SEMAINE
THE SENSITIVE AGENT PROJECT

D7c
Final report on standardisation activities

Date: 17 December 2010

Dissemination level: Public

ICT project contract no.	211486
Project title	**SEMAINE** **Sustained Emotionally coloured Machine-human Interaction using Nonverbal Expression**
Contractual date of delivery	*31 December 2010*
Actual date of delivery	*17 December 2010*
Deliverable number	D7c
Deliverable title	Final report on standardisation activities
Type	Report
Number of pages	9
WP contributing to the deliverable	WP 7
Responsible for task	Dirk Heylen (d.k.j.heylen@ewi.utwente.nl)
Author(s)	Dirk Heylen, Catherine Pelachaud, Marc Schröder
EC Project Officer	Philippe Gelin

Table of Contents

1 Executive Summary

 This report lists the activities carried out by the SEMAINE project partners on standardisation as part of Workpackage 7 on Dissemination and Standardisation (for further information on Dissemination Events see the yearly reports D0a, D0b and D0c).

Several members have been actively involved in defining mark-up languages that allow a technological component to represent and process data, and to enable interoperability between different technological components processing. The implementation of the SEMAINE system has taken into consideration compliance to these emerging standards where possible. This is detailed in Section 2.

In addition to direct contributions to standards, standards for evaluation and acoustic features were set in the field of speech emotion and general paralinguistics analysis. This is described in more detail in Section 3.

2 Standardisation Activities: markup languages

The goal of markup languages is twofold: to allow a technological component to represent and process data, and to enable interoperability between different technological components processing the data. Partners in the SEMAINE consortium have worked on standardisation through participation in international groups dealing with several components that are relevant for the system developed in SEMAINE. The state-of-the-art, activities and results are briefly listed below.

2.1 Function Markup Language and Behavior Markup Language

2.1.1 Description

SAIBA is concerned with: "the construction of a common framework for multimodal behavior generation that allows the researchers to pool their efforts and speed up construction of whole multimodal interaction systems. The overall framework is called SAIBA, and within it, two important interfaces are being defined: FML and BML. FML describes communicative function at the level of intent, regardless of surface behavior, whereas BML describes the surface form which then is realized by an animation engine."

2.1.2 Current state of the standard

After two workshops at AAMAS (2008 and 2009) in which researchers from the Intelligent Virtual Agent community were invited to contribute their ideas on the Function Markup Language, a third workshop on the topic was organized by David Traum and Stacy Marsella at the Institute of Creative Technologies at Marina del Rey (LA). The goal of the workshop was no longer to simply identify the issues pertaining to the definition of a common markup language that aims to specify the functional (meaning/ intentional) level of communication but to come up with the first XML example.

A first version of the Behaviour Markup Language has been specified and various sites (including CNRS and UT amongst them) are currently implementing behaviour realisers that are BML compliant.

2.1.3 Activities carried out

UT and CNRS were co-organisers of two workshops devoted to FML at AAMAS. One in 2008 and one in 2009. They participated in several working group meetings.
- UT and CNRS actively participated in discussion groups on BML and FML. The last BML face-to-face meeting was in Amsterdam at ACII (2009) and the last meeting on FML was in Los Angeles (see above).
- Besides face-to-face meetings active discussions take place through email.

The results were presented at several occassions.
- Presentations about FML were held at the two AAMAS workshops by both UT and CNRS members.
- A joint paper on FML by UT, CNRS and other members of the working group on FML was

presented at IVA 2009
- The outcome of the Los Angeles workshop was presented at IVA 2010.
- A presentation on FML was given at the SSPNET Workshop in Paris (October 2010).
- BML was presented at IVA (2008).

2.1.4 Tangible results

The proceedings of the AAMAS workshops on FML give an overview of the discussion on FML.
- D. Heylen, C. Pelachaud, R. Catizone, D. Traum (2009) Proceedings of *Towards a Standard Markup Language for Embodied Dialogue Acts* (AAMAS)

- D. Heylen, S. Kopp, S. Marsella, C. Pelachaud en H. Vilhjalmsson (2008) Proceedings of the *First Function Markup Workshop* (AAMAS)

The following papers were written by SEMAINE partners (in collaboration with others) on FML and BML.
- E. Bevacqua, Ken Prepin, Etienne de Sevin Radoslaw Niewiadomski and Catherine Pelachaud (2009) Reactive behaviors in SAIBA architecture. In Proceedings of *Towards a Standard Markup Language for Embodied Dialogue Acts.*
- D. Heylen, S. Kopp, S. Marsella, C. Pelachaud, H. Vilhjálmsson: The Next Step Towards a Function Markup Language (IVA 2009)
- D. Heylen, M. ter Maat: A Linguistic View on Function Markup Languages. In: Proceedings of *Towards a Standard Markup Language for Embodied Dialogue Acts.*
- M. Mancini, C. Pelachaud (2008) The FML-AMPL Language. In Proceedings of the First Function Markup Workshop (AAMAS 2008)
- H. Vilhjalmsson, N. Cantelmo, J. Cassell, N. E. Chafai, M. Kipp, S. Kopp, M. Mancini, S. Marsella, A.N. Marshall, C. Pelachaud, Zs. Ruttkay, K. R. Thorisson, H. vanWelbergen, and R. van der Werf (2007). The Behavior Markup Language: Recent Developments and Challenges.

2.1.5 More information

Besides in the papers mentioned above, information about the standards can be found at the following websites.

- The 2010 FML workshop site is: *http://projects.ict.usc.edu/nld/FML2010/*

- The official SAIBA-site: *http://www.mindmakers.org/projects/SAIBA*

2.2 Emotion Markup Language

2.2.1 Description

Emotion Markup Language (EmotionML) 1.0 is a markup language designed to be usable in a broad variety of technological contexts while reflecting concepts from the affective sciences. As for any standard format, the first and main goal of an EmotionML is twofold: to allow a technological component to represent and process data, and to enable interoperability between different technological components processing the data.

Use cases for EmotionML can be grouped into three broad types:

1. Manual annotation of material involving emotionality, such as annotation of videos, of speech recordings, of faces, of texts, etc;
2. Automatic recognition of emotions from sensors, including physiological sensors, speech recordings, facial expressions, etc., as well as from multi-modal combinations of sensors;
3. Generation of emotion-related system responses, which may involve reasoning about the emotional implications of events, emotional prosody in synthetic speech, facial expressions and gestures of embodied agents or robots, the choice of music and colors of lighting in a room, etc.

Interactive systems are likely to involve both analysis and generation of emotion-related behavior; furthermore, systems are likely to benefit from data that was manually annotated, be it as training data or for rule-based modelling.

2.2.2 Current state of the standard

EmotionML is under active development at the World Wide Web Consortium (W3C), currently as part of the MultiModal Interaction Working Group (MMI-WG). It is edited by SEMAINE coordinator Marc Schröder; active group members include .SEMAINE partner Catherine Pelachaud[1].

The standardisation effort for the Emotion markup language goes back to 2006. It started as a succession of two Incubator groups at the W3C (the Emotion XG and the Emotion Markup Language XG, respectively). The first group worked on use cases and requirements for EmotionML; the second group started working on a concrete syntax specification.

After the decision to go for developing EmotionML into a web standard in the form of a W3C Recommendation, the work continued within the MMI-WG. A First Public Working Draft was published in October 2009; a revised working draft was published in July 2010.

The group is now in the process of preparing the so-called Last-Call Working Draft (LCWD), which is the feature-complete specification which then undergoes a thorough review for consistency and implementability at the W3C. Subsequent stages of standardisation according to the W3C process are Candidate Recommendation, Proposed Recommendation, and finally Recommendation.

---

1 Previous work in the EmotionML Incubator group also involved SEMAINE partner Björn Schuller who participated as an Invited Expert. Due to W3C regulations, he was not able to continue working in the MMI-WG because his institution is not a W3C member.

2.2.3   Activities carried out

The SEMAINE project has supported the standardisation of EmotionML throughout the project's lifetime. This covers most of the duration of the Emotion Markup Language Incubator Group (12/2007-11/2008) and the work in the MMI-WG up to the end of the SEMAINE project lifetime.

Group meetings:

- The work was paced by regular phone meetings (approximately every two weeks),
- face-to-face meetings at the W3C technical plenaries in 2008 (Cannes) and 2010 (Lyon)

A W3C Workshop on Emotion Markup Language was organised by Catherine Pelachaud of CNRS ParisTech, in October 2010, which was aimed at receiving feedback from the community on the latest EmotionML specification. The workshop had 18 attendees from Telecom ParisTech, DFKI, Queens University of Belfast, Roma Tre University, University of Greenwich, Dublin Institute of Technology, Loquendo, Deutsche Telekom, Cantoche, Dwango, nViso, and W3C Team. The feedback from participants confirmed that the specification generally covers the user's needs and is scientifically sophisticated. Individual suggestions for improvement of functionality were proposed and are being taken into account in the preparation of the Last-Call specification.

2.2.4   Tangible results

The publications of the EmotionML-related activities are publicly available via the W3C website.

- W3C Emotion Incubator Group Final Report (July 2007):
  http://www.w3.org/2005/Incubator/emotion/XGR-emotion/
- Final Report of the Emotion Markup Language Incubator Group (November 2008): Elements of an EmotionML 1.0, http://www.w3.org/2005/Incubator/emotion/XGR-emotionml/
- Emotion Markup Language (EmotionML) 1.0 First Public Working Draft (October 2009): http://www.w3.org/TR/2009/WD-emotionml-20091029/
- Emotion Markup Language (EmotionML) 1.0 Working Draft (July 2010): http://www.w3.org/TR/2010/WD-emotionml-20100729/
- W3C Workshop on Emotion Markup Language (October 2010): http://www.w3.org/2010/10/emotionml/summary.html

2.2.5   More information

More information can be found in the publications mentioned above and the Multimodal Interaction Activity page at W3C: http://www.w3.org/2002/mmi/

3 Setting up competitions

In addition to direct contributions to standards, standards for evaluation and acoustic features were set in the field of speech emotion and general paralinguistics analysis by the SEMAINE initiated and lead organised first of their kind INTERSPEECH 2009 Emotion Challenge (Schuller et al., 2009) and its follow up - the INTERSPEECH 2010 Paralinguistic Challenge (Schuller et al., 2010). While the first attracted 15 participating sites with 9 accepted contributions for the comparison of individual engines for the recognition of 2 and 5 discrete emotions, the second attracted 32 registrations; 11 paper were accepted. In its Affect Sub-Challenge subjects' interest had to be estimated in continuous space on TUM's Audiovisual Interest Corpus as also used for interest assessment in the SEMAINE 3.1 release. In addition to providing standard data sets and evaluation protocols to the community, these Challenges provided standard acoustic feature sets for better comparability of machine learning algorithms: SEMAINE's openSMILE feature extractor was used to provide a reference set that could be used by all participants. The 2009 set comprised 384 features, the 2010 one 1582 - these are since commonly found as standard reference sets for diverse speech analysis tasks.

Further information can be found here:

- B. Schuller, S. Steidl, A. Batliner, F. Burkhardt, L. Devillers, C. Müller, S. Narayanan: "The INTERSPEECH 2010 Paralinguistic Challenge", Proc. INTERSPEECH 2010, ISCA, Makuhari, Japan, pp. 2794-2797, 2010.
- B. Schuller, S. Steidl, A. Batliner: "The Interspeech 2009 Emotion Challenge", Proc. Interspeech (2009), ISCA, Brighton, UK, ISSN 1990-9772, pp. 312-315, 2009.