Coding Guidelines for Affect Annotation of the AMI Corpus

The AMI Emotion Annotation Subgroup

No Institute Given

1 Introduction

The AMI corpus is a collection of multi-modal meeting recordings. The majority of meetings were elicited using a scenario whereby groups of four participants played different roles in a corporate design team. Your job is to watch videos of these meetings and annotate them with information about the “mental state” of the participants. This means that you continuously try to answer the question “what goes on in this person’s mind; is he happy, surprised, interested, bored?”.

This manual and the connected training sessions will make clear how you are going to do that. For clarity’s sake, we want to stress the following points:

1. Firstly, the notion “mental state of a person” is often loosely interpreted as “the feeling of a person”. We want to stress that we consider feeling in a broad sense to include not only typical emotional categories such as “irritated” or “amused” but also so called cognitive states such as “interested”, “distracted” or “puzzled”.

2. Secondly, we don’t expect you to be able to mindread what is going on with a participant. We do not ask you to psychoanalyse a person in depth but only to give a broad description of what you can reasonably assume to be the mental state of the person, based on what you observe.

The information that results from this annotation task will help us to learn more about human behaviour in meetings as well as to develop algorithms for the automatic recognition of aspects of this behaviour.

2 Task overview

For each annotation assignment, you will watch the video recordings of a meeting. In particular, you will have close-up recordings for every single participant available along with overview videos that allow you to see the behaviour of the participant in context of the other participants.

The task consists of two parts: first, defining “cuts” (segmentation points) in the video of a person at places where you see a distinct change in the mental state of this person, and second, to fill in a form that describes each segment that is thus created. In brief, the procedure is as follows:
You start watching the video and try to imagine what the mental state is of the person you are observing.

As soon as you notice a distinct change in the mental state you press a key that will mark a segment boundary. The video stops playing\(^1\). Note that this boundary will in fact be an “end-point”. The start of this particular segment is the end-point of the previous segment (or in case it is the first segment, the start will obviously be the start of the video).

You fill in a form that describes the mental state of the participant in the segment. This will include specifying the intensity, and the quality or evaluation of the mental state (whether it is a positive or a negative one). This will be explained in more detail below. You also have to choose one or more relevant category words from a predefined set that fits the mental state of the participant in the segment best.

You press the “continue” key. The video resumes playing at the beginning of the next segment and the annotation process reiterates.

In the next two section the segmentation (“cutting”) and annotation (filling in the form) part will be described in more detail. Next, the tool that will be used for the task will be explained, followed by a step-wise description of the annotation procedure and training.

### 3 Defining a change in the mental state

The basis of the annotation process is the definition of changes in the mental state of the participant in a video. This will be the most difficult aspect of the annotation task and we will explain our idea of what exactly a change in the mental state is here.

There are two types of change that can be noted and consequently should lead to the creation of a segment boundary: a change in mental state **type** and

---

\(^1\) You will be able to fine-tune the position where you want to place the boundary. Try to find a good balance between precise annotation and the amount of time that you spend on it.
a clear change in the intensity of the mental state. Being amused, annoyed, angry, happy or relaxed are examples of mental state types. It is important to note that we explicitly define one special “mental state”: the neutral state. Evidently, a neutral mental state is a construct that does not really exist, but we define it here as an observed mental state that does not have a clear type with some intensity level. Such a segment will thus be denoted as a neutral segment with intensity 0.

A change in mental state type will be observed for example when somebody has a neutral look for a while, and you observe that this person starts to look amused. Or if somebody has been looking very annoyed for a while, and then relaxes again to a neutral state. Note that the mental states changes have may fade-ins and fade-outs.

A change in mental state intensity will be observed for example when somebody has been looking vaguely annoyed for a while, and you observe that the person suddenly looks extremely annoyed and frustrated. In this example, a segment boundary will be placed when the intensity level starts to increase. The segment left of that boundary receive a type description “annoyed” with a relatively low intensity, and the segment right of the boundary a type description “annoyed” with the highest intensity value. Another example of an intensity change is that someone first looks amused and then starts laughing. The segment boundary is placed just before the laughing starts.²

4 Describing the mental state

After having created a segment boundary, it is time to describe the mental state in the segment of focus. You will be guided in this process by a form that asks you to rate the intensity and the evaluation of the mental state and to choose a category label from a list of predefined category labels that best fits the observed mental state of the meeting participant in the segment of focus.

Intensity Intensity will be rated on a three-point scale: (1) low intensity, (2) medium intensity and (3) high intensity. A neutral mental state will automatically receive the intensity level 0.

² There are cases where it is difficult to decide whether a type or intensity change should be marked as a segment boundary or not. In the example of the person that becomes amused (a smile develops, a smile remains for a while and gradually disappears) the choice will usually be to interpret this event as a single segment or “episode” which covers all three phases, instead of marking this as three separate segments. However, when the smile becomes a roaring laughter, the smile and the laughter should definitely be two separate segments with different intensity values. In order to get familiar with these kinds of decisions, the instruction demo’s contain segments with displays of fluctuations in the intensity of certain mental states that we consider to be minor differences that do not deserve further segmentation.
Evaluation

People constantly evaluate the events, their own and others actions in many ways. This can be judgements about they are good or bad in a moral, ethical sense; whether they are good or bad for the goals they are pursuing; whether what is being said is true or false, believable or unbelievable, etcetera. You will mark the evaluation of a segment 'negative' in case the negative aspects of evaluation seem predominant for the participant, and positive otherwise.

Sometimes positive and negative are not applicable, in which case you can choose to mark the “not applicable” option. During the training you will see a few examples of all three categories. ‘Positive’ might for example be appropriate when someone is laughing; ‘negative’ for example when someone is very angry.

Category labels

Finally, a label that describes the mental state of the participant in the segment best has to be chosen from a predefined list. If you feel that the segment is more adequately described by more than one label, you may add extra labels. We expect that you will not need more than a few labels per segment. The labels indicate quite general, diffuse categories. If you can immediately think of a label or description that is exactly to the point given a particular observed mental state, there is room to add this word. Also, if you think that the segment does not fit any of the categories indexed by the labels, you can think of a new label and add this. However, we want you to use the standard label set, that was defined especially for the meeting domain, as much as possible. This set is given in full detail in Section 5.

If you think that a segment is neutral, you can mark it so by clicking the ‘set neutral’ button.

5 Appendix: The label set of affect-related words

Neutral neutral – Nothing remarkable is happening.

Curious curious, interested, attentive, focussed – The participant shows special interest in a topic or issue.

Note: In many situations ”paying attention” is the neutral state (people are listening to others, for instance). These cases should be labelled as neutral.

Amused amused, cheerful, joking – The participant is clearly amused.

Distracted distracted, inattentive – The participant is not paying attention to the central issue in the meeting. The participant can be distracted by specific other things, or the participants mind may be wandering.

Bored bored – The participant is clearly bored with the proceedings of the meeting.

Confused confused, puzzled – There is something that the participant does not understand or that the participant cannot work out.
Uncertain uncertain, hesitant – The participant is not certain about something. ("I don’t know")

Surprised surprised – Something occurs or is said that the participant had not expected.

Frustrated frustrated, annoyed – The participant appears frustrated or annoyed about something.

Decisive decisive, certain, confident – The participant is decisive, or very certain and certain about something and shows this by being more assertive and resolute than normal. This may be about an issue in the meeting (e.g. giving an opinion of which the participant is very certain), or about the meeting process itself (e.g. decisively taking the meeting to a new point).

Disbelief disbelief, sceptis, doubt – The participant does not believe something, is e.g. sceptical whether an idea is good, a statement is true, a solution will work.

Dominant dominant, challenging – The participant shows dominant behaviour wrt someone else, e.g. the participant is commanding, controlling or persuading others.

Defensive defensive, apologetic – The participant reacts defensively to e.g. protect own ideas, or authority.

Supportive supportive, affirmative, agreeing, approval – The participant shows support for another participant, either wrt his or her contributions to the meeting (ideas, argumentation), or towards his or her presence in the meeting.

6 Organization of the work

This section describes the organization of the work, as it will continue after you have done the training described in Section 8.

6.1 Get assignments

The first thing to do is to find out which files have been assigned to you by referring to the appropriate work allocation table at the AMI emotion annotation internet site [http://wiki.idiap.ch/ami/EmoAnnTrial2WorkAllocationNL]. An assignment consists of a Meeting ID and the name of that participant in the meeting that has the focus for annotation.
6.2 Check the discussion page for new information.

Every time a new annotation session is started, it is important that you check the “emotion annotation discussion webpage” first, to see if there might be new information regarding the annotation task. This is crucial, as there may be a risk that your annotation work may be in vain as it does not apply to the standard. That would be a waste of your valuable time. The discussion page can be found here: [http://wiki.idiap.ch/ami/EmoAnnTrial2DiscussionNL].

6.3 Get video and audio files

Before you actually start annotating, you may have to copy the video and audio files\(^3\) for a meeting to your computer. If you are working on one of the HMI annotation laptops, those files will already be there.

6.4 Load next assignment in the AffectCoder

*Start the tool* Go to the desktop. Click on the link ‘Start AMI Annotation’. A list of programs will be shown; select ‘AMI Affect Coder’ and click on the ‘Run’ button.

*Select the meeting ID* of the meeting that has been assigned to you (see Figure 2) and press ‘OK’.

*Enter your annotator-ID* or select it from the list if it is already there (see Figure 2).

\(^3\) The video does not have audio attached to it automatically as in the experience of most people is normally the case. The video and audio stream are played synchronously in the annotation software.
**Prepare your workspace** Open the appropriate video window and audio window (see Figure 3). You will need at least the close-up video of the participant that has the annotation focus, plus the audio signal. If you want, you can also open the overview video to see the participant in context.

Now activate the annotation window for the participant of focus (menu ANNOTATE → name of assigned participant; menu ANNOTATE → name of layer). The annotation window shown in Figure 4 will appear.

![Image of annotation window](image.png)

**Fig. 3.** Open the relevant video signals and the annotation panel.
6.5 Begin annotating

Refer to the first few sections of this manual for the annotation procedure. There will be a training session to show in more detail how the AMI Affect Coder tool supports this task; a list of available commands can be found in the next section. For any questions or comments that you have, write them down at [http://wiki.idiap.ch/ami/EmoAnnTrial2DiscussionNL]. Before you start a new assignment, always read the latest additions to that page.

And finally, don’t forget to save often!

6.6 Check your work into CVS, update work allocation table

When you have finished an assignment, you close the tool, send the annotations to the central server (see below) and update the work allocation table that is used by the annotation coordinator to keep track of task completions. After that you can start a new assignment.

Storing the annotations on the server The annotation files are stored in a central location that has a back-up facility and a version control system (CVS).
This is done using ‘smartcvs’, which can be found on the desktop (or in menu
Start - All Programs - SmartCVS).

1. Start SmartCVS. The program will automatically look in the directory where
the annotations are stored if there are new annotation files.

2. Now you see files that already have been annotated by you or other anno-
tators below ‘name’. Below ‘Local State’ marks the CVS status. ‘Non-CVS’
means that the file has not been send to CVS yet (CVS is not aware of its ex-
istence). Normally you will see only those files that just have been annotated
and have not been send to CVS yet.\footnote{In the ‘View’ menu you can set the option ‘Unchanged Files’ to see all files that were
send to CVS. With this option set you can inspect the annotations that you already
made on the machine that you are working on}

3. Select the file(s) you want to send to CVS and click the icon with the “+”
sign (‘CVS add’), on the right above ‘Name’. In the menu that appears next
(named ‘Add’), click ‘OK’. Some messages will appear in the lower window
but you may ignore them. Behind the selected file the ‘Local State’ changes
from ‘Non-CVS’ to ‘Added’.

4. Push the icon on the left (‘CVS commit’), next to the “+” icon (or use
control-o). Now a menu appears where you can leave a log message. If some-
thing special has occurred you can write this down here, but normally you
may leave it blank.

5. Repeat step 3 and 4 until all your files have been send to CVS.

6. Now you are ready. You can close all programs and log out.

7.\textbf{IMPORTANT:} do NOT delete any file using CVS!

\textbf{Updating the progress tables} To give the annotation coordinator a good
overview of the progress of the work, you change the work allocation table to
indicate that you have finished the assignment. The assignment table on the wiki
page contains a column ‘status’. For unfinished assignments this column contains
a ‘-‘. You should change this into ‘D’ for each assignment that is finished. To
edit the table, click ‘edit’, find the appropriate row of the table and change the
‘-‘ into a ‘D’.
7 The tool reference

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>Continues annotating immediately after the previous annotated segment (or from the start of the video).</td>
</tr>
<tr>
<td>Stop</td>
<td>Stops annotating.</td>
</tr>
</tbody>
</table>

Creating a new segment

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cut</td>
<td>Places a segment cut at the current time point. The video will pause, and you should describe the affective content of the segment <strong>from the previous cut point up until the just created cut point</strong>.</td>
</tr>
<tr>
<td>Undo cut</td>
<td>Removes the cut that you just placed. The video will continue playing from the time of the old (removed) cut-point.</td>
</tr>
<tr>
<td>Delete</td>
<td>Deletes the segment that you just created. Only works if you did not already finalize the segment. The annotation tool will 'pause'; if you press 'start' again you continue annotating directly following the previous (now last) segment.</td>
</tr>
<tr>
<td>Finish</td>
<td>Deactivated, not needed.</td>
</tr>
</tbody>
</table>

Fine tuning the end time of the new segment

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shift end left</td>
<td>Shifts the cut that you just placed left by 0.1 seconds.</td>
</tr>
<tr>
<td>Shift end right</td>
<td>Shifts the cut that you just placed right by 0.1 seconds.</td>
</tr>
</tbody>
</table>

Replaying the new segment

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loop</td>
<td>Keeps replaying the current segment over and over again in the active video and audio players.</td>
</tr>
<tr>
<td>Play once</td>
<td>Replays the current segment once in the active video and audio players.</td>
</tr>
<tr>
<td>Pause</td>
<td>Stops replaying the current segment in the video and audio players.</td>
</tr>
</tbody>
</table>

Describing the affective content of the new segment

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set to Neutral</td>
<td>Describes the current segment as being 'neutral'. If a segment is set to neutral, you cannot set its intensity or evaluation, nor add any text labels to it.</td>
</tr>
<tr>
<td>Intensity</td>
<td>Alters the intensity of the affective content of the current segment: (1) low intensity, (2) medium intensity and (3) high intensity</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Alters the valence of the affective content of the current segment (negative, positive, or 'not applicable)</td>
</tr>
<tr>
<td>Text labels</td>
<td>Selects relevant text labels for the current segment.</td>
</tr>
</tbody>
</table>

Adding comments to the segment

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add Comment</td>
<td>Opens a text box in which you can add a textual comment to the current segment.</td>
</tr>
</tbody>
</table>

Finalizing a segment

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finalize and Continue</td>
<td>Finalizes the current segment. The annotation process will continue with the following segment.</td>
</tr>
</tbody>
</table>

Saving the corpus

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Save</td>
<td>Saves the new annotations. Please do this often.</td>
</tr>
</tbody>
</table>
8 Training

It is assumed that you have already read the explanation on the annotation task in chapters 1-4. This chapter will lead you through a training round. Please save your work often during the training (ctrl-s).

8.1 Start the training tool

Start the tool Go to the desktop. Click on the link ‘Start AMI Annotation’. A list of programs will be shown; select ‘AMI Affect TRAINING’ and click on the ‘Run’ button.

Enter your annotator-ID or select it from the list if it is already there (see Figure 5).

![Fig. 5. Select your annotator ID, open the relevant video signals](image)

Open audio and video windows as needed You can open new video windows by selecting the signals from the drop down list (see figure 5). During the training exercise, you will sometimes get the instruction to switch to a certain video. In that case you can close the current video signal and open the new one by selecting it from the list. The training starts with audio lapelmix and video ME.

The tool is now ready to start the training See figure 6 for how it looks. The window in the upper left corner (the training controller) will be used to browse step by step through the training video fragments in the training session. Please don’t skip ahead and watch all the segments, but rather wait till they are mentioned in the training instructions. The ‘previous’ and ‘next’ buttons are used to go forward and backward through the fragments; a text label will show at which fragment you are.
(a) The training tool when it is ready to start the training session

![Training interface](image)

- **Training Fragment 1**
  - Watch video fragment
  - Task 1: place a cut
  - Compare with correct example
  - Intensity: ...
  - Evaluation: ...
  - Labels: ...

- **Task 2: Describe content**
  - Watch video fragment

(b) The first steps of the training procedure summarized

**Fig. 6. Training**
8.2 The actual training

The first two tasks are explained in great detail. This is to make you familiar with how the training tool works. The first task will be to place a segment boundary in the first training fragment. The second task will be to describe the content of the segment left of that boundary. Figure 6 shows an overview of how the training starts.

Task one: placing the first cut

- Make sure that you have the audio signal opened, as well as the video for participant ME.
- Go to the first fragment in the training controller (top-left corner of program window).
- Play this fragment once.

The fragment that you have just watched contains one segment boundary. Your first assignment is to place a cut at exactly this boundary.

- Play this fragment again, by pressing ‘replay once’ in the training controller.
- The fragment will be replayed. When you are sure where the boundary should go, press the ‘cut’ button in the annotation panel. The video will then stop playing.

By placing the cut you have created a segment. The segment ends at the point where you pressed ‘cut’, and begins at the beginning of fragment 1. If you want to see the segment back you can ‘play once’ on the annotation panel. Do this to see what segment you have just created.

The next fragment in the training controller, fragment 2, contains the correct version of the segment that you would get if you placed the cut in the correct place (see Figure 6).

- Go to fragment 2 in the training controller. You can watch it by pressing ‘play once’ in the training controller.
- Compare what you see with the segment as you have created it.

If you do not agree with our assessment of the segmentation, feel free to set a comment text on the segment that you just created (click ‘Set comment’). Such feedback is always welcome, also during the remainder of the training. Remember: every segment can have exactly one comment. If you want to add another comment later on, you should extend the comment text, not replace it as your previous comment will then be lost.

Task two: describing the content of the first segment When you placed the cut in the first training fragment, parts of the annotation panel were enabled so you can describe the content of the new segment (see Figure 6). You can describe the segment as being ‘neutral’ (see section 4), or describe its intensity and evaluation and assign one or more text labels to it.
– Describe the content of the correct segment from task one.
– When you are ready, press ‘Check’. The tool will now show you how we think that the first segment should have been described. Compare this with how you described the first segment.
– When you are ready, press ‘finalize and continue’. The new segment is stored and the tool is ready to continue the annotation directly after the new segment.

You do not yet need to describe the contents of the segment that follows after the cut that you placed in task 1. If we want you to do that, we will show that right-side segment as a training fragment later on.

Please save your work often during the training (ctrl-s).

Task three: continuing the training The rest of the training follows the same pattern.

– Watch the next fragment from the training controller (now: fragment 3).
– Place a cut at the appropriate moment.
– Watch the subsequent fragment (now: fragment 4) from the training controller as well as the segment that you just created in the annotation panel, to compare them. If you have a remark about the difference between the two, add a comment in the annotation panel.
– Describe the content of the ‘correct’ new segment using the annotation panel.
– Compare the result with the ‘correct’ annotation by pressing ‘Check correct annotation’. Again, if you have a remark about the difference between the two, extend the comment in the annotation panel.
– Press ‘finalize and continue’.

Often there is a little bit of overlap between the different training segments, for example because the next training sample continues directly after the correct cut-point of the previous example. Not all training fragments are about the same person in the meeting. The table below shows which fragments belong to which video. Make sure that you have the correct video open!

<table>
<thead>
<tr>
<th>Training controller fragments</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1..14</td>
<td>ME</td>
</tr>
<tr>
<td>15..34</td>
<td>PM</td>
</tr>
</tbody>
</table>