

Abstractive Summarization of Voice Communications

Vincenzo Pallotta¹, Rodolfo Delmonte², and Antonella Bristot²

¹Department of Computer Science
Webster University, Geneva
Route de Collex 15
CH-1293 Bellevue, Switzerland
pallotta@webster.ch

²Department of Language Science
Università “Ca Foscari”
30123 – Venezia, Italy
delmont@unive.it

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

Keywords: summarization, dialogue, argumentation.

1 Introduction

We consider an open problem in Natural Language Processing (or Computational Linguistics) that is the generation of text summaries from voice communication that is spoken dialogs occurring between two or more people. 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 perfectly transcribed and that transcription is segmented into turns assigned to each speaking participants. While this assumption might seem restrictive, most of corpora of voice communications are of this kind. Moreover, the cost of transcribing speech is far lower than the cost of summarization. The problem of automatic transcription of voice communication is beyond the scope of this work. However, the algorithm we present has characteristics that make it robust with respect to speech disfluency, either naturally occurring or generated from recognition errors of automatic transcription systems.

Automatic summarization of voice communications would be a very useful tool that allows enterprises in extracting knowledge from meeting records and integrating this knowledge into corporate knowledge bases for future access by question-answering. For instance, one could ask the system questions like “why this decision was made” or “who rejected the proposal made by XX?” or “How the decision of

doing X impacted the progress of the project Y?” While very likely to be asked by user of meeting recording systems [1], these kinds of question are very hard to answer because one has to consider several meetings and know about the context of these meetings (e.g. the projects, the people, the roles, the duties, the agenda, the corporate policies and strategies). Moreover, it requires a deep understanding of the meeting situations such as its dynamics, the rules of order adopted, the specific language used, and the culture-specific rituals.

Producing automatic summaries of meeting would benefit the enterprise because it would turn implicit and tacit information into explicit, actionable knowledge. It would save time because people who did not attend the meeting could have a quick insightful glimpse of what happened during the meeting without having to replay the whole recording with a meeting browser [2]. In more general terms, this approach to summarization of spoken data would be also beneficial to cope with information overload arising from many audio-visual data on TV channels and on the Internet such as broadcasted news, talk shows, podcasts, webinars, etc.

1.1 Related Work

The problem of automatic speech summarization was initially investigated in the 80' in the context of several DARPA projects [3]. 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 [4, 5].

In fact, this method has severe limitations due to the intrinsic characteristics of the source data: conversations are not as coherent as ordinary narrative text (such as news or scientific articles) and obtaining a coherent text from conversations is practically impossible using the extractive approach. Any system that performs extractive summarization must be evaluated against human-annotated test datasets. As pointed out in [6], inter-annotator agreement is very low for this type of task, which makes test data nearly useless for evaluation. Intuitively, selecting salient content from conversations is a really difficult task and subjective selection of excerpts leads to fairly different results. Moreover, the proposed solutions for extractive summarization of meetings have already reached their qualitative upper bounds [7].

Abstractive summarization of dialogs has only been recently considered as an important challenge in several research projects among which the IM2¹ project, ICSI Meeting Recorder² project, the European FP6 AMI³ project, and the DARPA's CALO⁴ project. At the beginning, these projects mainly focused on the first (and simplest) part of the problem that is capturing and storing the multimodal events occurring during a voice communications and provide a multimedia record of the events. This record includes aligned media such as audio, video, speech transcription, pointers to document used, gestures, gaze and additional information such as the emotional status of the speakers (involved, bored, etc.). These types of information are indeed very useful for the subsequent step of analysis whose goal is to provide an

¹ <http://www.im2.ch/>

² <http://www.icsi.berkeley.edu/Speech/mr/>

³ <http://www.amiproject.org/>

⁴ <http://caloproject.sri.com/>

automatic summary of the meeting. On this side, some advances have been achieved such as extracting (very basic) “dialog acts” [8], detecting “agreement and disagreement” [9, 10], segmenting the meeting into “thematic zones” [11], and, only recently, detecting “decision points” [12, 13], “action items” [14, 15], and “subjective statements” [16]. Also in the framework of the AMI project, a group was investigating on the automatic argumentative structuring of multi-party dialogs [17]. Unfortunately, although very relevant, none of these works have directly addressed the problem of abstractive summarization of voice communication and most of the advances on the detection of rich content information from meetings were mainly exploited in improving *extractive* summarization, which, as said before, can be only improved up to a given (unsatisfactory) upper bound.

The problem of *abstractive* summarization of voice communications is apparently much harder to solve than its *extractive* version. It requires almost the full understanding of the source data [18]. Moreover, as also recognized by [19], in order to perform the generation of an abstractive summary it is necessary to classify participants’ contributions according to their informativeness and function in the conversation so that an appropriate ontology of the meeting could be adequately populated.

In [20] the problem of abstractive summarization of conversations is framed as a sentence compression problem. This work is an important step in filling the gap between *extractive* and *abstractive* summarization. However, they recognize the need of a sophisticated extraction and generation algorithm to achieve full abstractive summarization. In fact, their approach is not appropriate for generating abstracts of conversations where turns are already highly compressed.

We conclude our review of related work by looking at [21] where a language generation algorithm produces indicative abstracts of meeting transcriptions. The existing annotations of the AMI corpus data are used to create a content representation that is subsequently fed to a sentence planner. The semantic representation produced is then rendered into a surface form with surface generator. This work is in spirit very similar to ours with the notable exception that the semantic representation is not as rich as ours. Basically, only topics, very general meeting stages (opening, debating, closing), and temporal sequencing of topics discussed are used to generate the abstract.

As we will show in the next sections, using similar (or even simpler) generation techniques and with the information we are able to extract, we can provide much richer indicative abstracts. In order to extract the information that allows us to provide better abstracts we need a very powerful syntactic analyzer (i.e. a parser) and a mapping of the analysis onto semantic and pragmatic representations. This step is obtained through a discourse analyzer and an argumentative analysis algorithm.

1.2 Paper Outline

The paper is organized with a presentation of the system in section 2 which we evaluated on a manually annotated corpus of meeting transcriptions. In section 3 we present a preliminary generation system that provides summaries in form of memos. A final section is devoted to the conclusions and future work.

2 Automatic Argumentative Analysis

We start from transcribed turns in multi-party dialogs such as those available from corpuses built within the projects we described earlier. Each turn is assigned to a speaker and the transcription is assumed to be 100% accurate⁵.

Our solution to the problem of abstractive summarization of voice communications would be learning a general semantic/pragmatic structure for meeting dialogs data that could be specialized for different voice communication types such as brainstorming, design, formal meetings with rules of order, board meetings, 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 hearth of voice communications. Unfortunately, argumentative processes are not sufficiently lexically marked in dialogs to justify extractive summarization. For instance, extracting a turn containing the single word “no” cannot summarize a disagreement to a proposal about an issue. This action is the result of a complex, sometimes long exchange between two or more participants. By simply extracting some turns from the dialog some essential parts of the context are lost. In contrast, it would be much better to provide a description where an argumentative-like terminology is used such as, for instance, “X disagreed to the proposal made by Y on the issue raised by Z”. This paraphrase is made of words that do not occur in the dialog.

Our system is capable to recognize a small but relevant set of argumentative action that can be used to produce such description and thus creating a meaningful abstractive summary of the analyzed dialog. While the generation of the summaries might be quite standard given the extracted information, we will mainly focus in this paper on the recognition of argumentative events as the core component required for implementing an abstractive summarizer of voice communications.

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 co-reference;

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 make decisions in argumentative automatic classification algorithm.

⁵ We plan to relax the latter constraints in future implementations of our system and deal with output from automatic speech recognition systems.

2.1 Argumentative Structure – Issues and Theories

We consider the adoption of a structured representation of conversational content based on argumentation theory. The argumentative structure defines the different patterns of argumentation used by participants, as well as their organization and synchronization in the discussion. A dialog is decomposed into several stages such as *issues*, *proposals*, and *positions*, each stage being possibly related to specific aggregations of elementary dialog acts [22]. 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 (or ontology), the Meeting Description Schema [23]. 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 (i.e., 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.

2.2 Computing Argumentative Annotations

The core of our solution is based on adapting and extending GETARUNS [24, 25], a system for text understanding developed at the University of Venice. GETARUNS is organized as a pipeline that 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:

1. The lower module for parsing, where sentence strategies are implemented;
2. The middle module for semantic interpretation and discourse model construction which is cast into Situation Semantics;
3. The higher module where reasoning and generation takes place.

Automatic Argumentative Annotation (A³) 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:

- the Discourse Model, which include individuals, sets, classes, cardinality, properties related to entities by means of their semantic indices;
- information about 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;
- 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.

2.3 The A³ Algorithm

The A³ algorithm, described in details in [26], uses the following 21 discourse relations:

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

These relations 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 that are non-sentences and are not semantically interpretable.

The algorithm is outlined as follows:

1. It recovers Dialog Acts for each dialog turn as they have been assigned by the system⁶;
2. It recovers Overlaps as they have been marked during the analysis;
3. It produces a Polarity attribute for the Opinion label, 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 multiple candidates by looking at a hierarchy of argumentative labels graded for Pragmatic Relevance which establishes that, for instance, Question is more relevant than Negation, which is in turn more relevant than Raising an Issue, etc.

⁶ The dialog act set used by GETARUNS differs from that used in the ICSI corpus and described in [22]. Details are available in [26].

2.4 Experimental Results

When we started using the system to parse ICSI corpus of meetings [27], we realized that the generated semantic representation and the output of the syntactic parser were both inadequate. So we worked at the deficiencies we identified in an empirical manner. This approach made us 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.

ICSI dialogs are characterized by the need to argument in an exhaustive manner the topics to be debated which are the themes of each multiparty dialog.

The mean length of utterances/turns in each dialog we parsed was rather long. This makes ICSI dialogs hard to compute. From a count of number of words per turn, we came up with the following average figures:

- 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 after subtracting 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 number of people on the floor.

Our system has been used to parse the first 10 dialogs of the ICSI corpus for a total number of 98523 words and 13803 turns. This has been done to “train” the system: what happened was that, for the first 5 dialogs, we had to take care of failures. We also 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. For this reason, we then chose two random additional dialogs to test this second task.

We had one skilled linguist to provide a turn level annotation for argumentative labels: we don’t have any inter-annotator agreement measure in this case, even though we expect the annotation to be in line with experiments on the same subject presented in [28], namely a Kappa of 0.42 and a raw agreement of 62%.

On a total of 2304 turns, 2251 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%.

3 Abstract Summarization of Conversations

A complementary way to provide users with simplified access to conversational content is by means of summaries. Analyzing and summarizing conversations (or dialogues) is very challenging [29]. Many existing summarization techniques are tailored for the narrative genre and can hardly be adapted to the dialogue genre.

We advocate for abstractive summarization of conversational content. Abstractive summarization of narrative texts is typically based on sentence compression and/or

paraphrase [20]. Our approach differs from that of [21] because our indicative abstracts are not limited to exclusively thematic content. Instead, our abstractive summarization system generates descriptions of the conversation dynamics based on both thematic content and argumentative structure we are able to automatically extract as described in section 2.

The output of our system consists instead of a memo made of several sections, namely describing the conversational settings, participants' number and names, statistics about the interactivity of participants (e.g. the degree of competitiveness), the topics discussed, and the arguments. Arguments are grouped into episodes bound together by thematic cohesion. 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. The following is a (partial) example of a memo which is generated by our system from the analysis of an ICSI meeting (only turns and speaker with no additional annotations):

MEETING MEMO

GENERAL INFORMATION ON PARTICIPANTS

The participants to the meeting are 7.

Participants less actively involved are Ami and Don who only intervened respectively for 38 and 68 turns.

LEVEL OF INTERACTIVITY IN THE DISCUSSION

The speaker that has held the majority of turns is Adam with a total of 722 turns, followed by Fey with a total of 561.

The speaker that has undergone the majority of overlaps is Adam followed by Jane.

The speaker that has done the majority of overlaps is Jane followed by Fey.

Jane is the participant that has been most competitive.

DISCUSSION TOPICS

The discussion was centred on the following topics: **schemas, action, things and domain.**

The main topics have been introduced by the most important speaker of the meeting.

The participant who introduced the main topics in the meeting is: **Adam.**

The most frequent entities in the whole dialogue partly coincide with the best topics, and are the following: **action, schema, things, 'source-path-goal', person, spg, roles, bakery, intention, specific, case, categories, information, idea.**

ARGUMENTATIVE CONTENT

The following participants: **Andreas, Dave, Don, Jane, Morgan** expressed their dissent 52 times. However Dave, Andreas and Morgan expressed dissent in a consistently smaller percentage.

The following participants: **Adam, Andreas, Dave, Don, Jane, Morgan** asked questions 55 times.

The remaining 1210 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, Morgan** expressed acceptance 213 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.]

then he continues

[each of which have i.d.'s attached to it .]

[and then you can refer from other sections to those i.d.'s .]

[if you want to .]

[so that you start with - with a timeline tag .]

[timeline .]

[and_then, you, have, a, bunch, of, times, .]

[I don't ... I don't remember exactly what my notation was .]

Jane, overlapped by Don, raises the following issue

[oh; i, remember seeing an example of this .]

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.]

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 agrees

[right, .]

...

The generated memo outlines the extracted argumentative structure and can be used to generate a more succinct abstract where many details of the conversation are

hidden. For instance, explanations and other types of elaborations around proposals or issues can be omitted. Moreover, the details of agreements and disagreements (e.g. the reasons) can be omitted. Finally, overlaps (recognized as backchannels) can also be removed.

For the previous example, the episode 1 can be summarized in the following manner:

Don, Jane and Morgan discussed about “single timeline”. Jane raised an issue: “I remember seeing an example of this”. Don asked Jane to clarify and Morgan accepted Jane’s explanation. Morgan agreed and provided an explanation. Don disagreed..

From the above abstract it is possible to understand what type of interaction was ongoing between the three participants without having to look at the details. From this abstract one can already figure out that the topic “single timeline” was particularly hot, that no agreement was reached at that point, and that Jane’s opinion was shared by Morgan, but not by Don.

4 Conclusions

We have presented work carried out to extend and adapt a natural language understanding system designed for text analysis in order to make it fit for dialogue understanding as the core component of abstractive summarization of voice communications technology. We implemented a set of enhancements to cope with typical dialogue related problems, such as presence of non-sentential fragments, elliptical fragments interpretable as speech acts, massive presence of generic non-referential pronominal expressions, etc.

We implemented a number of additional components: an algorithm that takes care of overlaps and uses that information to split current utterances and temporally realign the conversational flow. We outlined the A³ algorithm for the automatic argumentative labelling and linking of conversation turns, which relies on discourse relations and other semantic markers determined by the semantic component of the system.

The system has been evaluated for the argumentative classification module. Results are very encouraging. However, we note that in that task, labels which may cause great uncertainty and are highly ambiguous (e.g. explain and justify), have been lumped together to facilitate the classification task.

4.1 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. We also need to design an evaluation framework for abstractive summaries of conversations because the current metrics used for evaluating extractive summaries are not well adapted (e.g. ROUGE). We agree with [30] that the evaluation of

abstractive summaries must be extrinsic, i.e. based on human judgement or indirectly observed as its helpfulness in performing a given task.

References

- [1] Pallotta, V., Seretan, V., Ailomaa, M.: 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), Prague, pp. 1008–1015 (2007)
- [2] Cremers, A.H.M., Hilhorst, B., Vermeeren, A.P.O.S.: What was discussed by whom, how, when and where? personalized browsing of annotated multimedia meeting recordings. In: Proceedings of HCI 2005, Edinburgh, UK, pp. 1–10 (2005)
- [3] Zechner, K.: Automatic Summarization of Open-Domain Multiparty Dialogues in Diverse Genres. *Computational Linguistics* 28(4) (2002)
- [4] Murray, G., Renals, S., Carletta, J.: Extractive summarization of meeting recordings. In: Proceedings of the 9th European Conference on Speech Communication and Technology, pp. 593–596 (2005)
- [5] Garg, N., Favre, B., Reidhammer, K., Hakkani-Tür, D.: ClusterRank: A Graph Based Method for Meeting Summarization. In: Proceedings of Interspeech 2009, Brighton, UK (2009)
- [6] Buist, A.H., Kraaij, W., Raaijmakers, S.: Automatic Summarization of Meeting Data: A feasibility Study. In: Proceedings of the 15th CLIN Conference (2005)
- [7] Reidhammer, K., Gillick, D., Favre, B., Hakkani-Tür, D.: Packing the Meeting Summarization Knapsack. In: Proceedings of the 9th International Conference of the ISCA (Interspeech 2008), Brisbane, Australia, pp. 2434–2437 (2008)
- [8] Ang, J., Liu, Y., Shriberg, E.: Automatic dialog act segmentation and classification in multiparty meetings. In: Proceedings of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Philadelphia (March 2005)
- [9] Galley, M., McKeown, K., Hirschberg, J., Shriberg, E.: Identifying Agreement and Disagreement in Conversational Speech: Use of Bayesian Networks to Model Pragmatic Dependencies. In: Proceedings of 42nd Meeting of the ACL, Barcelona, July 21–26 (2004)
- [10] Hillard, D., Ostendorf, M., Shriberg, E.: Detection of agreement vs. disagreement in meetings: Training with unlabeled data. In: Proceedings of HLT-NAACL 2003 (2003)
- [11] Hsueh, P., Moore, J.D.: Automatic topic segmentation and labeling in multiparty dialogs. In: Proceedings of Spoken Language Technology Workshop, pp. 98–101. IEEE, Los Alamitos (2006)
- [12] Hsueh, P., Moore, J.D.: Automatic Decision Detection in Meeting Speech. In: Popescu-Belis, A., Renals, S., Boulard, H. (eds.) *MLMI 2007*. LNCS, vol. 4892, pp. 168–179. Springer, Heidelberg (2008)
- [13] Hsueh, P., Moore, J.D.: What Decisions Have You Made: Automatic Decision Detection in Conversational Speech. In: Proceedings of HLT-NAACL (2009)
- [14] Niekrasz, J., Purver, M., Dowding, J., Peters, S.: Ontology-Based Discourse Understanding for a Persistent Meeting Assistant. In: Proceedings of the AAAI Spring Symposium Persistent Assistants: Living and Working with AI, Stanford (2005)
- [15] Murray, G., Renals, S.: Detecting action items in meetings. In: Popescu-Belis, A., Stiefelhagen, R. (eds.) *MLMI 2008*. LNCS, vol. 5237. Springer, Heidelberg (2008)

- [16] Raaijmakers, S., Truong, K., Wilson, T.: Multimodal subjectivity analysis of multiparty conversation. In: Proc. of EMNLP 2008, Honolulu, HI, USA (2008)
- [17] Rienks, R.J., Heylen, D., van der Weijden, E.: Argument diagramming of meeting conversations. In: Proceedings of the Workshop on Multimodal Multiparty Meeting Processing at the 7th International Conference on Multimodal Interfaces (ICMI), Trento, Italy, pp. 85–92 (2005)
- [18] Hahn, U., Mani, I.: The Challenges of Automatic Summarization. *IEEE Computer* 33(11), 29–36 (2000)
- [19] Murray, G., Carenini, C., Ng, R.: Interpretation and Transformation for Abstracting Conversations. In: Proceedings of Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL (HLT-NACL), Los Angeles, California, pp. 894–902 (June 2010)
- [20] Liu, F., And Liu, Y.: From Extractive to Abstractive Meeting Summaries: Can It Be Done by Sentence Compression? In: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, Suntec, Singapore, August 4, pp. 261–264 (2009)
- [21] Kleinbauer, T., Becker, S., Becker, T.: Indicative Abstractive Summaries of Meetings. In: Proceedings of 4th Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI 2007), Nagoya, Japan (2007)
- [22] Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Van Ess-Dykema, C., Meteeret, M.: Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech. *Computational Linguistics* 26(3), 339 (2000)
- [23] Pallotta, V.: Framing Arguments. In: Proceedings of the International Conference on Argumentation, Amsterdam, Netherlands (June 2006)
- [24] Delmonte, R.: Computational Linguistic Text Processing – Logical Form, Semantic Interpretation, Discourse Relations and Question Answering. Nova Science Publishers, New York (2007)
- [25] Delmonte, R.: Computational Linguistic Text Processing – Lexicon, Grammar, Parsing and Anaphora Resolution. Nova Science Publishers, New York (2009)
- [26] Delmonte, R., Bristot, A., Pallotta, V.: Towards Automatic Dialogue Understanding. In: Proceedings of the International Conference on Spoken Communication (GSCP 2009), Naples, Italy, February 23-25 (2009)
- [27] Janin, A., Baron, D., Edwards, J., Ellis, D., Gelbart, D., Morgan, N., Peskin, B., Pfau, T., Shriberg, E., Stolcke, A., Wooters, C.: The ICSI Meeting Corpus. In: Proceedings of IEEE/ICASSP 2003, Hong Kong, April 6-10, vol. 1, pp. 364–367 (2003)
- [28] Pallotta, V., Seretan, V., Ailomaa, M., Ghorbel, H., Rajman, M.: Towards an argumentative coding scheme for annotating meeting dialogue data. In: Proceedings of the 10th International Pragmatics Conference, IPrA 2007, Göthemburg, July 8-13 (2007)
- [29] Maybury, M.: Keynote on Searching Conversational Speech. In: Proceedings Workshop on Searching Spontaneous Conversational Speech Part of the ACM SIGIR 2007 Conference, Amsterdam, July 27 (2007)
- [30] Murray, G., Kleinbauer, T., Poller, P., Renals, S., Becker, T., Kilgour, J.: Extrinsic Summarization Evaluation: A Decision Audit Task. In: Popescu-Belis, A., Stiefelhagen, R. (eds.) MLMI 2008. LNCS, vol. 5237, pp. 349–361. Springer, Heidelberg (2008)