EmoTag: Automated Mark Up of Affective Information in Texts.

Virginia Francisco
Departamento de Ingeniería del Software e Inteligencia Artificial
Universidad Complutense de Madrid, Spain
virginia@fdi.ucm.es

Raquel Hervás
Departamento de Ingeniería del Software e Inteligencia Artificial
Universidad Complutense de Madrid, Spain
raquelhb@fdi.ucm.es

Abstract

This paper presents an approach to automated mark up of affective information in texts. The approach considers in parallel two possible representations of emotions: as emotional categories and emotional dimensions. For each representation, a corpus of example texts previously annotated by human evaluators is mined for an initial assignment of emotional features to words. This results in a List of Emotional Words (LEW) which becomes a useful resource for later automated mark up. EmoTag employs for the actual assignment of emotional features a combination of the LEW resource, the ANEW word list, WordNet for knowledge-based expansion of words not occurring in either and an ontology of emotional categories.

1 Introduction

The task of annotating text with specific labels indicating its emotional content or inclination is fundamental for any attempt to make computer interfaces respond in some way to the affective nature of the content they are handling. This is particularly true for research which attempts to produce synthesized voice with different emotional states, but it may also be applicable in other contexts, such as multimodal presentation, where colors, typography or similar means can be used to convey emotion.

In NLP is important to know the connection between emotions and words and linguistic structures in order to drive plot generation in automatic narrative generation. There are programs as MEXICA (Pérez y Pérez and Sharples, 2001) which employ the emotional links between characters to retrieve possible logical actions to continue the story while developing a narration. In this type of programs it will be interesting retrieve the words and the linguistic structures which adapt to the emotion that the story is trying to express with the actions retrieved to continue the story in a emotional way.

Emotions are not an easy phenomenon; there are a lot of factors that contribute to the generation of emotions. For Izard (Izard, 1971) a good definition of the word *emotion* must take into account: the conscious feeling of the emotion, the processes that appear in the nervous system and in the brain and the expressive models of the emotion.

There are a lot of theories about how many emotions there actually are. In these theories the number of emotions varies from two up to infinity (Ortony and Turner, 1990). There are two main approaches related to the number of emotions: emotional categories and emotional dimensions. The first approach is based on the use of emotion-denoting words (Cowie and Cornelius, 2003). The second approach is that there are two or three dimensions (Dietz and Lang, 1999) which represent the essential aspects of emotional concepts. Common labels for these dimensions include: *evaluation* (positive / negative) and *activation* (active / passive) which are the main dimensions; sometimes they are completed with the *power* dimension (dominant / submissive).

Many psychologists have claimed that certain emotions are more basic than others. Plutchik's (1980) postulates that there is a small number of basic, primary, or prototype emotions (*anger*, *anticipation*, *disgust*, *joy*, *fear*, *sadness* and *surprise*).

All other emotions are mixed or derivative states. Plutchik states that all emotions vary in their degree of similarity to one another and that each emotion can exist in varying degrees of intensity or levels of arousal. Ekman (1992) has focused on a set of six basic emotions that have associated facial expressions: *anger, disgust, fear, joy, sadness* and *surprise*. Those emotions are distinctive, among other properties, by the facial expression characteristic to each one. Izard (1977) determines that the basic emotions are *anger, contempt, disgust, distress, fear, guilt, interest, joy, shame* and *surprise*. The OCC Model (Ortony et al, 1988) has established itself as the standard model for emotional synthesis. It presents 22 emotional categories: *pride* - *shame, admiration* - *reproach, happy* - *resentment, gloating* - *pity, hope* - *fear, joy* - *distress, satisfaction* - *fear-confirmed, relief* - *disappointment, gratification* - *remorse, gratitude* - *anger* and *love* - *hate*. OCC Model considers that categories are based on valence reactions to situations constructed as: goal relevant actions and attractive or unattractive objects. Parrot (2001) presents a deeper list of emotions, where emotions were categorized into a short tree structure, this structure has three levels: primary emotions, secondary emotions and tertiary emotions. As primary emotions Parrot presents: *love, joy, surprise, anger, sadness* and *fear*.

The aim of this work is to present a system for the automated mark up of affective information in texts, EmoTag (Francisco and Gervás, 2006b). The last section discusses some ideas for future work which will improve the results obtained with the first approach.

2 Resources employed by EmoTag

This section presents a brief review of the existing resources used by EmoTag.

The *Affective Norms for English Words* (*ANEW*) (Bradley and Lang, 1999) is a set of normative emotional ratings for a large number of words in the English language. The goal is to have a set of verbal materials rated in terms of pleasure, arousal, and dominance. This data base of emotional words is content independent, a set of words have been showed to subjects and each word is rated giving a value for each emotional dimension. We use ANEW for marking up texts with emotional dimensions in order to look for the words which does not appear in our list of emotional words (LEW list) which is content dependent.

WordNet is a semantic lexicon for the English language. It groups English words into sets of synonyms called synsets. We used WordNet for knowledge-based expansion of words not occurring in LEW list (or ANEW list in the case of emotional dimensions). By means of WordNet we obtained synonyms, antonyms and hypernyms (words related with the original word).

Shallow Parsing Techniques are used for syntax analysis. This type of syntax analysis employs the binary relations between lexical units. There are a lot of automatic dependency analyzers for different languages: English, French, Swedish… The most successful is MINIPAR (Lin, 1998) which is used by EmoTag (Francisco and Gervás, 2006a) to determine the scope of negations appearing in the sentences, in order to take their effect into account. Nowadays there are no dependency analyzers in Spanish so we are studying the use of tools as MaltParser [1] in order to get a Spanish dependency analyzer automatically.

A *POS Tagger* marks up the words in a text with its corresponding part of speech, based on both its definition as well as its context. We used qtag[2] for English and Tree Tagger [3] for Spanish.

A *Stemmer* reduces inflected (or sometimes derived) words to their stem, base or root form. The stem does not need to be identical to the morphological root of the word, it is sufficient that related words map to the same stem. We need a stem in order to group the related words in our LEW list, that have the same emotional content. In the case of English we have used the stem given by MINIPAR and in the Spanish case we have used Snowball Spanish Stemmer[4].

3 Implemented Resources

This section presents a brief review of the resources that we have implemented in order to mark up texts with emotions.

[1]http://w3.msi.vxu.se/$\sim$nivre/research/MaltParser.html
[2] http://www.english.bham.ac.uk/staff/omason/software/qtag.html
[3] http://www.ims.uni-stuttgart.de/projekte/corplex/ TreeTagger/DecisionTreeTagger.html
[4] www.snowball.tartarus.org

3.1 Ontology of Emotions (OntoEmotions)

We have developed an ontology of emotional categories (Francisco et al, 2007). They are structured in a taxonomy that covers from basic emotions to the most specific emotional categories. This ontology is based on the emotional structures mentioned in Section 1, as basic emotions we have: *sadness, happiness, surprise, fear* and *anger*. We have adapted the Parrot model to these basic emotions, and we have integrated in this model all the emotions which appear in other models. Finally we have added all the emotion-denoting words of the English and Spanish language.

OntoEmotions is an ontology of emotional categories, that is emotion-denoting words such as h*appy*, *sad*, *fear*…, not an ontology of words and their relation with emotions as WordNet Affect.

Our ontology has two root concepts:

- Emotion: The root for all the emotional concepts which are used to refer to emotions. Each of the emotional concepts are subclasses of the root concept Emotion. Some examples of these subclasses are: Happiness, Sadness, Fear, Envy…

- Word: The root for the emotion-denoting words, the specific words which each language provides for denoting emotions. Our ontology is currently available for two different languages: English and Spanish. In order to classify the words into their corresponding language the root concept Word has two subclasses: *EnglishWord* and *SpanishWord*.

As instances of the *EnglishWord* and *SpanishWord* subclasses there are emotion-denoting words, which are all the words used for denoting *Surprise, Happiness, Indignation, Horror*… Each of these instances has two parents: a concept from the Emotion hierarchy (which indicates the type of abstract emotion denoted by the word) and a concept from the Word hierarchy (which indicates the language of the word).

Figure 1 shows a fragment of the ontology. In this fragment it can be seen how the words are related both to one emotional concept and to one word concept, for example the word *cheerfulness* is an instance of the emotional concept *Happiness* at the same time it is an instance of the word concept *EnglishWord*, which means that *cheerfulness* is an English word for denoting the emotion *Happiness*.

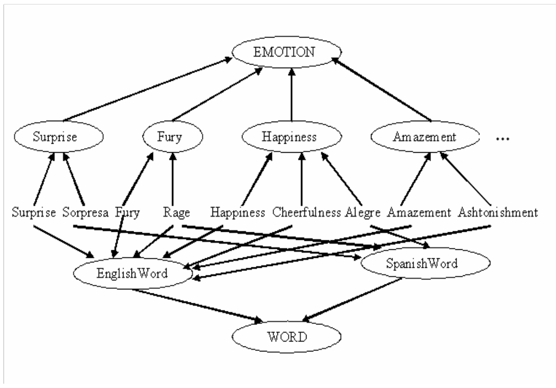

Figure 1. Fragment of the emotional ontology.

Given the semantics we have chosen for our ontology, two instances of the Word-concept can be considered to be synonyms if they are also instances of the same single Emotion-concept from the parallel Emotion subhierarchy. For example, in the figure above, we can find that the words *amazement* and *astonishment* are synonyms because they are both instances of the Emotion-concept *Amazement*.

From a given emotion-denoting word by means of our ontology we obtain the direct emotional concept associated to it as well as the more general emotional concept related to the direct emotional concept. We can get, too, the synonyms for an emotional word and the corresponding word in other language. For example, given the emotional word *grief*, we have *Grief* as direct emotional concept, *Distress*, *Sadness* and *Emotion* as general emotional concepts, *Agonía* as Spanish translation and *agony*, *anguish* and *sorrow* as antonyms.

3.2 List of Emotional Words (LEW List)

Our method for annotating text with emotions relies on a dictionary of word to emotion assignments (Francisco and Gervás, 2006b). This is obtained from a corpus of human evaluated texts by applying language analysis techniques. Similar techniques are later applied to assign emotions to sentences from the assignments for the words that compose them.

If we want to obtain a program that marks up texts with emotions, as a human would, we first need a corpus of marked-up texts in order to analyze and obtain a set of key words which we will use in the mark up process. Each of the texts which forms part of the corpus may be marked by more than one person because assignment of emotions is a subjective task so we have to avoid “subjective extremes”. In order to do that we obtain the emotion assigned to a phrase as the average of the mark-up provided by fifteen evaluators. Therefore the process of obtaining the list of emotional words involves two different phases: *evaluation method*, several people mark up some texts from our corpus; *extraction method*, from the mark-up texts we obtain the list of emotional words.

First we have to decide which texts are going to be part of our corpus. We decide to focus the effort on a very specific domain: fairy tales. This decision was taken mainly because generally fairy tales are intended to help children understand better their feelings, and they usually involve instances of the emotions that most children experiment on their way to maturity: *happiness, sadness, anger*…

Once the domain of the corpus' texts is established, the set of specific tales that we are going to work with must be selected. We have selected eight tales, every one of them popular tales with different lengths (altogether they result in 10.331 words and 1.084 sentences), in English and Spanish. The eight tales are marked up with emotional categories and emotional dimensions. We provide a list of different emotions in order to help the evaluators in the assignment of emotional categories and the SAM standard which can be seen in the Figure 2 in order to help them in the assignment of values for each dimension.

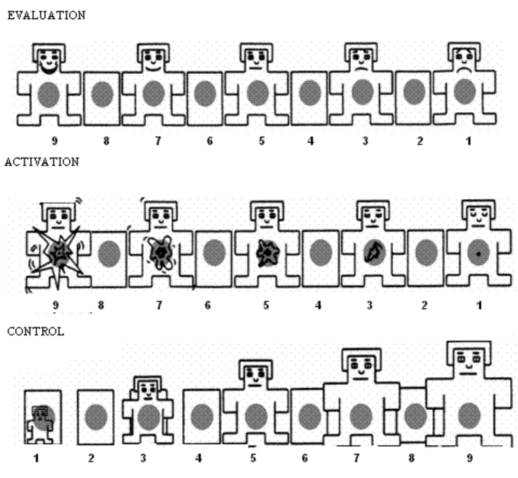

Figure 2. Dimensional scales according to SAM.

Based on the tales marked up by the evaluators we obtain two data base of words and their relation to emotional categories and emotional dimensions, one in English and other in Spanish.

Firstly we have to obtain the emotion mostly assigned to each sentence by the evaluators, in order to do that we distinguishes between emotional dimensions and emotional categories:

- Emotional dimensions: In order to get the reference value for each sentence in the text we get the average value for each of the emotional dimensions (*evaluation*, *activation* and *power*).
- Emotional categories: In order to get the reference value we carry out the following process: if at least half of the evaluators agree on the assignment of one emotion, this emotion is taken as the reference value for the sentence. In other case, we group the emotions in levels (according to the level of the emotional concept they refer to) then we obtain the related concepts for the emotion with the lower level. Once these new concepts are added to their corresponding levels, if we have any emotion supported by at least half of the evaluators we take it as the reference value. If two emotions supported by most of the evaluators we get the emotion with a lower level. In other case, we obtain the related concepts for the emotion with the next lower level and repeat this step in ascending order until we have an emotion supported by at least half of the evaluators.

Once we have the reference value for each sentence we carry out the following process in order to get the LEW list. Firstly we split the text into sentences which are processed with MINIPAR in order to obtain the words affected by a negation and with a tagger which assigns a part-of-speech tag to each word in a text. Every sentence is divided into words and with every word and its label we carry out the following process:

- Discard the words with a label which belong to our list of stop POS tags (conjunctions, numbers, determiners, existential there, prepositions…).
- Obtain the stem of the words.

- Insert the pair stem - label into the LEW list with its corresponding emotional value in the case of words not affected by negation and the opposite emotional value in the case of words affected by negation.

Once all the tales have been processed we extend our list with synonyms and antonyms which are look up in WordNet. For inserting related words into the database, the same emotional values of the original word are used in the case of synonyms and the opposite values are used in the case of antonyms.

For emotional dimensions LEW list stores for each pair (word's stem, label) the average value of *activation*, *evaluation* and *power* of that pair in all the analyzed texts. In the Table 1 it can be seen a fragment of the English LEW list for emotional dimensions:

Word's Stem	Label	Act.	Eval.	Power
Alarm	Adj.	6	3	4
Taste	Adj.	3	1	4
Sword	Noun	7,3	5	5,5

Table 1. Fragment of the LEW list for emotional dimensions.

For emotional categories LEW list stores for each pair (word's stem, label) the probability of this pair of been indicating of one of the categories. In the Table 2 it can be seen a fragment of the English LEW list for emotional categories:

Stem	Label	Grief	Sad	Happy	Neutral	…
Misery	Noun	20%	20%	0%	0%	…
Death	Noun	43%	7%	0%	29%	…
Dark	Noun	20%	40%	0%	0%	…
Marry	Adj.	0%	0%	75%	25%	…

Table 2. Fragment of the LEW list for emotional categories.

4 EmoTag

Our process classifies English and Spanish sentences into emotions (emotional categories and emotional dimensions). The first step is to perform sentence detection and tokenization in order to carry out our process based in the relation between words and different emotions. We carry out the process in the following way: we obtain by means of MINIPAR the words affected by negations and the POS tag for every word in the sentence, based on these tags and the stem of the words we decide the emotion of the sentence in the following way:

- If the tag associated with the word is in our label stop list we leave it out in other case we obtain the stem of the word.
- We look up the pair stem - label in the LEW list, if the word is present we assign to it the probability of carrying the emotions we are studying in the case of emotional categories or the value for the emotional dimensions in other case. If the word is affected by a negation we reverse the probability in the case of emotional dimensions and the value of each dimension in the case of emotional dimensions. Based on these emotional values of the words we obtain the final emotion of the sentence.
- If the word is not in the LEW list, in the case of emotional dimensions we look up it in the ANEW list.
- If the word is not in the LEW list in the case of emotional categories, and is not in the LEW list or the ANEW list in the case of emotional dimensions we look up the hypernyms of the word in WordNet, and look them up in the LEW list (and in the ANEW list in the case of emotional dimensions). The first appearance of a hypernym is taken and the emotional content associated to our original word and the new word is inserted in the LEW list for subsequent occurrences of these words in our tales.
- If none of the hypernyms appear in the LEW list we leave out the word and it does not take part in the mark up process.

Once all the words of the sentences have been evaluated, we have to obtain the emotion for marking up the sentence; the process is different in the case of emotional categories and emotional dimensions:

- Emotional dimensions: Once all the words of the sentences have been evaluated we obtain the average value of each emotional dimension.
- Emotional categories: Once all the words of the sentences have been evaluated we

add up the probability of each emotion of the different words, and we carry out the following process for each of the possible emotions: we process all the emotions in order to obtain the related emotional concepts (the parents of the emotion in the ontology), these related emotional concepts are added to the previous ones with the probability associated to the more specific concept. Then emotions are group by their corresponding level in the ontology and the emotion more general (with the lower level in the ontology) with the bigger probability is assigned to the sentence.

A sample part of a marked tale with emotional categories:

<anxiety>The knight faced the lioness. </anxiety>
<neutral>He fought she. </neutral>
<neutral>The knight threw the spear.</neutral>
…
<delight>She returned to the strong castle. </delight>
<happy>The knight and the princess lived happy ever afterward. </happy>

And a sample part of a marked tale with emotional dimensions:

<emo val=5.04 act=5.05 cont=5.04>A Fox once saw a Crow fly off with a piece of cheese in its beak and settle on a branch of a tree. </emo>
<emo val=5.09 act=4.84 cont=4.80>That's for me, as I am a Fox, said Master Reynard, and he walked up to the foot of the tree. </emo>
...
<emo val=4.52 act=5.03 cont=5.08>In exchange for your cheese I will give you a piece of advice for the future: </emo>
<emo val=3.41 act=4.85 cont=3.44>Do not trust flatterers. </emo>

5 Evaluation

Four tales took part in the tests. Tales are English popular tales with different number of words (from 153 words and 20 lines to 1404 words and 136 lines). Each of our four tales will be tagged first with the emotional dimensions, and then with the categories.

The data on emotional dimensions we have available for each tale are the values that each dimension takes for each sentence. To evaluate our tagger we have divided the evaluation according to the different dimensions. In order to get a measure of our tagger we have taken measures first from the evaluators' tales and then from our tagger's tales. For evaluator's tales we have, as reference data, the values assigned for each dimension and each sentence by the human evaluators. An average emotional score for each dimension of a sentence is calculated as the average value of those assigned to the corresponding dimension by the human evaluators. The deviation among these values is around 1.25, which indicates the possible range of variation due to human subjectivity. In the case of tagger's tales for each dimension, if the deviation of the tagger is less or equal to the average deviation among evaluators, we consider that the sentence is tagged correctly. Results indicate that the tagger is obtaining better results in terms of deviation from the average obtained by humans for the arousal and dominance dimensions, and comparable results in the case of valence. The graph in Figure 3 shows the success percentage.

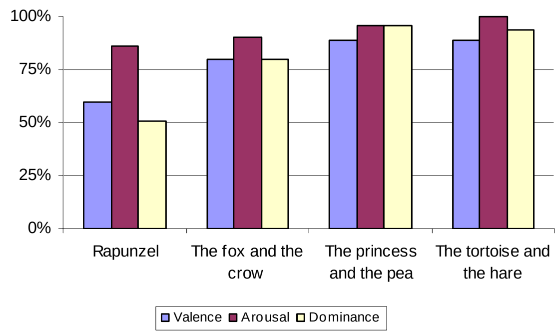

Figure 3. Success percentage in EmoTag for emotional dimensions.

The data on emotional categories we have available for each tale are emotional label for each sentence. For evaluator's tales we have noticed that the percentage of sentences on which the majority of the human evaluators - half of their number plus one - agrees on the assignment of an emotion is around 70%. This is an important data when it comes to interpreting the results obtained by our tagger. A reference value for the emotion of each phrase is obtained by choosing the emotion most often assigned to that sentence by the human evaluators. In the case of tagger's tales the reference value obtained in the evaluator's tales is used to compare with the results generated by our tagger. The graph in Figure 4 shows the percentages of success obtained for each tale.

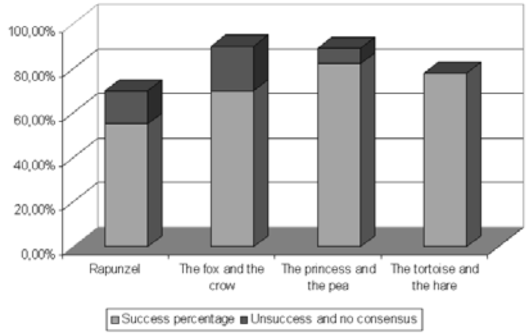

Figure 4. Succes percentage in EmoTag for emotional categories.

6 Future Work

At the moment EmoTag marks up texts taking into account only the *lexical emotion*, the emotion which derives from the words employed but this is not the only type of emotion that take part in the emotion perceived by the users when they are reading or listening a narration. There are mainly three types of emotions that we have to consider: *lexical emotion, semantic emotion* and *mood of the user*.

In order to improve the results we have obtained with EmoTag for the *lexical emotions* we will consider the following ideas:

- Use a finer granularity for emotional units. We have observed that very long sentences lead to confusion when assigning emotions. For example, we will consider subordinate sentences as emotional units inside a bigger unit, the general sentence.
- Use existing approaches for content analysis of textual data as “The General Inquirer'” (Kalin, 1966) which is a dictionary of words marked with different categories; seven of these categories could be interesting for our work: *EMOT* indicates if the word is related to emotion, *POSITIV* or *NEGATIV* marks words of positive or negative outlook, *POWER* which indicates a concern with power, control or authority or *SUBMIT* which connotes submission to authority or power or dependence to others, *ACTIVE* or *PASSIVE* which marked words as implying an active or passive orientation. Using these tools we can leak the words which are going to take part in our marked up process.
- Processing modifiers. When a modifier appears in a sentence, the emotion associated to that sentence should be increased or reduced.
- Processing modal verbs. When a modal verb appears in a sentence the words under their scope must be treating in a special way. For example, “care to sing” does not imply that the subject is singing so a possible solution could be to reduce the activation because the action is not being held.
- In the first approach of EmoTag only one emotion is attributed to a sentence, the one that is manifested most strongly. In some sentences a better emotional mark up could include not only the strongest emotion but other emotions.

The next step is to get the *semantic emotion* of a text. We have an ontology of concepts which represents the concepts which take part in a tale and we have ontology of emotions. In order to get the *semantic emotion* we have to link these two ontologies. This way, we will obtain the emotions related to the concepts that take part in the tale. An example of the relation between concepts (characters and actions) and emotion can be seen in (Francisco et al, 2006). In order to identify the links between a given sentence and the concepts in the domain we can apply text analysis techniques. We have already used techniques such as dependency analysis to identify key concepts of a sentence in order to build conceptual cases from a text corpus (Hervás et al, 2007). This is important, because depending on the semantic content the final emotion could differ from the lexical emotion. For example, the action “to die” is a sadness action, and in our LEW list is marked up this way, but if the subject of this action is the witch the action “to die” turn into a happy action.

Finally, we have the *user mood*. It seems obvious that the mood of the speaker plays a very important role in determination of the emotional inflection given to an utterance. In order to model this influence we have chosen to consider a representation of the mood of the user as an additional input, modelled along the same lines as the other emotional information. This can be taken into consideration when it comes to get the final emotion of a text that is going to be present to the user.

Once the three types of emotions are fixed for specific texts it is time to combine them and obtain the final value of emotion for each of the emotional units. We have to study which emotion has the main weight and how each of this type of emotion influence in the others. One way of combine these types of emotions could be, first obtain the lexical emotion, then obtain the emotion associated to the concepts and how the lexical emotion influences in the concepts, and finally modify this emotion with the user mood.

Acknowledgements

This research is funded by the Spanish Ministry of Education and Science (TIN2006-14433-C02-01 project) and a joint research group grant (UCM-CAM-910494) from the Universidad Complutense de Madrid and the Dirección General de Universidades e Investigación of the Comunidad Autónoma de Madrid.

References

M.M. Bradley and P.J. Lang. 1999. Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings. Technical Report c-1. The Center for Research in Psychophysiology, University of Florida.

R. Cowie and R.R. Cornelius. 2003. Describing the Emotional States that are Expressed in Speech. In Speech Communication Special Issue on Speech and Emotion.

R. Dietz and A. Lang. 1999. Affective Agents: Effects of Agent Affect on Arousal, Attention, Liking and Learning. In Proceedings of the Third International Cognitive Technology Conference, San Francisco.

P. Ekman. 1992. Are there Basic Emotions? Psychological Review, 99(3):550–553.

V. Francisco and P. Gervás. 2006a. Análisis de Dependencias para la Marcación de Cuentos con Emociones. In Proceedings of the XXII Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural, Zaragoza, Spain.

V. Francisco and P. Gervás. 2006b. Automated Mark Up of Affective Information in English Texts. In Proceedings of the International Conference on Text, Speech and Dialogue (TSD'06), Brno.

V. Francisco, R. Hervás, and P. Gervás. 2006. Two Different Approaches to Automated Mark Up of Emotions in Text. In Proceedings of the International Conference on Innovative Techniques and Applications of Artificial Intelligence (AI-2006), Cambridge.

V. Francisco, P. Gervás, and F. Peinado. 2007. Ontological Reasoning to Configure Emotional Voice Synthesis. In Proceedings of the First International Conference on Web Reasoning and Rule Systems (RR2007), Innsbruck, Austria.

R. Hervás, P. Gervás, and V. Francisco. 2007. Dependency Analysis and CBR to Bridge the Generation Gap in Template-Based NLG. In Proceedings of the Eighth International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2007), Mexico D.F.

C.E. Izard. 1971. The Face of Emotion. Appleton Century-Crofts, New York.

C.E. Izard. 1977. Human Emotions. Plenum Press, New York.

R. Kalin, S.N. Davis, and D.C. McClelland. 1966. The General Inquirer. MIT Press.

D. Lin. 1998. Dependency-Based Evaluation of Minipar. In Proceedings of the Workshop on the Evaluation of Parsing Systems, Granada, Spain.

A. Ortony and T.J. Turner. 1990. What's Basic about Basic Emotions? Psychological Review, 97:315–331.

A. Ortony, G.L. Clore, and A. Collins. 1988. The Cognitive Structure of Emotions. Cambridge University Press, New York.

W.G. Parrott. 2001. Emotions in Social Psychology: Essential Readings. PA: Psychology Press, Philadelphia.

R. Pérez y Pérez and M. Sharples. 2001. Mexica: A Computer Model of a Cognitive Account of Creative Writing. J. Exp. Theor. Artif. Intell., 13(2):119–139.

R. Plutchik. 1980. A general Psycho Evolutionary Theory of Emotion. Emotion: Theory, Research, and Experience.

K.R. Scherer. 1984. On the Nature and Function of Emotion: *A Component Process Approach*. Scherer and K.R. and Ekman P and editors, Erlbaum, Hillsdale, NJ.