Decision Support Systems 53 (2012) 742–753

Contents lists available at SciVerse ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Detecting implicit expressions of emotion in text: A comparative analysis

Alexandra Balahur *, Jesús M. Hermida, Andrés Montoyo

Department of Software and Computing Systems, University of Alicante, Apartado de correos 99, E-03080 Alicante, Spain

ARTICLE INFO

Available online 23 May 2012

Keywords:
Emotion detection
Sentiment analysis
Knowledge base
EmotiNet
ISEAR

ABSTRACT

Sentiment analysis is one of the recent, highly dynamic fields in Natural Language Processing. Most existing approaches are based on word-level analysis of texts and are mostly able to detect only explicit expressions of sentiment. However, in many cases, emotions are not expressed by using words with an affective meaning (e.g. happy), but by describing real-life situations, which readers (based on their commonsense knowledge) detect as being related to a specific emotion. Given the challenges of detecting emotions from contexts in which no lexical clue is present, in this article we present a comparative analysis between the performance of well-established methods for emotion detection (supervised and lexical knowledge-based) and a method we propose and extend, which is based on commonsense knowledge stored in the EmotiNet knowledge base. Our extensive evaluations show that, in the context of this task, the approach based on EmotiNet is the most appropriate.

1. Introduction

Research in affect has a long established tradition in many sciences—Linguistics, Psychology, Socio-Psychology, Cognitive Science, Pragmatics, Marketing or Communication Science. Recently, many closely related subtasks were developed also in the field of Natural Language Processing (NLP), such as emotion detection, subjectivity analysis and opinion mining (also known as sentiment analysis, attitude and appraisal analysis or review mining) [30]. All these research fields can be considered as part of the wider area of research in artificial intelligence (AI) of Affective Computing [37]. Among these tasks, sentiment analysis aims at detecting the expressions of sentiment in text and subsequently classifying them, according to their polarity (semantic orientation) into different categories (usually, positive and negative). The problem is defined by Pang and Lee (2008) as “the binary classification task of labeling an opinionated document as expressing either an overall positive or an overall negative sentiment.” According to the Webster dictionary,[1] “sentiment suggests a settled opinion reflective of one's feelings”, where the term feeling is defined as “the conscious subjective experience of emotion” [52], “a single component of emotion, denoting the subjective experience process” [45]. Expressions of sentiment are thus directly related to expressions of emotions in text. As such, as in the case of the latter, sentiments can be expressed directly (e.g. “I like Nokia phones.”), indirectly (e.g. “This phone is light as a feather.”) or implicitly, by describing a situation which points the reader towards a specific sentiment (e.g. “I paid 200€ for this phone and it broke in two days.”). Most of the research performed in the field of sentiment analysis and the related task of emotion detection has aimed at detecting explicit expressions of sentiment (i.e. situations where specific words or word combinations are found in texts). Nevertheless, the expression of emotion is most of the times not achieved through the use of emotion-bearing words [34], but indirectly, by presenting situations that based on commonsense knowledge can be interpreted in an affective manner [4,5]. A clear example of such cases can be seen in self-reported affect (textual descriptions of affect-eliciting situations), a technique that has been widely used in psychological experiments [46]. In these cases, the subjects are asked to describe events that made them experience different emotions, without necessarily mentioning the emotion itself. A corpus with such examples is ISEAR [International Survey of Emotional Antecedents and Reactions 46].[2] In a first effort to overcome the issue of emotion detection from texts in which little or no lexical clues exist to mark the presence of a specific emotion (i.e. presence of words such as “joy”, “happy”, “angry”, etc.), Balahur et al. [6] proposed a method to build a commonsense knowledge base (EmotiNet) storing situations that trigger emotions, based on the principles of the Appraisal Theories [42]. The main idea behind our approach, inspired by the Appraisal Theories, is that situations trigger emotions based on the result of the individual evaluation of their components, in accordance to “appraisal criteria”. In order to detect the values of such criteria, each such situation was represented in EmotiNet as a chain of actions, with their corresponding actors, objects, their properties and the associated emotion. We subsequently demonstrated that by using this resource, we are able to detect emotion from examples in ISEAR describing family-related situations in which little or no

* Corresponding author.
E-mail addresses: abalahur@dlsi.ua.es (A. Balahur), jhermida@dlsi.ua.es (J.M. Hermida), montoyo@dlsi.ua.es (A. Montoyo).

[1] http://www.merriam-webster.com/.

[2] http://www.unige.ch/fapse/emotion/databanks/isear.html.

0167-9236/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2012.05.024

explicit mention of affect is present. In the present article, we analyze the peculiarities of the ISEAR data employed in our previous evaluation [6] and comparatively evaluate the performance of our system using established supervised and lexical knowledge-based methods for emotion detection versus the use of EmotiNet as emotion detection resource. Subsequently, we extend the knowledge contained in EmotiNet and perform additional evaluations to demonstrate the appropriateness of the created knowledge base for the automatic detection of implicit expressions of emotion.

2. Related work

Emotion is a complex phenomenon, on which no definition that is generally accepted has been given. However, a commonly used definition considers emotion as "an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems (Information Processing, Support, Executive, Action, Monitor) in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism" [44]. The term feeling "points to a single component denoting the subjective experience process" [45] and is therefore only a small part of an emotion. Moods are less specific, less intense affective phenomena, product of two dimensions—energy and tension [51]. Sentiment is a personal belief or judgment that is not founded on proof or certainty.

These affect-related phenomena have traditionally been studied in depth by disciplines such as Philosophy or Psychology. However, due to the advances in computing and the growing role of technology in everyday life, the past decades have shown an increasing interest in building software systems that automatically process affect. In order for such systems to benefit from the knowledge acquired in social sciences, interdisciplinary methods have been proposed, which use the existing theoretical models as basis for engineering computational ones.

This section explores the state of the art in the three domains our present research is closely related to: approaches to emotion detection in artificial intelligence, appraisal models in Psychology and knowledge bases in NLP applications.

2.1. Emotion detection systems in artificial intelligence

In AI, the term "Affective Computing" was first introduced by Picard [36]. Although there were previous approaches in the 80's and 90's, in the field of NLP, the task of emotion detection has grown in importance together with the exponential increase in the volume of subjective data on the Web in blogs, forums, reviews, etc. Previous approaches to spot affect in text include the use of models simulating human reactions according to their needs and desires [15], fuzzy logic [49], lexical affinity based on similarity of contexts—the basis for the construction of WordNet Affect [48] or SentiWordNet [16], detection of affective keywords [41] and machine learning using term frequency [31,53]. The two latter approaches are the most widely used in emotion detection systems implemented for NLP, because they are easily adaptable across domains and languages. Other proposed methods include the creation of syntactic patterns and rules for cause–effect [27]. Significantly different proposals for emotion detection in text are given in the work by Liu et al. [25] and the recently proposed framework of "Sentic Computing" [10], whose scope is to model affective reaction based on commonsense knowledge. Danisman and Alpkocak [13] proposed an approach based on vectorial representations. The authors compute the set of words that is discriminatory for 5 of the 7 emotions in the ISEAR corpus and represent the examples using measures computed on the basis of these terms. Other approaches were proposed within the SemEval 2007 Task 14: Affective Text [47]. Here, five different teams proposed a wide variety of approaches for emotion detection from news headlines [2,11,21,22], both unsupervised, knowledge-based (i.e. using pre-compiled lists of emotion or sentiment-bearing words and rules), as well as statistical approaches (i.e. using unigram information from pre-compiled training examples). The best performing approach – CLaC [2] – used a knowledge-based approach, combining knowledge on the emotional labels of different words and rules for modifiers. Finally, an up-to-date survey on the models of affect and their AC applications is presented by Calvo and D'Mello [9].

2.2. Appraisal theories

The set of models in Psychology known as the appraisal theories claim that emotions are elicited and differentiated on the basis of the cognitive evaluation of the personal significance of a situation, object or event [14,18,20,29]. Thus, the nature of the emotional reaction can be best predicted on the basis of the individual's appraisal of an antecedent situation, object or event. In consequence, there is a need to contextualize emotional response, due to which the same situation can lead to different affective reactions and similar reactions can be obtained through different stimuli.

There are different explanations for the elements considered in the appraisal process [see 42], which are called appraisal criteria. Currently, there is no common set of such criteria, as different versions of the appraisal theory have defined their own list of such factors. However, Scherer [42] shows that the appraisal criteria proposed in the different theories do converge and cover the same type of appraisals.

Examples of such criteria are the ones proposed and empirically evaluated by Lazarus and Smith [23], organized into a four categories:

1. intrinsic characteristics of objects and events;
2. significance of events to individual needs and goals;
3. individual's ability to cope with the consequences of the event;
4. compatibility of event with social or personal standards, norms and values.

Scherer [42] proposed five different categories of appraisal (novelty, intrinsic pleasantness, goal significance, coping potential, compatibility standard), containing a list of 16 appraisal criteria (suddenness, familiarity, predictability, intrinsic pleasantness, concern relevance, outcome probability, expectation, conduciveness, urgency, cause: agent, cause: motive, control, power, adjustment, external compatibility standards, internal compatibility standards). He later used the values of these criteria in self-reported affect-eliciting situations to construct the vectorial model in the expert system GENESIS [43]. The system maintains a database of 14 emotion vectors (corresponding to 14 emotions), with each vector component representing the quantitative measure associated to the value of an appraisal component. The values for the new situations are obtained by asking the subject a series of 15 questions, from which the values for the appraisal factors considered (components of the vector representing the situation) are extracted. Subsequently, the label assigned to the emotional experience is computed by calculating the most similar vector in the database of emotion-eliciting situations.

The appraisal models defined in Psychology have also been employed in linguistics. The "Appraisal Framework" [26] is a development of work in Systemic Functional Linguistics [19] and is concerned with interpersonal meaning in text—the negotiation of social relationships by communicating emotion, judgment and appreciation. In a previous approach, we proposed EmotiNet [6], a knowledge base containing commonsense knowledge on situations (seen as "action chains") and their associated emotions. The main difference between EmotiNet and the annotations made within the scope of the Appraisal framework in Linguistics is that the approach employing EmotiNet is also capable of detecting implicit statements of appraisals. The research done within this field aims at labeling more direct statements of appraisal.

2.3. Knowledge bases for NLP applications

As far as knowledge bases are concerned, many NLP applications have been developed using manually created knowledge repositories such as WordNet [17], Cyc (http://www.cyc.org), ConceptNet [24] or

SUMO (Suggested Upper Merged Ontology, http://www.ontologyportal.org/). Some authors tried to learn ontologies and relations automatically, using sources that evolve in time—e.g. Yago [50] which employs Wikipedia to extract concepts, using rules and heuristics based on the Wikipedia categories. Other approaches to knowledge base population were by Pantel and Ravichandran [32], and for relation learning Berland and Charniak [7]. DIPRE [8] and Snowball [1] label a small set of instances and create hand-crafted patterns to extract ontology concepts.

It has been shown that a great advantage of using ontologies is the easiness with which they are employable and extendable with external sources of knowledge, as well as to other languages [3].

3. Motivation and contribution

To illustrate the difficulty of detecting implicit expressions of emotion, we will start with a series of examples.

Given a sentence such as (1) "I am sad", an automatic system should label it with "sadness". Given this sentence, a system working at a lexical level would be able to detect the word "sad" (for example using WordNet Affect) and would correctly identify the emotion expressed as "sadness". But already a slightly more complicated example – (2) "I am not sad" – would require the definition of "inverse" emotions and the approach would no longer be straightforward. In the second example, although emotion words are present in the text, additional rules have to be added in order to account for the negation. Let us consider another example: (3) "I failed my exam", which should be labeled with "sadness" as well. A system working at a lexical level would find in the text no word that is directly related to this emotion. A method to overcome this issue is proposed by Liu et al. [25] and Cambria et al. [10]. The main idea behind these approaches is to acquire knowledge on the emotional effect of different concepts. In this manner, the system would know that "failing an exam" is something that produces "sadness". These approaches solve the problem of indirectly mentioning an emotion by using the concepts that are related to it instead. However, they only spot the emotion contained in separated concepts and do not integrate their interaction or the context in which they appear. If the example we considered is extended as in (4) "I failed my exam because I did not study enough", the emotion expressed is no longer "sadness", but most probably "guilt". As it can be noticed, even if there are concepts that according to our general knowledge express a certain emotion, their presence in the text cannot be considered as a mark that the respective sentence directly contains that emotion. Finally, the same situations are associated with distinct emotion labels depending on properties of the actor, action or object (e.g. "The man killed the mosquito" versus "The man killed his wife", or "The kitten climbed into my lap" versus "The pig climbed into my lap", or "The dog started barking as I approached" versus "The dog started wagging his tail as I approached"). The properties of the actors, actions and objects, as we can see, are very important at the time of determining the emotional label of a situation. These properties actually translate into the different values of the appraisal criteria. Therefore, a resource aimed for emotion detection must include this information. The quantity of commonsense knowledge required is tremendous. However, most of it is already present in existing commonsense knowledge bases (e.g. CYC, SUMO or ConceptNet). Given a sufficiently flexible model of representation for action chains, the underlying ontology can be enriched with such knowledge from external sources, resulting in a deep semantic representation of the situations, from which emotion can be detected in a more precise way.

In the light of these considerations, we proposed and implemented EmotiNet [6]—a knowledge base (KB) for modeling affect based on the appraisal theories, which supports the following:

1. The automatic processing of texts to extract:
   - The components of the situation presented (which we denote by "action chains") and their relation (temporal, causal, etc.);
   - The elements on which the appraisal is done in each action of the chain (agent, action, object);
   - The appraisal criteria that can automatically be determined from the text (modifiers of the action, actor, object in each action chain);
2. The inference on the value of the appraisal criteria, extracted from external knowledge sources (characteristics of the actor, action, object or their modifiers that are inferable from text based on commonsense knowledge);
3. The manual input of appraisal criteria of a specific situation.

Following the analysis of the results of our initial evaluations [6], we could draw a series of conclusions that are relevant for improving the performance of our approach. The analysis of the results obtained motivated the contributions brought by the present work.

The first conclusion we could draw is that the data employed in our study has specific characteristics, which must be described in order to make the approach useful. In this sense, a first contribution of the present work is to analyze the characteristics of the corpus employed in our study. Subsequently, we analyze the characteristics of the corpus with respect to those of the existing lexical resources that are used for emotion detection in NLP—WordNet Affect [48] and the emotion categories (anger, anxiety, sadness) in the Linguistic Inquiry and Word Count [LIWC; 35].

The second conclusion of our initial study is that there is a clear limitation of the supervised and lexical methods that are currently being used for emotion detection. However, there is no clear evaluation that shows to what extent they are useful and what is the performance of using them in the context of emotion detection in general and, more specifically, on the data we use for evaluation. Thus, our second contribution resides in testing two methods for emotion detection: a) one that is supervised, using Support Vector Machines Sequential Minimal Optimization [SVM SMO; 38] learning with uni-, bi- and trigrams and similarity of the examples among themselves; and b) another that is lexicon knowledge-based, which uses SVM SMO, but taking into account only the emotion-related words found in WordNet Affect and LIWC.

Thirdly, another important conclusion of our previous evaluations was that additional knowledge is required in order for the commonsense-based approach using EmotiNet to be successful. By analyzing the results obtained, we could see that the type of data needed can be obtained from two different sources. The first one is the information on the situations represented (i.e. we require more commonsense knowledge from existing repositories). The second source of knowledge we require is related to the surface realization of the textual presentation (i.e. the same situation can be described using different linguistic expressions). Thus, the third important contribution of this work is the extension of the knowledge contained in EmotiNet with new sources subsequent evaluation of the new, extended resource.

Finally, we comparatively analyze the performance of the methods presented and discuss their advantages and limitations.

4. ISEAR—a corpus of self-reported affect: dataset analysis

Self-reported affect is the most commonly used paradigm in Psychology to study the relationship between the emotional reaction and the appraisal preceding it [44].

In ISEAR, a corpus containing examples of self-reported affect, the student respondents, both psychologists and non-psychologists, were asked to report situations in which they had experienced all of 7 major emotions (joy, fear, anger, sadness, disgust, shame, and guilt). In each case, the questions covered the way they had appraised the situation and how they reacted. An example of entry in the ISEAR databank is as follows: "My mother treated me unjustly as if I was a little child. I said mean things. I was unable to defend myself adequately." Each example is attached to one single emotion (e.g. "anger" in the case of the previous example).

Table 1
Characteristics of the ISEAR examples used in our experiments.

Emotion	#examples	#tok	#uniq tok	#w W	#w L	#uniq W	#uniq L
Anger	174	5074	879	141	70	35	31
Disgust	87	2291	554	61	32	24	14
Fear	110	3525	669	96	52	33	18
Guilt	223	6903	967	184	79	49	36
Joy	76	1894	437	51	3	20	3
Sadness	292	6360	847	181	117	46	28
Shame	119	3299	640	67	30	34	18

Legend
#examples—Number of examples.
#tok—Number of tokens in all examples.
#uniq tok—Number of unique tokens in all examples.
#w W—Number of words in examples found in WordNet Affect.
#w L—Number of words in examples found in LIWC.
#uniq W—Number of unique words in examples found in WordNet Affect.
#uniq L—Number of unique words in examples found in LIWC.

In order to have a homogenous starting base, we selected from the 7667 examples in the ISEAR database only the ones that contained descriptions of situations involving family members (i.e. cases where the emotional bond between the participants in the situation is known). This resulted in a total of 1081 examples, of which we employed 175 to construct the core knowledge in EmotiNet, and the rest – 895 examples – to test the approach by Balahur et al. [6].[3] In order to study to what extend existing lexical knowledge-based and statistical methods can successfully be employed for this task, we have analyzed the corpus characteristics: number of examples per emotion, number of tokens and number of unique tokens, and, additionally, the number of words found in two of the most relevant resources for emotion detection – WordNet Affect and LIWC.

Table 1 presents the characteristics of the 1081 examples we previously employed by Balahur et al. [6].

As can be seen from Table 1, the number of words in the examples we previously used to test our approach [6] that can be found in WordNet Affect or LIWC is very small compared to the total number of words in the ISEAR dataset employed. In the following section, we present a set of experiments aimed at testing the performance of some of the most widely used methods in the context of these examples.

5. Preliminary experiments using the ISEAR dataset

In our initial evaluation of EmotiNet [6], we employed the set of 1081 examples from the ISEAR corpus described in Section 4, from which 175 examples (25 per emotion) were removed, due to the fact that they had been used to model the core of knowledge.

In order to test the performance of alternative methods for emotion detection, we will consider, on the one hand, the whole set of 1081 examples (which we denote by set A), as well as the reduced set of 895 examples which has been employed to test EmotiNet (test set B). The 175 examples used to build the initial core of knowledge in EmotiNet will be denoted as set T.

5.1. Emotion detection in text using lexical similarity

The first experiment we performed aimed at assessing if the similarity of the lexica used in the examples is high enough in order to produce a correct classification of the emotions described. In order to assess the similarity, we computed the Lesk distance between all examples in test set A using Ted Pedersen's Statistics package.[4] Subsequently, each of the examples in this set was represented as a vector, whose components were the similarities with all texts in test set A. We applied SVM SMO and performed a ten-fold cross-validation. The results are presented in Table 2.

Table 2
Results of ten-fold cross validation using SVM SMO and inter-example similarity as features on test set A.

Emotion	Precision	Recall	*F*-Measure
Anger	0.353	0.414	0.381
Disgust	0.292	0.241	0.264
Fear	0.482	0.491	0.486
Guilt	0.462	0.386	0.421
Joy	0.439	0.474	0.456
Sadness	0.707	0.76	0.733
Shame	0.441	0.412	0.426

Comparing these results with the ones previously obtained in the approach using EmotiNet [6], we can see that this approach has a similar performance. However, the knowledge contained in EmotiNet is only the one extracted by modeling the initial core—i.e. test set T. Therefore, the only just comparison that can be done is by repeating the previous experiment, but computing the similarity of examples only with the ones in test set T, using test set T for training and classifying the 895 examples in test set B. The results of these experiments are reported in Table 3.

Table 3
Results of classifying test set B using SVM SMO and inter-example similarity with test set T.

Emotion	Precision	Recall	*F*-Measure
Anger	0.259	0.282	0.270
Disgust	0.132	0.061	0.083
Fear	0.265	0.086	0.129
Guilt	0.272	0.335	0.300
Joy	0.143	0.203	0.168
Sadness	0.512	0.583	0.545
Shame	0.263	0.238	0.250

As we can see from the results in Table 3, the performance when training only on the examples which in fact are used as initial knowledge in EmotiNet drop dramatically.

5.2. Emotion detection in text using affect lexica

In order to test the appropriateness of using existing lexical resources for this task (i.e. WordNet Affect and LIWC), we subsequently performed a series of experiments in which we represented the examples in test set A, B and T as vectors whose features accounted for the presence of words from the two lexical resources and then applied SVM SMO. Due to space limitations, we only present the results obtained when combining the two vocabularies. Table 4 presents the results obtained when performing a ten-fold cross-validation on test set A. Table 5 presents the results obtained when training on set T and testing on set B.

Table 4
Results of ten-fold cross-validation on test set A using SVM SMO and the presence of words in WordNet Affect and LIWC as features.

Emotion	Precision	Recall	*F*-Measure
Anger	0.610	0.284	0.388
Fear	0.712	0.33	0.451
Disgust	0.692	0.202	0.313
Guilt	0.559	0.293	0.385
Joy	0.895	0.218	0.351
Sadness	0.336	0.895	0.489
Shame	0.500	0.066	0.117

[3] In our previous study, we employed a Semantic Role Labeling system [28] to extract the main components of the situation—actor, action, patient. In the case of 11 examples, the result of extracting these three roles was void.

[4] http://www.d.umn.edu/tpederse/text-similarity.html.

Table 5
Results of classification using SVM SMO and the presence of words in WordNet Affect and LIWC as features for test set B using set T as raining.

Emotion	Precision	Recall	*F*-Measure
Anger	0.405	0.201	0.269
Fear	0.457	0.165	0.242
Disgust	0.933	0.175	0.295
Guilt	0.207	0.772	0.326
Joy	0.204	0.172	0.135
Sadness	0.667	0.188	0.293
Shame	0.243	0.085	0.126

Table 6
Results of classification using ten-fold cross-validation with SVM SMO and n-grams as features for test set A.

Emotion	Run									
	Unigrams		Bigrams		Trigrams		u+b		u+b+t	
	P	R	P	R	P	R	P	R	P	R
Anger	0.38	0.42	0.37	0.28	0.38	0.16	0.41	0.37	0.44	0.38
Disgust	0.50	0.49	0.49	0.32	0.59	0.17	0.54	0.41	0.57	0.09
Fear	0.67	0.75	0.45	0.77	0.39	0.85	0.55	0.80	0.62	0.36
Guilt	0.47	0.53	0.47	0.55	0.43	0.48	0.50	0.59	0.48	0.59
Joy	0.42	0.35	0.38	0.23	0.39	0.16	0.49	0.37	0.71	0.26
Sadness	0.60	0.38	0.54	0.26	0.54	0.17	0.70	0.30	0.50	0.82
Shame	0.28	0.16	0.40	0.07	0.38	0.03	0.41	0.13	0.51	0.29

In this case, we can see that there is a significant drop in performance and that the results are comparatively lower than the ones obtained using EmotiNet.

5.3. Emotion detection in text using supervised learning with N-gram features

Finally, in the following set of experiments we performed, we represented each example as a feature vector, whose values (0 or 1) accounted for the presence of unigrams, bigrams, trigrams (separately) and jointly (unigrams and bigrams—u+b; unigrams, bigrams and trigrams—u+b+t). We extracted these five different representations for test set A and performed a ten-fold cross-validation in each case (Table 6). Subsequently, we extracted these five different representations for set T and B, using T as training and B as test set (i.e. the presence of n-grams was computed based on the vocabulary in T). Results of these evaluations are presented in Table 7.

In Figs. 1 and 2, we can see a graphical comparison between the results obtained using the different methods presented (ten-fold cross-validation using test set A and classification using set T as training and set B as test, respectively) versus the method employing EmotiNet.

Table 7
Results of classification of test set B using SVM SMO and n-grams as features with set T as training.

Emotion	Run									
	Unigrams		Bigrams		Trigrams		u+b		u+b+t	
	P	R	P	R	P	R	P	R	P	R
Anger	0.30	0.37	0.33	0.13	0.12	0.01	0.32	0.26	0.34	0.15
Disgust	0.42	0.25	0.40	0.04	0.40	0.02	0.50	0.11	0.44	0.04
Fear	0.36	0.33	0.29	0.96	0.27	0.98	0.32	0.93	0.29	0.98
Guilt	0.53	0.16	0.54	0.07	0.50	0.06	0.57	0.12	0.70	0.07
Joy	0.50	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sadness	0.30	0.21	0.20	0.05	0.00	0.00	0.33	0.10	0.27	0.06
Shame	0.09	0.04	0.17	0.01	0.50	0.01	0.08	0.01	0.00	0.00

As we can see, the approach based on commonsense performs at a comparable level to the one using knowledge extracted from the entire set A. In the cases where the knowledge employed for training is equal to the one in the EmotiNet core, the difference in performance is significant, all the other methods performing much below EmotiNet.

The results of these evaluations show that the approaches working at the word level are not capable of accurately detecting and classifying emotions from examples as the ones described in the ISEAR corpus. The fact that little or no lexical clues of affect are present in the text makes the performance of these methods inapplicable for the task of emotion detection. In these cases, additional knowledge is needed on the affective meaning of the described situation, such as the one contained in EmotiNet.

Further on, we present new experiments we performed using EmotiNet for emotion detection.

6. Experiments with the ISEAR corpus using EmotiNet

EmotiNet [6] is a KB aiming to be a resource for detecting emotions in text. EmotiNet captures and stores emotional reaction to real-world situations in which commonsense knowledge plays a significant role in the affective interpretation. Within the KB, each situation is specified as chains of actions and their corresponding emotional labels from several situations in such a way that it facilitates the extraction of general patterns of appraisal. Action chains are sequences of action links, or simply actions, that trigger an emotion on one or more subjects. In addition, each specific action link is described with a tuple (actor, action type, patient, emotional reaction).

The process followed in the development of EmotiNet is comprised of the following stages: (1) the design of the EmotiNet ontology, which specifies the main concepts, properties and relations managed by the KB; (2) the extension and population of this ontology using the situations stored in ISEAR database; and (3) the expansion of the EmotiNet KB using existing commonsense knowledge bases—ConceptNet and other resources—VerbOcean [12].

6.1. Design of the ontology

The process of building the core of action chains of the EmotiNet KB started with the design of the EmotiNet ontology, which was specifically divided into three subprocesses:

1. Establish the scope and purpose of the ontology. The EmotiNet ontology captures and manages knowledge from three domains: kinship relations, emotions (and their relations) and actions (characteristics and relations between them).
2. Reuse knowledge from existing ontologies. Two existing ontologies from the Web were reused in the Emotinet core: the ReiAction ontology,[5] which represents actions between entities in a generic manner, and the family ontology,[6] which contains knowledge about family members and the relations between them.
3. Create the final ontology. This third stage involved the design of the Emotion ontology and the combination of the different knowledge sources into a single ontology. The Emotion ontology describes emotions and their relationships according to Robert Plutchik's wheel of emotion [39] and Parrot's tree-structured list of emotions [33]. In particular, this ontology includes different types of relations between emotions and a collection of specific instances of emotion (e.g. anger, joy or surprise). Finally, the mentioned ontologies were combined into the EmotiNet ontology by means of a set of new classes and relations that interconnect the components reused from them (Fig. 3).

[5] http://www.cs.umbc.edu/lkagal1/rei/ontologies/ReiAction.owl.
[6] http://www.dlsi.ua.es/jesusmhc/emotinet/family.owl.

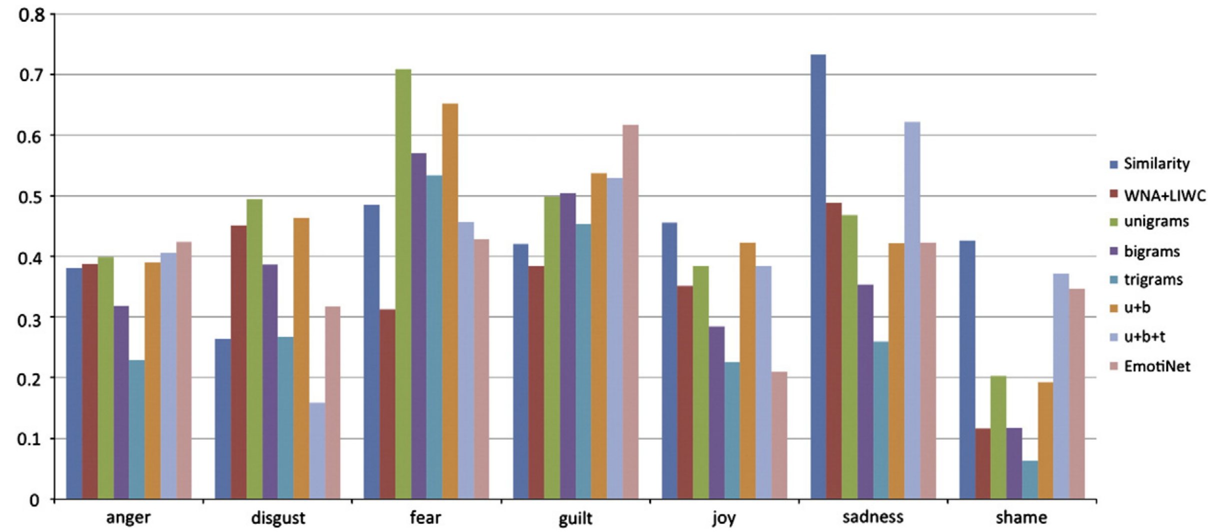

Fig. 1. Comparison between the *F*-measure results obtained when classifying test set A using the different methods presented versus EmotiNet.

6.2. Extension and population of the ontology

The extension and population of the EmotiNet ontology was carried out using the situations contained in the ISEAR database as examples (test set T, described in Section 4).

The first step was to extract the actions chains described in each of the documents in test set T. This task was supported by Semrol, the semantic role labeling (SRL) system introduced by Moreda et al. [28]. The initial core of knowledge in the EmotiNet KB needs 100% accurate information, and, therefore, for each situation or action, the agent, the verb and the patient (the surface object of the verb) were manually extracted from the output given by Semrol.

Further on, anaphoric expressions were automatically resolved using a heuristic selection of the family member mentioned in the text that is closest to the anaphoric reference and whose properties (gender, number) are compatible with the ones of the reference. In case of ambiguity, we choose the youngest, female member. Finally, the action links (triplets) were grouped and sorted into action chains. This process of sorting is determined by the adverbial expressions (e.g. “although”, “because” or “when”) that appear within the sentence, which actually specify the position of each action on a temporal line. A set of pattern rules were defined to establish the actual order of the actions surrounding these modifiers, i.e. which one happens prior to or after the current context.

Finally, using our combined emotion model as a reference, all the action links obtained were manually associated to one of the seven most basic emotions of ISEAR or to the neutral value, thus generating ordered sequences of 4-tuples (actor, action, object, emotion) or action chains.

Once this process was applied to the chosen documents, 175 action chains were obtained (ordered lists of tuples). In order to be included in the EmotiNet knowledge base, all their action links needed to be mapped to existing concepts or instances within the KB. When these did not exist, they were added to the KB. In EmotiNet, each tuple (actor, action, object, emotion) extracted has its own representation as an instance of the subclasses of Action. Each instance of Action is related to an instance of the class Feel, which represents the emotion felt in this action. Subsequently, these instances (action links) were grouped in sequences of actions (using instances of the class Sequence) ended by an instance of the class Feel, which determines the final emotion felt by the main actor(s) of the chain.

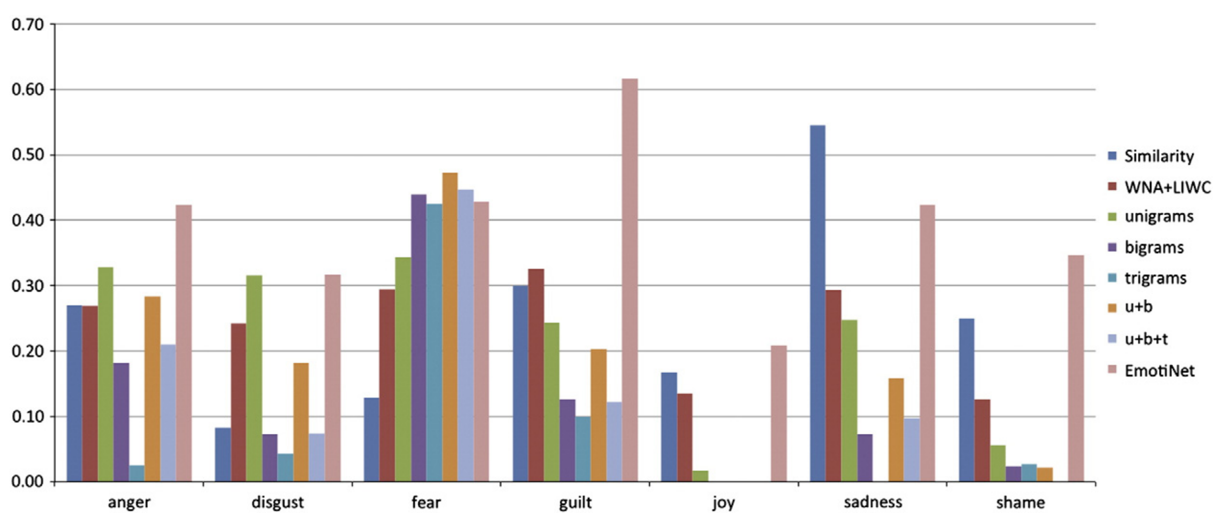

Fig. 2. Comparison between the *F*-measure results obtained when classifying test set B using the different methods presented versus EmotiNet.

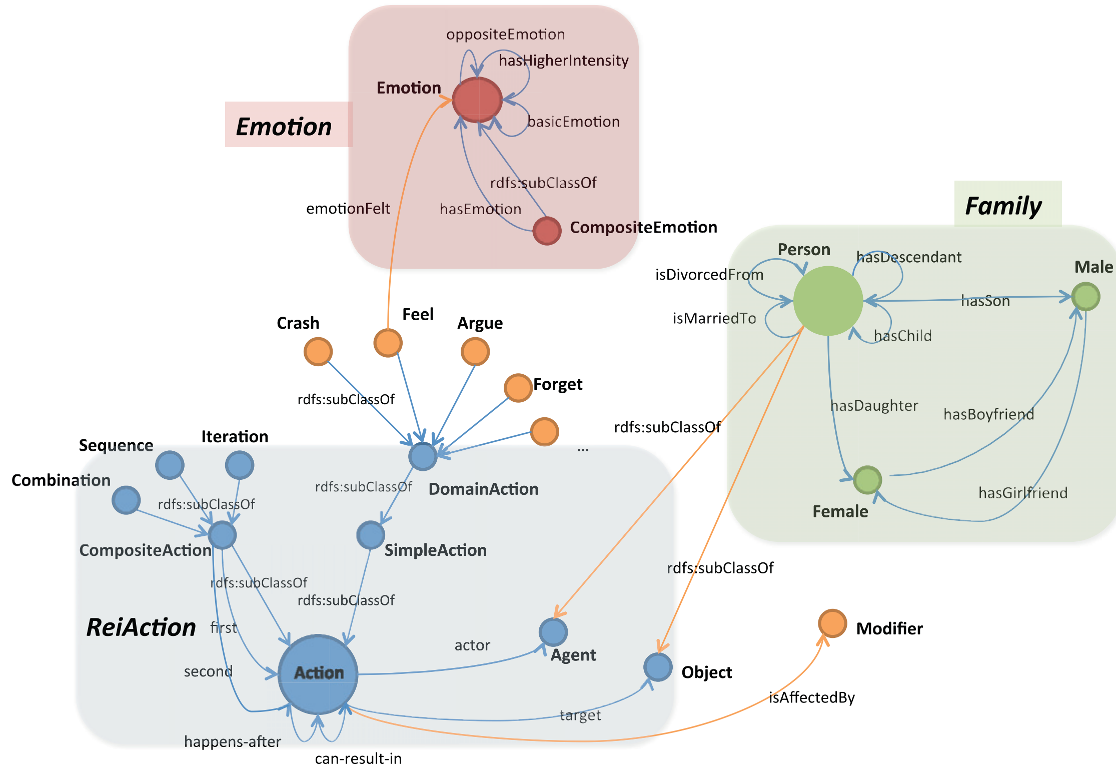

Fig. 3. Main concepts and relations of EmotiNet (RDF-like schema).

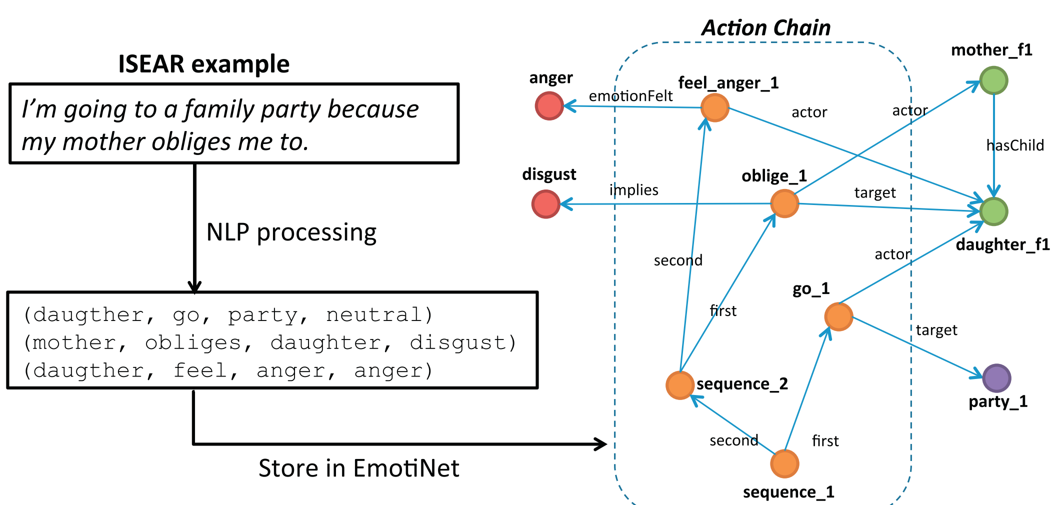

Fig. 4. RDF graph of an action chain and the process to obtain it.

Fig. 4 shows an example of a RDF graph of an action chain. The frameworks chosen to manage and store EmotiNet on a database were Jena and MySQL.[7]

6.3. Ontology expansion

In order to extend the coverage of the resource, the ontology was iteratively expanded with new types of actions and relations between actions from existing resources. This process is essential for EmotiNet, because it adds new elements that are not included in ISEAR and, thus, have not been considered before. In this manner, the degree of dependency between the resource and the initial set of ISEAR examples can be reduced and, as a consequence, the resource can be used to analyze unseen situations from the same domain (family) or other similar domains.

Specifically, this process extended the EmotiNet ontology by adding new actions that were similar to the ones included in the core. The new set of actions was obtained from three existing resources: VerbOcean [12], "Core" WordNet [40] and WordNet Affect. In order to effectively carry out the task, it was considered that verbs represent the essence of actions, so that the verbs contained in these resources can be considered as EmotiNet actions. New actions were included in EmotiNet as subconcepts of the class DomainAction and related to the initial

[7] http://jena.sourceforge.net/.

Table 8
Degree of overlap between resources, measured in terms of number of action types.

	# Actions	Overlap (# actions)				
		EmotiNet	VO	CWN	WNA	Unique
EmotiNet	143	*	83	100	7	28
VO	782	83	*	466	35	288
CWN	2230	100	466	*	51	1702
WNA	174	7	35	51	*	109

Unique: Number of actions only contained in a resource.
VO: VerbOcean.
CWN: "Core" WordNet.
WNA: WordNet Affect.

Table 9
Number of situations in which the emotion was correctly detected and number of situation for which an emotion was found in the first collection of experiments (1a–e).

Emotion	Total	Run									
		EmotiNet		EN + VO		EN + WNA		EN + CWN		EN + All	
		Ok	Result	Ok	Result	Ok	Result	Ok	Result	Ok	Result
Anger	174	72	132	84	157	72	132	73	150	86	159
Disgust	86	28	73	40	82	28	73	31	78	40	82
Fear	110	24	80	19	90	24	80	24	86	20	90
Guilt	222	61	199	61	211	61	199	60	206	60	212
Joy	76	33	55	30	58	33	55	34	57	32	59
Sadness	292	70	178	71	197	70	178	70	197	70	206
Shame	119	51	99	52	113	51	99	51	104	52	113

EmotiNet action set by means of a new ontology relationship: similarAction. Each resource defines the similarity between actions using different mechanisms. VerbOcean explicitly contains and manages the relationship of similarity (called similar) between verbs. "Core" WordNet and WordNet Affect follow the same structure as WordNet, i.e. extracting similar verbs is reduced to obtaining those verbs that are in the same synset. Due to this fact, the mapping between the similarAction EmotiNet relations and the mechanisms employed in the rest of the resources is direct. The reason for using two different versions of WordNet is that each of them is aimed for a specific application and thus contains different collections of verbs. Instead of using the whole WordNet, with its known problems of ambiguity and granularity, these reduced versions can provide a simplified view of the most frequently used verbs with their usual semantics for different tasks.

Table 8 shows a comparison between the resources used to expand the EmotiNet ontology and the ontology itself. It also illustrates the degree of overlap that exists between each resource, in order to clarify the contribution of each resource to the resulting ontology. Note that the column Unique contains the number of actions that are uniquely present in that specific resource and not included in the rest.

6.4. Experiments with EmotiNet

In the set of experiments carried out with EmotiNet, we assessed the performance of the task of emotion detection in text using EmotiNet and analyzed the impact of the different resources used in its expansion on the final results. These experiments were divided into two collections and were aimed at improving the performance of the results we previously obtained using EmotiNet [6]. The processes carried out in each group are illustrated in Fig. 5:

a) Experiments using the EmotiNet action chains. In the first collection of experiments, once the action chains are extracted from the input texts, we compute their similarity with those contained in EmotiNet. The resulting emotion has the same label as the EmotiNet action chain with the highest similarity score. When an action found in the text is not contained in EmotiNet, we use the ontology relations to the actions imported from VerbOcean (VO), "Core" WordNet (CWN) and WordNet Affect (WNA).
b) Experiments using the EmotiNet emotion tags. This second set of experiments is based on the use of the "implies" relationship, which associates an action to the possible emotions felt by the agents of that action. We have performed different experiments in which we have automatically annotated the actions contained in EmotiNet using two existing resources: LIWC and WordNet Affect. After extracting the action chains from a new real situation (described in a text), we used these annotations to obtain the emotions associated to each of its individual action links using EmotiNet.

The following subsections describe each of the experiments and the obtained results.

6.4.1. Assessing the impact of extending EmotiNet with other resources

In the first collection of experiments, we calculated the similarity between the action chains obtained from the ISEAR corpus and the action chains contained in EmotiNet. Each experiment (or run) used different EmotiNet relationships to obtain similar actions in case the exact action was not contained in the initial version of EmotiNet. Each type of EmotiNet relationship links the original EmotiNet actions to the actions imported from one or more specific resources, i.e. VerbOcean (similar relation), "Core" WordNet (CWN_similar relation) and WordNet Affect (wna_relation relation). Specifically, five experiments were designed and executed: (1a) use the initial core of EmotiNet, which establishes a baseline for the rest (EmotiNet); (1b) use similar actions from

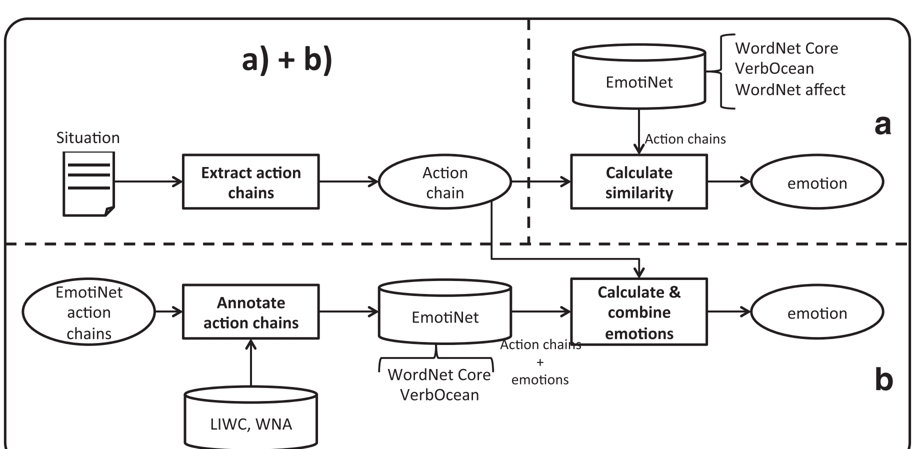

Fig. 5. Activities carried out during the experiments.

Table 10
Precision and recall for each run of the first collection of experiments (1a–e).

Emotion	Run									
	EmotiNet		EN + VO		EN + WNA		EN + CWN		EN + All	
	P	R	P	R	P	R	P	R	P	R
Anger	54.54	41.37	53.50	48.27	54.54	41.37	48.66	41.95	54.08	49.42
Disgust	38.35	32.55	48.78	46.51	38.35	32.55	39.74	36.04	48.78	46.51
Fear	30.00	21.81	21.11	17.27	30.00	21.81	27.90	21.81	22.22	18.18
Guilt	30.65	27.47	28.90	27.47	30.65	27.47	29.12	27.02	28.30	27.02
Joy	60.00	43.42	51.72	39.47	60.00	43.42	59.64	44.73	54.23	42.10
Sadness	39.32	23.97	36.04	24.31	39.32	23.97	35.53	23.97	33.98	23.97
Shame	51.51	42.85	46.01	43.69	51.51	42.85	49.03	42.85	46.01	43.69

Table 11
Average precision and recall for each run of the first collection of experiments (1a–e).

Emotion	Average Run—EmotiNet	
	Precision	Recall
Anger	53.06	44.48
Disgust	42.80	38.83
Fear	26.25	20.18
Guilt	29.52	27.29
Joy	57.12	42.63
Sadness	36.84	24.04
Shame	48.81	43.19
Average	42.06	34.38

VerbOcean (EN+VO); (1c) use similar actions from "Core" WordNet (EN+CWN); (1d) use similar actions from WordNet Affect (EN+WNA); (1e) use similar actions from all the resources (EN+All).

The results obtained in this set of experiments over the ISEAR corpus, described in the previous sections, are illustrated in Tables 9–11 and in Fig. 6. Table 9 shows the number of situations for which the experiment correctly resolved its associated emotion (columns "Ok" for each run). The columns called "Result" contain the number of cases in which it was possible to obtain a result, i.e. to associate an emotion to the input situation. Table 10 contains the values of precision and recall from the same group of experiments and Table 11 shows the average values of the same parameters. The line chart illustrated In Fig. 6 depicts the *F*-measure values for each run.

6.4.2. *Assessing the impact of annotating the action links of EmotiNet with existing resources*

As mentioned before, in the second collection of experiments, we used EmotiNet and a different method for detecting emotions from text. This method obtained the emotions associated to each action link and further on combined them by means of the Emotion core of the EmotiNet ontology through a voting process. In order to retrieve the emotion associated to each action link, the method employs the infer relationship from EmotiNet. In order to carry out this collection of experiments, we previously generated different versions of EmotiNet. In each of these versions, the infer action–emotion relations were automatically extracted using two well-known resources: a) the LIWC dictionary, and, more specifically, three word categories from it, i.e. Anx (LIWC code 128), Anger (LIWC code 129), Sad (LIWC code 130); and b) WordNet Affect. However, these do not cover all the emotions considered by ISEAR. LIWC only contains words associated to anxiety (as a subtype of fear), anger and sadness, and the elements of WNA are only related to five emotions: anger, disgust, fear, joy and sadness. As in the first collection of evaluations, these experiments were carried out using the initial EmotiNet core and the relations of action similarity, in this case, for VerbOcean and "Core" WordNet.

Specifically, five experiments were designed and executed: (2a) use the initial core of EmotiNet annotated with LIWC (LIWC); (2b) use the initial core of EmotiNet annotated with WNA (WNA); (2c) use the initial core of EmotiNet annotated with LIWC and WNA (LIWC+WNA), including the emotions of both resources without resolving possible conflicts; (2d) use the initial core of EmotiNet annotated with LIWC and WNA and similar actions from VerbOcean (L+WA+V); (2e) use the initial core of EmotiNet annotated with LIWC and WNA and similar actions from "Core" WordNet (L+WA+CW); (2f) use the initial core of EmotiNet annotated with LIWC and WNA and similar actions from VerbOcean and "Core" WordNet (L+WA+V+CW).

The results for this second collection of experiments are shown in Table 12 ("P" stands for "Precision" while "R" for "Recall") (Fig. 7). The column "Avg. (2c–f)" contains the average values for the experiments (2c) to (2f) and graphically represented as far as F-Measure is concerned in Fig. 7.

As expected, the combined approach (LIWC+WNA) has a better coverage of the emotions, leading to a higher recall. However, after performing a shallow analysis of the results, we concluded that, in some cases, the performance of the method is higher when EmotiNet is annotated only with the LIWC approach. Therefore, we carried out additional experiments in order to evaluate the performance of EmotiNet when it is only annotated using LIWC: (2g) use the initial core of

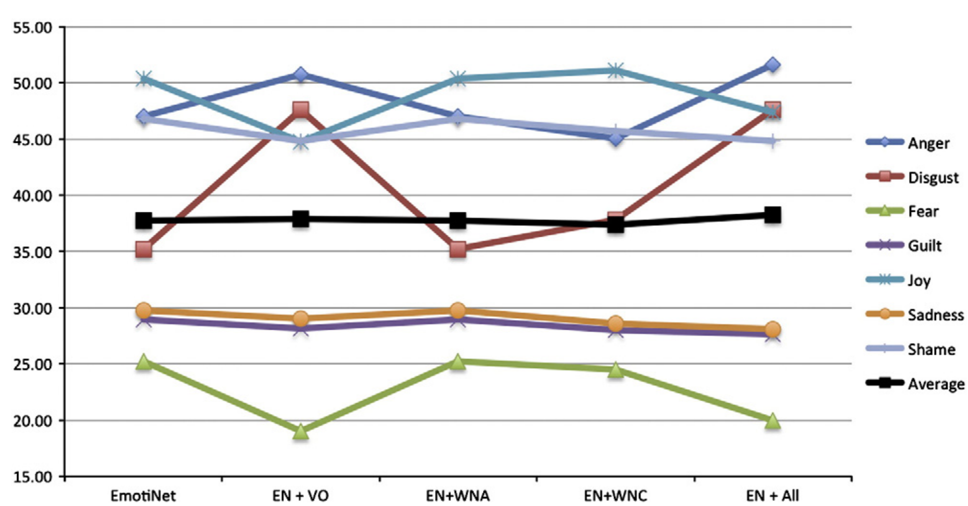

Fig. 6. *F*-measure values for each run and emotion of the first collection of experiments (1a–e).

Table 12
Precision and recall for each run of the second collection of experiments (2a–f).

Emotion	Run					
	LIWC		WNA		LIWC + WNA	
	P	R	P	R	P	R
Anger	64.70	12.64	16.21	3.44	36.20	12.06
Disgust	0.00	0.00	37.50	10.46	32.14	10.46
Fear	34.78	7.27	0.00	0.00	25.80	7.27
Joy	0.00	0.00	50.00	5.26	44.44	5.26
Sadness	78.04	10.95	11.42	1.36	48.52	11.30

	L + WA + V		L + WA + CW		L + WA + CW + V		Avg. (2c–f)	
	P	R	P	R	P	R	P	R
Anger	32.75	21.83	39.75	18.96	36.80	26.43	36.38	19.82
Disgust	10.76	8.13	18.18	9.30	8.82	6.97	17.48	8.72
Fear	11.53	8.18	15.68	7.27	10.71	8.18	15.93	7.73
Joy	57.57	25.00	82.35	18.42	75.67	36.84	65.01	21.38
Sadness	32.03	14.04	42.38	13.35	30.21	14.38	38.29	13.27

EmotiNet annotated with LIWC and similar actions from VerbOcean (LIWC + VO); (2h) use the initial core of EmotiNet annotated with LIWC and similar actions from "Core" WordNet (LIWC + CWN); (2i) use the initial core of EmotiNet annotated with LIWC and similar actions from VerbOcean and "Core" WordNet (L + V + CW).

The results are depicted in Table 13 ("P" stands for "Precision" and "R" for "Recall"). In order to compare the performance of using the combined approach (LIWC + WNA) to the one using LIWC, Fig. 8 illustrates a bar chart representing the precision, recall and *F*-measure values for each run.

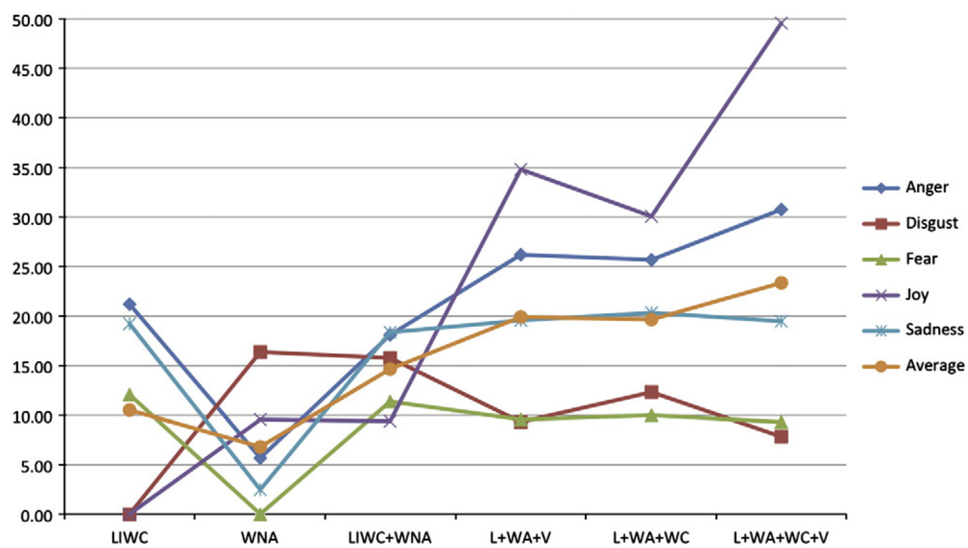

Fig. 7. *F*-measure values for each run and emotion in the second collection of experiments (2a–f).

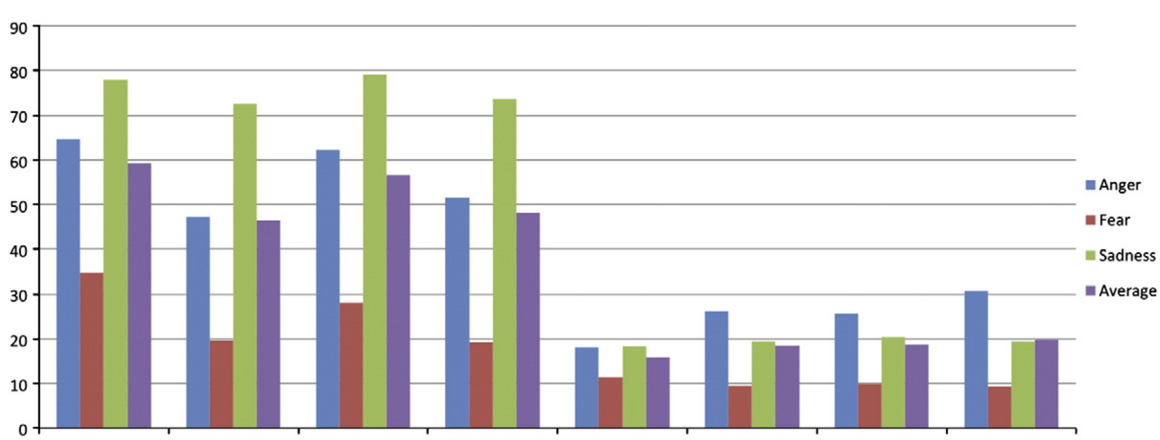

Fig. 8. *F*-measure values for each run and emotion in the second collection of experiments (2a–i).

Table 13
Precision and recall for each run of the experiments performed with LIWC (2a,g–h).

Emotion	Run									
	LIWC		LIWC + VO		LIWC + CWN		L + VO + CWN		Average	
	P	R	P	R	P	R	P	R	P	R
Anger	64.70	12.64	47.25	24.71	62.26	18.96	51.48	29.88	56.42	21.55
Fear	34.78	7.27	19.60	9.09	28.12	8.18	19.29	10.00	25.45	8.64
Sadness	78.04	10.95	72.52	22.60	79.16	13.01	73.68	23.97	75.85	17.63
Average	59.17	10.29	46.46	18.80	56.51	13.38	48.15	21.28	52.57	15.94

Table 14
Results from the combination of experiments (1e) and (2f).

Emotion	Total (#)	Correct (#)	Result (#)	Precision (%)	Recall (%)	*F*-Measure (%)
Anger	174	100	159	62.89	57.47	60.06
Disgust	86	43	83	51.80	50.00	50.88
Fear	110	27	95	28.42	24.54	26.34
Guilt	222	60	214	28.03	27.02	27.52
Joy	76	46	59	77.96	60.52	68.14
Sadness	292	98	206	47.57	33.56	39.36
Shame	119	52	114	45.61	43.69	44.63
Average	154.14	60.86	132.86	48.90	42.40	45.27

6.4.3. Experiments on the combinations of best-performing methods

Finally, we decided to perform another experiment, which combines the two methods with the best performance (in terms of average *F*-measure) from the first and second collection of experiments, i.e. (1e) `EN+All` and (2f) L + WA + V + CW. For the cases in which the methods obtained different values, the final value that was assigned was that from experiment (1e). In this last evaluation, we did not consider the LIWC experiments, because the vocabulary contained in these resources does not cover the majority of emotions.

The results for this last experiment are represented in Table 14 and Fig. 9.

7. Discussion and conclusions

From the results obtained in the initial evaluation of EmotiNet [6], as well as the experiments with EmotiNet versus well-established methods we have presented herein, we can conclude that the task of emotion detection from texts such as the ones in the ISEAR corpus (where little or no lexical clues of affect are present) can be best tackled using approaches based on commonsense knowledge. In this sense, EmotiNet, apart from being a precise resource for classifying emotions in such examples, has the advantage of being extendable with external sources, thus increasing the recall of the methods employing it.

With the extensive evaluations we have performed, we have shown that by using EmotiNet, even with a small quantity of knowledge, we obtain comparable results to the methods that employ supervised learning on a much greater training set or lexical knowledge. From the comparisons among the different settings and experiments, we can conclude that the approach using EmotiNet is valid and represents a method that is appropriate for the detection of emotions from contexts where no affect-related words are present.

Nonetheless, much remains to be done to fully exploit the capabilities of EmotiNet. We showed that the approach has a high degree of flexibility, i.e. new information can be easily introduced from existing commonsense knowledge bases, such as ConceptNet [6], mainly due to its internal structure and degree of granularity. Further on, the recall of the resource can also be increased by adding supplementary information on the surface realization of the texts that are analyzed. As such, we have shown that by adding knowledge from lexical resources (WordNet Affect, LIWC), we were able to further increase the performance of the approach using EmotiNet.

The error analysis we performed shed some light on the causes of error of the system. The first finding is that extracting only the action, verb and patient semantic roles is not sufficient. New elements, such as the modifiers, should be added in order to increase the expressivity of the resource and make it capable of distinguishing among situations in which the surface realization is very similar, but in which the presence of a modifier notably changes the nuance of the text and consequently, its emotional label. Therefore, such modifiers should be included as attributes of the concepts identified in the roles.

A further source of errors remained the lack of knowledge on specific actions. Therefore, the knowledge in EmotiNet must be further extended using existing knowledge bases or applying automatic methods that have been proven successful in other approaches for knowledge base population.

Finally, other errors were produced by NLP processes and propagated at various steps of the processing chain (e.g. SRL, coreference resolution). Some of these errors cannot be eliminated; however, a thorough study should be performed in order to detect to what extend these issues can be solved by using alternative NLP tools. Additionally, alternative (manual) evaluations should be performed to assess the quality of the imported knowledge, without taking into consideration the errors introduced in the processing and evaluation chain.

Future work aims at extending the model by adding properties to the concepts included, so that more of the appraisal criteria can be introduced in the model, testing new methods to assign affective value to the concepts and adding new knowledge from sources such as CYC. Additionally, we intend to expand the knowledge in EmotiNet to other

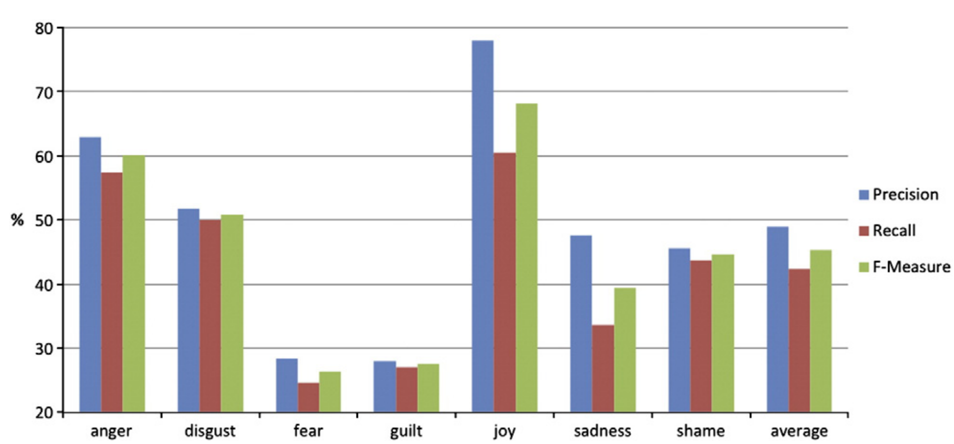

Fig. 9. *F*-measure values for the combined methods.

languages and domains, making it a reliable resource for emotion detection from any type of text.

Acknowledgements

This paper has been supported by the Spanish Ministry of Science and Innovation (grant no. TIN2009-13391-C04-01), by the Spanish Ministry of Education under the FPU Program (AP2007-03076), and by the Valencian Ministry of Education (grant no. PROMETEO/2009/119 and ACOMP/2010/288).

References

[1] E. Agichtein, L. Gravano, Snowball: Extracting relations from large plain-text collections, in: 85–94 (Ed.), Proceedings of the 5th ACM International Conference on Digital Libraries (ACM DL), 2000.
[2] A. Andreevskaia, S. Bergler, Clac and clac-nb: knowledge-based and corpus-based approaches to sentiment tagging, Proceedings of the 4th International Workshop on Semantic Evaluations. SemEval '07, Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, pp. 117–120, URL, http://dl.acm.org/citation.cfm?id=1621474.1621496.
[3] S. Aviv, A. Gal, Enhancing portability with multilingual ontology-based knowledge management, Decision Support Systems 45 (3) (2008) 567–584.
[4] A. Balahur, A. Montoyo, Applying a culture dependent emotion triggers database for text valence and emotion classification, Proceedings of the AISB 2008 Convention "Communication, Interaction and Social Intelligenc", 2008.
[5] A. Balahur, R. Steinberger, Rethinking opinion mining in newspaper articles: from theory to practice and back, Proceedings of the first workshop on Opinion Mining and Sentiment Analysis (WOMSA 2009), 2009.
[6] A. Balahur, J.M. Hermida, A. Montoyo, Building and exploiting emotinet, a knowledge base for emotion detection based on the appraisal theory model, IEEE Transactions on Affective Computing 99 (2011) (PrePrints).
[7] M. Berland, E. Charniak, Finding parts in very large corpora, Proceedings of ACL 1999, 1999, pp. 57–64.
[8] S. Brin, World-extracting patterns and relations from the wide web, Proceedings of the 1998 International Workshop on Web and Databases (WebDB 1998), 1998, pp. 172–183.
[9] R.A. Calvo, S. D'Mello, Affect detection: an interdisciplinary review of models, methods and their applications, IEEE Transactions on Affective Computing 1 (1) (2010).
[10] E. Cambria, E. Hussain, C. Havasi, C. Eckl, Affective space: blending common sense and affective knowledge to perform emotive reasoning, Proceedings of the 1st Workshop on Opinion Mining and Sentiment Analysis (WOMSA), 2009.
[11] F.-R. Chaumartin, Upar7: a knowledge-based system for headline sentiment tagging, Proceedings of the 4th International Workshop on Semantic Evaluations. SemEval '07, Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, pp. 422–425, URL, http://dl.acm.org/citation.cfm?id=1621474.1621568.
[12] T. Chklovski, P. Pantel, Verbocean: mining the web for fine-grained semantic verb relations, Proceedings of EMNLP 2004, 2004, pp. 33–40.
[13] T. Danisman, A. Alpkocak, Feeler: emotion classification of text using vector space model, Proceedings of the AISB 2008 Convention "Communication, Interaction and Social Intelligence", 2008.
[14] J. De Rivera, A structural theory of the emotions, Psychological Issues, Monograph 40 10 (4) (1977).
[15] M. Dyer, Emotions and their computations: three computer models, Cognition and Emotion 1 (1) (1987) 323–347.
[16] A. Esuli, F. Sebastiani, Determining the semantic orientation of terms through gloss analysis, Proceedings of CIKM 2005, 2005, pp. 617–624.
[17] C. Fellbaum, WordNet: An Electronic Lexical Database, MIT Press, Cambridge, MA, 1998.
[18] N. Frijda, The Emotions, Cambridge University Press, 1986.
[19] M.A.K. Halliday, An Introduction to Functional Grammar, Edward Arnold, London, 1994.
[20] P.N. Johnson-Laird, K. Oatley, The language of emotions: an analysis of a semantic field, Cognition and Emotion (3) (1989) 81–123.
[21] P. Katz, M. Singleton, R. Wicentowski, Swat-mp: the semeval-2007 systems for task 5 and task 14, Proceedings of the 4th International Workshop on Semantic Evaluations. SemEval '07, Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, pp. 308–313, URL, http://dl.acm.org/citation.cfm?id=1621474.1621541.
[22] Z. Kozareva, B. Navarro, S. Vázquez, A. Montoyo, Ua-zbsa: a headline emotion classification through web information, Proceedings of the 4th International Workshop on Semantic Evaluations. SemEval '07, Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, pp. 334–337, URL, http://dl.acm.org/citation.cfm?id=1621474.1621546.
[23] R.S. Lazarus, C.A. Smith, Knowledge and appraisal in the cognition–emotion relationship, Cognition and Emotion (2) (1988) 281–300.
[24] H. Liu, P. Singh, Conceptnet: a practical commonsense reasoning toolkit, BT Technology Journal 22 (2004).
[25] H. Liu, H. Lieberman, T. Selker, A model of textual affect sensing using real-world knowledge, Proceedings of IUI 2003, 2003.
[26] J.R. Martin, P.R. White, Language of Evaluation: Appraisal in English, Palgrave Macmillan, 2005.
[27] S. Mei Lee, Y. Chen, C. Huang, Cause event representations of happiness and surprise, Proceedings of PACLIC 2009, 2009.
[28] P. Moreda, B. Navarro, M. Palomar, Corpus-based semantic role approach in information retrieval, Data and Knowledge Engineering (DKE) 61 (3) (2007) 467–483.
[29] A. Ortony, G. Clore, A. Collins, The Cognitive Structure of Emotions, Cambridge University Press, New York, 1988.
[30] B. Pang, L. Lee, Opinion mining and sentiment analysis, Foundations and Trends in Information Retrieval 2 (1–2) (2008).
[31] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up? Sentiment classification using machine learning techniques, Proceedings of EMNLP 2002, 2002, pp. 79–86.
[32] P. Pantel, D. Ravichandran, Automatically labeling semantic classes, Proceedings of HLT-NAACL-04, 2004, pp. 321–328.
[33] W.E. Parrott, Emotions in Social Psychology, Psychology Press, 2001.
[34] J. Pennebaker, M.R. Mehl, K. Niederhoffer, Psychological aspects of natural language use: our words, our selves, Annual Review of Psychology (54) (2003) 547–577.
[35] J. Pennebaker, R. Booth, M. Francias, Linguistic Inquiry and Word Count. LIWC 2007, 2007.
[36] R. Picard, Affective Computing. Vol. Technical Report, MIT Media Laboratory, 1995.
[37] R. Picard, Affective Computing, MIT Press, Cambridge, 1997.
[38] J.C. Platt, Fast Training of Support Vector Machines Using Sequential Minimal Optimization, MIT Press, Cambridge, MA, USA, 1999, pp. 185–208, URL, http://dl.acm.org/citation.cfm?id=299094.299105.
[39] R. Plutchik, The nature of emotions, American Scientist 89 (4) (2001) 344.
[40] Princeton University, "Core" WordNetURL, http://wordnetcode.princeton.edu/standoff-files/core-wordnet.txt2011.
[41] E. Riloff, J. Wiebe, T. Wilson, Learning subjective nouns using extraction pattern bootstrapping, Proceedings of the Conference on Natural Language Learning (CoNLL), 2003, pp. 25–32.
[42] K. Scherer, Handbook of Cognition and Emotion, John Wiley & Sons Ltd., 1989 Ch. Appraisal Theory.
[43] K.R. Scherer, Studying the emotion-antecedent appraisal process: an expert system approach, Cognition and Emotion 7 (3–4) (1993).
[44] K. Scherer, Toward a dynamic theory of emotion. The component process of affective states, Cognition and Emotion 1 (1) (2001).
[45] K. Scherer, What are emotions? And how can they be measured? Social Science Information 3 (44) (2005) 695–729.
[46] K. Scherer, H. Wallbott, The ISEAR Questionnaire and Codebook, 1997.
[47] C. Strapparava, R. Mihalcea, Semeval-2007 task 14: affective text, Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007), Association for Computational Linguistics, Prague, Czech Republic, June 2007, pp. 70–74, URL, http://www.aclweb.org/anthology/S/S07/S07-1013.
[48] C. Strapparava, A. Valitutti, WordNet-Affect: an affective extension of WordNet, Proceedings of LREC, vol. 4, 2004, pp. 1083–1086.
[49] P. Subasic, A. Huettner, Affect analysis of text using fuzzy semantic typing, IEEE Transactions on Fuzzy Systems (9) (2000) 483–496.
[50] F. Suchanek, G. Kasnei, G. Weikum, Yago: a core of semantic knowledge unifying WordNet and Wikipedia, Proceedings of the World Wide Web Conference 2007, 2007, pp. 697–706.
[51] R.E. Thayer, Calm Energy: How People Regulate Mood With Food and Exercise, Oxford University Press, N.Y., 2001.
[52] G. Van den Bos, APA Dictionary of Psychology, American Psychological Association, Washington, DC, 2006.
[53] J. Wiebe, E. Riloff, Creating subjective and objective sentence classifiers from unannotated texts, Proceedings of the 6th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing), 2005, pp. 73–99.

Dr. Alexandra Balahur obtained her PhD in Computer Science (CS) from the University of Alicante (2011). She also obtained her Diploma of Advanced Studies in CS from the same university (2009). Alexandra is a CS graduate of the "Al. I. Cuza" University of Iasi, Romania (2007). Her main fields of interest are sentiment analysis, Emotion Detection, Information Extraction and Textual Entailment. She is the author of over 40 scientific publications, in book chapters, international reviews and conference proceedings. She is/was a member of the implementation team in different national and international projects containing interdisciplinary research on the impact of technology on social phenomena. She has co-organized the two editions of the Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, held with ECAI 2010 and ACL-HLT 2011.

Jesús M. Hermida received a 5-year degree in Computer Engineering from the University of Alicante, Spain (2007). He also obtained the Diploma of Advanced Studies in the area of Software and Computing Systems at the same university (2009). At present, he is a PhD student at the University of Alicante, supported by the FPU program of the Spanish Ministry of Education. His current research interests include the application of Semantic Web technologies, techniques and resources (especially ontologies) to different areas, such as Natural Language Processing, model-driven development of Web applications, Rich Internet Applications or knowledge representation in the health domain.

Prof. Dr. Andrés Montoyo is professor of Databases at the Technical School and of Ontology Design in Natural Language Processing and Semantic Web at the Masterate in Computing Technologies at the University of Alicante. He received his Bachelor and Master's degree in Computer Science from the Polytechnic University of Valencia and his PhD in Computer Science from the University of Alicante, Spain. His research interests include Information Extraction, Word Sense Disambiguation, Opinion Mining, Ontologies, and the Semantic Web. He is the author of more than 70 scientific publications in international journals and conferences on many topics, including sentiment analysis and opinion mining. He has been involved both in national and international research projects and co-organized several conferences and workshops on these topics.