Generating Abstractive Summaries of Conversations through Automatic Argumentative Analysis

Vincenzo Pallotta[1] and **Rodolfo Delmonte**[2]

Abstract. Abstractive summarization of conversations is a very challenging task that requires full understanding of the utterance/s contained in a dialog turn, their roles and relationships in the conversation. We present an efficient system, derived from a full-fledged text analysis system, which performs the necessary linguistic analysis of turns in conversations and provides useful argumentative labels to build synthetic abstractive summaries.

1 Introduction

The generation of text summaries from transcribed spoken dialogs occurring between two or more people is considered an open problem in Natural Language Processing (or Computational Linguistics). Voice communications can happen either during face-to-face encounters or when parties are remotely connected and using communication services such as VoIP, teleconferencing or just telephones. In this work, we start from the assumption that voice communication can be transcribed accurately and that it is segmented into turns assigned to each speaking participant. While this assumption might seem restrictive, most corpora of voice communications are of this kind. Moreover, the cost of transcribing speech is far lower than the cost of summarization.

Automatic summarization of conversations would be a very useful tool that allows enterprises in extracting knowledge from many sources (e.g. web forums, social media, meeting records) and integrating this knowledge into corporate knowledge bases for future content-based access. For instance, one could easily find answers to questions like "why this decision was made" or "who rejected the proposal made by X?" or "How the decision of doing X impacted the progress of project Y?". It is apparent that answering this kind of questions requires a deep understanding of the conversational situation such as its dynamics, the rules of order adopted, the specific language used, and even culture-specific rituals.

1.1 Related Work

The problem of automatic summarization of conversations was initially investigated in the 80' in the context of several DARPA projects [20] focusing on meeting data. Automatic summarization of meetings has been typically approached in a so-called "extractive" fashion that is by extracting excerpts of the dialogs and by assembling them into a hopefully coherent text [12]. This method has severe limitations due to the intrinsic characteristics of the source data: dialogs are not as coherent as an ordinary narrative text (such as news or scientific articles) and obtaining a coherent text from dialog turns is practically impossible using the extractive approach. Moreover, the proposed solutions for extractive summarization of meetings have already reached their qualitative upper bounds [17].

Abstractive summarization of dialogs has been recently only partially addressed by some research project among which the European FP6 AMI, and the DARPA's CALO project. On this side, some advances have been achieved such as extracting (very basic) "dialog acts" [1], detecting "agreement and disagreement" [7, 5], segmenting the meeting into "thematic zones" [8], and, only recently, detecting "decision points" [9, 10], and "action items" [13]. At the University of Twente, also in the framework of the AMI project, a group was investigating on the automatic argumentative structuring of meeting dialogs [18].

Unfortunately, although very related, none of these projects directly address the problem of abstractive summarization of conversations and all the advances on the detection of useful content information from meeting data are mainly exploited to improve extractive summarization, which, as said before, can be only improved up to a given (unsatisfactory) upper bound.

The problem we address here, instead, is that of abstractive summarization of conversations. This problem is apparently much harder to solve than the extractive one since it requires almost the full understanding of the source data [6].

1.2 Argumentative Analysis of Dialogs

We start from transcribed turns in conversations. Our solution to the problem of abstractive summarization of conversations would be learning a general semantic/pragmatic structure for dialogs data that could be specialized for different conversations types (e.g. meetings, debates, focus groups, forums and interviews).

We were looking at pervasive pragmatic phenomena in dialogs and we realized that most of encountered dialogs include argumentative processes. Basically, exchange of opinions, questions-answers, negotiations and collaborative decision-making are at the heart of voice communications. Unfortunately, these phenomena are not sufficiently lexically marked in the dialogs to justify extractive summarization. For instance, a disagreement to a proposal about an issue cannot be summarized as just the turn "no" occurring in the dialog. This action is the result of a complex, sometimes long exchange between two or more participants.

By simply extracting some turns from the dialog is often insufficient to fully render the situation. In contrast, it is necessary to provide a paraphrase where some argumentative terminology is used such as, for instance, "Mr. X disagreed to the proposal made by Mr. Y on the issue raised by Mr. Z". Our system is capable to recognize a small, but highly relevant set of general argumentative action that

[1] Webster University Geneva, Switzerland, pallotta@webster.ch
[2] University of "Ca Foscari" - Venice, Italy, delmont@unive.it

can be used to produce such paraphrases thus creating a meaningful abstractive summary of the analyzed conversations. While the generation of the summaries is quite trivial given the extracted information, we will focus in this paper on the recognition of argumentative events as the core of the technology required to implement an abstractive summarizer of voice communications.

The paper is organized with a presentation of the system in section 2; then some data and an evaluation in section 3. A final section will be devoted to conclusions.

2 Automatic Argumentative Annotation

Computing semantic representations for argumentative annotation of conversations is a particularly hard task which – when compared to written text processing - requires the following additional information to be made available:

- adequate treatment of fragments;
- adequate treatment of short turns, in particular one/two-words turns;
- adequate treatment of first person singular and plural pronominal expressions;
- adequate treatment of disfluency, thus including cases of turns made up of just such expressions, or cases when they are found inside the utterance;
- adequate treatment of overlaps;
- adequate treatment of speaker identity for pronominal coreference;

In addition, every dialog turn must receive one polarity label, indicating negativity or positivity, and this is computed by looking into a dictionary of polarity items. This can subsequently used to decide on argumentative automatic classification.

2.1 Argumentative structure – issues and theories

As shown by [19], tracking argumentative information from meeting discussions is of central importance for building summaries of project memories since, in addition to the "strictly factual, technical information", these memories must also store relevant information about decision-making processes [14].

We consider the adoption of a deeper structured representation based on argumentation theory. The argumentative structure defines the different patterns of argumentation used by participants in the dialog, as well as their organization and synchronization in the discussion. A conversation is decomposed into several stages such as *issues*, *proposals*, and *positions*, each stage being possibly related to specific aggregations of elementary dialog acts. Moreover, argumentative interactions may be viewed as specific parts of the discussion where several dialog 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 adopted an argumentative coding scheme, the Meeting Description Schema (MDS) from [15]. 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.

3 Computing Argumentative Annotations

The core of our solution is based on adapting and extending GETARUNS [2, 3], a system for text understanding developed at the University of Venice. GETARUNS is organized as a pipeline which includes two versions of the system: the Partial and the Deep GETARUNS. The Deep version, used in this work, is equipped with three main modules:

- a lower module for parsing, where sentence strategies are implemented;
- a middle module for semantic interpretation and Discourse Model (DM)l construction which is cast into Situation Semantics;
- a higher module where reasoning and generation takes place.

GETARUNS, has a highly sophisticated linguistically based semantic module which is used to build up the DM. Semantic processing is strongly modularized and distributed amongst a number of different submodules which take care of Spatio-Temporal Reasoning, Discourse Level Anaphora Resolution, and other subsidiary processes like Topic Hierarchy which cooperate to find the most probable antecedent of coreferring and cospecifying referential expressions when creating semantic individuals. These are then asserted in the DM, which is then the sole knowledge representation used to solve nominal coreference. The system has been adapted to cope with some peculiar phenomena in dialogs such as overlap and short turns or fragments. Overlaps are an important component of all spoken dialogue analysis. In all dialogue transcription, overlaps are treated as a separate turn from the one in which they occur, which usually follows it. This is clearly wrong from a computational point of view. For this reason, when computing overlaps we set as our first goal that of recovering the temporal order. This is done because:

- overlaps may introduce linguistic elements which influence the local context;
- eventually, they may determine the interpretation of the current utterance;

For these reasons, they cannot be moved to a separate turn because they must be semantically interpreted where they temporally belong. In addition, overlaps are very frequent. The algorithm we built looks at time stamps, and everytime the following turn begins at a time preceding the ending time of current turn it enters a special recursive procedure. It looks for internal interruption in the current turn and splits the utterance where the interruption occurs. Then it parses it split initial portino of current utterance and continues with the overlapping turn. This may be reiterated in case another overlap follows which again begins before the end of current utterance. Eventually, it returns to the analysis of the current turn with the remaining portion of current utterance.

Fragments and short turns are filtered by a lexical lookup procedure that searches for specific linguistic elements which are part of a list of backchannels, acknowledgements expressions and other similar speech acts. If this procedure succeeds, no further computation takes place. However, this only applies to utterances shorter than 5 words, and should be made up only of such special words. No other linguistic element should be present apart from nonwords, that is words which are only partially produced and have been transcribed with a dash at the end. The two remaining cases are the following:

- graceful failure procedures for ungrammatical sentences, which might be fullfledged utterances but semantically uninterpretable due to the presence of repetitions, false starts and similar disfluency phenomena;
- failure procedures for utterances that are constituted just by disfluency items and no linguistically interpretable words.

We also implemented a principled treatment of elliptical utterances and contribute one specific speech act or communicative act. They may express agreement/disagreement, acknowledgements, assessments, continuers etc. All these items are computed as being complements of abstract verb SAY which is introduced in the analysis, and has as subject, the name of current speaker.

Finally, Automatic Argumentative Annotation is carried out by a special module of the GETARUNS system activated at the very end of the computation of the each dialog. This module takes as input the complete semantic representation produced by the system. The elements of semantic representation we use are the following ones:

- all facts in Situation Semantics contained in the Discourse Model, which include individuals, sets, classes, cardinality, properties related to entities by means of their semantic indices;
- facts related to spatio-temporal locations of events with logical operators and semantic indices;
- vectors of informational structure containing semantic information at propositional level, computed for each clause;
- vectors of discourse structure with discourse relations computed for each clause from informational structure and previous discourse state;
- dialog acts labels associated to each utterance or turn following ICSI classification;
- overlaps information computed at utterance level;
- topic labels associated to semantic indices of each entity marked as topic of discourse;
- all utterances with their indices as they have been automatically split by the system.

To produce Argumentative annotation, the system uses the following 21 *Discourse Relations labels* computed by GETARUNS in the higher module:

statement, narration, adverse, result, cause, motivation, explanation, question, hypothesis, elaboration, permission, inception, circumstance, obligation, evaluation, agreement, contrast, evidence, hypoth, setting, prohibition.

These are then mapped onto five general argumentative labels:

ACCEPT, REJECT/DISAGREE, PROPOSE/SUGGEST, EXPLAIN/JUSTIFY, REQUEST.

In addition, we use the label DISFLUENCY for all those turns that contain fragments which are non-sentences and are not semantically interpretable. Full details about the argumentative annotation algorithm are provided in [4]. The argumentative annotation algorithm is outlined as follows:

1. It recovers Dialog Acts for each dialog turn as they have been assigned by the system;
2. It recovers Overlaps as they have been marked during the analysis;
3. It produces an Opinion label which we call Polarity, which can take one of two values: Positive or Negative;
4. It produces a list of Hot Spots (i.e. a set of turns in sequence where the interlocutors overlap each other frequently) and builds up Episodes (i.e. a set of turns in which a single speaker "arguments" his/her topics without interruption). Episodes may occasionally be interrupted by overlaps or by short continuers, backchannel or other similar phenomena by other speakers without however grabbing the floor;
5. It assigns a set of argumentative labels, one for each clause. The system then chooses the label to associate to the turn utterance from a hierarchy of argumentative labels graded for Pragmatic Relevance which establishes that, for instance, Question is more relevant than Negation, which is more relevant than Raising an Issue, etc.

We are also able to evaluate the degree of collaboration vs. competitiveness of each participant in the conversation and make a general statements like the ones we include here below taken from the summaries produced for a dialog of our test corpus.

3.1 Abstractive Summary Generation

We present here the structure and the core content of the abstractive summary (Memo) that is generated automatically by the system. The Memo is made of five sections of paraphrased summary, touching upon different conversational features that are relevant to the general understanding of the conversation.

The first section reports general information regarding the *participants and their level of involvement* in terms of number of turns and actual presence in the conversation. In this meeting, for instance, two of the participants joined late and contributed less.

> *GENERAL INFORMATION ON PARTICIPANTS*
> *The participants to the meeting are 6. Participants less actively involved are Adam and Andreas who only intervened respectively for 9 and 72 turns.*

The second section reports about *levels of interactivity* and is based on the use of time stamps and the computation of overlaps. We compute number of turns effectively held by each speaker, then the number of overlaps done by each interlocutor. These numbers allow us to define levels of competitiveness in the conversation.

> *LEVEL OF INTERACTIVITY IN THE DISCUSSION*
> *The speaker that has held the majority of turns is Don with a total of 512 turns, followed by Morgan with a total of 456 turns, followed by Jane with a total of 263.*
> *The speaker that has undergone the majority of overlaps is Morgan followed by Jane. The speaker that has done the majority of overlaps is Morgan followed by Jane. Morgan is the participant that has been most competitive. This participant: Andreas, only intervened after turn no. 1091*

The third section reports lexical chains that highlight the *main topics* of the conversation.

> *DISCUSSION TOPICS AND LEXICAL CHAINS*
> *The discussion was centered around the following topics: format, file and tool.*
> *The main topics have been introduced by the second most important speaker of the meeting.*
> *The participant who introduced the main topics in the meeting is: Morgan.*

The most frequent entities in the whole dialogue partly coincide with the best topics, and are the following: format, utterance, representation, annotations, information, p_file, equals, ATLAS, prosodic, database, data, diff, a_p_i_, tool, sequence, files, external, boundaries, whole, versions, timeline, text, speaker, rising, file_format, 'Adam', transcripts, prosody, phrase, perl, lattice, idea, channels.

The fourth section highlights *speakers' attitudes t*owards topics and they are computed from pragmatic information collected from idiomatic expression and sentiment analysis. It uses five values: Positive, Negative, None, Acceptance, Suspension, where Suspension indicates an expression of doubt or some modality and None refers to simple declaratives without opinion specified.

SPEAKERS ATTITUDES TOWARDS THE TOPIC "format"
We report here below the lexical chains associated to the most frequent entities and topics discussed in the dialogue. The lexical chain headed by the word "format", is present in the following turns, headed by a given predicate and introduced by a given speaker, with a given attitude.

TNo. 45, Pred. develop, Spk. Don, Att. Positive
TNo. 205, Pred. have, Spk. Don, Att. Positive
TNo. 230, Pred. like, Spk. Jane, Att. Positive
TNo. 279, Pred. see, Spk. Morgan, Att. Negative
TNo. 300, Pred. describe, Spk. Don, Att. None
TNo. 345, Pred. have, Spk. Jane, Att. Positive
TNo. 490, Pred. have, Spk. Don, Att. Negative
TNo. 523, Pred. be, Spk. Don, Att. Positive
TNo. 569, Pred. find, Spk. Morgan, Att. Suspension
TNo. 583, Pred. be, Spk. Don, Att. Suspension
TNo. 605, Pred. do, Spk. Jane, Att. Positive
TNo. 773, Pred. add, Spk. Morgan, Att. Suspension
TNo. 775, Pred. be, Spk. Don, Att. Suspension
. . . .
TNo. 1207, Pred. be, Spk. Jane, Att. Acceptance

The last fifth section highlights the *argumentative content* of the conversation, is made of two parts. The first is an overview of the argumentative activity of participants; the second is an outline of the argumentative structure detected for each thematic episode. Each turn is paraphrased according to its recognized argumentative function and linked through a discourse relationship to previous turns.

ARGUMENTATIVE CONTENT
The following participants: Andreas, Dave, Don, Jane and Morgan, expressed their dissent 58 times; however Dave, Andreas and Morgan expressed dissent in a consistently smaller percentage.
The following participants: Adam, Andreas, Dave, Don, Jane and Morgan, asked questions 55 times.
The remaining 1204 turns expressed positive content by proposing, explaining or raising issues However Adam, Dave and Andreas suggested and raised new issues in a consistently smaller percentage.
The following participants: Adam, Andreas, Dave, Don, Jane and Morgan, expressed acceptance 217 times.

EPISODE ISSUE No. 1 In this episode we have the following argumentative exchanges between Don, Jane and Morgan:
Don raises the following issue: [um so - so it definitely had that as a concept .]
then he, overlapped by Morgan, continues: [so th ... it has a single timeline .]
Morgan accepts the previous explanation: [uhhuh, .]
Then Don elaborates the following explanation: [and then you can have lots of different sections.]

EPISODE ISSUE No. 2 In this episode we have the following argumentative exchanges between Don, Jane and Morgan:
Jane, overlapped by Dave, suggests the following explanation [actually we - we use a generalization of the - the sphere format .] [yeah so there is something like that .] [but it's um probably not as sophist... .]
Don asks the following question [well what does H.T.K. do for features ?]
then he, overlapped by Jane, continues [or does it even have a concept of features ?]
Jane provides the following answer [they h... it has its own ... .]
then she , overlapped by Morgan, continues [I mean Entropic has their own feature format that's called like s... s.d. or some s... s.f. or something like that.]
Morgan accepts the previous explanation [yeah.]

EPISODE ISSUE No. 3 In this episode we have the following argumentative exchanges between Don, Jane and Morgan:
Don, overlapped by Morgan, provides the following explanation [you know we'll do something where we - some kind of data reduction where the prosodic features are sort ... uh either at the word level or at the segment level .]
Morgan agrees [right, .]
then he continues [or - or something like that .] [they're not going to be at the phone level and they're n- ... not going to be at the frame level when we get done with sort of giving them simpler shapes and things .] [and so; the main thing is just being able ... .] [well I guess the two goals .]
Don disagrees with the previous explanation [um one that Chuck mentioned is starting out with something that we don't have to start over that we don't have to throw away if other people want to extend it for other kinds of questions .]
Morgan agreees [right, .]

It's important to notice the use of pronouns in the paraphrase and the insertion of discourse markers to indicate the continuation by the same speaker. Formulating adequate paraphrases may sound awkward sometimes and this is due to the inherent difficulty of the task.

4 Experimental Results

We started using the system to parse ICSI corpus of meetings [11] and we realized that the semantic representation and the output of the parser were both inadequate. By looking at errors, we became aware of the peculiarities of spoken dialog texts such as the ones made available in ICSI corpus, and to the way to implement solutions in such a complex system. These dialogs are characterized by the need to argument in an exhaustive manner the topics to be debated which are the theme of each multiparty dialog. The following are some statistics that characterize the corpus:

- percent of turns made of one single word: 30%
- percent of turns made of up to three words: 40%
- average number of words x turn overall: 7
- average number of words x turn minus short utterances: 11

These values correspond to those found for PennTreebank corpus where we can count up to 94K sentences for 1M words – again 11 words per sentence. In analyzing ICSI, we found turns with as much as 54 words depending on the topic under discussion and on the people on the floor. The system has been used to parse the first 10 dialogs of the ICSI corpus for a total number of 98523 words and 13803 turns. We had to tune all the modules and procedures carefully. In particular, the module for argumentative automatic classification was incrementally improved in order to cover all conventional ways to express agreement.

We had one skilled linguist to provide a turn level annotation for argumentative labels: we don't have any agreement measure in this case, even though we expect the annotation to be in line with current experiments on the same subject [16]. In Table 1 we report data related to the experiment of automatic annotation of argumentative categories.

	Correct	Incorrect	Total Found
Accept	662	16	678
Reject	64	18	82
Propose	321	74	395
Request	180	1	181
Explain	580	312	892
Disfluency	19		19
Total	1826	421	2247

Table 1. Accuracy of automatic argumentative annotation

On a total of 2304 turns, 2247 have received an argumentative automatic classification, with a Recall of 97.53%. We computed Precision as the ratio between Correct Argumentative Labels/Found Argumentative Labels, which corresponds to 81.26%. The F-score is 88.65%.

5 Conclusions

We have presented work carried out to extend and adapt a language analysis system designed for text understanding in order to make it fit for argumentative analysis as the core technology for abstractive summarization of conversations. Summaries are generated by aggregating the statistics of the analysis, by highlighting the topic discussed and by segmentating the content of conversation into argumentative episodes. The turns of the episodes are then paraphrased according to their argumenative functions highlighting their discourse relationship (e.g. replies and elaborations).

For this purpose, we implemented a module that computes Argumentative automatic classification labels out of a small set, on top of discourse relations and other semantic markers determined by the semantic component of the system. The system has been evaluated for the argumentative classification module and results are very encouraging.

5.1 Future work

Of course we intend to complete the analysis of all dialogues contained in the ICSI corpus and refine our algorithms. In particular we still need to work at the level of DECISION labelling, and to improve the discrimination of really argumentative from pragmatically irrelevant utterance, a choice that in some cases is hard to make on an automatic basis. We would like to formally assess the robustness of the system by applying the algorithm to the output of automatic transcription systems and evaluate its degree of degradation.

References

[1] J Ang, Y. Liu, and E. Shriberg, 'Automatic dialog act segmentation and classification in multiparty meetings', in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (2005).

[2] R. Delmonte, *Computational Linguistic Text Processing Logical Form, Semantic Interpretation, Discourse Relations and Question Answering*, Nova Science Publishers, New York, 2007.

[3] R. Delmonte, *Computational Linguistic Text Processing Lexicon, Grammar, Parsing and Anaphora Resolution*, Nova Science Publishers, New York, 2009.

[4] R. Delmonte, A. Bristot, and V. Pallotta, 'Towards automatic dialogue understanding', in *Proceedings of the International Conference on Spoken Communication (GSCP09)*, Naples, Italy, (2009).

[5] M. Galley, K. McKeown, J. Hirschberg, and E. Shriberg, 'Identifying agreement and disagreement in conversational speech: Use of bayesian networks to model pragmatic dependencies', in *Proceedings of 42nd Meeting of the Association of Computational Linguistics (ACL)*, (2004).

[6] U. Hahn and I. Mani, 'The challenges of automatic summarization', *IEEE Computer*, **33(11)**, 29–36, (2000).

[7] D. Hillard, M. Ostendorf, and E. Shriberg, 'Detection of agreement vs. disagreement in meetings: Training with unlabeled data', in *Proceedings of HLT-NAACL*, (2003).

[8] P. Hsueh and J.D Moore, 'Automatic topic segmentation and labeling in multiparty dialogs', in *Proceedings of IEEE Spoken Language Technology Workshop*, pp. 98–101, (December 2006).

[9] P. Hsueh and J.D. Moore, 'Automatic decision detection in meeting speech', in *Proceedings of the workshop on Machine Learning for Multimodal Interfaces (MLMI)*, pp. 168–179, (2007).

[10] P. Hsueh and J.D. Moore, 'What decisions have you made: Automatic decision detection in conversational speech', in *Proceedings of HLT-NAACL*, (2009).

[11] A. Janin, D. Baron, J. Edwards, D. Ellis, D. Gelbart, N. Morgan, B. Peskin, T. Pfau, E. Shriberg, A. Stolcke, and C. Wooters, 'The icsi meeting corpus', in *Proceedings of IEEE/ICASSP*, pp. 364–367, (2003).

[12] G. Murray, S. Renals, and J. Carletta, 'Extractive summarization of meeting recordings.,', in *Proceedings of the 9th European Conference on Speech Communication and Technology*, pp. 593–596, (2005).

[13] J. Niekrasz, M. Purver, J. Dowding, and S. Peters, 'Ontology-based discourse understanding for a persistent meeting assistant', in *Proceedings of the AAAI Spring Symposium Persistent Assistants: Living and Working with AI.*, (2005).

[14] V. Pallotta, H. Ghorbel, A. Ballim, A. Lisowska, and S. Marchand-Maillet, 'Towards meeting information systems: Meeting knowledge management', in *Proceedings of ICEIS*, pp. 464–469, (2004).

[15] V. Pallotta, H. Ghorbel, G. Coray, and P. Ruch, 'An argumentative annotation schema for meeting discussions', in *Proceedings of the LREC 2004 international conference,*, (2004).

[16] V. Pallotta, V. Seretan, and M. Ailomaa, 'User requirements analysis for meeting information retrieval based on query elicitation', in *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL 2007)*, pp. 1008–1015, (2007).

[17] K. Riedhammer, D. Gillick, B. Favre, and D. Hakkani-Tür, 'Packing the meeting summarization knapsack', in *Proceedings of the 9th International Conference of the ISCA (Interspeech 2008)*, pp. 2434–2437, (2008).

[18] R.J. Rienks, D. Heylen, and E. van der Weijden, 'Argument diagramming of meeting conversations', in *Proceedings of the Workshop on Multimodal Multiparty Meeting Processing at the 7th International Conference on Multimodal Interfaces (ICMI)*, pp. 85–92, (2005).

[19] D. Rosemberg and J.A.A. Silince, 'Common ground in computer-supported collaborative argumentation', in *Proceeding of the Workshop on Computer Supported Collaborative Argumentation for Learning Communities (CSCL99)*, (1999).

[20] K. Zechner, 'Automatic summarization of open-domain multiparty dialogues in diverse genres', *Computational Linguistics*, **28(4)**, (2002).