Detecting Implicit Expressions of Affect from Text using Semantic Knowledge on Common Concept Properties

Alexandra Balahur, Hristo Tanev
European Commission Joint Research Centre
Institute for the Protection and Security of the Citizen
Via E. Fermi 2749, 21027 Ispra (VA), Italy
E-mail: {alexandra.balahur, hristo.tanev}@jrc.ec.europa.eu

Abstract

Emotions are an important part of the human experience. They are responsible for the adaptation and integration in the environment, offering, most of the time together with the cognitive system, the appropriate responses to stimuli in the environment. As such, they are an important component in decision-making processes. In today's society, the avalanche of stimuli present in the environment (physical or virtual) makes people more prone to respond to stronger affective stimuli (i.e., those that are related to their basic needs and motivations – survival, food, shelter, etc.). In media reporting, this is translated in the use of arguments (factual data) that are known to trigger specific (strong, affective) behavioural reactions from the readers. This paper describes initial efforts to detect such arguments from text, based on the properties of concepts. The final system able to retrieve and label this type of data from the news in traditional and social platforms is intended to be integrated Europe Media Monitor family of applications to detect texts that trigger certain (especially negative) reactions from the public, with consequences on citizen safety and security.

Keywords: emotion detection, media monitoring, argument mining, emotion triggers, common-sense knowledge, decision support

1. Introduction Paper

Anger, fear, sadness, disgust, happiness or surprise. Any human being is able to relate to these emotions and give examples of situations when they can be felt and the possible manners in which they can be expressed. However, the mechanisms of emotion are in no way simple or straightforward to understand, explain or mimic.

Until recently, emotions were considered separate from cognition (Forgas, 2008). The latter was considered in relation to rational thinking, reasoning, rigor, while affective phenomena were thought to "overwhelm or subvert rational mental processes" (Elster, 1985). Recent studies have shown that emotion and cognition actually go hand in hand, regulating our reactions and ensuring our well-being and balance, being an important part of the decision-making process (Forgas, 2008). Given the importance of these aspects, emotions have been studied intensely in the past years.

In Natural Language Processing (NLP), the task that deals with the detection and classification of texts according to the polarity of the opinions they express is called Sentiment Analysis (SA). In spite of the dynamics of the field and the high amount of research that has been done under its umbrella in the past decade, systems dealing with sentiment analysis have only reached performance levels that are around 80% accuracy at most (reaching almost 90% in the systems that are employing deep learning and that are applied to product review classification, a more straightforward task). Of course, these levels are influenced by the types of text, the language and the language style. Apart from the inherent difficulties posed by the need for linguistic processing resources and tools, an important explanation for the limited performance of emotion detection systems is in itself given by the approaches used. While most automatic sentiment analysis systems employ linguistic knowledge or supervised learning, human affect is expressed not only at the level of emotion-bearing words or phrases, but at the level of concepts that based on common-sense knowledge remind of or trigger affective reactions (e.g. "being fired", "virus infection", "receiving a present", "going to a party", etc.).

Detecting such factual statements that trigger emotions is an important step towards discovering the intended effect of arguments used in the media. Here, journalistic style requires that no subjective evaluations are used. Nonetheless, journalists can elicit subjective evaluations even when using only factual statements, because these statements trigger in the readers an emotional reaction. An example from recent press is the "refugee crisis", covered in traditional and social media. Here, many factual statements present in the news relate to refugees being "not vaccinated", "not educated"; other arguments refer to the already-existing financial crisis and the fact that tackling the refugee issue requires a great financial effort. These arguments give rise to fear from the audience, especially when they are repeated over many sources, since they evoke the possibility to get ill, to have their safety and security threatened, to suffer the consequences of lack of finances, etc. Other applications are related to "hate speech", which is on many occasions very subtle. While there is a thin line between freedom of speech and hate speech, it is useful to detect texts possibly containing the latter.

The final system able to detect and label such data from the news in traditional and social platforms could be integrated Europe Media Monitor[1] family of applications to detect texts that trigger certain (especially negative) reactions from the public, with consequences on citizen safety and security.

[1] http://emm.newsbrief.eu

2. Background and Motivation

Motivated by the insight that emotions are often triggered by factual information in text, in our previous work (Balahur et al., 2012), we proposed a method to build a knowledge base – ***EmotiNet*** - in which to store situations that based on commonsense knowledge about them trigger emotions. After studying the relevant literature in Psychology, the theoretical model that best approximated this was found to be given by the Appraisal Theory Models (De Rivera, 1977; Frijda, 1986; Ortony et al., 1988; Johnson-Laird and Oatley, 1989). This set of models claim that emotions are elicited and differentiated on the basis of the cognitive evaluation of the personal significance of a situation, object or event based on *"appraisal criteria"* (intrinsic characteristics of objects and events, significance of events to individual needs and goals, individual's ability to cope with the consequences of the event, compatibility of event with social or personal standards, norms and values). The goal of our work is also to increase the quality of the Europe Media Monitor by improving its sentiment detection capabilities.

In order to perform this task automatically, the criteria should ideally be detected and classified automatically. In our previous experiments (Balahur et al., 2013a; Balahur et al., 2013b), we employed the ISEAR database of self-reported affect to model situations that trigger emotions in the people that experiment them. Further to these experiments, we have shown that these situations can be decomposed into triples of (subject, action, object) and that the emotional response to the situation can be dependent on the characteristics and properties of the concepts in the actor or object role (e.g. "The monkey climbed the tree" does not produce the same emotional effect as "The crocodile climbed the tree" or "The man killed the fly" does not produce the same emotional effect as "The man killed the lion", because the two actors/objects have very different properties).

In this article, we propose an algorithm to learn the connection between the intrinsic characteristics/ properties of objects and the emotions they trigger. In the following sections, we described the algorithm we employed and the results we obtained evaluating the accuracy of the concepts and properties discovered. Such factual statements that trigger emotions is an important step towards discovering the intended effect of arguments used in the media.

3. Learning emotion-related properties of entities and phenomena

The goal of the algorithm which we implemented was to learn intrinsic properties of entities, actions and phenomena which trigger certain emotions. These properties constitute one of the appraisal criteria in the Appraisal Theory Models. In particular, we considered the emotions **fear**, **anger**, **joy** and **disgust**. Emotion-related properties are crucial for in-depth analysis of sentiments transmitted through the text implicitly. For example, the Twitter message ***"I don't ever eat or drink after anyone because germs. yea."*** implies the fear emotion, brought forward because of the possible diseases which germs can transmit. However, for a software system to detect that, it should be able to conclude that germs are capable of spreading diseases and people are afraid of things which can spread diseases. The learning algorithm which we put forward here is a step towards the implementation of this kind of reasoning. In particular, we learn from ConceptNet (Singh et al., 2002) properties of objects, actions and phenomena which cause certain emotions. Consequently, these properties can be used to detect expression of emotions. This can be done first via mentioning of entities or phenomena having these properties, e.g. "I have seen a snake" ("snake" has the property "can bite" which is related to fear) or second, through mentioning this property directly – "It can bite me". Clearly, a proper reasoning mechanism necessary to infer an emotion goes beyond the detection of a property, however learning the relevant properties is an important step towards the implementation of this reasoning.

In designing our learning algorithm, we were led by the following rationale: entities and phenomena which typically trigger some emotion, but otherwise are semantically different, will have in common properties which are semantically related to this emotion. For example, the emotion "fear" is caused by things like "dark alleys", "storms" and "roller coasters". All the three concepts have the property "dangerous". Since these words are semantically different apart from their relation to the "fear" emotion, their common properties are most probably related to this emotion. Consequently, "dangerous" may be considered semantically related to the emotion "fear".

The algorithm, which follows this rationale, uses Twitter to collect status messages, expressing certain emotions, distributional clustering to put together semantically similar words and finally, it exploits the semantic network ConceptNet, in order to extract properties common to these word classes.

ConceptNet is a semantic network which uses the database from the Open Mind Common Sense project based at the MIT Media Lab. The purpose of OMCS was to build and utilize large collections common sense knowledge from contributions of thousands of Web users. Words are related between each other with properties such as "causes", "causes desire", etc. In our experiments, we used a subset of these relations, namely: "causes", "causes desire", "has property", and "capable of". We chose these properties on the basis of empirical observations.

Our algorithm works as follows:

1. For each emotion under consideration, we manually create patterns which are likely to co-occur with words, designating entities and phenomena which trigger this emotion, for example "afraid of [EMOTION]", "scared of [EMOTION]", "I feel [EMOTION]", "makes me happy", "makes me sick", "makes me angry", "surprises/d me", "bewilders/ed me", "makes me scared", "makes me sad", "makes me

cry".

2. We scan Twitter for tweets, mentioning these patterns and extract the words appearing in the position of the slot. Then, we consider the slot fillers with highest co-occurrence rate, where we measured the co-occurrence using pointwise mutual information. For example, for the "fear" emotion we will have words like "spiders", "bugs", "rejection", "failure", etc.

3. The slot fillers, extracted in the previous step, are clustered using distributional similarity. We use as features for this clustering the words and bigrams adjacent to the slot fillers. In order to extract the features, we used the multilingual lexical learning tool Ontopopulis/Lexiclass (Tanev et al. 2009) and for clustering we used agglomerative clustering, implemented in CluTo . In our case, "spiders" and "bugs" will be clustered together, since they share many similar contexts. On the other hand, "rejection" and "failure" will finish in two separate clusters.

4. Finally, we check for each word cluster what properties it is related to in ConceptNet and we take those properties, which appear in 2 or more clusters and have the same values. In our example the clusters containing "rejection" and "failure" have in common the property "CapableOf" with value "upsetting_people", consequently our algorithm will learn the property-value pair "CapableOf -> upsetting_people".

4. Preliminary experiments

We have collected from Twitter data about 4 emotions, namely joy, anger, fear and disgust, based on the patterns described. Then, we followed the algorithm above in order to extract properties and their values, related to the considered emotions. In point 2 of the algorithm we used the most frequent 500 slot fillers, which are then clustered. Table 1 shows the parameters and the outcome of the experiments. The irrelevant property-value pairs are marked with a star.

Emotion	Number of clusters	Number of the obtained property values	% of the relevant property values	Extracted ConceptNet properties and values
joy	150	8	75%	HasProperty → good HasProperty → beautiful HasProperty → pleasant HasProperty → fun *HasProperty → dangerous HasProperty → romantic HasProperty → good_for_your_health *HasProperty → edible
anger	113	15	67%	*HasProperty → good HasProperty → bad HasProperty → dangerous *HasProperty → female *HasProperty → beautiful HasProperty → dress_herself *HasProperty → read_book HasProperty → powerful *HasProperty → fly HasProperty → important HasProperty → very dangerous HasProperty → corner_criminal HasProperty → tail_suspect HasProperty → direct_traffic Causes → death
fear	219	21	48%	HasProperty → bad HasProperty → make_person_upset *HasProperty → small *CapableOf → fly CapableOf→spread_disease *CausesDesire→go_somewhere *HasProperty → transparent CapableOf→surprise_person CapableOf→sting *HasProperty→cool HasProperty→loud CapableOf→buzz CapableOf→carry_disease *CapableOf → be_pet CapableOf → scare_person *HasProperty → blue *HasProperty → fun CapableOf → hurt *HasProperty → small *CapableOf → run

				CapableOf → make_person_upset CapableOf → bite Causes → death
Dis-gust	193	24	29%	*CausesDesire→ drink *CapableOf→sleep HasProperty→smelly HasPropety→sticky HasProperty→bitter HasProperty→ addictive HasProperty→ bad_for_health CapableOf→ divide_family HasProperty→smelly *HasPropety→cute *HasProperty→ dangerous *HasProperty→ sweet *HasPropety→green *HasProperty→blue *HasProperty→ natural *HasProperty→ corner_criminal *HasProperty→good *HasProperty→cold *HasProperty→hot *HasProperty→bad *HasProperty→ direct_traffic *HasPropety→ tail_suspect *HasProperty→ white *Cause→death

Table 1. Properties of emotion-triggering concepts extracted from ConceptNet

The accuracy is highest for joy. However we have learnt the smallest number of property-value pairs with respect to the other emotions. We have the lowest accuracy for the disgust emotion. In general, the property-value pairs we learn are not many, but we could increase the coverage of our method by considering more slot fillers (currently, we use the top 500).

These preliminary experiments show that EmotiNet can be enriched using concepts gathered based on their properties. We have decided to further extend the process by using synonyms of the properties. In such a way, we can automatically discover new concepts with similar properties that are expressed in a different manner. In order to define the synonym terms, we employ WordNet (Miller, 1995). For each of the properties found that are related to specific concepts triggering an emotion, we sought their synonyms and manually sifted through them to find the most appropriate ones (i.e. that would keep the intended meaning). In Table 2, we show the synonyms that were chosen for this task.

E-motion	Extracted ConceptNet properties and values	Synonyms of properties
joy	HasProperty → good HasProperty → beautiful HasProperty → pleasant HasProperty → fun HasProperty → romantic HasProperty → good_for_your_health	**Good**-> beneficial, agreeable, pleasing, right **Beautiful** -> delightful, exciting **Pleasant** ->enjoyable **Fun** ->amusing, comic, comical, funny, laughable, mirthful, risible **Romantic**->amatory, amorous, romantic, wild-eyed **Healthy**-> intelligent, levelheaded, sound
anger	HasProperty → bad HasProperty → dangerous HasProperty → powerful HasProperty → important Causes → death	**Bad**->tough, unfit, unsound, risky **Dangerous**->grave, grievous, serious, severe, life-threatening **Powerful**->brawny, hefty, muscular, powerful, sinewy, potent **Important**-> significant, important, crucial, authoritative **Deadly**->deathly, mortal, venomous, virulent, pernicious, pestilent
fear	HasProperty → bad HasProperty → make_person_upset CapableOf→ spread_disease CapableOf→ surprise_person CapableOf→sting HasProperty→loud	**Bad** -> tough, unfit, unsound, risky **Upsetting**-> disconcerting, upsetting **Contagious**-> catching, communicable, contagious, contractable, transmissible, transmittable **Surprising**-> unexpected **Stinging**-> biting, burning **Loud**-> brassy, cheap, flash, flashy, garish, gaudy, gimcrack, loud, meretricious, tacky, tatty, tawdry, trashy, forte

	CapableOf→buzz	**Buzzing**-> abuzz, buzzing
	CapableOf→ scare_person	**Scary**-> chilling, scary, shivery, shuddery
	CapableOf → hurt	**Hurtful**-> deleterious, hurtful, injurious
	CapableOf → bite	**Biting** -> barbed, biting, nipping, pungent, mordacious
	Causes → death	**Deadly** -> deathly, mortal, venomous, virulent, pernicious, pestilent
disgust	HasProperty→smelly	**Smelly**-> fetid, foetid, foul, foul-smelling, funky, noisome, smelly, stinking, ill-scented
	HasPropety→sticky	**Sticky**-> gluey, glutinous, gummy, mucilaginous, pasty, sticky, viscid, viscous
	HasProperty→bitter	**Bitter**-> acerb, acerbic, acid, acrid, bitter, blistering, caustic, sulfurous, sulphurous, virulent, vitriolic, acrimonious
	HasProperty→ addictive	**Addictive** -> addictive, habit-forming
	HasProperty→ bad_for_health	**Unhealthy**-> insalubrious, unhealthful, unhealthy

Table 2. Synonyms of properties of emotion-triggering concepts extracted from WordNet

In our future experiments, we can try to extract from ConceptNet the words which have properties, related to certain emotions and include them in the training set, in this way we can implement a bootstrapping schema.

Next iterations

Subsequently, we performed a new search in Twitter using as pattern fillers a manually selected subset of the extended set of properties. An example for the clusters obtained for the "joy" emotion is given in Table 3.

1: sweet heart; adulthood
2: a good day; this photo
3: your success; my company; your big day; this picture;
4: the money; his time; your holidays; my time; my money; your time
5: your day; the journey; my day; your meal; your stay; the ride; your night; your trip; our day; your weekend; my own company; your evening
6: cake; salad
7: my freedom; the final days; a great day; the happy times; the happy moment
8: life; music; games; sun; peace
9: beautiful; pretty; simple; sweet; sunny; pleasing
10: shopping; camp; dinner; company; girl
11: helping; staying; watching; reading; learning; talking; sleeping; spending; bringing; writing; working; playing; hanging

Table 3. Clusters of concepts that trigger "joy" and have similar properties, automatically discovered from tweets

5. Discussion and Conclusions

Human emotions are difficult to detect and classify, especially because most of the times, they are not expressed directly, through affect-related words. In many cases, emotions are expressed through concepts and descriptions of situations that based on our common-sense knowledge, trigger specific emotions.

Automatically detecting emotions is an even harsher issue to tackle, since it requires methods to gather and generalize the common-sense knowledge based on which emotions can be detected and classified from text.

In this paper, we showed how ConceptNet can be employed to extract properties of concepts in order to relate them to the emotions they trigger. Despite the relatively small amount of properties that can be exploited, the knowledge that we have gathered has allowed us to infer a large number of new concepts that trigger emotions. Preliminary inspections confirm thus that the affective link between these terms is given by the shared properties, which in turn translate into emotional reactions from their experiencers.

In the light of these findings, we will exploit additional information contained in knowledge bases such as Open Cyc[2] or ontologies such as SUMO, to further expand, based on shared properties, our knowledge base with new concepts.

The algorithm we proposed in this paper will allow us to properties, situations and events, related to persons and then detect implicit sentiment expressed in the news towards these people. For example, we could learn that "signed a peace pact" or "won a medal" or "was sentenced", "committed a crime" suggests indirectly certain judgement and consequently, a sentiment towards this person. This can be exploited to mark some sentences in the news stories captured by the Europe Media Monitor as outstanding and deserving more attention.

6. References

Balahur, A., Hermida, J.M., Montoyo, A. (2012). Building and Exploiting EmotiNet, a Knowledge Base for Emotion Detection Based on the Appraisal Theory Model. IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 88-101, Jan.-March 2012.

[2] http://sw.opencyc.org/

Balahur, A., Hermida, J.M., Montoyo, A., Muñoz, R. (2013a). Detecting implicit expressions of affect in text using EmotiNet and its extensions, Data & Knowledge Engineering, Volume 88, November 2013, pp. 113-125.

Balahur, A., Hermida, J.M., Tanev, H. (2013b). Detecting Implicit Emotion Expressions from Text Using Ontological Resources and Lexical Learning. In New Trends of Research in Ontologies and Lexical Resources, pp 235-255, Springer Verlag.

De Rivera, J. (1977).A Structural Theory of the Emotions. Psychological Issues, vol. 10, no. 4.

Elster, J. (1985). Sadder but wiser? Rationality and the emotions. Social Science Information, 24, 375-406.

Forgas, J.P. (2001). Introduction: Affect and Social Cognition. In Handbook of Affect and Social Cognition. Taylor& Francis.

Frijda, N. (1986). The Emotions. Cambridge Univ. Press,

Johnson-Laird, P.N., Oatley, K. (1989). The Language of Emotions: An Analysis of a Semantic Field. Cognition and Emotion, vol. 3, pp. 81-123, 1989.

Miller, G.A. (1995). WordNet: a lexical database for English. Commun. ACM 38, 11 (November 1995), 39-41.

Ortony, A., Clore, G.L., Collins, G. (1988). The Cognitive Structure of Emotions. Cambridge Univ. Press.

Pease, A., Niles, I. & Li, J. (2002). The Suggested Upper Merged Ontology: A Large Ontology for the Semantic Web and its Applications. In Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web.

Scherer, K., Wallbott, H. (1997). The ISEAR Questionnaire and Codebook. Geneva Emotion Research Group, 1997.

Singh, P., Lin, T., Mueller, E., Lim, G., Perkins, T., and Zhu, W.L. (2002). Open Mind Common Sense: Knowledge Acquisition from the General Public. In On the Move to Meaningful Internet Systems, 2002 - DOA/CoopIS/ODBASE 2002 Confederated International Conferences DOA, CoopIS and ODBASE 2002, Robert Meersman and Zahir Tari (Eds.). Springer-Verlag, London, UK, UK, 1223-1237.

Tanev, H., Zavarella, V., Linge, V., Kabadjov, M., Piskorski, J., Atkinson, M., Steinberger, R. (2009). Exploiting Machine Learning Techniques to Build an Event Extraction System for Portuguese and Spanish. In LINGUAMATICA (NLP journal for Iberian Languages), vol.2.