Abstract

This paper reports on the main issues arisen during the development and test of a coding scheme for the argumentative annotation of meeting discussions. A corpus of meeting discussions has been collected in the framework of a research project on multimodal dialogue analysis and a coding scheme has been proposed. Annotations have been gathered by a set of annotators with different skills in argumentative discourse analysis and the reliability of the coding schema has been assessed against standard statistical measures.

1 Introduction

Interaction through meetings is among the richest human communication activities. Multimodal multi-party dialogs can be audio-visually recorded and stored in a multimedia repository. Recording meetings therefore implies the storage and the structuring of a large set of heterogeneous information scattered over time and media. The raw data format from the various recording devices is not directly usable for the creation of indexes, or for the content-based access to the relevant parts of the meetings. The data needs to be analyzed and annotated in order to provide thematic access to the meeting recordings and support for the process of querying for relevant information.

In this paper, we propose a coding scheme that allows us to highlight the argumentative structure of meeting discussions. The coding scheme evolved over time in order to cope with problems arisen during the annotation of data. The coding scheme we propose is aimed at reconstructing the linked sequence of argumentative actions that are common in meeting discussions and that highlight the main effort of participants to solve conflicts of opinion in order to end up with some agreed decisions.

The practical motivations behind this study are supported by a user study on queries for the meeting domain (Pallotta et al. 2007). The user study showed that potential users of a meeting query tool would ask questions about the argumentative processes that might occur during the discussions as well as how and why certain decisions have been taken.

1.1 Meeting Information Systems

Corporate meetings contain a wealth of tacit knowledge, which might play a central role in the construction of corporate or project memories. As remarked by (Corrall, 1998), Meetings provide opportunities for face-to-face contacts and electronic interaction, fostering learning groups and holding 'best practice’ sessions. They are part of a large repository of knowledge, which includes network and discussions and remain tacit knowledge unless manually processed and stored as meeting minutes.

The importance of tracking collaborative argumentation of discussion meetings has a central importance for building what are called ”Project Memories” in (Rosemberg and Silince, 1999). The construction of projects memories is similar to the annotation of meetings by their argumentation structure since it highlights not only "strictly factual, technical information", but also relevant information about the decision-making process.

As part of a large Swiss research project IM2, we studied models for representing knowledge that could be extracted from meetings records. Meetings can be of various types depending on the pur-
pose and the context where they take place. We are particularly interested in modeling discussion or debate meetings. This type of meeting has typically two clearly defined dimensions: decision making and conflict resolution.

1.2 Goals of the paper

The main goal of the paper is to look for a better understanding of the discourse structure of meeting discussion from the argumentative perspective.

From the study of collected queries (Pallotta et al. 2007), we discovered that users are not only interested in finding passages where a given topic has been addressed by the speakers but also how the discussion has been carried out in terms of the issues that have been raised and how conflict of opinions have been resolved. That is why we decided to “frame” meeting discussions in terms of argumentative categories. While it is not always the case, meeting discussions are very often directed toward the decision-making goal. The decision making process can be modeled, among others, from an argumentative perspective. Taking this perspective allows us to structure the content of meeting discussions and possibly define procedures for answering the frequent types of questions asked by the users.

It is crucial to notice that in this paper we focus on the specific domain of debate meetings. By debate meetings we refer to those meetings where an outcome is envisaged, whether it is a decision, a plan, a schedule, an agreement in a negotiation, etc. Other types of meetings (e.g. information, briefing, task reporting or task assignment) show little or no argumentation. Also, we do not consider here informal conversations where the topic of discussion does not require deep involvement of the participants in a task and where no conflicts of opinion and no conflict resolution processes emerge.

1.3 Outline of the paper

The paper is organized as follows. In section 2, we outline the principles of our multi-dimensional data model for structuring and annotating meeting data. Our data model contains several levels of abstractions including one based on argumentation on which we will focus. In section 3, we will show that argumentative structuring of meeting discussion is a difficult but feasible task by evaluating the inter-annotator agreement against data annotated using the Meeting Description Schema (Pallotta et al. 2004). Section 4 will conclude the paper with some discussion about the lesson learned and future work.

2 Meeting Data Models

The IM2 meeting data model proposed in (Marchand-Maillet, 2003), assumes that a meeting is part of a project, held in a given meeting room that will capture the data. Typically, the meeting room will help in recording meeting data (video and audio). In parallel, these will be stored as data associated to the meeting, as will the documents (e.g. distributed during the meeting).

The meeting is driven by an agenda that relates to topics, which then become meeting topics. The agenda, i.e. a composition of topics related (or not) to the project, will be stored likewise. A meeting may be structured in terms of its temporal activities. Their delimiters are said to be meeting milestones that will form key points for the structuring of the meeting. Examples of these milestones are decision points and topic change points. They all relate to a specific time point, and their combination forms meeting timeframes that will be referred to during information search and labeled using meeting activity titles. Finally, minutes, as a set of notes (taken or not during the meeting) and the text transcripts will also be stored in this repository. Text transcripts can be linguistically annotated following the shallow dialogue model proposed in (Armstrong, 2003).

More generally, the data model required to store information about a meeting should take into account the following issues:

- it must be instantiated by the information extracted from the analysis of the recorded meetings;
- it must be indexed for the subsequent retrieval;
- it contains temporal information;
- it is described by concepts.

Due to the dynamic nature of a meeting, we need to consider at least the following classes of entities:

- Individuals and objects
- Events
• Complex entities (aggregated structures).

Meeting data annotations can be conceptually organized at four different semantic levels:

• Physical layer: the rough material annotated with absolute time and possibly with low-level extracted features without any semantic interpretation;

• Factual layer: Information extracted from the rough material, e.g. from utterances, video-shots, prosodically separated speech units. This information can be retrieved on the basis of the annotation of each unit;

• Thematic layer: Information built by clustering together basic units of factual information into larger semantic units. For instance a set of dialogue turns may be structured into an episode which is a higher-level abstraction of dialogue;

• Rhetorical layer: abstract relations between events or more complex event structures (e.g. episodes) of the meeting.

2.1 Meeting annotation

We might consider the analysis of meeting discussions from different perspectives. On one hand, we are interested in studying the occurring linguistic phenomena and in looking at correlations between the use of particular language patterns and a specific dialogue situation. On the other hand, we might look at the semantic structure of meeting dialogues, abstracting from their linguistic realization. If our goal is to exploit these types of analysis in a Question-Answering System, we realize that the first type of analysis allows us to provide only a description of dialogues in terms of local intra-sentential phenomena like, for instance, the use of specific linguistic markers in the presence of given dialogue acts. The second type of analysis allows us to determine several correlated abstract structures which are independent of their linguistic realization and which are aimed at providing a qualitative description of the phenomenon, as a whole. We intend to perform the latter kind of analysis and store the result in form of annotations of the meeting transcripts.

2.2 Levels of annotation

Annotating a meeting entails the production of the content’s meta-descriptions. Given a collection of time-stamped audio-video sequences, the transcription of the dialogues is of primary importance. It can be done manually by transcribers or automatically by a speech recognition system with obvious differences in accuracy. The quality of the transcription, in terms of word-error rates, is a crucial issue on which all of the subsequent processing depends. It is apparent that a manual transcription provides us with reliable sources of information, while the performance of a speech recognition engine is very likely to be negatively influenced by the frequent noise situations that a meeting presupposes.

After the transcription of the audio recordings, a first level of annotation concerns the description of various meta-data about static meeting information such as date, location, participants, shared documents, etc.

The second level of annotation is based on the shallow dialogue model, proposed in (Armstrong, 2003). This model provides a simple logical structure of dialogues based on:

• a dialog is a non empty set of *episodes*; a new episode is identified by a topic shift.

• an episode is a non empty set of *turns*; a new turn is introduced by a speaker change.

• a turn is a non empty sequence of *utterances*. Each turn is annotated by one or more dialogue acts highlighting the communicative function of an utterance.

The third level of annotation is the discursive annotation which first structures the dialogues into utterances and assigns to each utterance a set of dialogue acts highlighting its communicative function. It is at this level that we consider it relevant to annotate the argumentative structure of meeting discussions.

2.3 Argumentative structure

In addition to the shallow dialogue model, we consider the adoption of a deeper structured representation based on argumentation theory. The main limitation of the shallow dialogue model is that a single utterance may have multiple communicative functions and that there is no trace of participants’
intentions. In multi-modal multi-party conversations, for instance, there are other types of communicative actions besides utterances, e.g. agreement by applause, disagreement by gesture or facial expressions. Also silence might express an agreement after a question like:

*Do you have something to object?*

Moreover, the shallow model does not take into account the formation of opinions by hearers about speakers, and, more important, it does not highlight the social behavior of the participants, nor their role in the deliberation process.

In order to overcome the above limitation, we propose to consider meeting dialogues from the Collaborative Decision Making (CDM) perspective. In CDM, a meeting is defined as a multi-party (multi-agent) decision making process: a collaborative process, where agents follow a series of communicative actions in order to establish a common ground on the dimension of the problem.

The main four dimensions of CDM process are:

- an overall task issue;
- a set of alternative proposals;
- a set of arguments in favor or against each proposals;
- a collection of choice criteria (perspectives and preferences) settled upon the participants;
- a decision (or evaluation) function that combines criteria to judge the alternatives.

This definition focuses on the processes, which take place during meetings and how these processes contribute to the accomplishment of a joint goal. In order to capture the above dimensions, we then adopted and extended a suitable argumentative model of discussions, namely the IBIS model proposed by (Kunz and Rittel, 1970). The IBIS model provides us with an abstract description of the discussion’s rationale by outlining the important points discussed, the conflicts arisen and, hopefully solved, and the decisions that have been made.

IBIS captures and highlights the main lines of a meeting discussion in terms of what issues have been discussed, what alternatives have been proposed to solve issues, and finally on what positions (accept or rejected) were taken by participants on the stated proposals.

![Figure 1. The IBIS argumentation process](image)

The construction of the IBIS structure is a dynamic process occurring during a meeting. In (Gordon and Karacapilidis, 1999), the process of proposing and arguing over alternatives is modeled by a state transition graph shown in Figure 1.

This simple state-chart describes a strict protocol, which constrains the interaction between participants, but hardly scales up from computer-mediated discussion to real life unconstrained meetings. The main reason lies on the fact that the IBIS model is intended as a tool for guiding and constraining the meeting discussions in order to make them more efficient and productive (Conklin 2006).

The IBIS model abstracts from the dynamics of the discussion, which needs to be modeled as well in order to extract the IBIS structures from meeting events. Relevant meeting events are special types of Dialogue Acts that have an argumentative force. This type of Dialogue Acts (Bunt 1979) called Argumentative Acts, are backward-looking acts with forward-looking expectations (Goffman 1981).

Modeling the dynamics of argumentation also means dealing with multimodal knowledge about the discussion events, since some argumentative acts can be executed with other modalities (e.g. agreement by silence or by applause, disagreement by laughing).

### 2.4 Related Work

Among other approaches to argumentative structuring we mention (Delannoy, 1999) who provided an Argumentation Mark-Up language consisting of a set of XML tags to pre-process monologues focusing on argumentative rhetorical relations, in order to build summaries. We believe that this model is not sufficient for meeting dialogues, since it only highlights argumentative rhetorical relations...
in monologues and does not allow us annotating the interactive construction of arguments.

Within the Adjacency Pairs model (Shegloff & Sacks 1973), the importance of tracking agreement and disagreement in discussions has been recognized also in (Galley et al., 2004; Hillard, Ostendorf, and Shriberg, 2003). Although these methods have the great advantage of being automatic, they only partially help in reconstructing the argumentative information we need in order to answer real user queries.

Within the Information State Update Dialogue Model, a new argumentative perspective has been introduced by Larsson’s notion of Issue-Under-Negotiation (IUN) (Larsson, 2002). This notion naturally extends the Ginzburg’s notion of Question-Under-Discussion (Ginzburg 1999X) to the new class of negotiation dialogues where argumentative threads are seen as pertaining to particular Issues, modeled as questions on the IUN stack. Introduce moves introduce new issues, Proposals introduce possible alternative answers thereto, Acceptances or Rejections remove those alternatives. This model is very similar in spirit to the IBIS model and is also grounded in a fully-fledged dialogue theory. This model has been adopted by (Niekrasz et al. 2005) for the real-time reconstruction of an argumentative structure by overhearing discussions in design meetings. Finally, (Rienks and Verbree 2006) propose the Twente Annotation Schema that is based on fewer categories but more relation types being inspired by the Rhetorical Structure Theory (Mann and Thompson 1988).

3 The construction of a meeting discussion corpus

The design of a QA system for accessing meeting records requires the construction of an annotated corpus and thus the definition of precise annotation guidelines, the transcription and annotation of a large number of recorded meetings, including their semantic/conceptual annotation and the annotation of the dialogue and argumentative structure. A dialogue-type-specific ontology can be used to represent and store the information related to the purpose and nature of the dialogue, regardless of the task domain in which the dialogue takes place. This type of ontology expresses conceptual information about the roles of participants and the type of dialogue acts that take place between them.

3.1 Meeting Description Schema (MDS)

When using the IBIS mark-up labels, a meeting is decomposed into several stages such as issues, proposals, and positions, each stage being possibly related to specific aggregations of elementary dialogue acts. Moreover, argumentative interactions may be viewed as specific parts of the discussion where several dialogue acts are combined to build such an interaction; as for instance, a disagreement could be seen as an aggregation of several acts of reject and accept of the same proposal. From this perspective, we elaborated an argumentative coding scheme, the Meeting Description Schema (Palotta et al. 2004), which takes into account the different stages (or episodes) defined by the IBIS model and extend the concept of adjacency pairs (Shegloff & Sacks, 1973) to relate these episodes to each other and to the corresponding argumentative function.

The final goal is to produce annotation guidelines that define a set of mark-up labels and the rules for their application. These guidelines will be the basic reference for human annotators to generate coherent argumentative annotations of meeting discussion (see Annex).

3.2 Basic elements of MDS

In MDS, the argumentative structure of a meeting is composed of a set of topic discussion episodes (a discussion about a specific topic). In each discussing topic, there exists a set of issue discussion episodes.

An issue is generally a local problem in a larger topic to be discussed and solved. Participants propose alternatives, solutions, opinions, ideas, etc. in order to achieve a satisfactory decision. Meanwhile, participants either express their positions and standpoints through acts of accepting or rejecting proposals, or by asking questions related to the current proposals. Hence, for each issue, there is a corresponding set of proposals episodes (solutions, alternatives, ideas, etc.) that are linked to a certain number of related positions episodes (for example a rejection to a proposed alternative in a discussing issue) or questions and answers.

The rules in Table 1 illustrate the definition of MDS. The symbol ‘\(=\)’ corresponds to the “replies-to” relation and ‘\(<;:\)’ corresponds to the “elaborates” relation.
The manual annotation of meeting data (e.g. transcriptions aligned with audio and video) should be supported by a suitable annotation tool with a user-friendly interface for segmenting the dialogue into episodes, assigning argumentative categories to each episode and organizing their structure by linking related episodes. The tool should also gradually check the consistency of the annotated argumentative structure with the rules fixed within the MDS so as to assist the user to generate coherent annotations.

<table>
<thead>
<tr>
<th>Issue</th>
<th>&lt;= Suggest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggest</td>
<td>&lt;= RequestExplanation</td>
</tr>
<tr>
<td>Suggest</td>
<td>&lt;= RequestJustification</td>
</tr>
<tr>
<td>Suggest</td>
<td>&lt;= Agree</td>
</tr>
<tr>
<td>Suggest</td>
<td>&lt;= Decide</td>
</tr>
<tr>
<td>RequestExplanation</td>
<td>&lt;= Explanation</td>
</tr>
<tr>
<td>RequestJustification</td>
<td>&lt;= Justification</td>
</tr>
<tr>
<td>Explain</td>
<td>&lt;= Agree</td>
</tr>
<tr>
<td>Justify</td>
<td>&lt;= Agree</td>
</tr>
<tr>
<td>Suggest</td>
<td>&lt;=: Explain</td>
</tr>
<tr>
<td>Agree</td>
<td>&lt;=: Explain</td>
</tr>
<tr>
<td>Disagree</td>
<td>&lt;=: Justify</td>
</tr>
<tr>
<td>Decide</td>
<td>&lt;=: Explain</td>
</tr>
<tr>
<td>&lt;=: replies-to</td>
<td></td>
</tr>
<tr>
<td>&lt;=: elaborates</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. MDS grammar

From an annotator’s point of view, more general categories can be refined into sub-categories if the annotator is able to identify finer-grained episodes. This means that the annotation can be done in several stages, following either a top-down (i.e. from general to specific categories) or bottom-up strategy (i.e. by identifying basic argumentative acts such as "Agree" or "Disagree" and then looking at embedding episodes for them).

MDS allows us to build hierarchical argumentative structures such as, for instance, the one shown in Figure 2 taken from an excerpt of a discussion about buying microphones. In this example, turns are annotated with above argumentative categories and linked together with the “replies-to” and “elaborate” relation. It is easy to check that this annotation allows a QA system to answer why-questions such as:

Why did B reject the proposal of A to buy head-mounted crown microphones?

Unfortunately, such annotation tool does not exist and we performed the annotation campaign with a simpler tool. For this purpose we used an ordinary spreadsheet containing the (previously segmented and time-stamped) discussion turns (with different color for each speaker) in the first column, the MDS categories and the argumentative relations in the remaining columns, as shown in Figure 3. The annotators can browse the audio-video recording of the meeting in another window. This type of annotation tool is not optimal (error prone) but it has the advantage of having a well-known interface requiring almost no training. Moreover, the data in tabular form are easy to analyze with statistical tools.

3.3 The Meeting Corpus

We annotated four meetings from the IM2 repository taken from different meeting collections. We wanted to show that different types of meetings
have different argumentative content. We then selected meetings on different topics and different settings.

The first meeting (ICSI-bmr02) is taken from the ICSI corpus (Janin et al. 2003) and it is an academic/research staff meeting. The second one (ISSCO37) is a simulated scenario, where the meeting is about furnishing a room in a faculty. The participants are members of the IM2 project. The third one (IS1008) is a simulated design meeting and the fourth one (MovieClub) is a simulated meeting for setting up the schedule of a movie club with members of IM2 project.

Only the first meeting is real while the other ones are scripted. However, depending on the type of content (i.e. business, leisure, technical) we noticed different levels of argumentative activity as shown in Table 2, which confirmed our hypothesis that in real discussions the level of involvement is higher than in simulated ones, and that technical content is less prone to debate.

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Argumentative Turns</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICSI-bmr02</td>
<td>44%</td>
</tr>
<tr>
<td>ISSCO-37</td>
<td>40%</td>
</tr>
<tr>
<td>IS-1008</td>
<td>35%</td>
</tr>
<tr>
<td>MovieClub</td>
<td>36%</td>
</tr>
</tbody>
</table>

Table 2. Argumentative Activity in Meetings

### 3.4 Inter-annotator agreement

In the framework of the evaluation of the proposed annotation schema (MDS), we dispose of a corpus of 4 meetings annotated by at least two annotators (maximum four). The Kappa score (Carletta 1996) is then calculated to measure the degree of agreement of the annotators. Two annotators agree on annotating one turn if there is at least one common category assigned to that turn. If there is an agreement on more categories, only one is counted. Similarly, if there is a disagreement on another category, the agreement is kept.

In this paper we report only the evaluation regarding the assignment of argumentative categories without considering the rhetorical links (i.e. “replies-to” and “elaborates”). We performed three types of evaluation.

The first evaluation considers 8 argumentative categories: Suggest/Propose, Agree (Accept), Disagree (Reject), Request Explanation, Request Justification, Explain, Justify, Decide. Turns can be assigned to one or more of these categories. This is necessary when one turn contain more than one argumentative action such as the following turn that contains both Agree and Explain actions:

“Yeah. It's only right right at the end where you find out really what's going on and it makes you rethink everything that happened beforehand.”

We evaluated the global Kappa score (i.e. for all annotations of all meetings, regardless of the annotators) and for individual meeting and pairs of annotators. We report in Table 3 the global Kappa score and the average Kappa for each of the four annotated meetings and for all the annotations (all the meetings). Following (Landis and Koch 1977), the “All meetings”, ICSI-bmr02 and MovieClub fall into the “moderate agreement” class (.41-.60), while the remaining ones can be interpreted respectively ISSCO-37 as fair agreement (.21-.40) and IS-1008 as slight agreement (0-.20).

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Kappa-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICSI-bmr02</td>
<td>0.45</td>
</tr>
<tr>
<td>MovieClub</td>
<td>0.41</td>
</tr>
<tr>
<td>ISSCO-37</td>
<td>0.38</td>
</tr>
<tr>
<td>IS-1008</td>
<td>0.19</td>
</tr>
<tr>
<td>All meetings</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Table 3. Inter-annotator agreement per meeting

We show in Table 4 the confusion matrix relative to all annotations. This table is very insightful since it shows us that Suggest and Justify are the most confused categories. From the confusion matrix it is easy to derive raw agreement by considering only the figures on the diagonal divided by the number of annotated turns (5139). The raw agreement is then 0.31.

<table>
<thead>
<tr>
<th></th>
<th>Sugg</th>
<th>Agr</th>
<th>Dis</th>
<th>ReqE</th>
<th>ReqJ</th>
<th>Exp</th>
<th>Just</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugg</td>
<td>695</td>
<td>14</td>
<td>52</td>
<td>26</td>
<td>19</td>
<td>40</td>
<td>191</td>
<td>72</td>
</tr>
<tr>
<td>Agr</td>
<td>34</td>
<td>67</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>Dis</td>
<td>37</td>
<td>4</td>
<td>130</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>35</td>
<td>15</td>
</tr>
<tr>
<td>ReqE</td>
<td>34</td>
<td>2</td>
<td>5</td>
<td>86</td>
<td>5</td>
<td>5</td>
<td>16</td>
<td>26</td>
</tr>
<tr>
<td>ReqJ</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>114</td>
<td>13</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Exp</td>
<td>59</td>
<td>24</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>117</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td>Just</td>
<td>127</td>
<td>6</td>
<td>46</td>
<td>15</td>
<td>11</td>
<td>20</td>
<td>320</td>
<td>59</td>
</tr>
<tr>
<td>Dec</td>
<td>45</td>
<td>2</td>
<td>12</td>
<td>26</td>
<td>8</td>
<td>3</td>
<td>41</td>
<td>69</td>
</tr>
</tbody>
</table>

Table 4. Confusion Matrix for all meetings

For the second evaluation we added a new category for those turns that contain no argumentative
actions. Considering this new label, which was implicitly assigned by annotators when skipping the turn, we obtained the all meeting Kappa score of 0.42 and a raw agreement score of 0.61.

If we collapse all the argumentative categories into a single one and we consider also the non-argumentative category, we get a Kappa score of 0.54 and a raw agreement of 0.77. The interpretation of this last result is that discriminating between argumentative and non-argumentative is a quite feasible task.

The last evaluation was done by clustering the categories that we considered intuitively more ambiguous: explanation and justification. We collapsed into two single categories those pertaining to explanations and justifications, namely Explain/Justify and Request-for-Explanation/Justification. For the evaluation of all meeting including the “non-argumentative” category we obtained a Kappa score of 0.42 and a raw agreement of 0.62.

4 Conclusions

In this paper we have described a coding scheme for the argumentative annotation of meeting discussions, the Meeting Description Schema. We evaluated the reliability of this scheme by annotating a corpus of meeting transcriptions and assessing the inter-annotator agreement using the Kappa statistics. From the obtained results, we can conclude that the coding scheme is moderately reliable and at the current stage needs to be refined in order to obtain better agreement. However, we noticed that some disagreement might be due to our under-constrained guidelines that do not provide clear criteria for discriminating between possible categories.

As a future work, we will evaluate inter-annotator agreement on the structural dimension of annotation (rhetorical links). We also expect to improve the guidelines by providing a decision tree to annotators based, among others, on linguistic features such as those proposed by (Polanyi 2006).

From a practical point of view, a better annotation tool will be implemented on top of the NOMOS annotation tool (Niekrasz and Gruenstein 2006). Finally, an extended coding scheme based on the Meeting Discussion Ontology (Pallotta 2006) will be used for the next annotation campaign.

References


ii An adaptation of the NOMOS tool (Niekrasz and Gruenstein 2006) for argumentative annotation is under development, but still not available for large scale annotation campaigns.
Annex: Annotation Guidelines

1. Task

The annotator is requested to assign argumentative categories to turns (=utterances) in the meeting dialogue. An argumentative action made of more than one turn is called Argumentative Segment. Notice that a one-turn argumentative action is a special case of an argumentative segment. Also, notice that one segment can sometimes belong to more than one category. In this case, assign all corresponding categories to this segment.

The annotator is also requested to assign (backward) links between two argumentative segments. There are two types of links:

1. "Replies to": links two segments whose turns belong to different speakers (e.g. “Agree” Replies to “Suggest”).
2. "Elaborates": links two segments that are typically (but not necessarily) spoken by the same person (e.g. “Justify” Elaborates “Disagree”).

In Excel, you will assign a numeric (progressive) identifier to each turn which is part of an argumentative segment, not necessarily adjacent turns, e.g.:

A: blab la (Suggest, 1)
A: blab la (Suggest, 1)
B : hmm
A : bla (Suggest, 1)
B: No no…. (Disagree, 2) (Replies to, 1)

For those argumentative segments that have a link to some previous segment either through the “replies to” or “elaborates” relation, the identifier of the referred segment is specified in the appropriate cell.

More concrete examples are given in meeting ISSCO-35.xls (FurnitureMeeting):

Example 1: Suggestion – explanation – request for explanation – explanation

Turn 43: Suggestion/proposal/idea (ID:2)
Turns 44-47: Explanation (ID:3) elaborates on suggestion (ID:2)
Turn 48: Request for an explanation (ID:4) replies to the explanation (ID:3)
Turn 49: Explanation (ID:5) replies to the request (ID:4)
See also a similar example for turns 65-86.

Example 2: Proposal – disagreement – justification - proposal

Turn 490-491: Proposal (ID:29)
Turn 504-508: Disagreement (ID:30) replies to proposal (ID:29)
Turn 510-513: Justification (ID:31) elaborates disagreement (ID:30)
Turn 514: Proposal (ID: 32) replies to disagreement (ID:31)

Example 3: Proposal – decision

Turn 689: Proposal (ID:33)
Turn 712: Decision (ID:34) replies to proposal (ID:33)
IMPORTANT: Only utterances that can be considered to belong to the below listed categories should be annotated. Not every utterance in the meeting is argumentative, e.g. private communication, floor grabbing, floor assignment, etc.! If you are not sure about a segment, it’s better NOT to annotate it.

2. Argumentative Categories
   - Suggestion/Proposal/Idea
   - Opinion
   - Agree
   - Disagree
   - Justify
   - Explain
   - Request justification
   - Request explanation
   - Decide

* Suggestion/Proposal/Idea

In this category we collapse all contributions that have the role of providing something new in the discussion about a given topic, or alternatively, to trigger a discussion about a new topic (raising a new issue).

* Opinion

This category represents contributions that have the role of declaring a speaker’s attitude to a given issue. Opinions may (explicitly) refer to an earlier suggestion. For instance: “Concerning buying a new computer, I think Apple is very good.” This relation is expressed with the “Replies to” link.

*Agree (=Accept)

An agreement is a contribution where the participant expresses his/her support to another participant's suggestion/proposal/idea or opinion, and should therefore be linked to this segment through the “replies to” relation. Notice that agreements can be followed (or contain in the same turn) an explanation of why the speaker agrees. In this case the “Explanation” MUST be annotated and linked to the “Agree” through the “elaborates” link. E.g.

A: “Yes, I think it’s a great idea because then we can use the space for several purposes.”

This turn might get the annotations: (Agree, 21) (Explain, 22), (22 elaborates 21)
*Disagree (=Reject)*

This is a contribution that expresses disagreement with another participant's opinion or suggestion/proposal/idea. Disagreement might be followed by (or contain in the same turn) a justification of why the speaker disagrees. The same rule applies as for “Agree” and “Explain”, i.e. that the “Justification” should be linked to the “Disagree” through the “elaborates” relation.

*Justify*

A contribution that justifies either a standpoint (expressed by “Opinion” or “Suggestion/Proposal/Idea”) or a disagreement (expressed by a “Disagree” segment). It should be linked to its antecedent through the “elaborates” relation. If some participant explicitly requested for a justification before the actual justification, then the “Justify” segment should also be linked to this “Request justification” segment through the “replies to” relation AND to the segment for which it counts as a justification through the “elaborates” relation.

*Explain*

A contribution that explains the reasons of a “Suggestion/Proposal/Idea” or elaborates the reasons of an “Agreement”. It should be linked to its antecedent through the “elaborates” relation. Also, if there is an explicit “Request explanation”, the “Explain” segment should be linked to this segment with the “replies to” relation AND to the segment for which it counts as an explanation through the “elaborates” relation.

*Request justification*

A contribution that requests a justification for a standpoint (Suggestion/Proposal/Idea or Opinion) or for a “Disagreement”. It should be linked through the “Replies to” relation.

*Request explanation (or clarification)*

A contribution that requests an explanation or clarification for a Standpoint (Suggestion Proposal/Idea or Opinion) or for an Agreement (rare). Also, it may be linked to another explanation in case the referred explanation is perceived as insufficient to the participant. It should be linked through the “replies to” relation.

*Decide (e.g. actions items or task assignment)*

This type of contribution explicitly mentions which Suggestion(s)/Proposal(s) has(-ve) been accepted. Also it includes decisions about Action Items (e.g. what will happen in the future concerning issue X) and Task Assignments (e.g. who will do what for when). “Decision” can be linked (but not necessarily) to Suggest/Propose/Idea through the “elaborates” relation.