (WN) (Fellbaum, 1998). WN is a lexical dictionary with word senses and descriptions. In order to enrich the WN resource, it has been linked with different lexical resources such as WordNet Domains (WND) (Magnini and Cavaglia, 2000) a lexical resource containing the domains of the synsets in WordNet, SUMO (Niles, 2001) an ontology relating the concepts in WordNet, WordNet Affect (WNA) an extension of WN where different synsets are annotated with one of the six basic emotions proposed by Ekman (1999), SentiWordNet (Esuli and Sebastiani, 2006) a lexical resource where each synset is annotated with polarity, Semantic Classes (SC) (Izquierdo *et al.*, 2007) a set of Base Level Concepts (BLC) based on WN, etc. The usage of these resources allows the tackling of NLP tasks from different points of view, depending on the resource used.

Our approach proposes using different semantic dimensions according to different resources. In order to achieve this, we use the Integration of Semantic Resources based on WordNet, which we explain in the next section and the Semantic Classes (SC).

2.1 Integration of Semantic Resources based on WordNet (ISR-WN)

ISR-WN (Gutiérrez *et al.*, 2010b) is a new resource that allows the integration of several semantic resources mapped to WN. In ISR-WN, WordNet 1.6 or 2.0 is used as a core to link several resources: SUMO, WND and WNA. As Gutiérrez *et al.* (2010a) describe, the integrated resource allows navigate inside the semantic network.

2.2 Semantic Classes (SC)

The Semantic Classes resource (Izquierdo *et al.*, 2007) consists of a set of Base Level Concepts (BLC) from WN obtained before applying a bottom-up process using the chain of hypernym relations. For each synset in WN, the process selects as its Base Level Concept the first local maximum, according to the relative number of relations. As a result, a resource with a set of BLCs linked semantically to several synsets is obtained.

In order to apply the multidimensionality that ISR-WN and SC provide, we have analyzed related approaches like (Magnini *et al.*, 2002; 2008), (Vázquez *et al.*, 2004), (Villarejo *et al.*, 2005), (Zouaq *et al.*, 2009) and others that take into account semantic dimensionality. Then, we have decided to use Relevant Semantic Trees (Gutiérrez *et al.*, 2010a) because it is an approach capable of being applied over several dimensions (resources) at once.

2.3 Relevant Semantic Trees (RST)

RST (Gutiérrez *et al.*, 2010a) is a method able to disambiguate the senses of the words contained in a sentence by obtaining the Relevant Semantic Trees from different resources. In order to measure the association between concepts in each sentence according to a multidimensional perspective, *RST* uses the Association Ratio (*AR*) measure (Vázquez *et al.*, 2004). Our purpose is to include the Multidimensional Semantic Analysis into the Opinion Analysis using *RSTs*.

In order to evaluate our approach the rules and corpus that concern the English monolingual subtasks from MOAT were used.

2.4 English monolingual subtasks

In these tasks the participants were provided with twenty topics. For each one of the topics, a question was given with a short and concise query,

the expected polarity of the answer and the period of time. For each of the topics, a set of documents were assigned and they had to be splitted into sentences for the opinionated and relevance judgements and into opinion units for the polarity, opinion target and source tasks. In this work, we describe twelve runs for the opinionated, relevance and polarity judgement tasks.

3 WSD method

We propose an unsupervised knowledge-based method that uses the RST technique combined with SentiWordNet 3.0 (Esuli and Sebastiani, 2006) to tackle 3 of the monolingual English tasks proposed in the NTCIR 8 MOAT. In this approach WN 2.0 version is used.

The aim of this method is to obtain a *RST* of each sentence and then associate the *RST* with polarity values. The process involves the following resources: WND, WNA, the WN taxonomy, SUMO and Semantic Classes (SC). Because of SC does not have a tree structure we simply obtain the Relevant Semantic Classes. Subsequently, we determine the polarities collected for each label of each *RST* obtained according to the analyzed sentence. Our proposal involves four steps presented on sections 3.1, 3.2, 3.3 and 3.4.

3.1 Obtaining the Relevant Semantic Trees

In this section, we use a fragment of the original RST method with the aim of obtaining Relevant Semantic Trees of the sentences. Notice that this step must be applied for each resource.

Once each sentence is analyzed, the *AR* value is obtained and related to each concept in the trees. Equation 1 is used to measure and to obtain the values of Relevant Concepts:

$$AR(C, f) = \sum_{i=1}^n AR(C, f_i); \quad (1)$$

Where:

$$AR(C, w) = P(C, w) * \log_2 \frac{P(C, w)}{P(C)}; \quad (2)$$

In both equations *C* is a concept; *f* is a sentence or set of words (*w*); *f_i* is the *i*-th word of the sentence *f*; *P* (*C*, *w*) is the joint probability distribution; *P* (*C*) is the marginal probability.

In order to illustrate the processing steps, we will consider the following example: “*But it is unfair to dump on teachers as distinct from the*

educational establishment”. Using the WND resource, we show the manner in which we obtain the RST.

The first stage involves the lemmatization of the words in the sentence. For the example considered, the obtained lemmas are:

Lemmas [*unfair*; *dump*; *teacher*, *distinct*, *educational*; *establishment*]

Next, each lemma is looked up in ISR-WN and it is correlated with the WND concepts. Table 1 shows the results after applying Equation 1 over the example.

| Vector | | | |
|--------|----------------|------|----------------|
| AR | Domain | AR | Domain |
| 0.90 | Pedagogy | 0.36 | Commerce |
| 0.90 | Administration | 0.36 | Quality |
| 0.36 | Buildings | 0.36 | Psychoanalysis |
| 0.36 | Politics | 0.36 | Economy |
| 0.36 | Environment | | |

Table 1. Initial Concept Vector of Domains

After obtaining the Initial Concept Vector of Domains we apply Equation 3 in order to obtain the Relevant Semantic Tree related to the sentence.

$$AR(PC, f) = AR(ChC, f) - ND(IC, PC) \quad ;(3)$$

Where:

$$ND(IC, PC) = \frac{MP(IC, PC)}{TD} \quad ;(4)$$

Here *AR(PC, f)* represents the *AR* value of *PC* related to the sentence *f*; *AR(ChC, f)* is the *AR* value calculated with equation 1 in case of *ChC* was included in the Initial Vector, otherwise is calculated with the equation 3; *ChC* is the Child Concept of *PC*; *ND* is a Normalized Distance; *IC* is the Initial Concept from we have to add the ancestors; *PC* is Parent Concept; *TD* is Depth of the hierarchic tree of the resource to use; and *MP* is Minimal Path.

Applying the Equation 3, the algorithm to decide which parent concept will be added to the vector is shown here:

```
if (AR(PC, f) value > 0) {
  if ( PC had not been added to vector)
    PC is added to the vector with AR(PC, f) value;
  else PC value = PC value + AR(PC, f) value; }
```

The result after processing is shown in Table 2. This vector represents the Domain tree associated to the sentence. After the Relevant Semantic Tree is obtained, the Factotum Domain is eliminated

from the tree. Due to the fact that Factotum is a generic Domain associated to words that appear in general contexts it does not provide useful information and experimentally we confirmed that it introduced errors; so we eliminate it (Magnini and Cavaglia, 2000).

| Vector | | | |
|--------|----------------|------|--------------|
| AR | Domain | AR | Domain |
| 1.63 | Social_Science | 0.36 | Buildings |
| 0.90 | Administration | 0.36 | Commerce |
| 0.90 | Pedagogy | 0.36 | Environment |
| 0.80 | Root_Domain | 0.11 | Factotum |
| 0.36 | Psychoanalysis | 0.11 | Psychology |
| 0.36 | Economy | 0.11 | Architecture |
| 0.36 | Quality | 0.11 | Pure_Science |
| 0.36 | Politics | | |

Table 2. Final Domain Vector

3.2 Obtaining the Positive Semantic Trees

In order to obtain the Positive Semantic Trees (*PST*) of the sentence, we will follow the same process described in section 3.1. In this case, the *AR* values will be replaced by the polarity value pertaining to the analyzed sense. The polarity is obtained from the SentiWordNet 3.0 resource, where each given sense from ISR-WN for WordNet version 2.0 is mapped to WordNet version 3.0. Hence, we can find each given sense from ISR-WN in SentiWordNet 3.0 and obtain the respective polarities. This new value will be called Positive Association (*PosA*). The *PosA* value is calculated using Equation 4 .

$$PosA(C, f) = \sum_{i=1}^n PosA(C, f_i); \quad (4)$$

Where:

$$PosA(C, w) = \sum_{i=1}^n PosA(C, w_i); \quad (5)$$

Where *C* is a concept; *f* is a sentence or set of words (*w*); *f_i* is a *i*-th word of the sentence *f*; *PosA* (*C*, *w_i*) is the positive value of the sense (*w_i*) related to *C*.

The *PosA* is used to measure the positive value associated to the leaves of the Semantic Trees where Concepts are placed. Subsequently, using the same structure of *RST* we create new Semantic Trees without *AR* values. Instead, the leaves with Concepts of this new Semantic Trees will be annotated with the *PosA* value.

Later, to assign some Positive value to the parent Concepts, each parent Concept will accumulate the positive values from child Concepts. Equation 6 shows the bottom-up process.

$$PosA(PC) = \sum_{i=1}^n PosA(ChC); \quad (6)$$

Where *PC* is the Parent Concept; *ChC* is the Child Concept of *PC*; and *PosA(ChC)* represents the positive value of the *ChC*.

3.3 Obtaining the Negative Semantic Trees (NST)

In this phase, we repeat the step described in Section 3.2, but for negative values. Table 3 shows the *PST* and *NST* obtained from the example.

| Vectors Pos-Neg | | | | | |
|-----------------|-------|----------------|-------|-------|--------------|
| PosA | NegA | Domain | PosA | NegA | Domain |
| 0.00 | 1.00 | Social_Science | 0.00 | 0.00 | Buildings |
| 0.00 | 0.00 | Administration | 0.00 | 0.50 | Commerce |
| 0.00 | 0.00 | Pedagogy | 0.00 | 0.00 | Environment |
| 0.00 | 0.00 | Root_Domain | 0.375 | 0.375 | Factotum |
| 0.00 | 0.00 | Psychoanalysis | 0.00 | 0.00 | Psychology |
| 0.00 | 0.50 | Economy | 0.00 | 0.00 | Architecture |
| 0.375 | 0.375 | Quality | 0.00 | 0.00 | Pure_Science |
| 0.00 | 0.00 | Politics | | | |

Table 3. Final Domain Vectors Pos-Neg

As we can see, the analyzed sentence is more linked to the *Social_Science* domain and it accumulates a negative value of 1 and a positive value of 0. This indicates that the sentence is more negative than positive.

3.4 Obtaining polarities of the sentences

In this step, we concentrate on detecting which polarity is more representative according to the Semantic Trees obtained for each resource (dimension). For that, we combine the *RST* with *PST* and *RST* with *NST*. Depending on the obtained results we classify the sentence as Positive, Negative or Neutral. Before performing this step, we have to normalize the three types of Semantic Trees (*RST*, *PST* and *NST*) for each dimension to work with values between 0 and 1.

Our main goal is to assign more weight to the polarities related to the most relevant Concepts in each Relevant Semantic Tree. Equation 7 shows the steps followed in order to obtain the positive semantic value.

$$ACPosA(RST, PST) = \sum_{i=1} RST_i * PST_i; \quad (7)$$

Where $ACPosA$ is the Positive Semantic Value of the analyzed sentence obtained for one Dimension, RST is the Relevant Semantic Tree sorted with the format: $RST [Concept| AR]$; PST is the Positive Semantic Tree sorted according RST structure with format: $PST [Concept|PosA]$; RST_i RST_i is the i -th AR value of Concept i ; PST_i PST_i is the i -th $PosA$ value of the concept i .

In order to measure the negative semantic value ($ACNegA$), we employ a similar equation replacing PST with NST . After obtaining the semantic opinion requirements, we evaluate our approach over three of the tasks proposed in the NTCIR 8 MOAT, for the monolingual English setting.

3.5 Judging sentence opinionatedness

The ‘‘opinionated’’ subtask requires systems to assign the values YES or NO to each of the sentences in the document collection provided. This value is given depending on whether the sentence contains an opinion (Y) or it does not (N). In order to tackle this task, we analyze the PST and NST of all dimensions (WN, WSD, WNA, SUMO and SC). After reviewing the $PSTs$ and $NSTs$ if at least one Concept has assigned a value distinct from zero the result will be ‘‘YES’’ in other cases will be ‘‘NO’’.

3.6 Determining sentence relevance

In the sentence relevance judgement task, the systems have to decide whether a sentence is relevant to the given question or not (Y|N). We assume that the given question is related to each sentence per topic if it has a RST 50% similar (the similarity is obtained by quantity of Concept labels that match). The analyzed sentence is relevant only if the PST and the NST values of all dimensions that are taken into account contain at least a positive or a negative value.

3.7 Polarity and topic-polarity classification

The polarity judgment task requires the systems to assign a value of ‘‘ POS ’’, ‘‘ NEG ’’ or ‘‘ NEU ’’ (positive, negative or neutral) to each of the sentences in the documents provided.

Our proposal consists of accumulating the $ACPos$ values and $ACNeg$ values of all Dimensions

and comparing them. These accumulated values will be named $ACPosD$ and $ACNegD$ respectively. In case $ACPosD > ACNegD$ the assigned value is POS , if $ACPosD < ACNegD$ the assigned value is NEG , otherwise, the assigned value is NEU .

4 Evaluation and analysis

In this section we concentrated on measuring the influence of each Dimension (resource) taken separately and jointly in our proposal. Also, we have compared our results with the best results obtained by the participant systems in the NTCIR 8 MOAT competition.

4.1 Influence of each dimension

In this section, we present the results of the three tasks described above using the combination of all dimensions and using each of the resources separately. Moreover, we describe the experiments we have performed. Exp1: Combining all Dimensions (WND, WNA, WN taxonomy, SUMO and SC). Exp2: Using WNA. Exp3: Using WND. Exp4: Using SC. Exp5: Using SUMO. Exp6: Using WN taxonomy. The results are presented in Table 4.

| Exp | Opinion | | | Relevance | | | Polarity | | |
|-----|---------|------|-------------|-----------|------|-------------|----------|------|-------------|
| | P | R | F | P | R | F | P | R | F |
| 1 | 20.6 | 87.8 | 33.3 | 78.8 | 86.8 | 82.6 | 39.4 | 34.5 | 36.8 |
| 2 | 23.8 | 57.2 | 33.6 | 77.9 | 55.8 | 65.1 | 39.7 | 22.2 | 28.5 |
| 3 | 22.6 | 69.5 | 34.1 | 79.4 | 69.2 | 74.0 | 40.3 | 27.5 | 32.7 |
| 4 | 20.1 | 88.5 | 33.3 | 78.8 | 87.3 | 82.3 | 39.7 | 34.9 | 37.2 |
| 5 | 21.3 | 86.5 | 34.2 | 79.0 | 85.8 | 82.3 | 40.6 | 33.7 | 36.8 |
| 6 | 21.1 | 87.6 | 34.1 | 78.8 | 86.6 | 82.5 | 40.5 | 34.2 | 37.1 |

Table 4. Results on each task. Precision (P), Recall (R) and F-Measure (F).

As we can see, the best results are obtained in Experiment 4 and 6, which use the WN taxonomy and SC to obtain the RST , PST and NST . However, the other experiments results are similar in performance level. This indicates that our proposal can be successfully applied to opinion mining tasks.

4.2 Influence of the semantic dimensions without normalizing the vector

In order to prove that the value normalization introduces noise, we performed the same experiments without normalizing vectors. In Table 5, we show in bold font the F-Measure obtained

that constitutes an improvement to previous results. It is important to remark that not normalizing the vectors helps the Polarity Classification task. All the experiments presented in Table 5 improved the previous results and the SC obtained one of the best results for the Polarity and the Relevance task.

| Exp | Opinion | | | Relevance | | | Polarity | | |
|-----|---------|------|-------------|-----------|------|-------------|----------|------|-------------|
| | P | R | F | P | R | F | P | R | F |
| 7 | 20.1 | 88.5 | 33.3 | 78.8 | 87.3 | 82.8 | 39.7 | 34.9 | 37.2 |
| 8 | 23.3 | 61.1 | 33.7 | 78.4 | 60.0 | 68.0 | 42.3 | 25.5 | 31.8 |
| 9 | 21.9 | 77.9 | 34.2 | 79.2 | 77.3 | 78.2 | 39.4 | 30.5 | 34.4 |
| 10 | 20.6 | 87.7 | 33.4 | 78.9 | 86.7 | 82.6 | 44.6 | 38.9 | 41.6 |
| 11 | 20.6 | 85.0 | 33.2 | 78.5 | 83.6 | 81.0 | 44.6 | 37.7 | 40.9 |
| 12 | 20.5 | 85.5 | 33.1 | 78.7 | 84.4 | 81.5 | 43.7 | 37.0 | 40.1 |

Table 5. Results without normalized vectors. Precision (P), Recall (R) and F-Measure (F).

4.3 Comparison with other proposals

In this section, we present a comparison between our proposal and the best participating systems in NTCIR 8 MOAT. In the sentence opinionatedness judgement task, the only systems that obtained better results compared to our proposal are UNINE (Zubaryeva and Savoy, 2010) and NECLC systems. These systems obtained F-measure values of 40.1% and 36.52% respectively. These results are not so far from our results, with the simple difference of 5.9% and 2.32% respectively.

In comparison to our proposal, UNINE is based on selecting terms that clearly belong to one type of polarity compared to the others and the value types of polarities are defined summing the count number of terms that tend to be overused in positive, negative and neutral opinionated sentences possibilities (Zubaryeva and Savoy, 2010). The opinionated score is the sum of *Positive Scores* and *Negative Scores* for each selected term. The score of non-opinionated sentences is computed as a sum of *Objectivity Score* for each selected term, divided by the number of words in the sentence. Our proposal neither takes into account the detection of relevant terms, nor the objective scores. UNINE also obtained better results than us in the Polarity task; we think that the combination of this proposal with ours could obtain better results. Taking into account that both proposals use Features Extraction we could combine not only Lexical Features but also Semantic Features.

In the Polarity task we could obtain similar results to the first run of UNINE system around 37% of F-measure but with results some distance of the best system that obtained a 51.03% of F-measure. For the relevance task, our proposal obtained a difference of 3.22% as far as F-measure is concerned from the best result of all runs submitted by the National Taiwan University (NTU). So, our proposal could be located around the first places among the three tasks mentioned.

5 Conclusion and further works

In this paper our research was focused on solving a recent problem stemmed from the availability of large volumes of heterogeneous data which provides different kind of information. We have conducted an analysis of how the scientific community confronts the tasks related to Opinion Analysis. One of the most used approaches is to apply Features Extraction and based on this idea, our proposal is to apply Semantic Features Extraction based on Relevant Semantic Trees. With our proposal we are able to associate the polarities presented on the sentences with Concept Semantic Trees. Thus, the Semantic Trees allow the classification of sentences according to their opinionatedness, relevance and polarity, according to MOAT competition. The obtained results were compared with the best results obtained on this competition achieving values very close to the best systems. Several experiments were conducted applying vector normalization and without normalization to know which semantic dimension performed better.

After a comparative analysis with the systems which results were not improved, we propose as further work to include the lexical features extraction in our proposal. We have planned to use Latent Semantic Analysis and other techniques to do this work.

Acknowledgements

This paper has been supported partially by Ministerio de Ciencia e Innovación - Spanish Government (grant no. TIN2009-13391-C04-01), and Conselleria d'Educació - Generalitat Valenciana (grant no. PROMETEO/2009/119, ACOMP/2010/288 and ACOMP/2011/001).

References

- Alexandra Balahur, Ester Boldrini, Andrés Montoyo and Patricio Martínez-Barco. 2010. The OpAL System at NTCIR 8 MOAT. In *Proceedings of NTCIR-8 Workshop Meeting*: 241-245. Tokyo, Japan.
- Alexandra Balahur and Andrés Montoyo. 2009. A Semantic Relatedness Approach to Classifying Opinion from Web Reviews. *Procesamiento del Lenguaje Natural*, 42:47-54.
- Andrea Esuli and Fabrizio Sebastiani. 2006. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. In *Fifth international conference on Lenguaje Resources and Evaluation* 417-422.
- Amal Zouaq, Michel Gagnon and Benoit Ozell. 2009. A SUMO-based Semantic Analysis for Knowledge Extraction. In *Proceedings of the 4th Language & Technology Conference*. Poznań, Poland.
- Bernardo Magnini and Gabriela Cavaglia. 2000. Integrating Subject Field Codes into WordNet. In *Proceedings of Third International Conference on Language Resources and Evaluation (LREC-2000)*: 1413--1418.
- Bernardo Magnini, Carlo Strapparava, Giovanni Pezzulo and Alfio Gliozzo. 2002. Comparing Ontology-Based and Corpus-Based Domain Annotations in WordNet. In *Proceedings of the First International WordNet Conference*: 21-25 Mysore, India.
- Bernardo Magnini, Carlo Strapparava, Giovanni Pezzulo and Alfio Gliozzo. 2008. Using Domain Information for Word Sense Disambiguation. In *Proceedings of the First International Conference on Emerging Trends in Engineering and Technology (icetet 2008)*: 1187-1191. Nagpur, India.
- Christiane Fellbaum. 1998. WordNet. An Electronic Lexical Database. *The MIT Press*.
- Guo-Hau Lai, Jyun-Wei Huang, Chia-Pei Gao and Richard Tzong-Han Tsai. 2010. Enhance Japanese Opinionated Sentence Identification using Linguistic Features: Experiences of the IISR Group at NTCIR-8 MOAT Task. In *Proceedings of NTCIR-8 Workshop Meeting*: 272-275. Tokyo, Japan.
- Hatzivassiloglou, Vasileios and Janyce Wiebe. 2000. Effects of Adjective Orientation and Gradability on Sentence Subjectivity. In *International Conference on Computational Linguistics (COLING-2000)*.
- Ian Niles. 2001. Mapping WordNet to the SUMO Ontology. *Teknowledge Corporation*.
- Janyce Wiebe, Theresa Wilson and Claire Cardie. 2005. Annotating Expressions of Opinions and Emotions in Language. In *Kluwer Academic Publishers*: Netherlands.
- Luis Villarejo, Lluís Márquez and German Rigau. 2005. Exploring the construction of semantic classifiers for WSD. In *Sociedad Española para el Procesamiento del Lenguaje Natural*, 35: 195-202.
- Olena Zubaryeva and Jacques Savoy. 2010. Opinion Detection by Combining Machine Learning & Linguistic Tools In *Proceedings of NTCIR-8 Workshop Meeting*: 221-227. Tokyo, Japan.
- Paul Ekman. 1999. Handbook of Cognition and Emotion. *Handbook of Cognition and Emotion*: John Wiley & Sons, Ltd.
- Rubén Izquierdo, Armando Suárez and German Rigau. 2007. A Proposal of Automatic Selection of Coarse-grained Semantic Classes for WSD. *Procesamiento del Lenguaje Natural*, 39:189-196.
- Rudi L. Cilibrasi and Paul M.B. Vitányi. 2007. The Google Similarity Distance. *IEEE Transactions On Knowledge And Data Engineering*, 19(3).
- Soo-Min Kim and Eduard Hovy. 2006. Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text. In *In Proceedings of workshop on sentiment and subjectivity in text at proceedings of the 21st international conference on computational linguistics/the 44th annual meeting of the association for computational linguistics (COLING/ACL 2006)*: 1-8. Sydney, Australia.
- Soo-Min Kim and Eduard Hovy. 2005. Identifying Opinion Holders for Question Answering in Opinion Texts. In *Proceedings of AAI-05 Workshop on Question Answering in Restricted Domains*.
- Sonia Vázquez, Andrés Montoyo and German Rigau. 2004. Using Relevant Domains Resource for Word Sense Disambiguation. In *IC-AI'04. Proceedings of the International Conference on Artificial Intelligence*: Ed: CSREA Press. Las Vegas, E.E.U.U.
- Yoan Gutiérrez, Antonio Fernández, Andrés Montoyo and Sonia Vázquez. 2010a. UMCC-DLSI: Integrative resource for disambiguation task. In *Proceedings of the 5th International Workshop on Semantic Evaluation*: 427-432. Uppsala, Sweden.
- Yoan Gutiérrez, Antonio Fernández, Andrés Montoyo and Sonia Vázquez. 2010b. Integration of semantic resources based on WordNet. In *XXVI Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 45: 161-168. Universidad Politécnica de Valencia, Valencia, Spain.