



Emotion Ontology for Context Awareness

Franck Berthelon, Peter Sander

► To cite this version:

Franck Berthelon, Peter Sander. Emotion Ontology for Context Awareness. Coginfocom 2013 - 4th IEEE Conference on Cognitive Infocommunicaitons, Dec 2013, Budapest, Hungary. hal-00908543

HAL Id: hal-00908543

<https://hal.science/hal-00908543>

Submitted on 24 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Emotion Ontology for Context Awareness

Franck Berthelon
Laboratoire I3S
Sophia-Antipolis, France
Email: berthelo@i3s.unice.fr

Peter Sander
Laboratoire I3S
Sophia-Antipolis, France
Email: sander@polytech.unice.fr

Abstract—We present an emotion ontology for describing and reasoning on emotion context in order to improve emotion detection based on bodily expression. We incorporate context into the two-factor theory of emotion (bodily reaction plus cognitive input) and demonstrate the importance of context in the emotion experience. In attempting to determine emotion felt by another person, the bodily expression of their emotion is the only evidence directly available, eg, “John looks angry”. Our motivation in this paper is to bring context into the emotion-modulating cognitive input, eg, we know that John is a generally calm person, so we can conclude from expression (anger) plus context (calm) that John is not only angry, but that “John must be furious”. We use a well known interoperable reasoning tool, an ontology, to bring context into the implementation of the emotion detection process. Our emotion ontology (EmOCA) allow us to describe and to reason about philia and phobia in order to modulate emotion determined from expression. We present an experiment suggesting that people use such a strategy to incorporate contextual information when determining what emotion another person may be feeling.

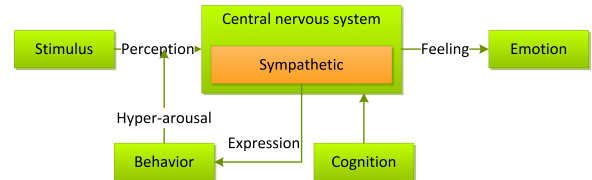
I. INTRODUCTION

Detecting emotions in others is something that most people can do naturally, and emotion determination by computer is a field of increasing interest, see for instance articles on affective computing[1] or cognitive infocommunications[2]. That there is no way to have direct access to another person’s emotions somewhat complicates the task; instead we can only interpret the person’s outward expressions. Among evidence for determining emotions felt by others, facial expressions play an important part[3], and are what we are concerned with in this article. However, people can also incorporate other less immediate information, for example knowing that John is generally calm whereas Jean-Claude is volatile will temper our appreciation of what they are really feeling even though their outward expressions may be similar.

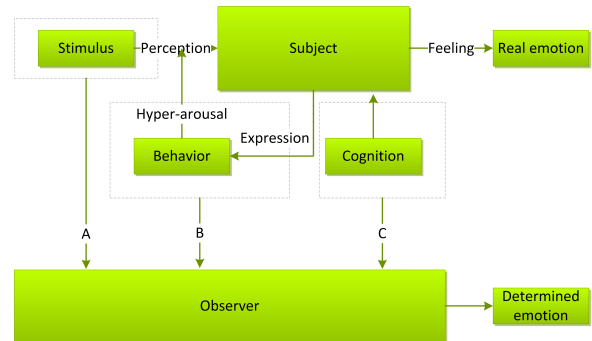
This paper investigates a computational tool for incorporating such context into emotion determination, and presents a study showing how context can influence a person’s determination of emotion felt by someone else. We present work on using context for computational emotion determination elsewhere[4].

There exist numerous and often contradictory theories on emotion[5], [6], [7]; we are particularly motivated by the *two-factor theory*[8]. This theory proposes that after perceiving a stimulus (internal eg, souvenir or external eg, real world object), a person first experiences an immediate bodily reaction, and secondly a cognitive process is launched which uses the bodily expression and cognitive valuation to define their emotion, see Figure 1a.

An observer trying to determine the subject’s emotion typically only has access to the bodily expression, eg, they



(a) The two-factor theory proposes that after a stimulus a person experiences an immediate bodily reaction, followed by a cognitive valuation. These elements combine to produce their emotion



(b) An observer typically has only access to the subject’s physical expression in order to determine their emotion (determination A). We hypothesize that if the observer also has access to the stimulus (B), and to a limited aspect of the subject’s cognitive valuation (C), then each of A, B, C, or combinations thereof, can give rise to the observer determining a different emotion felt by the subject

Fig. 1: Two-factor model of emotion

might see that the subject “looks angry”. In this paper, we assume that the observer also has access to some of the subject’s cognitive input as well, in the form of knowing something about a particular aspect of the subject’s personality, see Figure 1b. We hypothesize that having access to such contextual information about the subject can lead the observer to modulate the emotion that they determine the subject to be feeling¹.

Using such context (B and C in Figure 1b) is hard because of the amount of user-specific knowledge and data needed, such as a complete enumeration of their likes and dislikes. In this paper, we propose a computational tool, viz, an emotion ontology, to model contextual information. An ontology has the advantage of being able to draw on and organize existing

¹We refer to the person feeling an emotion in response to a stimulus as “the subject”. We refer to the cognitive being (person or artificial system) detecting the subject’s emotional state as “the observer”.

information – that is indeed its purpose[9] – and an ontology can be associated with an inference engine to enable reasoning about the information. The ontology can be used to temper emotion determination based on facial expression alone. For example, we see that Bob is expressing fear, we see that Bob is looking at a spider, and since we know that Bob suffers from arachnophobia we conclude that he is feeling *really* afraid.

We start by discussing the use of the ontology to describe and to reason about context. We then propose a modelization of a restricted emotion ontology for phobia and philia. Finally, we carry out an experiment to illustrate the effect of using this particular context in people’s estimation of emotion felt by others.

A. Background

This paper is a part of work toward a computational emotion detection process[4]. In a simple emotion detection architecture, processing might consist of the following steps: sensor input → interpretation → classification. Figure 2 shows our architecture, which uses roughly the same decomposition, but instead of a sensor and classification algorithm, we propose a regression algorithm followed by contextualisation with an ontology. Here, we are concerned with the top part of the figure, passing from a continuous dimensional model (see below), via an ontology, to a categorical model. Basically, this means going from sensor measurements to an emotion label such as *joy* or *disgust*.

Our goal with this architecture is to abstract our detector away from the specific sensors used to determine emotion, eg, facial expression, heart rate measurement, etc. At the same time, we propose combining bodily expression and contextual information to improve emotion detection. Our hypothesis is that we perceive emotion in dimensional terms[10] but that we speak in categorial terms[11] (see below for more detail). Which means that people can feel an infinite nuance of emotion but are limited to a discrete expression through a finite vocabulary.

II. EMOTION ONTOLOGY FOR CONTEXT AWARENESS

Using emotion context is a complex task, the more so due to the potentially huge resources required to model every possible stimulus able to elicit an emotion. However, we can find a ready-made helpful tool in the domain of web semantics[12], which is aimed at structuring data over the web in order to allow sharing and reasoning on it. An *ontology*[9] is a way to model semantic relations between concepts, with an associated inference engine providing reasoning on the data. As well, an ontology is capable of interoperability with multiple knowledge bases. This characteristic is very important for our purposes because we are searching for a relation between context and user knowledge without needing a complete description of it. So we want to use this capacity to describe how the context can impact emotion experience and reuse descriptions of already-defined concepts to find semantic relations. Our prototype context is restricted to only philia and phobia; our aim in this article is to show the usability of an ontology to describe this restricted context. We intend to extend this EmOCA to more complex contexts in future work.

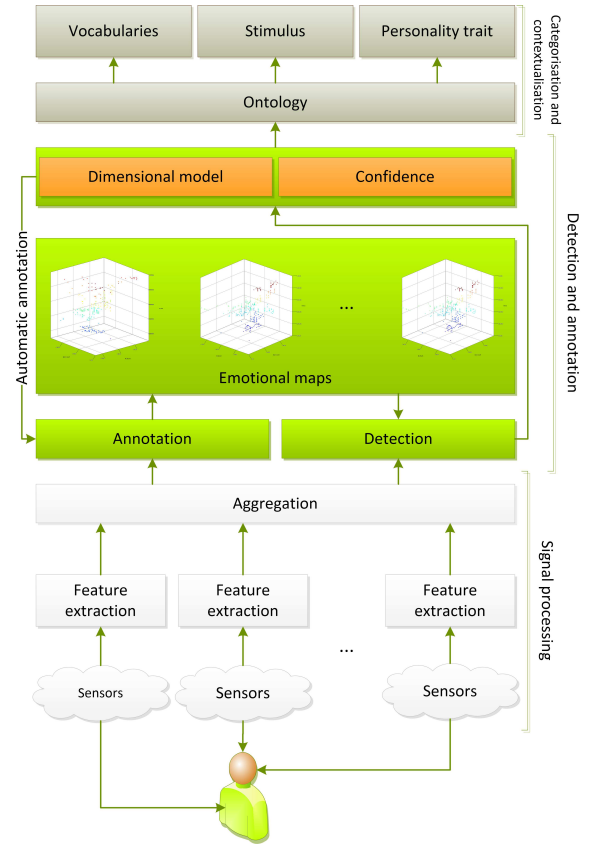


Fig. 2: Emotion processing chain, from sensor measurements to a dimensional model, then via an ontology to a categorial model. Only the latter step is considered in this paper

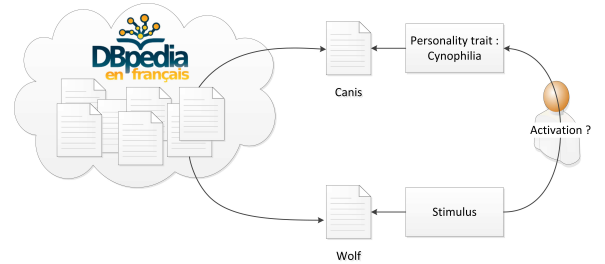


Fig. 3: A use case of our emotion ontology

Figure 3 shows an example of the kind of relations we are searching for during emotion detection. In this example, a dog lover perceives a wolf. In our ontology, cynophilia is described as a concept related to the dbpedia² concept of canis. When the person sees the wolf, we describe the stimulus as a concept in relation with the dbpedia concept of wolf. Now, we can reason on the semantic links between wolf and canis as described by the knowledge base of dbpedia. In this case, there is a relation of subsumption between wolf and canis so we can infer that the cynophilia trait is activated. This first attempt may be somewhat simplistic but we can in further work improve reasoning capabilities using semantic distance [13].

²<http://dbpedia.org/About>

A possible scenario where emotion recognition modulated by contextual information can be useful might be the following. Consider remote surveillance of seniors living at home in the (not too distant) future when today’s computer-savvy 40- or 50-somethings have aged. In the course of his interactions with the virtual world Ted frequently becomes angry or frustrated at some “stupid web site” because it “doesn’t work right”. But there is no cause for worry because the system knows that Ted is generally pretty excitable. On the other hand, when Alice suddenly starts getting angry for the same reasons, the system rapidly alerts a care-giver that something may be worth looking into because it knows that Alice is generally competent and calm. It reasons that, given her age and personality, the sudden change in emotional reaction might have a physiological cause, perhaps a cerebrovascular accident, or a complication due to her known diabetes?

A. Modelization

Our modelization is divided into two interconnected parts. The first, *categorisation*, aims to map continuous measures into a controlled vocabulary. The second, *contextualisation*, aims to use context to temper emotion expression as a function of a person’s contextual knowledge.

1) *Categorization*: As describe in section I-A, we use a regression algorithm to produce continuous measures along the Russel valence/arousal emotion model[14]. To improve understanding, we would like to convert these continuous measures into an intelligible label. However, among psychologists there is no single generally accepted set of emotion labels - numerous vocabulary terms exist. We have to choose a set, so in our ontology we start by representing Ekman’s six basic emotions[15]. Our goal is not to propose a unique mapping from dimensional model continuous measures to categorial model labels; we would rather leave to experts the possibility of describing the semantic relations between these two kinds of model.

We propose the possibility of describing a label as an interval in a dimensional model. For the purpose of our work, we describe only the semantic relation between Russel’s and Ekman’s models. We show in our modelization that the ontology can be extended with all kinds of dimensional and categorial models. Figure 4 shows two different ways to define an interval in our Russel model. By using both cartesian and polar coordinates, we increase the flexibility to determine how an emotion can be described. A polar description uses radius to define intensity, which simplifies understanding the description, while cartesian description eases the decomposition of emotional space.

2) *Contextualisation*: The second aim of the ontology is to allow use of contextual information about the subject to weight the continuous measurements of the dimensional model. For the work described here, we are only interested in one personality trait - phobia / philia. Contextual information is used to identify which phobia or philia are activated and to retrieve the impact of these traits on emotion. We describe a personality trait as a concept in relation with a generic concept that can be retrieved from any knowledge base compatible with our ontology. Similarly, we define stimulus as a concept in relation with any other concept of a distinct ontology. For

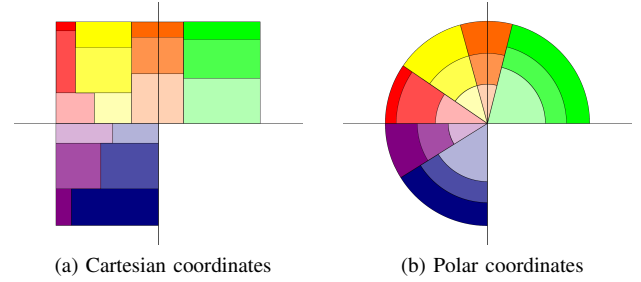


Fig. 4: Two different ways to describe intervals. Colour corresponds to valence[14], the quality of the emotion, and colour shade corresponds to arousal, the intensity of the emotion

example, we can define arachnophobia as a personality trait in relation with the concept of arachnid from dbpedia, and a stimulus like black widow in relation with the concept of “latrodectus”. Now with those two definitions, we can infer a subsumption relation between arachnid and “latrodectus” and retrieve the impact of arachnophobia on emotion.

Emotional impact defines the weighting of the dimensional measure for each dimension. So within Russel’s model, we can say that arachnophobia decreases valence and increases arousal, which pushes the emotion toward increased fear.

3) *Overview*: The use case described above is modeled in our emotion ontology. Figure 5 shows a graphic representation of our ontology where the modelization is divided into three main parts. The first one describes the observed person and the perceived stimulus. The second part describes the impact between personality traits and emotion description, and the last part describes the mapping from a dimensional model to a categorial one.

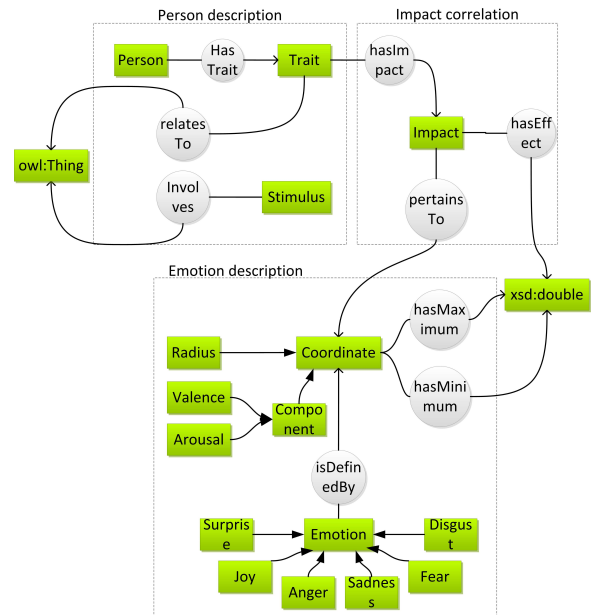


Fig. 5: A use case of our emotion ontology

The current implementation for this ontology uses RDFs³ and RDF⁴ languages. For this demonstration we don't need high levels of expressivity, but in a next step we can switch to a more expressive language such as OWL if we need to extend context to other things like personality traits (the big five[16]), or mood[17]. A full version of EmOCA is available for development⁵.

B. Reasoning

The knowledge base is used to structure personality traits and their semantic relations with stimulus and impact on dimensional measures. We now propose a way to reason on this information with the CORESE inference engine[18] and a SPARQL request⁶. Our aim with this request is to determine which personality trait is activated and how it can impact measurement weighting. Figure 6 shows an overview of how our request works. The top boxes contain inputs needed by the request to apply context reasoning to the categorisation process.

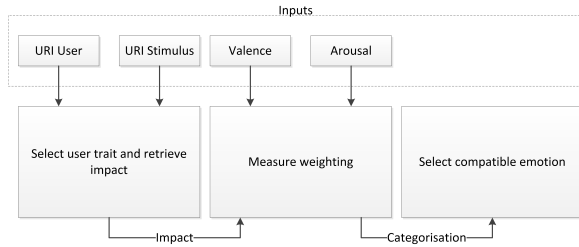


Fig. 6: Schematisation of the SPARQL request

With 4 inputs (user, stimulus and dimensional measures), the SPARQL request retrieves the personality trait activated by the stimulus and the impact on emotion expression measurement. Weighting is directly applied on the measures inside the request. Finally, the request maps dimensional measures into categorial labels.

III. USING EMOTION CONTEXT

We have described above how we developed our ontology and how we can use it for reasoning. We now propose an experiment to demonstrate the relevance of context in the emotion detection process. We show that giving an observer some knowledge about a subject (the subject's context in expressing an emotion) can affect the observer's appreciation of what emotion the subject is feeling.

A. Experiment

1) *Overview:* Our experiment aims to determine the impact of contextual information on emotion detection. Based on Schachter and Singer's two-factor theory[8], we use context to influence emotion detection. In this paper we restrict context to philia and phobia. The experiment was carried out on a group of people, based on answering questions while viewing

a set of images extracted from a custom video used in the learning step for our emotional maps (see figure 2). The images and questions were presented in the form of a two-step questionnaire.

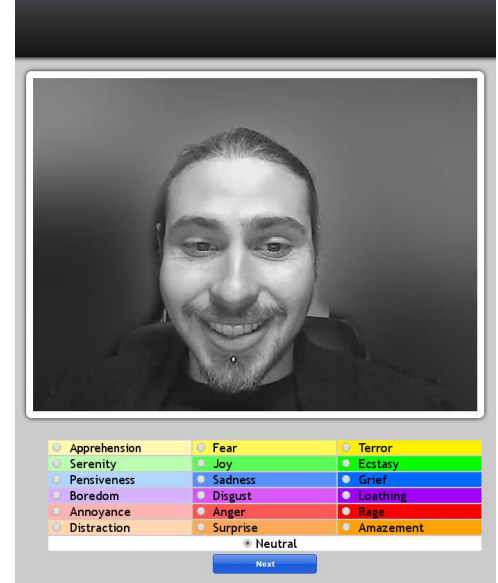


Fig. 7: Screenshot of the first step of the questionnaire; no context information

The first step of the emotion-testing experiment aims to recognize emotion from facial expression without any information about context. Figure 7 shows one question of this step. We present an observer with a picture of a subject (actually an actor to obtain canonical expressions) expressing an emotion, and a selection of choices containing Ekman's six basic emotions[15] divided into three intensities each, eg, fear is divided into anxiety, fear, terror. As a control, in the sequence we randomly present Ekman's images to measure observers' ability to recognize emotion.

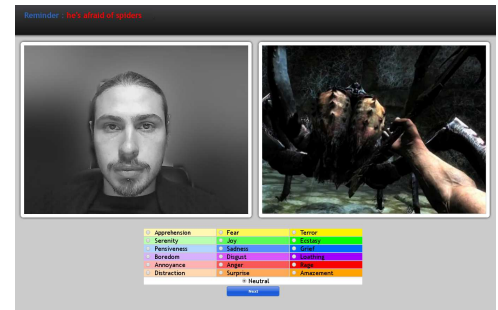


Fig. 8: Screenshot of the second step of the questionnaire; presentation includes context information

The second step of the questionnaire uses the same images of the subject's facial expressions, combined with a picture representing the stimulus shown to the subject and a personality trait such as "he's afraid of spiders" attributed to the subject; this is the context in which the subject experiences an emotion. Figure 8 shows the interface with the subject's picture, the stimulus picture, the personality trait, and the selection choices.

³<http://www.w3.org/TR/rdf-schema/>

⁴<http://www.w3.org/RDF/>

⁵<http://ns.inria.fr/emoca/>

⁶<http://www.w3.org/TR/rdf-sparql-query/>

2) *Experiment*: The group of participating observers was composed of six men and four women between 20 and 36 years old, divided equally into two groups. The first group (the “phobe” group) believed that the subject was arachnophobic, and the second group (the “phile” group) that the subject liked bears. In the first step of the questionnaire, each observer in each group was presented randomly with all six of Ekman’s pictures, one for each of the basic emotions, and 32 pictures of the subject’s expression of the same emotions. In the second step, only the pictures of the subject were presented, along with the supposed stimulus and the subject’s character trait. To test phobia, the group believing the subject to be arachnophobic saw the spider stimulus; to test philia, the group who believed that the subject liked bears was presented with the bear stimulus.

3) *Hypothesis*: We define two hypotheses for this experiment. The first is that observers would change their choices for the same facial expression when stimulus and character trait were also presented, meaning that context is used to modulate emotion detection. Observers could change the emotion type (fear to disgust for example) or its intensity (joy to ecstasy). We wanted to determine whether only the emotion intensity changes, or the emotion label itself.

The second hypothesis pertains to our work on computational emotion determination systems. It says that there is enough variability in how people see the effects of context that this becomes a parameter in such a system rather than a fixed constant. Therefore, in our ontology we do not fix phobia and philia effects for all cases – they must be determined case-by-case through a user-specific calibration process (the interested reader is referred to [4]).

B. Results

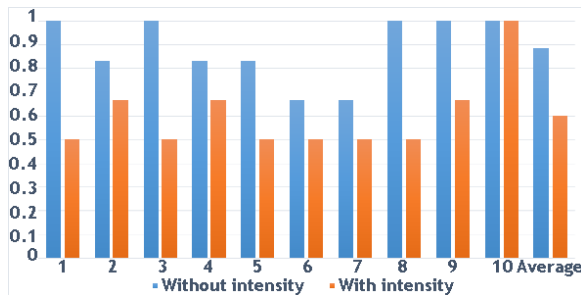


Fig. 9: Performance of each observer. The blue bars represent the percentage of correct matches without considering intensity; the orange bars are matches using intensity

1) *No context*: From the first step of the questionnaire, the first result is to set a control which allows us to determine how good each participant is at recognizing emotion in order to eliminate those incapable of distinguishing emotions sufficiently well for our purposes. For this, we use only the results given for Ekman’s images. Figure 9 shows results for our observers and reveals an average recognition rate. We consider the score in two different ways:

- without considering intensity differences for the same emotion, eg, anxiety, anger and rage would be the

same emotion – all observers scored at least 66% correct choices;

- taking into account emotion intensity differences, eg, the above emotions are considered as different – all observers scored at least 50% correct choices. Ekman’s images were considered to display mid-intensity emotions.

We considered this control result acceptable and all observers continued on to the next step of the experiment.

2) *Using context*: The second step involved presenting the subject images with context, and comparing the observers’ performance with the subject images from the first step.

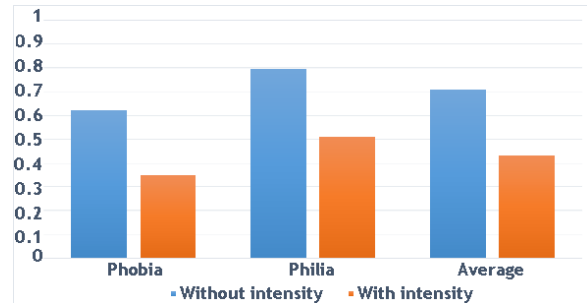


Fig. 10: Similarity between the two steps per user. The blue bars represent the percentage of pictures labeled consistently without considering context. The orange bars represent the percentage of pictures labeled consistently taking context into account. The leftmost group represents results for the phobe group, the phile group is in the middle, and the average over all observers is on the right

For each observer, we compute the similarity between labels selected with and without context information. Results are shown by group in Figure 10. The leftmost blue bar shows that on average the phobe group maintained a consistent choice (same emotion with and without context) in more than 60% of the images, discounting intensity. The orange bar is the same, but including intensity. The middle block is for the phile group, and the right-hand block is the average of all observers.

A lower level of this score shows a higher impact of the context in emotion recognition - less consistency and hence more change with contextual information. For our ten observers, we found that they generally kept the same kind of emotion: more than 60% for phobia and little less than 80% in the case of philia. But, when we consider emotion intensities the score is around 35% for phobia and around 50% for philia. This suggests that observers use both intensities and emotion type when they are aware of the context. Overall, observers change their emotion choice more than 55% of the time. Context does seem to have a significant effect.

Figure 11 shows the percentage of agreement for an expression with context over all observers. On the left and center, respectively the phobia and philia groups, 50% of the observers give the same annotation instead of the first step of the questionnaire where 65% of observers are in agreement without considering context.

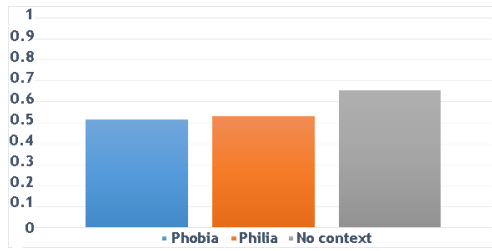


Fig. 11: Agreement between observers with the emotion conveyed by the subject in each image. The blue bar represents percentage agreement in the images with a spider stimulus, the orange bar shows the bear stimulus, the grey bar without context

C. Discussion of experiment

First, we can deduce from the first step of this questionnaire that emotion recognition from images of facial expressions is not necessarily easy! However, what we wanted to study with this questionnaire is the importance of the emotional context in the interpretation of emotion. On the scale of our experiment, we can say that there is a significant difference in interpretation when the observers were aware of the emotional context, ie, they know something about the subject whose emotions they are determining. Moreover, this difference is moderate if we do not consider the emotional intensity – the impact of phobia or philia when a person is faced with a corresponding stimulus does not often change the emotion label but rather only changes its intensity. On average, the proposed facial expression emotion remains the same in more than 70% of the cases. However, considering intensity of the emotion, labels stay unchanged in less than 45% of the cases. This observed behavior allows us to highlight the importance of weighting emotion with context after the interpretation of facial expressions.

In addition, the type of personality trait described changes the interpretation of emotions. We observed a greater impact on the recognition of emotions from phobia than from philia. However, the calculation of the weighting is not straightforwardly obvious. The last result suggests that it is difficult to define a fixed generic impact of a phobia or philia. The observers failed to agree on the emotion felt when they were aware of the context. Only about half of the observers reached an agreement considering the context, whereas they were 65% to agree without considering the context. This suggests that the definition of the impact of philia or phobia on the emotion felt depends on the individual and that the ontology weighting for this effect needs to be calibrated instead of hardwired as a generic configuration.

IV. CONCLUSION

Globally, we demonstrated how ontology can be a useful tool for using context in the emotion determination process, and suggested experimentally how humans use context to interpret emotion expressions. The two most important characteristics we used to argue that ontology is a good tool in that case are: interoperability with other knowledge bases and reasoning capability through an inference engine. In the case in this paper, the first feature allows concentrating uniquely

on the description of semantic relations between philia/phobia and their impact on emotion. The second feature allows finding semantic relations between two concepts not necessarily described in our knowledge base and makes philia or phobia explicit. The simple ontology is but a first step toward a contextual emotion ontology used within a computational emotion determination system. In this paper, we limited our focus to philia and phobia to demonstrate the usability of ontology, and in further work we will increase contextual knowledge to take in more personality traits.

While are not yet hazarding an opinion on whether artificial cognitive systems will recognize emotions as well (or as poorly) as people, we are however convinced that this line of research will shed some light on how people do it.

ACKNOWLEDGMENT

The authors wish to thanks Stonetrip company and region of PACA for their support.

REFERENCES

- [1] R. W. Picard, "Affective computing," *Pattern Recognition*, 1995.
- [2] P. Baranyi and A. Csapo, "Cognitive infocommunications: Coginfo-com," in *11th IEEE International Symposium on Computational Intelligence and Informatics*, 2010.
- [3] P. Ekman and W. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Malor Books, 2003.
- [4] F. Berthelon, "Modélisation et détection des émotions à partir de données expressives et contextuelles," Ph.D. dissertation, University of Nice-Sophia Antipolis, In progress.
- [5] W. James, "What is an emotion?" *Mind*, 1884.
- [6] W. B. Cannon, "The james-lange theory of emotions: A critical examination and alternative theory," *American Journal of Psychology*, vol. 39, pp. 106–124, 1927.
- [7] K. R. Scherer, "Emotions as episodes of subsystem synchronization driven by nonlinear appraisal processes," 2000.
- [8] S. Schachter and J. Singer, "Cognitive, social and physiological determinants of emotional state," *Psychological Review*, 1979.
- [9] T. R. Gruber *et al.*, "Toward principles for the design of ontologies used for knowledge sharing," *International journal of human computer studies*, vol. 43, no. 5, pp. 907–928, 1995.
- [10] A. Mehrabian and J. A. Russell, *An approach to environmental psychology*. the MIT Press, 1974.
- [11] R. Plutchik, "The nature of emotions," *American Scientist*, 2001.
- [12] T. Berners-Lee, J. Hendler, O. Lassila *et al.*, "The semantic web," *Scientific american*, vol. 284, no. 5, pp. 28–37, 2001.
- [13] A. Budanitsky and G. Hirst, "Semantic distance in wordnet: An experimental, application-oriented evaluation of five measures," in *Workshop on WordNet and Other Lexical Resources*, vol. 2, 2001.
- [14] J. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, 1980.
- [15] P. Ekman, "Universals and cultural differences in facial expressions of emotion," in *Nebraska symposium on motivation*. University of Nebraska Press, 1971.
- [16] L. R. Goldberg, "An alternative" description of personality": the big-five factor structure," *Journal of personality and social psychology*, vol. 59, no. 6, p. 1216, 1990.
- [17] S. Mardaga, O. Lalouaux, and M. Hansenne, "Personality traits modulate skin conductance response to emotional pictures: An investigation with cloningers model of personality," *Personality and individual differences*, vol. 40, no. 8, pp. 1603–1614, 2006.
- [18] O. Corby and C. Faron Zucker, "The kgram abstract machine for knowledge graph querying," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, vol. 1. IEEE, 2010, pp. 338–341.