IDIAP COMMUNICATION

Meeting Data Collection Specifications

Iain McCowan [a]   Daniel Gatica-Perez [a]
Samy Bengio [a]

IDIAP–Com 03-10

December 2003

Dalle Molle Institute for Perceptual Artificial Intelligence • P.O.Box 592 • Martigny • Valais • Switzerland

phone +41 − 27 − 721 77 11
fax +41 − 27 − 721 77 12
e-mail secretariat@idiap.ch
internet http://www.idiap.ch

[a] IDIAP

1 Overview

This document specifies a process for collecting a new corpus of meetings in the IDIAP Smart Meeting Room. This document is a working draft that is expected to be updated and augmented throughout the data collection process.

This follows from an earlier data collection effort that resulted in a corpus of 60 scripted meetings (30 train, 30 test), each of 5 minutes duration (now available at mmm.idiap.ch).

The current data collection effort aims to address some of the limitations of the previous corpus, as well as to cater for a richer variety of research tasks.

Some specific motivations for collecting a new corpus of meetings include the need for :

- **More natural data** : Meeting scenarios in the previous corpus were highly constrained and unrealistic.
- **Statistically significant results** : In order to properly train and test statistical models, more data is required in terms of number of meetings and number of basic events happening during a meeting.
- **Richer research tasks** : We wish to research richer sets of events occurring in meetings, particularly making more use of the visual modality.

To give natural data, the technique proposed in this document is to generate a meeting based on *scenarios*, rather than scripts. Scenarios are given to meeting participants (or actors) before a meeting to inform them about the context and purpose of the meeting.

2 Meeting and Corpus Requirements

1. In total, there should be approximately 200 meetings.
2. There should be a variable number of people per meeting (3-5), although the majority should be 4 person meetings.
3. The meeting duration should be long enough for natural behaviour and allow for an appropriate number of events to occur. For this we suggest durations of 15-60 minutes.
4. Some people in the database should have consistent roles across meetings.
5. Recording should start before participants enter the room and end after they leave.
6. Meeting recordings should continue if real-life interruptions (mobile phones, people needing to leave, technical support) occur. Similarly, if meetings diverge or finish without achieving their goals, this is okay.
7. Meetings may be stand-alone or multi-session (e.g. projects).
8. Artifacts (documents, pens, laptops, water, coffee) should be included in meetings.
9. Attention should be paid to keeping a balanced corpus in terms of gender and cultural background of participants.
10. Some meetings should contain only native English speakers, some only non-native, and some a mixture. Possibly this may be useful also at the accent level (Australian, Indian, French, German).
11. Care should be taken to include meetings with non-technical content.
12. The look of the room should vary across the corpus, using pictures, notice-boards, plants, etc.

13. Some meetings should have agendas.

14. At the end of the meeting, participants should be encouraged (but not obliged) to allow copies to be made of any notes/documents they took into the meeting, or wrote during the meeting.

3 Meeting Scenarios

3.1 Real Meetings

1. **The Meeting Recording Project (20 meetings):** Those involved in recording this meeting corpus can meet to discuss progress and take decisions on issues that arise. The meeting would be informally chaired by the corpus administrator, who would be reporting on progress to researchers. Use of an agenda is encouraged for these meetings.

2. **Creation of a Reading Area ($\approx$ 5 meetings):** There is interest in creating a reading area at IDIAP. Staff and students could meet to discuss requirements, location, furnishings, etc. A series of meetings would include initial brainstorming with different students, drafting a first proposal, presenting the proposal to management, coming up with a final proposal based on feedback, and a follow-up meeting once the room has been created. A particular person should be responsible for the project, and should chair the meetings. This scenario should include artifacts in the form of plans, drawings, and involve people drawing on paper or whiteboard. The presentation to management could then have these diagrams as slides, as well as a equipment lists and an estimated budget. Use of an agenda is encouraged for these meetings.

3. **Course Design (3-7 meetings):** Seniors could meet to propose a course to be taught at the university. A first meeting could cover initial brainstorming and discussion of the idea. Following this, participants can be asked to come back with some specific input (documents, slides) for a subsequent discussion meeting at a given date. A third meeting would then take decisions to come up with the course outline. Another idea would be to include sub-group meetings : a main committee with one representative from each group (speech, vision, machine learning, interaction) would hold the above meetings, and further meetings would be organised for example for the speech representative to discuss with other speech staff. These meetings should all be chaired.

4. **Conference Report ($\approx$ 5 meetings):** When people return from conferences/workshops, they should hold a meeting to present what they learnt, how their presentation was received, and discuss emerging trends. As people at IDIAP already write an email report of the conference, this could be used as a presentation slide or document to guide the discussion.

5. **Book Club ($\approx$ 10 meetings):** People could form a book club where they decide on a common (non-technical) book, read it in their spare time and meet to discuss it. This could happen each week where they discuss an agreed number of chapters, or every few weeks to discuss entire books. Depending on the number of interested people, multiple groups could be formed. An effort should be made to recruit people outside work to participate in these meetings. This meeting would be informal and unchaired, and involve coffee and cookies.

6. **Drafting Job Description ($\approx$ 2 meetings):** Staff members from the different groups will get together to discuss the needs for a new employee (e.g., a new research engineer), brainstorm about the desired profile, and to agree on a concrete job description to be posted. The typical scenario could include either two short meetings (one to brainstorm and assign somebody to come up with a draft, and another one to refine and finish the task), or only one longer meeting.

7. **Technical Reading Group ($\approx$ 10 meetings):** Staff members (students or researchers) will get together to decide on organizing a reading group on a technical theme of common interest,

to review and discuss relevant papers on the field, every one/two weeks for X weeks. Meetings will be informal; one person (the organizer) will be the general moderator, and one person at each time will be the presenter. Everyone should have read the article beforehand, and bring a copy (with their notes) to the meeting. The presenter can use slides if they think appropriate. Note that, as the meetings are being recorded and distributed, participants should be advised in advance not to make disrespectful comments regarding the author and their work. Of course some technical criticisms of the work are to be expected, and are acceptable.

8. **Presentation Rehearsals:** Students or researchers will rehearse a presentation (e.g. TAM, conference dry-run) in front of three or four people. Free discussion during or after the presentation. To encourage participation, presenters will be given the chance to watch themselves off-line to improve their presentation skills.

9. **Torch/Skorch Design Meetings (≈ 10 meetings):** The design teams for the Torch/Skorch software libraries could meet periodically to discuss problems, design issues, releases, etc. Use of an agenda is encouraged for these meetings.

10. **ISSCO Meetings (≈ 10-20 meetings):** ISSCO have offered to visit IDIAP to record a number of meetings. For ISSCO to specify.

3.2 Artificial Scenario Meetings

1. **Plagiarism:** As defined in appendix of 'Group Discussion as Interactive Dialogue or as Serial Monologue'.

2. **Investors Creating Company:** Definition of this to be obtained.

For this category of meetings, other scenarios will be recorded if they are specified by research partners.

4 Annotation

Each meeting should be annotated for

1. speech recognition,

2. speaker turns, and

3. event languages (see list below).

Speech recognition and speaker turn transcriptions should ideally be sub-contracted to a specialist company, or possibly temporary dictation secretaries. This should be refined and checked for quality.

For the event languages (listed below), an on-line program should be used to allow the recording supervisor to annotate as much as possible in real-time (during meeting recording sessions). 'The Observer' has been suggested as an ideal software for this purpose, as well as for subsequent off-line annotations of the event languages.

Offline, a minimum of two people should annotate each meeting for each set of events using such a software tool. For events that vary on a scale (here level of interest and mood), a numerical rating should be given at specified intervals, such as once every 15 seconds. For other events (here group focus of attention, enriched turn-taking and multi-modal dialog acts), the annotation tool could be used just to mark the boundaries between events.

Once this is done, annotations should be analysed for agreement between annotators. For the scale-based events, this can be done by calculating the correlation coefficient between each pair of annotators. For the other events, this should be done using kappa statistic. A precise methodology to do this will be specified.

Once the agreement between annotators has been analysed, they need to be combined in some way. For scale-based, this could be the mean of the annotators (potentially after normalising each annotator by their mean and variance). For the others, a majority-vote rule can be used (necessitating a minimum of three annotations in cases of disagreement between annotators).

4.1 Group Event Languages

Group Focus-of-Attention

- standing presentation.
- seated presentation.
- (standing) whiteboard.
- seated participant doing monologue.
- notes (reading or writing)
- seated dictating (one speaking, others writing+silent) - rare in our meetings
- standing dictating (one speaking, others writing+silent) - rare in our meetings
- unfocused (none of the above)

Enriched turn-taking

- monologue
- dialogue (discussion between two people only)
- group discussion
- floor (discussion controlled by one person)
- group silence

Multimodal dialog acts A small set of events based on dialog modelling, but related to the group. This set of events still needs to be defined.
Mood (emotion)

- positive (e.g. happiness)
- neutral (relaxed)
- negative (e.g. anger, frustration)

Level of interest

- engaged (focussed)
- neutral
- disengaged (bored)

5 Schedule

We propose the following basic schedule for data collection, assuming a **T0** of December 2003.

Commencement of Data Collection (T0): Meetings will be recorded according to this specification. This will be overseen by an administrator, responsible for the definition of actual scenarios, and managing logistics of the data collection and annotation.

Review of Data Collection and Annotation Procedures (T0+2): Based on the first recordings, the data collection and annotation specification will be reviewed and potentially revised.

Completion of Data Collection (T0+6)

Completion of Annotation (T0+8)