EMOXG internal progress report: state of the draft specification of EmotionML

Marc Schröder, 30 September 2008

The following is a snapshot of the discussion towards a first full draft specification of mandatory requirements for an Emotion Markup Language. Its purpose is to structure discussion during the final Incubator phase and to serve as a skeleton for the Final Report.

ISSUE NOTE: The suggested format does not yet take into account implementability in XML Schema; the format may need to be revised in order to make it possible to Schema-validate the markup.

Table of Contents

Emotion Core

Core 1. Type of emotion-related phenomenon

It was agreed that the type of emotion-related phenomenon (emotion proper, mood, interpersonal stance, attitude, etc.) will be specified in an externally defined set of possible values, similar to the "set" attribute in Core 2.

DISCUSSION NEEDED: There is not yet a suitable name for the attribute indicating the name of the set of allowed types. For the moment, the examples below use "emotion-related-state-types", which is clearly too long.

Two options for specifying the type have been identified, but there is no consensus yet which option should be used.

Option 1: type indicated by the name of the element:

```
<mood emotion-related-state-types="Scherer">
      <category set="everyday" name="relaxed" confidence="0.8" />
</mood>
```

Option 2: type indicated by an attribute "type":

```
<emotion type="mood" emotion-related-state-types="Scherer">
      <category set="everyday" name="relaxed" confidence="0.8" />
</emotion>
```

DISCUSSION NEEDED: There was some concern about using the tag name "emotion" in option 2, given the fact that it is one of the possible types of emotion-related states. Some group members would prefer to use a generic word such as "affect". Others prefer "emotion" because it is easier to understand for the non-expert.

Core 2. Emotion categories

There was agreement to represent emotion category as an element with two attributes:

- set, containing to a pre-defined name of a set of categories;
- name, the name of the category, which must be contained in the named set.

Syntax example:

```
<emotion>
      <category set="everyday" name="pleasure"/>
</emotion>
```

Because of its simplicity, the above representation was preferred over alternative suggestions to change the name of the XML element to represent different category sets.

DISCUSSION NEEDED: It is unclear at the moment how, and where, the set of allowed categories should be defined. Maybe there should be a repository of defined sets somewhere.

Core 3. Emotion dimensions

There was agreement to represent dimensions as a two-level element structure:

- a container element <dimensions>
    - with a mandatory attribute "set" containing a pre-defined name of a set of dimensions

- a number of distinct child elements, whose names must be contained in the named set
    - with a mandatory attribute “value” containing a Scale Value.

Example using discrete Scale Values:

```
<emotion>
      <dimensions set="Arousal-and-Valence">
            <arousal value="very much"/>
            <valence value="slightly positive"/>
      </dimensions>
</emotion>
```

Example using continuous Scale Values:

```
<emotion>
      <dimensions set="Arousal-and-Valence">
            <arousal value="0.9"/>
            <valence value="0.2"/>
      </dimensions>
</emotion>
```

DISCUSSION NEEDED: It is unspecified at the moment how to indicate whether a given dimension is unipolar or bipolar. Maybe this should be done wherever the set is defined.

Core 4. Appraisals related to the emotion

The syntax of appraisal markup should be the same as for Core 3, using the container element <appraisals>.

```
<emotion>
     <appraisals set="Scherer">
         <novelty value="(unipolar-scale)"/>
         <intrinsic-pleasantness value="(bipolar-scale)"/>
                  ...
         <goal-conduciveness value="(unipolar-scale)"/>
      </appraisals>
</emotion>
```

Core 5. Action tendencies

The syntax of action tendency markup should be the same as for Core 3, using the container element <action-tendencies>.

```
<emotion>
     <action-tendencies set="Frijda">
         <approach value="(unipolar scale)"/>
         <avoidance value="(unipolar scale)"/>
         <being-with value="(unipolar scale)"/>
         ...
     </action-tendencies>
</emotion>
```

ISSUE NOTE: It remains to be clarified what level of abstraction is appropriate for a default set of action tendencies. For example, a tendency to “eat food” may be appropriate for a human, but for a robot, the more generic “consume energy” may be more appropriate.

Core 6. Multiple and/or complex emotions

In the absence of the (currently optional) requirement Core 9 “Regulation”, a specification of

multiple/complex emotions can only be preliminary.

DISCUSSION NEEDED: There is currently no agreement whether it is necessary to make it explicit that emotions are complex. The latest state of the discussion is here: http://www.w3.org/2005/Incubator/emotion/minutes/2008-09-04.html#item02. In summary:

- Some participants are in favour of making it explicit that an affect is part of a complex affect, other participants do not see the need.
- Everybody agrees that individual affects should retain their individual time stamps.
- We need to make it clear what would be the added value of making it explicit that there is a complex affect, and who would benefit from that.

Core 7. Emotion intensity

Intensity is to be represented as a separate element, with a mandatory attribute "value".

Example:

```
<emotion>
      <intensity value="0.1" confidence="0.8"/>
      <category set="everyday" name="boredom" confidence="0.1"/>
</emotion>
```

This example shows the reason why intensity should be a separate element: it is possible to indicate confidence separately for the intensity. The example shows a high confidence that the intensity of the emotion is low, but only a very low confidence that the emotion may belong to the category "boredom".

Core 8. Emotion timing

See Links 2 for a discussion of time stamp issues. Here the time should be the time of the emotion itself; in Links 2, it is the time of external things such as a trigger or an observable behaviour.

Meta-information about emotion annotation

Meta 1. Confidence / probability

confidence is an attribute containing a Scale Value.

DISCUSSION NEEDED: We need to clarify which elements can be annotated with a confidence.

Meta 2. Modality

Links to the "rest of the world"

For “Links” information, there was agreement on the basic principle that we should be inspired from existing specifications as in EMMA and/or SMIL, but include only those concepts that we need, and include them into our namespace.

DISCUSSION NEEDED: Most of this is not yet fixed. More discussion is needed; latest contributions are at http://syntheticspeech.de/links2worldExamples2.html and http://lists.w3.org/Archives/Public/public-xg-emotion/2008Aug/att-0031/00-part

Links 1. Links to media

There was agreement that a link to media should provide:

- an attribute containing the URI of an object we point to;
- optionally, a mime type to indicate the type of data residing at that URI.

Many but not all group participants favoured the attribute name “uri” for the URI.

DISCUSSION NEEDED: it needs to be clarified if links to media should be allowed without any semantics associated to them (Links 3), and if so, what the meaning of that would be.

Links 2. Position on a time line in externally linked objects

It was agreed that we need to be able to specify relative timing, i.e. time counted from the beginning of a media file.

DISCUSSION NEEDED: we need to clarify if we also want to require absolute time. This was first discussed at http://www.w3.org/2005/Incubator/emotion/minutes/2008-07-31.html#Links2

DISCUSSION NEEDED: two different formats for relative time are available:

- EMMA format (offset in milliseconds from start of media file) and
- SMIL format (human readable, e.g. 00:10.5 = 10.5 seconds = 10 seconds and 500 milliseconds – see http://www.w3.org/TR/SMIL3/smil-timing.html#Timing-ClockValueSyntax).

DISCUSSION NEEDED: it remains to be clarified where to indicate timing; e.g., below <emotion> for Core 8, and as a sub-element of <context> for Links 1+3?

DISCUSSION NEEDED: suggestion by Felix to provide a <timing> element with the following optional attributes:

- start: defaults to 00:00
- end: no default
- duration: no default
- anchor: id of a media element, defaults to first media element found or parent element.

```
<timing anchor="audio1" start="00:03" end="00:09" duration="="00:06"/>
```

DISCUSSION NEEDED: suggestion by Ian to provide a substructure of a <timing element>:

```
<emoml:timing>
      <emoml:onset start="00:00:01:00" duration="00:00:04:00" />
      <emoml:hold  start="00:00:05:00" duration="00:00:02:00" />
      <emoml:decay start="00:00:07:00" duration="00:00:06:00" />
</emoml:timing>
```

DISCUSSION NEEDED: Should it be possible to indicate relative time in “samples” counted from the beginning of the media object, or is that too system-centric? See discussion at http://www.w3.org/2005/Incubator/emotion/minutes/2008-09-18.html#item3.

Links 3. The semantics of links to the "rest of the world"

DISCUSSION NEEDED: The suggestion by Felix and Ian was to use a <context> element with the following mandatory attributes:

- set, a set of possible semantic roles;
- role, a concrete semantic role, e.g. “experiencer”, “behaviour”, “trigger”, or “target”;
- uri, a pointer to the object/entity.

DISCUSSION NEEDED: is set required, or do we fix the list of semantic roles?

DISCUSSION NEEDED: is the uri attribute the same as Links 1, or is it different? If so, how?

Global metadata

Global 0. A generic mechanism to represent global metadata

DISCUSSION NEEDED: Two options were proposed for representing metadata at the document level, see http://www.w3.org/2005/Incubator/emotion/minutes/2008-09-18.html#item1.

Option 1: Individual key-value pairs. Example:

```
<emotionml>
      <meta set="myMetaTags" name="speakerOrigin" value="bavaria"/>
      <emotion>
      ...
      </emotion>
</emotionml>
```

Option 2: A parent element with arbitrary subelements, as in EMMA (http://www.w3.org/TR/emma/#s4.1.4):

```
<emotionml>
      <meta>
            ... custom metadata, e.g. using RDF...
      </meta>
      <emotion>
      ...
      </emotion>
</emotionml>
```

Cross-cutting issues

Scale values

Scale values are needed in Core 3, Core 4, Core 5, Core 7, and Meta 1. The spec should use a common format for scale values in all these places.

Conceptually, two types of scales are needed:

- unipolar (conceptually, from “nothing” to “a lot”)
- bipolar (conceptually, from “very negative” to “very positive”)

Practically, two representations are needed:

- numerical ([0,1] for unipolar scales, [-1,1] for bipolar scales; these intervals are to be understood as recommended ranges, but the language should allow the user to go beyond them in order to generate an “exaggeration” effect);
- discrete n-point scales. There was a tendency to prefer five-point scales (http://www.w3.org/2005/Incubator/emotion/minutes/2008-08-21.html#item3)

DISCUSSION NEEDED: There are two options for discrete five-point scales:

- verbal scales, such as “very negative – negative – neutral – positive – very positive”;
- abstract scales, such as “-- - 0 + ++”

It seems difficult to find generic wordings for verbal scales which fit to all requirements. Using custom wording, on the other hand, is difficult to reconcile with the “pluggable” nature of Core 3, Core 4 and Core 5.

Sampling mechanism

This is to represent episodes with changing Scale Values (e.g. valence, intensity...). Wherever a numeric Scale Value can be used, it should be possible to describe its change over time using a sample-based trace notation.

```
<emotion>
    <dimensions>
        <arousal>
            <trace freq="100Hz" samples="0.3 0.3 0.4 0.3 0.3 0.3 0.2 0.3 0.2"
                       confidence="0.3 0.3 0.4 0.3 0.3 0.3 0.2 0.3 0.2" />
        </arousal>
    </dimensions>
</emotion>
```

When a trace is used, it is possible (but optional) to indicate a trace of the confidence alongside with the trace of the scale itself. If present, the confidence must use the same sampling frequency as the content scale. If a global confidence for the entire annotation is to be annotated, it can be done on the enclosing element:

```
<emotion>
    <intensity confidence=”++”>
        <trace freq="100Hz" samples="0.3 0.3 0.4 0.3 0.3 0.3 0.2 0.3 0.2" />
    </intensity>
</emotion>
```