ArsEmotica: Emotions in the Social Semantic Web

Matteo Baldoni
Dipartimento di Informatica
Università degli Studi di Torino
baldoni@di.unito.it

Cristina Baroglio
Dipartimento di Informatica
Università degli Studi di Torino
baroglio@di.unito.it

Viviana Patti
Dipartimento di Informatica
Università degli Studi di Torino
patti@di.unito.it

Paolo Rena
Dipartimento di Informatica
Università degli Studi di Torino
paolo.rena@gmail.com

ABSTRACT

This work presents ArsEmotica, an application software that aims at extracting a rich emotional semantics (i.e. not limited to a positive or a negative reception) of tagged resources. We worked on a subset of artworks supplied by the art portal ArsMeteo, and focussed the sentiment analysis on the collections of tags related to artworks. Our approach exploits and combines multilingual lexicons (MultiWordNet), affective, and sentiment lexicons (WordNet-Affect, Senti-WordNet) with an ontology of emotional categories (OntoEmotion), enriched with over four hundred Italian emotional words referring to the about eighty-five concepts of the ontology. ArsEmotica uses the ontology to identify tags directly referring to emotions, while potentially affective tags can be emotionally annotated by the user, in a pure Web 2.0 approach. From all these different information sources, ArsEmotica ranks the emotions related to the artworks.

Categories and Subject Descriptors

H.4.m [**Information Systems Applications**]: Miscellaneous; J.5 [**Computer Applications**]: Arts and Humanities

General Terms

Human Factors

Keywords

Semantic web, folksonomies, ontologies, emotions

1. INTRODUCTION

Social networks and platforms promote the participation of users in many ways, stimulating the expression of opinions about the contents inserted by other users, by "Like/Dislike" annotations, star-rating systems, tag-based annotation and

navigation, and so forth. This huge amount of data is a precious information source about perceptions, trends, and feelings, and a lot of research is being carried on to identify ways for extracting meaningful information from these data. In this context, one emerging research field is Affective Computing, aimed at automatically recognizing the users' emotions by analyzing their behavior. Since emotions are often related to appreciation, the feeling of users are an important feedback that can support decisional tasks. Such techniques, however, return very simple feedbacks, i.e. the appreciation of a society of users given in terms of positive and negative reception.

This work faces a more complex task: the *identification of the prevalent emotions* associated by the users of a Social Network to *tagged* resources. The output we aim at supplying is not merely a general appreciation but rather amounts to a set of specific concepts, that emerge as the most significant in capturing the emotions of the users towards the resource at issue, each enriched with a *score* expressing the strength of that emotion. Thus, we propose an approach to sentiment analysis which presents two main novelties. First, the extraction of the emotional content from text is driven by an ontology of emotional categories where emotions are structured in a taxonomy, resulting in something richer than a polarized appreciation. Second, we focus on very essential form of textual resources: collections of tags. In many applications, users use tags to supply *concise reviews*. In these cases, it is therefore possible to leave aside the complications due to text analysis (often aimed at identifying the key words inside text) and focus on the emotional analysis. This happens mainly for artistic domains, where resources represent artworks, movies, books. Moreover, for languages like Italian, for which there are few affective lexical resources that are freely available for the development of new application softwares, tags can be the only processable information source that can be used.

Of course, not all tags have an affective value. To identify those bearing an emotional content, we integrated various Semantic Web and NLP tools. Specifically: multi-lingual and Italian computational lexicons [9] and affective lexicons (WordNet-Affect [13] and SentiWordNet [2]) with the aim of identifying (i) words that directly refer to emotional concepts (e.g. *paura*, fear), (ii) words that indirectly deliver an emotional content (e.g. *infinito*, infinite). After this analysis, it is possible to apply algorithms to identify and rank the emotions related to the tagged resources. The algorithm

that we have implemented reasons on the ontology structure and is inspired by the one described in [6].

2. BACKGROUND

The research, that we have been carrying on, strictly embraces or is in some way related to different technologies and topics, including the Social Semantic Web, computational lexicons (MultiWordNet [9], WordNetAffect [13], SentiWordNet [2]), studies on emotions in social, psychological, and computational contexts, and Sentiment Analysis. Let us briefly overview the background. One of the key problems in the Web is how to index resources so as to efficiently and effectively retrieve them. Ontologies and folksonomies are two ways for indexing resources: the former are to be designed by knowledge engineers, while the latter are spontaneously produced by the *tagging* activity of the members of a community. There are many attempts to reconcile folksonomies with ontologies [4, 3], for instance by inducing ontologies from folksonomies or by matching terms. The reason is that the Social Web aims at combining the ease of use, typical of its platforms, with the advantages deriving from a formal semantics, i.e. interoperability, data/service integration, personalization, better recommendations [12].

In this context, the identification of the feelings of a community or of its single members is receiving an increasing attention, as an indicator of the appraisal of topics, people, situations, resources, trends. Hence the development of opinion Mining [7], of ontologies of emotions and of W3C markup language proposals like Emotion ML [10, 11]. However, there are still few applications that use the most advanced results in Semantic Web technology to deal with emotions. Moreover, most of the approaches use emotional ontologies where emotions are individual isolated units (e.g. WordAffect). Such considerations motivated our focus on the ontology of emotions in [5], an OWL ontology where emotions are structured and organized in levels, trying to integrate the results of the most recent psychological models. OntoEmotion provided a very useful starting point for our aim of relating emotions to tags. The study of emotions is particularly relevant in artistic domains where it is a common experience for users to share the feelings kindled by an artwork with friends. As a consequence, a software which, starting from tags freely associated to resources, can extract a rich emotional semantics, has many possible and interesting applications. For instance, by creating new, emotion-driven navigations of the items shown by on-line museum portals, or by creating apps for portable devices, on top of which creating new ways of participating to art exhibitions. Most of current portals allow users to express their appreciation on artworks by means of star-rating systems. Our ontology-driven approach could provide users with new means for expressing and sharing emotions stirred by artworks.

3. ARSEMOTICA

This section describes *ArsEmotica*, the application we developed to identify the emotions stirred by a resource. A pre-processing phase filters tags so as to eliminate flaws like spelling mistakes, badly accented characters, and so forth. Figure 1 reports the three main steps that characterize the computation after the pre-processing:

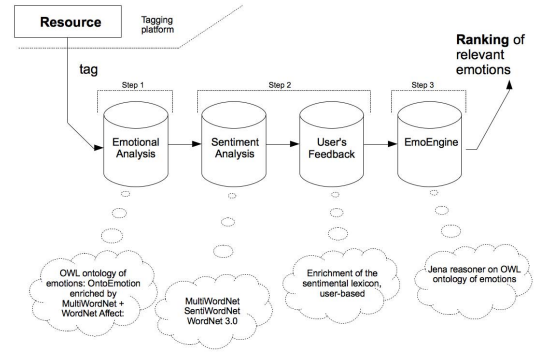**Step 1: Checking tags against the ontology of emo-**

Figure 1: Overall architecture.

tions. It amounts to check whether a tag belongs to the ontology of emotions.

Step 2: Checking tags with SentiWordNet. Tags that do not correspond to terms in the ontology are further analysed by means of *SentiWordNet*, to distinguish *objective* tags, which do not bear an emotional meaning, from *subjective* and, therefore, affective tags. The latter will be presented to the user to get a feedback on which emotional concept they deliver.

Step 3: Ranking of Prevalent Emotions. Based on data collected in the previous steps, the tool ranks the emotions associated by the users to the resource.

The following sections explain how the extraction of the emotional semantics is performed trough the three steps.

3.1 The ontology of emotions

The first step of ArsEmotica checks if tags are "emotion-denoting" words, directly referring to some concepts of an ontology of emotions. We started from *OntoEmotion*, an emotional ontology [5] developed at Univesidad Complutense de Madrid. The ontology is written in OWL and structures emotional categories in a taxonomy which contains 87 emotional *concepts*, including basic emotions (*Sadness*, *Happiness*, *Surprise*, *Fear* and *Anger*) as well as more specific emotional categories. The taxonomic structure basically refers to the psychological model by Parrot [8], adapted to these five basic emotions, and integrated with all the emotions which appear in other well-established models. In particular, under each basic emotion we have a number of levels which may vary.

OntoEmotion was conceived for categorizing emotion-denoting *words*. So each class was populated by instances, consisting in words from different languages. The ontology has two root concepts: *Emotion* and *Word*. *Emotion* is the root for all the emotional concepts. *Word* is the root for the emotion-denoting words, and originally had two subclasses: *EnglishWord* and *SpanishWord* populated with about 250 words. Intuitively, concepts *Emotion* hierarchy represent language-independent emotions corresponding to common experiences in life. Instances in the ontology represent the words provided by specific languages for referring to emotions, For example, the word *rage* is both an instance of the concept *Fury*, and an instance of the concept *EnglishWord*, which means that *rage* is an English word for denoting fury.

Since the tags used by the ArsMeteo community (our case

study) are mainly Italian words, we enhanced the ontology by adding a new subclass *ItalianWord* to *Word*. Our approach for semi-automatically populating the ontology with Italian emotion-denoting words relies on the use of the multilingual lexical database MultiWordNet, in which the Italian WordNet is strictly aligned with Princeton WordNet 1.6., and its affective domain WordNet-Affect, a well-used lexical resource that contains information about the emotion that the words convey. The identified terms were checked by human experts of the Italian language.

WordNet is a lexical database, in which nouns, verbs, adjectives and adverbs (lemmas) are organized into sets of synonyms (synsets), representing lexical concepts. After choosing one representative Italian word for each emotional concept in the ontology, we used it as an entry lemma for querying the lexical database. The query returns the word synsets, representing the senses of that word, and are labeled by unique identifiers. Each synset is then processed by using WordNet-Affect. When a synset bears affective information, all the synonyms belonging to that synset are imported in the ontology as Italian emotion-denoting words. This allows us to automatically enrich the ontology with many synonyms of the representative words, and to filter out synsets which not convey affective information. The resulting ontology contains more than 450 Italian words referring to the 87 emotional categories of OntoEmotion. ArsEmotica uses the ontology during the Emotional analysis for checking if a given tag denotes some emotional category. If yes, the tag is immediately classified as "emotional". The information collected during this phase is stored in a DB table as a set of triples (t, e, s), meaning that tag t is related to emotion e with a strength value $s \in [0, 100]$. When a tag is emotional, the strength will be 100. So, for example, since the word "affanno" (breathlessness) is an instance of "anxiety", the corresponding triple will be: ('affanno', 'Ansia', 100).

3.2 Sentiment analysis and User feedback

Tags can convey affective meaning also *indirectly*. As observed in [14], some words can be emotional for someone due to her individual story. In other cases the affective power is part of the collective imagination (e.g. words like "war"). As a consequence, it seems appropriate to involve the community in the definition of such indirect affective meanings. ArsEmotica enables users to give their personal opinion on the emotional meaning of tags which have not been recognized as directly referring to an emotional category.

When the set of tags to consider is too large (as in Arsmeteo, where the average number of tags for each artworks is over than ten), the activity of evaluating them can bore the users. Thus, we pre-process them and select the most promising ones by using SentiWordNet 3.0, a lexical resource for opinion mining where each synset of Princeton WordNet 3.0 is associated three numerical scores $Pos(s)$, $Neg(s)$ and $Obj(s)$ which indicate how neutral (Obj) or affective (Pos and Neg, in a positive and negative sense respectively) the synset is. Each score ranges in interval $[0, 1]$. Their sum is 1 for each synset.

Sentiment analysis on tags is performed on each tag that does not belong to the ontology of emotions, by the steps sketched in Figure 2 . Since SentiWordNet was created for English, we use MultiWordNet to align the Italian lemma corresponding to the tag to the English lemma. Moreover since SentiWordNet annotates version 3.0 of Princeton Word-
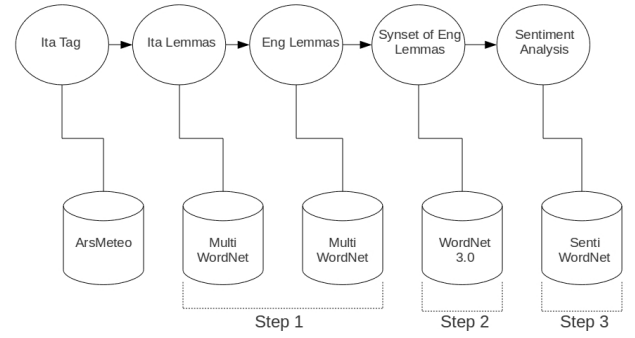

Figure 2: Steps of the sentiment analysis on tags.

Net while MultiWordNet is based on 1.6, we query such newer lexical database for English to reach the synsets for the words sentimentally analyzed by SentiWordNet and retrieve the affective power of the tag-word. As an example, consider the Italian tag *infinito*. MultiWordNet returns many synsets, having different corresponding English lemmas: 'infinite', 'infinitive' and 'inexhaustible'. If we use these English words for querying WordNet 3.0, we obtain the list of all of their meanings. Now, we can query SentiWordNet and get the results of the sentiment analysis, which are stored in a database:

(**00028651**), Pos(s): 0, Neg(s): 0;
(**01007354**), Pos(s): 0.125, Neg(s): 0.5;
(**01008745**), Pos(s): 0, Neg(s): 0;
(**00301951**), Pos(s): 0, Neg(s): 0;
(**00005718**), Pos(s): 0.25, Neg(s): 0.375;
(**01008289**), Pos(s): 0.375, Neg(s): 0.25;

The objectivity of a word is simply measured as: $1 - (Pos(s) + Neg(s))$. It returns values in the interval $[0, 1]$, 1 meaning the term is objective, 0 that it conveys some strong sentimental (positive or negative) meaning. Different senses of the same term can have different opinion-related properties and scores. Only when for no sense of the term has a relevant score, we conclude that it is probably mainly descriptive. When at least one meaning of the term has a relevant sentimental score, we can suppose that it will stir some emotion, thus we will offer the word to the users of the community for an evaluation. For instance, since *infinito* has a relevant sentimental score for some of its senses, it is proposed to the users for the sentimental evaluation.

For all tags resulting potentially affective, like *infinito*, the user will be free to indicate one or more emotions from the emotional categories of the ontology with a strength value which intuitively represents the user's measure for the semantic affinity of the term with the chosen emotional category. Again, a set of triples (t, e, s) is collected.

3.3 Getting the most relevant emotions

Once the analysis of the tags associated to a resource is finished, it is possible to rank the emotions associated by the users to the resource and to compute the prevalent emotion. The implemented algorithm relies on the taxonomic structure of the ontology and is inspired by the one described in [6], where was used to emotionally mark up a sentence by relying upon the emotions assigned to the words that compose it. Intuitively, the algorithm allows to select the *most specific*

emotion which represents the affective information related to the artwork from the probability (score) that each of its tag has of indicating different emotions. The basic steps are: (1) processing the emotional concepts appearing in the triples (t, e, s), so as to identify also those emotions in the ontology that are related of the ones appearing in the triples. The identified emotional concepts can be organized into *layers* by following the parent-child relationship (in this phase a Jena Reasoner has been applied to the collected triples) (2) starting from the leaves and moving upward towards the root, compose and propagate the scores.

3.4 Case Study

The developed prototype was tested against a corpus of tagged multimedia artworks, from the web platform *ArsMeteo* (http://www.arsmeteo.org [1]). 47 sample artworks and of 842 tags were analysed (see http://www.giorgiovaccarino. it/mostre/emozioni.html). To this aim, we designed a simple interface: the user chooses an artwork from a menu; button *tag* lists on the right of the artwork's preview all the tags associated to it by the ArsMeteo community; by pressing in sequence the other buttons it is possible to perform the analysis. *Tag to Emo* controls which tags belong to the ontology, storing the corresponding triples. ArsEmotica visualize below the artwork's preview the affective tags and the emotions they refer to, Figure 3. For instance, by applying the *emotional analysis* to the ArsMeteo artwork "Bianca e il suo contrario", four tags are identified as emotion-denoting words. In particular the tag "bene" (good) refers to emotion "Love", while "tristezza" (sadness) refers to "Sadness" and "Gloom". By pressing *Tag to Senti* all the objective tags will be discarded. By the last button the emotional engine

Figure 3: Emotion-denoting tags for an artwork

can be activated to rank the emotions of the analyzed artwork. By applying the *emotional analysis only* we obtained the emotional meaning of 14 images out of 47.

4. CONCLUSIONS AND FUTURE WORK

We presented a software solution that, by combining state of the art lexicons and libraries, allows both the population of an ontology of emotions with Italian emotion-bearing terms and the extraction of the prevalent emotions from the set of tags associated to a resource. The extracted information is richer than a polarized appreciation.

Given appropriate NLP tools, its use could be extended also capturing the latent emotions behind textual comments. Validating our approach with real users of the ArsMeteo

community will be part of the future work. One of the challenges is to motivate users to annotate with emotional concepts from our ontology tags which can convey an indirect affective meaning, by devising appealing interfaces that involve the user as in games. An alternative could be to integrate the use of automatic techniques, e.g. the one proposed in [3], for findings the proper associations of terms having an emotional value with ontological concepts.

5. REFERENCES

[1] E. Acotto, M. Baldoni, C. Baroglio, V. Patti, F. Portis, and G. Vaccarino. ArsMeteo: artworks and tags floating over the planet art. In *Proc. of HT*, pages 331–332. ACM, 2009.

[2] S. Baccianella, A. Esuli, and F. Sebastiani. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proc. of LREC*, 2010.

[3] M. Baldoni, C. Baroglio, A. Horváth, V. Patti, F. Portis, M. Avilia, and P. Grillo. Folksonomies meet ontologies in ArsMeteo: from social descriptions of artifacts to emotional concepts. In *FOMI*, 2008.

[4] F. Dotsika. Uniting formal and informal descriptive power: Reconciling ontologies with folksonomies. *Int. J. of Info. Manag.*, pages 407–415, 2009.

[5] V. Francisco, P. Gervás, and F. Peinado. Ontological reasoning to configure emotional voice synthesis. In *Proc. of RR*, volume 4524 of *LNCS*. Springer, 2007.

[6] V. Francisco, F. Peinado, R. Hervás, and P. Gervás. *Semantic Web Approaches to the Extraction and Representation of Emotions in Texts*. NOVA Publishers, 2010.

[7] B. Pang and L. Lee. *Opinion Mining and Sentiment Analysis*. Now Publishers Inc., 2008.

[8] W. Parrot. *Emotions in Social Psychology*. Psychology Press, Philadelphia, 2001.

[9] E. Pianta, L. Bentivogli, and C. Girardi. MultiWordNet: developing an aligned multilingual database. In *Proc. of the 1st Int. Conf. on Global WordNet*, 2002.

[10] M. Schröder, P. Baggia, F. Burkhardt, A. Oltramari, C. Pelachaud, C. Peter, and E. Zovato. Emotion Markup Language (EmotionML) 1.0. W3C Working Draft, 2010. http://www.w3.org/TR/emotionml/.

[11] M. Schroeder, H. Pirker, M. Lamolle, F. Burkhardt, C. Peter, and E. Zovato. Representing emotions and related states in technological systems. In *Emotion-Oriented Systems*, pages 369–387. Springer, 2011.

[12] N. Shadbolt, T. Berners-Lee, and W. Hall. The semantic web revisited. *IEEE Intelligent Systems*, 21(3):96–101, 2006.

[13] C. Strapparava and A. Valitutti. WordNet-Affect: an affective extension of WordNet. In *Proc. of LREC*, volume 4, pages 1083–1086, 2004.

[14] C. Strapparava, A. Valitutti, and O. Stock. The affective weight of lexicon. In *Proc. of LREC*, pages 1–83, 2006.