Emotions ontology for collaborative modelling and learning of emotional responses

Rosa Gil, Jordi Virgili-Gomá, Roberto García, Cindy Mason

Universitat de Lleida, Spain

Computer Science Department, Stanford University, USA

ABSTRACT

Emotions-aware applications are getting a lot of attention as a way to improve the user experience, and also thanks to increasingly affordable Brain–Computer Interfaces (BCI). Thus, projects collecting emotion-related data are proliferating, like social networks sentiment analysis or tracking students' engagement to reduce Massive Online Open Courses (MOOCs) drop out rates. All them require a common way to represent emotions so it can be more easily integrated, shared and reused by applications improving user experience. Due to the complexity of this data, our proposal is to use rich semantic models based on ontology. EmotionsOnto is a generic ontology for describing emotions and their detection and expression systems taking contextual and multimodal elements into account. The ontology has been applied in the context of EmoCS, a project that collaboratively collects emotion common sense and models it using the EmotionsOnto and other ontologies. Currently, emotion input is provided manually by users. However, experiments are being conducted to automatically measure users' emotional states using Brain–Computer Interfaces.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The emotional dimension of the interaction of humans with computers was considered for a long time a marginal factor (Brave & Nass, 2002). However, human beings are eminently emotional, as their social interaction is based on the ability to communicate their emotions and to perceive the emotional states of others. In this sense, emotions must be taken into account when implementing computing systems as a mechanism to improve the user experience.

Consequently, the study of emotion as a key faction in Human–Computer Interaction has gained a lot of traction resulting in well-established research and applications areas like affective computing. This paradigm deals with detecting, interpreting and responding to user's emotions when developing systems and devices (Picard, 2000).

Interest in this area is driven by a wide spectrum of promising applications, such as virtual reality, smart surveillance or perceptual interfaces (Tao & Tan, 2005). More recently, emotion-awareness has been highlighted as a key issue in online learning environments, especially in Massive Open Online Courses (MOOCs). In this case, massiveness makes it impossible to maintain a direct contact between instructors and learners that facilitates adapting the learning experience to learners needs while keeping their engagement.

This, among other factors, seems to explain the big dropout rate of most MOOCs, usually above 90% (Yang, Sinha, Adamson, & Rosé, 2013). Therefore, techniques like sentiment analysis have been applied to try to anticipate students' dropout (Wen, Yang, & Rosé, 2014). However, it is not enough to anticipate it, mechanisms should be provided to improve engagement and minimise it.

Therefore, the objective, in the case of MOOCs but also other Human–Computer interactions, should be to model, track and influence emotion through the concept of “emotional affordance” (Cheng, 2014), defined as the characteristics of a situation that influence emotions, as perceived by the people in that situation.

However, there are a great variety of theoretical models of emotions and there are different technologies that can be used for their implementation. Although there are many common properties, emotions are not universal: they are differently expressed in different cultures and languages, while many emotional properties are individual.

There is rarely a one-size-fits-all solution for the growing variety of computer users and interactions (Obrenovic, Garay, López, Fajardo, & Cearreta, 2005). Therefore, emotion-aware applications should be designed in a flexible way so a wider class of users can
use them. In this way, personalisation is necessary for more efficient interaction, better tuning and acceptance of developed systems.

There is a broad terminology related to affective states in human beings. The term “emotion” tends to be used in a broad sense, especially in technological contexts. Scherer (Scherer, 2000) proposed a number of taxonomies for these affective states. The original list was later modified and redefined in Douglas-Cowie and et al. (2006). This new list includes: Attitudes, Established emotion, Emergent Emotion (full-blown), Emergent Emotion (pressed), Moods, Partial emotion (topic shifting), Partial emotion (simmering), Stance towards person, Stance towards object/situation, Interpersonal bonds, Altered state of arousal, Altered state of control, Altered state of seriousness and Emotionless.

This paper focuses on Emergent Emotion (full-blown), instead of a global taxonomy of affective states. This is made to reduce the complexity of proposed domain so it is easier to deal with. Besides, for the same reason, focus is mainly devoted to emotion detection and expression systems instead of modelling internal emotion processing in humans.

In Douglas-Cowie et al. (2006), Emergent Emotion (full-blown) is defined as “states where the person’s whole system is caught up in the way they react to a particular person or situation”. It involves aspects such as:

- Distinctive positive or negative feelings about the people or situations involved.
- Impulses to act or express yourself in particular ways and avoid others.
- Distinctive changes in your body, for instance in your heart rate or tendency to sweat.
- Emotion does not last very long – it comes on quite quickly, and dies down reasonably soon (unless there is something very unusual happening).

Our contribution is to overcome the above mentioned limitations referring to the lack of flexibility when personalising affective computing applications by providing a generic ontology for describing detection and expression systems related with emotions, while taking contextual and multimodal elements into account. The ontology is proposed as a way to develop an easily implementable formal model, as it is based on the Web Ontology Language (OWL) standard (W3C OWL Working Group, 2012). However, the knowledge about emotions in the ontology can be used by affective computing applications with independence of underneath technology.

The proposed ontology in this paper is based on a generic model geared towards capturing the entities that take part in the Emergent Emotion process (López, Gil, García, Cearreta, & Garay, 2008). The model is presented in Section 3 and it has been formalised as an ontology following a classical ontology engineering methodology (García, 2006). The ontology has been developed to be as agnostic as possible regarding existing emotion theories. This way, developers of affective resources are not tied to a given theoretical approach.

In this sense, developers can use different theoretical approaches in the same ontological and technological framework. For instance, the categorical theory of emotions (Ekman, 1984), the dimensional one (Lang, 1979) or the one based on appraisal (Scherer, 1999).

Consequently, the ontology can help implementing emotion-aware applications based on a wider range of theoretical approaches. This flexibility and wide applicability of EmotionsOnto is in great part due to the fact that it is capable of modelling contexts. In order to do that, the DOLCE upper ontology (Gangemi, Giarrino, Masolo, Oltramari, & Schneider, 2002) is reused and extended, particularly the Description and Situation concepts.

Descriptions correspond to the representations for the situations, which then trigger and are associated with emotions. Moreover, in order to cope with the enormous range of different situations that might need to be associated with emotions, they are modelled using the building blocks provided by FrameNet (Scheffczyk, Baker, & Narayan, 2008). It is a big lexical database, with more than 10,000 word senses, structured following Frame Semantics (Fillmore, 2006). Frames fit really well with situations modelling as they try to explain words meaning by building a description of a type of event, relation, or entity and the participants in it.

The resulting ontology, EmotionsOnto, is validated in a real-world application that manages affective information, the Emotions Common Sense (EmoCS) initiative. It is inspired by the Open Mind Common Sense1 initiative, but integrating emotional feedback together with the common sense sentences contributed by users. In order to help understanding and automatising users input to guide emotions-aware applications, EmoCS provides dynamic forms with autocomplete features that guide the user while building structured descriptions of the situations to be associated with emotions, as detailed in Section 4. Moreover, we have started to explore the use of Brain–Computer Interfaces to automate emotional input gathering, as also reported in Section 4.

The rest of the paper is structured as follows. The next section presents several theories and concepts relevant for describing emotions, together with several topics related to ontologies and emotions. Then, the conceptual model underlying the proposed EmotionsOnto is introduced in Section 3. Then, the EmoCS application scenario is presented in Section 4. Finally, conclusions and future work conclude this paper in Section 5.

2. Related work

Emotion is a complex topic and findings of different areas, such as anthropology, psychology, and biology, are included in its wide-ranging discussion. In the field of psychology, definitions of emotion have been proposed with different theoretical orientations. In this sense, theories of emotions proposed by cognitive psychology are a useful starting point in order to describe emotion. Although several cognitive models of emotions exist, the most commonly ones used in affective computing area are the categorical (Ekman, 1984), dimensional (Lang, 1979) and appraisal (Scherer, 1999).

Lang (1979) also proposed analysing emotions according to three systems involved in their expression and detection: Subjective or verbal information (i.e. reports about perceived emotions described by users), Behavioural (i.e. facial and postural expressions, speech paralinguistic parameters), and Psychophysiological answers (such as heart rate, galvanic skin response –GSR–, and electroencephalographic response).

The subjective, behavioural and physiological correlates of emotions should be taken into account when possible. The correlations among the three systems could help computers to interpret ambiguous emotions. In that sense, more specific models or classifications that describe the components of each system of expression can be found in the literature and selected according to the particular case. These examples include acoustic correlates of speech (Scherer, 1986), verbal (Bradley, Lang, & Cuthbert, 1997) or facial expressions (Lang, 1979).

The emotional memory arisen from the experience of the individual and the cultural surroundings (also called socialised emotion) also has an influence on affective states in humans as well. Sociology of emotions has typically examined how affect arises,

---

linking emotions to particular types of interactions (Goffman, 1956). Besides, as the emotional answer is often socialised, it does not necessarily correspond to a pure emotional answer and it can mask real affective states.

It is noteworthy that, generally speaking, research has paid little attention to context in affective computing area (Cowie, Douglas-Cowie, & Cox, 2005). Context is inescapably linked to modality, and emotion is strongly multimodal as emotional cues may appear in various different channels. However, not all types of emotional cues tend to be available together, as context can affect relevant or accessible emotional cues.

For instance, Devillers, Abrilian, and Martin (2005) explained that emotional behaviour models require representing multiple levels involved in emotional processes: the emotional context, the emotion itself and associated multimodal behaviours. In that work, some appraisal descriptors derived from the appraisal model (Scherer, 1999) such as time-of-event were added in the context work, can deal with. Building an upper ontology is a very complex process and thus it is recommended to reuse existing upper ontologies instead of elaborating a full conceptualisation for the concepts in a specialised ontology.

Regarding the use of emergent emotion in real world applications, Lim and Aylett (2009) describe a novel emergent emotion model for a context-aware guide with emotions and personality that accompan-ies visitors touring an outdoor attraction.

2.1. Ontologies related with emotion modelling

Different ontologies have been proposed in literature with the aim of modelling emotion and affect related issues. These ontologies are normally focused on analysing concrete areas. For instance, in text analysis area, Mathieu (2005) presented a semantic lexicon in the field of feelings and emotions. This lexicon is described with an ontology. Words in the lexicon are emotionally labelled as positive, negative and neutral.

Emotional annotation has also been used in WordNet ontology (Fellbaum, 1998), producing the WordNetAffect extension (Strapparava & Valitutti, 2004). With the support of ontology technologies, users can retrieve information in a semantic manner (Chi, Peng, & Yang, 2007). A primary course of ontology building is relative to concept development. Focusing on speech, Galunov, Lobanov, and Zagoruiok (2004) present an ontology for speech signal recognition and synthesis where emotion is taken into account. On the other hand, focusing on the context, Benta, Rădui, and Creminé (2007) present an ontology based representation of the affective states for context aware applications which allows expressing the complex relations that are among the affective states and between these and the other context elements.

Although these kinds of unimodal approaches have relevance in their respective fields, they lack properly expressing the multimodal nature of human emotions. In this sense, multimodal ontologies for describing emotion have been proposed. For instance, Obrenovic et al. (2005) describes an ontology based on emotional cues that uses media properties from different sources to model emotion.

3. Contribution

The proposed ontology in this paper is based on a generic model geared towards capturing the entities that take part in the Emergent Emotion process. The model is summarised in Fig. 1. This model has been formalised as an ontology following a classical ontology engineering methodology (Garcia, 2006). The ontology, and the model it is based on, tries to be as emotions-theory agnostic as possible. The objective is to develop an ontology flexible enough to accommodate existing emotions theories, such as the ones presented in the literature review.

The underlying Emergent Emotion model is formalised using Semantic Web tools (Lytras & Garcia, 2008), more concretely the Web Ontology Language (OWL) (Lacy, 2005). OWL makes it possible to attain a great level of expressivity while producing a web ontology that can be easily shared through the web and thus be opened to third party extensions.

The different parts of the model have been developed using the primitives provided by OWL. The main building blocks are classes, which represent concepts in the model, and properties, which represent the relations among the concepts. The first step has been to model all the ovals in Fig. 1 as OWL classes. Therefore, there are classes such as EmergentEmotion, Description, Memory, Perception, and Sensor. For the relations among these concepts in the model, OWL object properties have been generated, i.e. hasInput, hasOutput, triggers, etc.

The ontology is completed with some axioms that restrict the kind of things that these properties can link. In OWL, these axioms are called OWL Restrictions and, in the context of a class, they specify to objects of what class does that property link to when applied to objects of the source class. For instance, the ontology contains a restriction that specifies that the triggers property when applied to the class Description points to objects of type EmergentEmotion.

The specification of the ontology (EmotionsOnto henceforth) in OWL format is available online.6 The previous formalisation of the Emergent Emotion model helps building an ontology that facilitates computerised emotions management. However, it provides little semantics apart from those explicitly present in the model. For instance, the ontology provides just some information about what a Sensor is. In order to enrich the ontology, we have taken existing upper ontologies into account.

Upper ontologies are very generic ontologies, about concepts like object or process, that settle down the ontological foundations about what is there in the world (Sowa, 1999). Consequently, they provide very basic and fundamental semantics about the kind of things that a more specialised ontology, like the one proposed in this work, can deal with. Building an upper ontology is a very complex process and thus it is recommended to reuse existing upper ontologies instead of elaborating a full conceptualisation for the concepts in a specialised ontology.

In order to do so, DOLCE (Gangemi et al., 2002) has been chosen, that stands for Descriptive Ontology for Linguistic and Cognitive Engineering, because it fits really well with the underlying considered cognitive aspects. In order to build the conceptual model. In fact, the Description and Situation concepts in this ontology have been reused from DOLCE and extended. These concepts provide a framework for representing contexts, which are the entities associated to EmergentEmotions in our model.

DOLCE also provides generic concepts that have been used in order to contextualise those in EmotionsOnto. First of all, there is Event that generalises any occurring thing in our model. There are some concretisations, i.e. Process, an event considered in its evolution, and Action, an event with at least one agent that participates in it.

On the other hand, there are objects. PhysicalObject has been used in order to contextualise concepts like Sensor, which we have detailed further in the ontology with artificial and biological sensors, and more specifically with human-like senses. SocialObject is the generalisation for Description and Situation, but also for

---

6 Emergent Emotion Ontology (EmotionsOnto), http://rhizomik.net/ontologies/emotionsonto.
Verbal, a kind of EmotionExpressionSystem together with Behavioural systems that have been specified as Actions.

All these relationships among EmotionsOnto and DOLCE concepts are shown in Fig. 2, where DOLCE concepts are coloured in grey. Moreover, the figure also shows additional concepts, apart from those shown in the model, that concretise concepts like EmotionExpression System or Sensor. Additionally, the ontology also includes the different kinds of Context identified during the conceptualisation process. SocialContext and EnvironmentalContext are modelled using Situation. On the other hand, PersonalContext is based on the Interface concept that includes both the EmotionExpression System and the Sensor concepts.

Although DOLCE provides the building block for modelling context, i.e. Description and Situation, concrete means that allow modelling the descriptions for the situations that trigger and are associated with emotions are necessary. However, to develop an ontology capable of dealing with the huge range of situations that might be associated with emotions is out of the scope of the ontology. Consequently, we have selected an existing ontology that provides such a wide scope called FrameNet (Scheffczyk et al., 2008).

FrameNet has been selected because it is better suited for modelling context as situations. FrameNet is based on the frame-modelling paradigm. A frame is a schematic representation of a typical situation (e.g. eating, removing, classifying, etc.) together with a list of the kinds of participants, properties and other conceptual roles that are seen as components of that situation. Moreover, it can be easily connected with DOLCE, as it can be noted in Fig. 2, where the concept Frame appears as a subclass of Description. Consequently, we can accomplish a smooth integration of DOLCE and FrameNet in the context of EmotionsOnto.

For instance, in order to model the situation “John enters the room” using FrameNet, it is possible to use the “enter.v” lexical unit, which belongs to the “Arriving” frame. This frame defines a set of Frame Elements (FEs) and some of them might be used in order to model the participants and properties of this situation. The Frame Element “Theme” is associated with the object that moves, in this case “John”. On the other hand, the FE “Goal” is associated with where the Theme ends up as a result of the motion, in this case “the room”.

4. Evaluation

The EmotionsOnto has been applied in the context of the Emotions Common Sense project (EmoCS), inspired by the Open Mind Common Sense (OMCS) initiative (Speer, 2007). OMCS aims to collaboratively build a database of common sense, while EmoCS does the same but for emotional common sense. This is done by providing a user interface that captures emotional input from users, who can associate common sense sentences, images and audio files to a set of predefined emotional states.

Another important difference between EmoCS and OMCS is that, while OMCS tries to capture common sense knowledge from user input that is basically unrestricted sentences, EmoCS restricts user input to structured representations based on FrameNet and associated emotional input. The main objective is to increase the functionality of the current OMCS website by including emotional aspects in its design. Moreover, an additional objective has been to facilitate processing users input. This is why, in EmoCS, users are required to provide structured input, which is based on FrameNet and linked to the EmotionsOnto.

As shown in Fig. 3, and as it can be observed in the EmoCS site, the user input form includes an autocomplete feature that guides the user. Once the user start typing about the kind of action she or he is willing to talk about, the autocomplete feature checks against the FrameNet ontology for candidates. When the user picks the desired action, the corresponding FrameNet frame is explored to detect the relevant frame elements and complete the form so it is shaped following the frame structure.

Finally, the user provides the emotional input using the slider in the right part of Fig. 3. Overall, when the form is submitted, a representation like the one in Fig. 4 is generated and associated to a set of emergent emotions. The FrameNet-based description and the emotional input can be also associated to uploaded images or referenced through an URL. In this case the image constitutes the situation and the description is about what is shown in the image.

The input is stored in a RDF store capable of mixing Web ontologies and RDF-based facts that also provides reasoning and


4 EmoCS, http://rhizomik.net/emocs (currently working just with Firefox).

5 Resource Description Framework, http://www.w3.org/RDF.
query answering. It can be used then to feed emotional-aware applications.

First, the application models emotional responses using EmotionsOnto so, afterwards, the systems based on the collected emotional knowledge are capable of responding to user emotional changes. For instance, if it is detected sadness, an emotion-aware system might respond by playing songs and displaying images that have been associated to a happy user response, as shown in Fig. 5.

It is important to note that the ontology does not help the system recognise the emotional responses, but to model the emotions and associate them to descriptions of situations and other parts of the Emergent Emotion Model. However, though this can be learnt from user interaction, our experience shows that the system requires some bootstrapping emotional knowledge that makes it useful and thus motivates user to continue using it. This is why

---

**Fig. 2.** EmotionsOnto core in the context of the DOLCE upper ontology.

**Fig. 3.** EmoCS input form.

**Fig. 4.** FrameNet based description of a situation generated from EmoCS input.
we have started exploring automated ways of annotating situation descriptions based on Brain–Computer Interfaces, like detailed in the next section.

### 4.1. Brain computer interface for emotion sensing

Brain Computer Interface (BCI) is a discipline that mainly came into common use with disabled users. Currently BCI is used in many different areas from marketing to learning systems through control robots, games or emotion detection (Gürkük & Nijholt, 2012).

Electroencephalography (EEG) provides a means of accessing and recording neural activity, allowing a computer to retrieve and analyse information from the brainwave patterns produce by thought. EEG has a much greater temporal resolution (allows millisecond-accuracy), however functional magnetic resonance imaging (fMRI) provides a much finer spatial resolution.

The most commercially widespread devices are Emotiv EEG and Neurosky. We have chosen Emotiv EEG because it features more sensors and thus provides more EEG data. Emotiv’s EEG Neuroheadset, shown in Fig. 6, is equipped with 14 saline sensors for the following EEG 10–20 standard scalp locations: AF3, AF4, F3, F4, F7, F8, FC5, FC6, P3 (CMS), P4 (DRL), P7, P8, T7, T8, O1 and O2 (Stern, 2013).

Prior to use, all felt pads on top of the sensors have to be moisturised with a saline solution. Emotiv headset is aimed at the gaming market and it is not classified as a medical device. It comes with some processing tools that can detect facial movements, emotional states and imagined motor movement (Emotiv’s Expressiv, Affectiv and Cognitiv suites respectively).

The Affective Suite has been used to monitor users’ emotional states while they were shown pictures or videos illustrating different situations, previously annotated using EmotionsOnto. The emotional states detected by Emotiv headset are then modelled using EmotionsOnto, concretely as the emergent emotions triggered by the descriptions associated to the situations shown to users.

Unfortunately, as shown in Fig. 7, the output generated by Emotiv Affectiv Suite is not available outside the provided SDK. Therefore, emotional annotations based on the headset where generated manually after analysing the Affectiv Suite screen output. Moreover, the emotion output is constrained to just 4 measures that do not conform to any established categorisation of emotions. The affective suite has four emotional states that are calculated depending on the frequency of the different brainwaves: “Excitement”, “Frustration”, “Meditation” and “Engagement/Boredom”. Emotiv admits that the names of the emotional states may not accurately reflect an emotion. For instance, the “Excitement” measure might increase both if the user feels fear or surprise. Therefore, it is not easy to distinguish them.

Fortunately, communities like OpenVibe or OpenEEG have emerged and provide more data rich software and applications suitable for research purposes. Data is sent from Emotiv EEG headset in encrypted form via Bluetooth. However, raw data can be unencrypted and collected by a Python library for Emotiv EEG.

Currently, our focus is on processing this raw data to at least detect the 6 emotions identified by Ekman: anger, disgust, fear, happiness, sadness and surprise (Ekman, 1984). This is work in progress that is detailed next, in the conclusions and future work section.

### 5. Conclusions and future work

Affectivity in human beings is a very complex term where a lot of multidisciplinary research has been performed from different fields. Providing a computerised basis to perform a knowledge base that allows dealing with affective related concepts such as emotions requires a ground. Performed work tries to provide such ground. In this paper we present a generic model for describing emotions and their detection and expression systems taking contextual and multimodal elements into account. The model is formalised as an easily computerised ontology.

The ontology has been developed to be as totally agnostic as possible regarding existing emotion theories. Although this fact
can seem to be surprising, it is made due to allow independence from theoretical approaches, so developers of affective resources are not tied to a given theoretical approach. In this sense, different theoretical approaches can be used by the developers.

Thus, it makes the ontology valid for a wide range of proposed theoretical approaches and applications domain, like students’ engagement in MOOCs. It is remarkable that context has received little attention regarding emotion-aware application development. This work takes this concept into consideration as a necessary component for modelling emotion. In this sense, proposed ontology is based on the definition of relevant contextual elements.

Proposed ontology does also have implications regarding emotion-aware applications development. On the one hand, it allows integrating emotion-related context and multimedia elements. On the other hand, defining sensors in the way performed allows making use of them as user interface elements for inputs (such as text through the keyboard, utterances through the microphone) and outputs (such as emotional pictures in user interfaces, text output on the screen or embodied avatars).

They can be linked with emotion expression and recognition systems. This approach has proven to be very useful for describing how emotion-aware applications work. Finally, it must be highlighted that being performed ontology developed using standards like OWL, it is also valid for a wide range of software development technologies and environments and thus can be used as a basis to engineer emotion-aware applications.

The ontology has been applied in the context of the Emotions Common Sense project, inspired by MIT’s Open Mind Common Sense. It enriches that initiative by providing also emotional input and by providing more structured user input based on.
EmotionsOnto and other ontologies that can be more easily put into practice to build emotions-aware applications.

Future work from the EmotionsOnto perspective focuses on extending the ontology beyond Emergent Emotion. The first extension considered is to model affective states in humans in order to make the ontology capable of modelling more complex aspects of human affectivity. This will make possible to model users bearing above-mentioned affective states in mind. These enriched user models enable including aspects related with user disabilities and developing applications even more adapted to their needs. Finally, the inclusion of social context in the ontology allows exploring emotion in computerised social environments such as social networks.

On the other hand, from the evaluation of EmotionsOnto in the context of EmoCS, future work aims to provide automatic emotion annotations for situation descriptions based on the data collected by the Emotiv EEG neuroheadset. Currently, we are collecting raw data from EEG sensors and processing it. Training sets, associating situations to expected emotions and direct user feedback, are being user together with machine learning algorithms to train a system capable of producing Ekman emotions out of the EEG data.

Moreover, neuroheadset measurements are going to be complemented with other kinds of sensor to increase robustness. A preliminary study of the state of the art shows that measuring skin conductance gives a quite reliable measure of stress that can complement EEG readings to better categorise the corresponding emotions (Ekman, 1952).

Acknowledgements

This work has been partially supported by the research project IndAGUs (Spanish Government TIN2012-37826-C02) Infrastructures for Sustainable Open Government Data with Geospatial Features.

References


