Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση

Τόμ. 1 (2016)

10ο Πανελλήνιο και Διεθνές Συνέδριο «Οι ΤΠΕ στην Εκπαίδευση»

Detecting emotions in Greek sentences: towards an open-source affect-awareness framework

Ioannis Panatsias, Hippokratis Apostolidis, Stergios Tegos, Thrasyvoulos Tsiatsos, Stavros Demetriadis

Detecting emotions in Greek sentences: towards an open-source affect-awareness framework

Ioannis Panatsias, Hippokratis Apostolidis, Stergios Tegos, Thrasyvoulos Tsiatsos, Stavros Demetriadis
ipanatsia@csd.auth.gr, aposti@csd.auth.gr, stegos@csd.auth.gr, tsiatsos@csd.auth.gr, sdemetri@csd.auth.gr
Department of Informatics, Aristotle University of Thessaloniki, Greece

Abstract

Considering the highly important role of automated emotion recognition in the development of affective intelligent systems, this study investigates a real-time sentence-level emotion recognition method. More specifically, this paper describes the construction process of an open-source library that can be used for the detection of emotions from sentences comprising Greek or English words. The proposed library was developed, based on the Synesketch software library, utilizing Ekman's basic emotion theory. Greek sentences are classified to six emotional types defined by Ekman (happiness, sadness, anger, fear, disgust and surprise).

Keywords: affective intelligent system, emotion recognition, affective computing.

Introduction

Over the past decades, the growth of computer technology has eliminated the physical distances and enhanced the human-machine communication. Human- computer interaction (HCI) is a research area of increasing interest and a major role in the development of technological products. HCI development is growing rapidly expanding its influence in many aspects of everyday life. One of these aspects is the interpretation of textual messages via emotion recognition. This method has been widely implemented in an emerging area, the field of affective computing (AC). References to ''affective computing" (Picard, 1997) and ''emotional design" (Norman, 2004) denote the integration of emotional intelligence to the design and the implementation of computer systems. These systems may become more natural for humans to understand and use (Picard, 2003).

This research focuses initially on sentence level emotion recognition, but our ultimate aim is the application of this technology to educational activities and the support of learning tools. Emotions related to academic learning, teaching, taking exams and achievement (e.g. enjoyment of learning, pride of success, or test-related anxiety) are proposed to be called as academic emotions (Pekrun et al., 2002). Defining academic emotions in this way, they are commonly used with terms such as academic motivation and academic self-concept (Pekrun et al., 2002).

Our current research effort focuses on the development of an open-source library based on Synesketch software library (Krcadinac et al., 2008; Synesketch, 2013) that can be used to detect emotions from textual data in real time. Considering that there are already systems (Krcadinac et al., 2013; Liu et al., 2003) detecting emotions in English text, we aim to utilize Ekman's emotional model (Ekman, 1973) in order to build an agile sentence-level emotion recognition system for the Greek language.

In this paper, we outline the theoretical background of our research and present a pilot study conducted to gather our initial word datasets. Moreover, we describe the development process of the first version of a Greek affect lexicon comprising numerous sets of words along with their respective emotional weights.

Related Work

Emotion recognition constitutes an interdisciplinary field of great interest. Emotion recognition in learning process and emotional intelligence (EI) in teaching, are very important and they influence the learning outcome (Armour, 2012; Silver, 1999).

Emotion Detection in Learning Environments

The extension of cognitive theory to explain and exploit the role of affect in learning is still in its infancy (Picard et al., 2004). Nevertheless, pioneering learning environments give emphasis on detecting the emotional state of students during the learning process. The main purpose of this technique is to improve learning outcome which can be achieved through the comprehension of students' emotions; the more positive sentiments a student has, the less is needed to change the interaction in the learning environment.

A typical example is the study of Shen et al. (2009). It describes an affective e-Learning model which combined the emotions of learners with a pervasive e-Learning platform; the target is to improve learning experience by adjusting learning material so that it would be compatible with the students' emotional situation. Similarly, Sidney D' Mello et al. (2007) propose an intelligent tutoring system (Auto Tutor) which incorporates learner's affective states into its pedagogical strategies.

The presented work attempts to apply student learning emotion awareness to the student himself/herself (one to one relationship), to the many to one teaching relationship and to collaboration relationships (many to many relationships). Thus, emotion detection from written messages (chat, forums) can be displayed to the student himself/herself, to the tutor and to the team-mates of the group of students in a collaborative learning environment.

Related Work and Affect Lexicons

Apart from educational purpose, emotion recognition studies orientate primarily in constructing lexicon resources which contain emotional words (words that correlate with emotional states). WordNet (Miller, 1995) usually constitute a starting point for building affect lexicons. It is a lexical database that groups words into sets of synonyms (synsets) and contains a number of relations among these synonym sets. It is widely used by applications for text analysis. We used the Greek edition (Greek WordNet, 2013; Greek WordNet in RDF, 2013) of WordNet as we parse Greek sentences as well.

Carlo Strapparava et al. (2004) describe a method that selects a subset of synsets from WordNet and assigns to it one or more affective labels, such as emotion, mood, attitude, sensation etc. in order to provide an emotional determination. Similarly, Das & Bandyopadhyay (2010) report the process of developing an affect lexicon for the Bengali language. According to the study, lexicon is organized in six basic Ekman's model emotions and affect words are updated, using equivalent synsets of SentiWordNet. Moreover, Poria et al. (2012) describe a process in which SentiNet and WordNet-Affect resources are being merged by assigning emotion labels to concepts, eliminating in this way the disadvantage of numerical polarity scores of SentiNet.

Keyword-spotting approach

A widespread method for performing emotion recognition from text is to detect emotional keywords. The term of emotional keywords is used for words with significant relation to certain emotional types. In general, a keyword-spotting system parses an input sentence and extracts the emotion by processing words that describe a sense. In some cases, certain numerical weights are being assigned to the associative words. One main issue on this approach remains the lack of adequate direct words. Direct emotion words are those that directly refer to an emotional state (i.e., joy, fear). Words that are indirectly related to an emotional state are called indirect emotion words (Strapparava et al., 2006). So, many researchers choose to assign a coefficient value to indirect words and expand their lexicon (Strapparava et al., 2006).

Landowska Agnieszka (2013) proposes a recognition algorithm, which is inspired from Synesketch project but based on different keyword lexicons. One of the first methods that utilize the keyword-spotting approach is the Affective Reasoner by Elliott (1992), which searches for an appropriate keyword in a text and uses a lexicon of unambiguously affective words.

Research motivation

Expanding on the Synesketch project, which is an open-source library for textual emotion recognition from English sentences, our research motivation is to investigate the development of an open-source affective-awareness library that supports the Greek language. Apart from the part that relates to the user emotional state elicitation, user affect interpretation and analysis, we are also interested in exploring the potential benefits of using such a library in real-time educational applications, such as conversational agent systems engaging in natural language interaction with the learners (Tegos et al., 2014).

Pilot study

A pilot study was conducted in order to gather all the data required for the development of an affective lexicon.

Participants

The participants of the study were 364 university students. All students were Greek native speakers and their age varied from 18 to 22 years old.

Material

The materials of this study included an online questionnaire as well as a WordNet MySQL database, which emerged after processing the publicly available XML files of the Greek WordNet [15]. The questionnaire asked participants to report a list of words (from 5 to 10) perceived as strongly associated with each of the six basic emotion types (happiness, sadness, anger, fear, disgust, surprise). The structure of the WordNet MySQL database derived from the processing of the WordNet XML files. It should also be noted that a prototype Visual Basic application was used to parse the answer sheets of all participants and calculate the frequency of each word detected.

Results

3032 Greek words gathered from the replies to the questionnaire were assigned into the six main emotional states of Ekman's model. The frequency of each word illustrated its emotional valence. Figure 1 reveals the most frequently mentioned words related to the Happiness emotion. It is noteworthy that a word could be associated with more than one emotion. Figure 2 presents the distribution of words within the emotional sets.

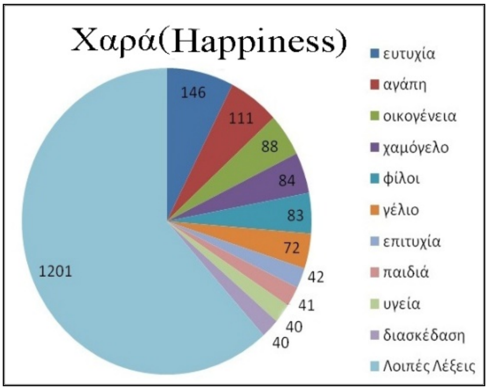

Figure 1. The frequencies of words associated with the Happiness emotion

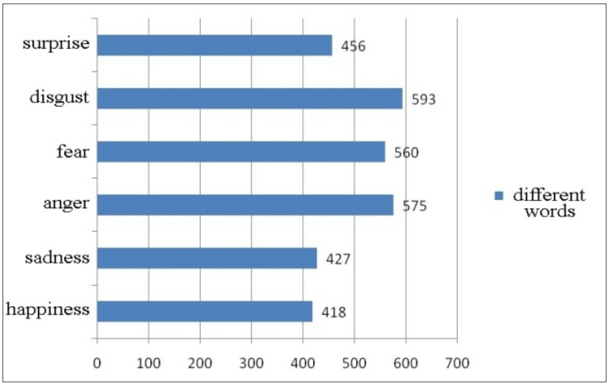

Figure 2. The number of unique words contained in each emotional set

An open source Greek affect lexicon

In this section, we present our initial steps towards the implementation of an affective-awareness library for the Greek language. Following the results of our pilot study, we proceeded to construct an affective Greek lexicon based on the method proposed by the Synesketch project. The main idea of the Synesketch approach is that words convey a strong emotional connotation, which can be used as a starting point for collecting their lexical "relatives" from WordNet. Following the method proposed by Synesketch project we created sets of words for every basic emotion. The words of each set were collected from the participants' responses. We retained only the words mentioned three or more times. The initial corpus that gathered by participants' responses, constitutes of 3032 Greek words. We kept a filtered list of 665 words. Using these initial sets of words we performed a search for synonyms, based on the Greek WordNet synsets. A synset is defined as a set of words within WordNet that are semantically equivalent. More specifically, we generated our lexicon through an algorithm (Krcadinac et al., 2013) which is described in detail bellow:

1. Six sets of emotional synsets S_k, k ∈ E and a word set W are created. Initially, all sets are empty. E is the set of six emotional types of Ekman model; E = {h, sd, a, f, d, su} (h – happiness, sd – sadness, a – anger, f – fear, d – disgust, su – surprise). A Greek emotional keyword is assigned to every one of the six emotional types (Krcadinac et al., 2013).
2. WordNet is searched for synsets of words from the initial sets of emotional keywords Vk, k ∈ E. These initial synsets are added to Sk, k ∈ E, set of emotional synsets for a given emotional type k (Krcadinac et al., 2013).
3. The previous step is repeated d times. In each iteration l (l = 1, 2, …, d), the algorithm searches into WordNet for synsets that are relative to the synsets from the S_k, via WordNet's pointer type SimilarTo. The extended synsets are added to S_k, k ∈ E. However, since these synsets are obtained indirectly, they are attached a penalty coefficient p, which is computed by the following equation:

   $$p_{kj} = 0.1 * l,\ k \in \{h, sd, a, f, d, su\},\ j = 1, 2, \ldots.., q_{ki} \quad (1)$$

   q_{ki} is the number of emotional synsets for the given word i and the given emotional type k (Krcadinac et al., 2013).
4. When all synsets are acquired, words from the synset sets S_k, k ∈ E, are added to the final set of words, W. The total number of words in W is m (Krcadinac et al., 2013).
5. The emotional weights w_{ki}, k ∈ E, i = 1, 2, . . ., m, are calculated for each word from W. For each word, the algorithm collects all the synsets in WordNet the word belongs to. For a given word i, the number of all synsets from WordNet is n_i. The emotional weight for each word and for each emotional type is calculated as a quotient between the number of emotional synsets (of a given emotional type) and the number of all synsets the word belongs to, diminished by using the average penalty of all its emotional synsets. This can be formally expressed in the following manner (Krcadinac et al., 2013):

   $$w_{ki} = \frac{q_{ki}}{n_i}\left(1 - \frac{\sum_{j=1}^{q_{ki}} p_{kj}}{q_{ki}}\right) = \frac{1}{n_i}\left(q_{ki} - \sum_{j=1}^{q_{ki}} p_{kj}\right) \quad (2)$$

The word lexicon formed this way consists of 386 words. This number is getting much bigger if we think that each of these words is stemmed according to Greek grammar rules and then is used for the real-time sentence based processing. In Greek language there is a large amount of common-root words which can be considered with the same emotional weights. A sample of affect lexicon is depicted in Table 1.

Automated textual emotion recognition

There are many real-time applications which are based on this library. For example, in Krcadinac et al. (2013) study is described a visualization system, which transfers a sequence of sentences into a generic animation. One of the main purposes of such applications is the expanding of human on-line communication as a result of stimulating user emotions.

In our research we developed a real-time Greek sentence-based emotion recognition application based on Emote library (Emote, 2011; Fulton, 2011). This application processes Greek sentences utilizing rule based logic and Greek affect lexicon.

Table 1. A sample of typical words included in affect lexicon with their emotional weights.

Word	H	Sd	A	F	D	Su
surprise	0.9	0	0	0	0	0.45
loss	0	0.9	0	0.9	0	0
sickness	0	0.6	0	0.6	0	0
disagreement	0	0	0.36	0	0	0
admiration	0	0	0	0	0	0.9

Discussion and conclusion

This article presents a broad approach of a Greek affect lexicon being accessed by a real-time emotion recognition application, which receives Greek or English sentences as input and classifies them to the Ekman's model six basic emotions (Ekman, 1973).

This research could be extended by enriching Greek affect lexicon. So, future expansion of our research will be directed in the enrichment of the lexicon. A larger amount of datasets produced from more participants' responses in conjunction with the iteration of lexicon building process could restrict the error probability. We will also try to evaluate the system reliability through a comparison between human and system emotional recognition efficacy; the more application results converge in human assessment about the emotion of input text, the more the system is reliable. Emotion recognition approach could be a considerable support in educational process. Our Greek sentence-based emotion recognition application could be applied in Web 2.0 or interaction learning tools such as chat, forum, blog, wiki and conversational agents.

Acknowledgments

The authors would like to thank Dimitrios Gousios whose contribution to pilot study procedure was valuable.

References

Agnieszka, L. (2013). Affect-awareness framework for intelligent tutoring systems. *2013 6th International Conference on Human System Interactions (HSI)*, 540–547.

Armour, W. (2012). *Emotional intelligence, student engagement, teaching practice, employability, ethics curriculum*, 8(2004), 4–10.

Clark Davidson Elliott. (1992). *The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System.* Ph.D. Dissertation. Northwestern University, Evanston, IL, USA.UMI Order No.GAX92-29901.

Das, D., & Bandyopadhyay, S. (2010). *Developing Bengali WordNet Affect for Analyzing Emotion.Iccpol-2010.* Retrieved from http://www.dasdipankar.com/Main_files/ICCPOL_2010.pdf

D'Mello, S., Graesser, A., & Picard, R. W. (2007).Toward an affect-sensitive autotutor. *IEEE Intelligent Systems*, 22(4), 53–61.

Greek WordNet (2013).Retrieved 15 of December 2015 from http://wordnet.okfn.gr/.

Greek WordNet in RDF, (2013). Retrieved 10 of December 2015 from http://wordnet.okfn.gr/ .

Ekman, P. (1973). Cross-cultural studies of facial expression. *Darwin and facial expression: A century of research in review*, 169-222.

Elliott, C. (1992). *The Affective Reasoner: A Process Model of Emotions in a Multi-agent System*. PhD thesis, Northwestern University, May 1992. The Institute for the Learning Sciences, Technical Report No. 32.

Emote (2011). Retrieved 15 of December 2015 from https://github.com/jessefulton/emotivetext.com/tree/master/src.

Krcadinac, U., Pasquier, P., Jovanovic, J., &Devedzic, V. (2013).Synesketch: An Open Source Library for Sentence-Based Emotion Recognition. *IEEE Transactions on Affective Computing*, 4(3), 312–325.

Liu, H., Selker, T., & Lieberman, H. (2003).Visualizing the affective structure of a text document. *CHI '03 Extended Abstracts on Human Factors in Computing Systems* - CHI '03, 740.

Miller, G. a. (1995).WordNet: a lexical database for English. *Communications of the ACM*, 38(11), 39–41.

Norman, D. A. (2004).Emotional Design : Why We Love (or Hate). *Everyday Things*.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002).Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. Educational Psychologist, 37(2), 91–105.

Picard, R. W. (1997). *Affective Computing*. MIT Press, (321), 1–16.

Picard, R. W. (2003). Affective computing: challenges. *International Journal of Human-Computer Studies*, 59(1–2), 55–64.

Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., …Strohecker, C. (2004). Affective learning - a manifesto. *BT Technology Journal*, 22(4), 253–268.

Poria, S., Gelbukh, A., Cambria, E., Yang, P., Hussain, A., &Durrani, T. (2012). Merging SenticNet and WordNet-Affect emotion lists for sentiment analysis. *2012 IEEE 11th International Conference on Signal Processing*, 1251–1255.

Shen, L., Xu, J., Zhang, B., Matsoukas, S., &Weischedel, R. (2009). *Effective Use of Linguistic and Contextual Information for Statistical Machine Translation*. Emnlp-2009, 72–80.

Silver, M. A. (1999). Emotional intelligence and legal education. *Psychology, Public Policy, and Law*, 5(4), 1173–1203.

Strapparava, C., & Valitutti, A. (2004, May). WordNet Affect: an Affective Extension of WordNet. In *LREC* (Vol. 4, pp. 1083-1086).

Strapparava, C., Valitutti, a., & Stock, O. (2006). The affective weight of lexicon. *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, 423–426.

Synesketch (2013). Retrieved 15 December 2015 from http://synesketch.cvs.sourceforge.net/viewvc/synesketch/.

Tegos, S., Demetriadis, S., & Tsiatsos, T. (2014). A Configurable Conversational Agent to Trigger Students' Productive Dialogue: A Pilot Study in the CALL Domain. *International Journal of Artificial Intelligence in Education*, 24(1), 62–91.