Towards Automatic Minuting of Meetings

Anna Nedoluzhko and Ondřej Bojar

Charles University, Institute of Formal and Applied Linguistics
{nedoluzko, bojar}@ufal.mff.cuni.cz

Abstract: Many meetings of different kinds will potentially benefit from technological support like automatic creation of meeting minutes. To prepare a reasonable automation, we need to have a detailed understanding of common types of meetings, of the linguistic properties and commonalities in the structure of meeting minutes, as well as of methods for their automation. In this paper, we summarize the quality criteria and linguistic properties of meeting minutes, describe the available meeting corpora and meeting datasets and propose a classification of meetings and minutes types. Furthermore, we analyze the methods and tools for automatic minuting with respect to their use with existing types of datasets. We summarize the obtained knowledge with respect to our goal of designing automatic minuting and present our first steps in this direction.

1 Introduction

Meeting minutes keep a record of what was discussed at the meeting. Meeting minutes is a written document used to inform participants and non-participants of what happened during the meeting. The problem with meeting minutes is that it takes much time to write them down properly. Considering different kinds of meetings and minutes, we can observe that most of meetings will potentially benefit from technological support like automatic creation of meeting minutes.

We suggest to develop automatic minuting, i.e. the summarization of dialogue transcripts into a compact structured form.

Taking into account the wide variety of real meetings, we believe that the most effective way is to structure minutes according to a meeting agenda, which is generally prepared manually by the organizers before the meeting. The audio recordings of the meetings can be transcribed using speech recognition techniques. Then the foreseen minuting software will automatically recognize and extract important information from meeting transcripts and classify it to the pre-defined "slots" in the agenda (such as, for example, "annotation strategy", "conference in Paris", "next meeting timing"). This is a very complex task that requires a thorough understanding of meetings structure with respect to what kinds of minutes they have, as well

as orientation in methods which can be used for meeting summarization.

In this paper, we prepare the theoretical and descriptive basis for automatic creation of minutes. Our main objective is to suggest a reasonable classification of meetings (Section 2.1), meeting minutes (Section 2.2), available meeting datasets (Section 3) and methods that can be used for meeting dialogue summarization (Section 4). In Section 5, we summarize the obtained knowledge with respect to our goal of designing automatic minuting and in Section 6 we present our first steps in this direction.

2 Meetings and Minutes Description

2.1 Types of Meetings

There are many different kinds of meetings carried out for different purposes. Every meeting is unique, but there are some common types of meetings, which can be distinguished according to aspects such as primary meeting goals, key participant roles or common challenges.

Different studies of meetings have been conducted by anthropologists, psychologists, sociologists, political scientists or business administrators (see e.g. [43, 10, 22, 14] and plenty of others). The handbooks about meeting organization begin with Robert's Rules of Order [41], which were first published in 1876. In the last edition [42], meetings are classified according to the timing and regularity (into *regular, special, adjourned* and *annual*) and according to confidentiality (into *executive* and *public* sessions). Moreover, *electronic* meetings are distinguished as a special type.

The classification may be also based on other aspects. For example, authors in [6] speak about *formal* and *informal* meetings, each type having specific linguistic features, such as different pronominal choice or modes of personal address (choosing between 'I' and 'we', 'you' and 'the people' and so on). Attorney [1] also distinguishes so-called paper meetings with minutes, when participants informally agree on specific corporate actions and minutes are prepared as though the decision were approved at a real meeting. The language aspects of meetings, search for coherence and sense-making are analysed in [6], where a cross-linguistic and cross-cultural comparison of Italian and British meetings has been provided. The authors point out the specificity of *non-native multi-party* meetings, where substantial parts of communication may be devoted to metalanguage details. The form of meetings—and minutes respectively—is related to cultural concepts.

Moreover, they present a cross-linguistic investigation of such pragmatic phenomena as for example, the pronominal choice or modes of personal reference.

The number of meeting types in online resources[1] varies from four to sixteen. Most lists include five meeting types and choose within business meetings, decision making meetings, information sharing meetings, status update meetings, planning meetings, innovation meetings, problem solving meetings, team-building meetings, workshops or conferences. What happens in reality, however, is that meetings can actually fit into several of these categories.

The analysis of existing classifications and real meetings which are available to the authors reveals many factors which may affect how the meeting is organized, and, subsequently, which kind of agendas and minutes it needs. Such factors are, e.g., the intention of the meeting, its format, the size of the group, regularity, information density, content and context of the meeting, participation styles, the expected audience, etc. However, it appears that such factors as meeting content, location, face-to-face vs. remote or even the group size do not really affect the core goals and format of the meetings.

2.2 Meeting Minutes

Meeting minutes are recorded in many different ways. The formats can vary according to the personal style of the minutes writer, national, group or domain preferences, according to the degree of meetings' formality, regularity, length and so on. Moreover, the format, style and content requirements for meeting minutes may vary depending on the meeting intention. It means that, for instance, an interdepartmental decision making meeting would look very differently from an informal idea generation meeting of close colleagues. Some note takers use standard templates for recording minutes, which are offered by websites[2] or suggested in [1, 2], etc. In this section, we present the description and a rough classification of meeting minutes.

What minutes definitely include. According to a variety of handbooks like [1, 2] or [42], the minutes should always include: (i) the name of the organization (ii) date, time and location of the meeting, (iii) a list of the attendees. The official meetings should also contain the signature of the chairperson. Robert's Rules of Order [42] also suggest some other official requirements such as a statement confirming that the organization's regular presiding officer and secretary are present and mentioning of whether the previous meeting's minutes were read and approved. The informative part of the minutes depends on the content of the meeting. Generally, it contains an indication of the content under discussion, what needs to be done as a result of the meeting, decisions made during the meeting and voting results.

One of the main issues about the meeting minutes is that they should contain mainly a record of what was done at the meeting, not what was said by the members [42]. Also, the minutes are not the right place for future action items or to-do lists.

Structure of minutes. Regarding the organization information in the minutes, Rigley [40] distinguishes *notes of meeting* by which each action proposed or reported is laid down by a numbering or bullet system, more detailed *narrative minutes* written as text, *resolution minutes* where only decisions are recorded and *action minutes* which distribute responsibilities between participants and are usually written in two or three columns.

Minutes can be also categorized as agenda-based and informal meeting minutes, which summarize decisions taken and follow-up actions and responsibilities, but do not necessarily contain all kinds of information prescribed for the official minutes. Another possibility of minutes classification is structuring into three categories: expressed ideas, achieved conclusions and next steps.

Linguistic features of minutes. As far as we have observed, there are no precise linguistic restrictions to meeting minutes. However, minutes are supposed to ensure brevity and clarity, so that they are easy to read. Therefore, the corresponding websites recommend to write minutes in so-called basic plain English. The relevant language characteristics for the meeting minutes creation are, writing in the same tense throughout the minutes, using the simplest words that are appropriate, and avoiding jargon and legalese, using verbs rather than abstract nouns like "consideration", "approval" or "clarification", writing in active rather than passive phrases, not using acronyms, keeping sentences short and using bullet points and numbered lists where appropriate. Furthermore, especially for the minutes, it is recommended to (1) avoid inflammatory or personal observations, (2) use as few adjectives or adverbs as possible and (3) avoid using people's names.

3 Available Datasets for Automatic Minuting

To create reliable automatic minuting, we need to have some training and test data of meetings and minutes. However, there is a significant disproportion between the number and domain variety of real meetings and available open datasets which can be used for this purpose. Meetings are being held all over the world thousands times a day, but we can hardly use them, because the transcripts and minutes are mostly publicly unavailable. The exceptions are mostly in the political domain, because politicians are obliged to make their meetings open to the public. For this reason, our non-political datasets will be relatively small. The brief description of them is given below.

The **AMI Meeting corpus**[3] [28] contains 100 hours of meeting discussions, two thirds of which are scenario

---

[1]See, for example, https://blog.firstagenda.com/the-4-most-important-types-of-meetings.

[2]https://lessmeeting.com/

[3]http://groups.inf.ed.ac.uk/ami/corpus/

Table 1: Summary of English meeting corpora.

Dataset	Length	Transcripts	Minutes
AMI Corpus (scenario and non-scenario)	100h (70h+30h)	manual	partly
ICSI Corpus	70 h	manual	no
NIST Meeting Corpus	20 h	manual	no
ISL Corpus	10 h	ASR	no

meetings which had been played (acted out) for creating the corpus. The AMI corpus contains both audio and video signals and text transcripts. It also contains a wide range of annotations such as dialogue acts and topic segmentation, named entities, extractive and abstractive summaries and text minutes, which are extremely helpful for our purposes.

The meetings contained in **the ICSI corpus**[4] [23] are for the most part regular meetings of computer science working teams. The corpus contains 70 hours of recordings in English (for 75 meetings collected in Berkeley during the years 2000-2002). The speech files range in length from 17 to 103 minutes and involve from 3 to 10 participants. Interestingly, the corpus contains a significant proportion of non-native English speakers, varying in fluency from nearly-native to challenging-to-transcribe. All audio files are manually transcribed.

The **NIST Meeting Room Corpus** [30] contains 20 hours of meeting recordings (in English) in a room equipped with five cameras. Its analysis in [24] suggests that detected peaks (conversation overlaps, and other changes at different temporal scales) can be useful in summarization and indexing of meetings.

The **ISL Corpus**[5] [9] contains 10 hours of recordings (19 meetings) linked to transcripts from meetings conducted in a special conference room in the Language Technologies Institute at Carnegie Mellon University.

Finally, there is a large number of Parliament and other available political meetings in the official meeting and minutes of European and UK parliaments, Agriculture Dialogue Groups[6] etc. With some data processing, they can be transformed into valuable datasets. There are also data available in other languages, for example, for Czech, there are more than one thousand hours of meetings with available transcripts and minutes from the Czech Parliament[7] and Prague City Hall meetings.[8]

A summary of English meeting corpora can be found in Table 1.

The typology of the meetings which are present in the referred corpora is not addressed in their descriptions very thoroughly. The most detailed information can be found about the scenario meetings in the AMI corpus (ca. 70 hours[9]). The participants of these meetings play the roles of employees in an electronics company developing a new type of television remote control. These are regular face-to-face decision making meetings which have clear goals and the organized structure. The meetings follow the pre-prepared agenda and the manual hand-written minutes are created after the meeting. Among other collected meetings (including the corpus which is being created by the authors of this paper) business and decision making meetings significantly prevail. The meetings however differ in the formality (political meetings are much more formal than the sessions of small business teams). Some of the meetings can be classified as information sharing or status update meetings, the minority can be considered as planning or problem solving ones. The collected data does not include workshops or conferences, nor any team-building meetings. Almost all meetings are multi-party, and non-native English speakers prevail in the data.

4 Methods for Meeting Summarization

The methodologies that are applied to solve meeting summarization problems are numerous and can be seen from different perspectives. In what follows, we observe them from three aspects: focused – unfocused (Section 4.1), extractive – abstractive (Section 4.2) and supervised – unsupervised (Section 4.3).

4.1 Decision-Focused Meeting Summarization

One aspect of meeting or dialogue summaries is their focus. If there is no focus or specific objective the system will try to collect the most relevant utterances to create the summary. Otherwise, if there is a special interest in certain parts like proposed ideas, supporting arguments or decisions, the system will try to first identify those parts of text that form types of utterance and then create the summary.

Literature observations reveal that the interest of most of the focused dialogue summarization works is in the decisions. This emphasizes the fact that decisions are the most essential outputs of meetings, which has indeed been outlined in several works like [5] or [25]. Authors in [16]

[4]http://groups.inf.ed.ac.uk/ami/icsi/
[5]https://ca.talkbank.org/access/ISL.html
[6]https://ec.europa.eu/agriculture/civil-dialogue-groups_en
[7]http://public.psp.cz/en/sqw/hp.sqw
[8]http://www.praha.eu

[9]http://groups.inf.ed.ac.uk/ami/corpus/

model dialogue structure to automatically detect decisions in multi-party meetings. They label each utterance according to the role it plays in decision-making by using SVM (Support Vector Machines) classifier in a hierarchical way. First, sub-classifiers are trained and used to detect the class (e.g., issue, resolution or agreement) of each DDA (Decision Dialogue Act). Later, a super-classifier is utilized to identify decision sub-dialogues. In [15] (follow-up work of the above authors) they parse decision-related dialogue utterances using Gemeini, an open-domain parser described in [13]. The candidate phrases that are generated are further analysed using SVM which filters those that would likely fit in the summary of the discussion.

Authors of [8] solve the problem in two steps. First, they distinguish between dialogue acts that describe the issue and those that describe its resolution. For this, they make use of DGMs (Directed Graphical Models) which according to their experiments outperform hierarchical SVMs when non-lexical features are used. In the second step, they extract words/phrases from the issue and resolution utterances by analysing the later with a semantic parser. The several candidates the parser produces are later processed with a SVM regressor which selects the best. The SVM regression model is enriched with a powerful semantic-similarity feature computed using WordNet as knowledge source.

In [48] we find an unsupervised framework that considers meeting dialogue summarization as an information extraction task. The authors adapt the relation learner of [11] with new features and use it to identify relations between decision cues and decision content of dialogue acts. The content output is a set of indicator-argument decision relations that form the basis of the decision summary. They show that this approach outperforms unsupervised extractive summarization methods and is highly promising.

In [49] authors propose a domain-independent summary generation framework. They first perform content selection using a classifier which identifies potential summary phrases. Next, they employ an overgenerate-and-rank strategy to produce and rank candidate summary sentences. The redundancy reduction process outputs the full meeting summary. Their evaluation reveals that the proposed system outperforms the state-of-the-art supervised extraction-based methods.

4.2 Extractive and Abstractive Summarization Strategies

The utilized text summarization strategy is another way of looking at meeting summarization research works. Summarization methods are generally either extractive or abstractive. Extractive methods only select suitable parts (sentences, words or phrases) from the document or the transcript, while abstractive methods can produce an arbitrary text as the summary. Pure extractive approaches seem very common in the literature.

Authors in [34] focus their entire efforts in the term weighting part which is essential for some of the most important extractive summarization schemes like MMR (Maximal Marginal Relevance). They report that their novel weighting metric ($SU \cdot IDF$) outperforms $TF \cdot IDF$.

In [35] we find another extractive approach that tries to overcome speech recognition errors in meeting transcripts. They try MMR and LSA contrasting them with supervised feature-based approaches that use lexical and prosodic features. They conclude that the feature-based approaches perform worse because of the difficulty to find the best feature collections.

A different extractive approach is the one in [21] where the semantic similarity measures of utterances and the whole dialogues are compared to find out which of the utterances carries important and relevant content for the summary. WordNet is used as a knowledge base for the semantic computations.

An extractive and fully feature-based approach is the work in [32] where authors present a set of generic conversational features for locating the most relevant sentences for the summary. They use linear SVM as classifier and report that their approach is portable in various conversational texts like meetings, emails, etc.

Interactive systems with user feedback such as [31] have also been proposed. The summarizer of this system is conceived as an agent that learns to better identify which relevant utterances to extract by interacting with the user. The advantages it offers are adaptability in different domains and the possibility to work even when a small initial source of data is available.

There are also studies that utilize both extractive and abstractive or neither extractive nor abstractive summarization. In [47] we find a complex framework that starts by clustering all decision-related dialogue acts (DAs). This creates certain clusters for each decision that was made. They later perform DA-level summarization by selecting the most important DAs from each cluster and join them to form a preliminary summary. SVM and LDA are used to further compress at a token-level. Finally, they add discourse context by augmenting the DA clusters of each decision with non-decision related DAs from the dialogue. This way the summary is more abstractive, which makes it concise and readable.

Another work that combines extractive and abstractive approaches for better meeting summaries is [36]. The authors start from human annotated extractive summaries and apply sentence compression to improve the readability. Different sentence compression methods like integer programming [12] or a filler phase detection module and the lexicalized Markov grammar-based approach [19] are explored. Their results indicate that sentence compression is promising for producing abstractive summaries. Similarly, authors in [18] start by finding the most valuable features for identifying and extracting the most informative and relevant DAs. In the second step, they try to in-

crease the abstraction degree of the extractive summaries by including "meta" DAs in which the speakers refer to the meeting itself. They conclude that the "meta" DAs are indeed very helpful and create more coherent and informative meeting summaries.

Authors in [33] compare extractive and abstractive dialogue summaries from a user (reader) perspective and argue that abstractive and concise summaries are usually favored over extractive ones. According to them, a weakness of extractive summaries is that the user does not understand why the extracted phrases are important. They build a summarizer which first maps sentences to a conversation ontology of decisions, action items, sentiments etc. It later identifies message patterns that abstract over several sentences and aggregate them to produce the summary. Authors conduct a user survey which reveals that their automatic summaries are better than the pure extractive ones.

Going in this direction (from extractive to abstractive summaries) some researchers have created fully abstractive systems. They were mostly inspired by similar developments in close tasks such as text summarization of news articles where the power the encoder-decoder framework based on RNNs is utilized [44, 37, 45]. In [4] they first split meeting dialogues into several topic segments. The most important phrases in each segment are identified using a classifier and merged to form a one-sentence summary. The dependency parses of each segment are combined to form a directed graph. ILP (Integer Linear Programming) is used to select the most informative sub-graph and produce the one-sentence summary of each topic segment, reaching to the summary of the entire meeting.

An even more complete pipeline is presented in [29] where they cluster the sentences and create an entailment graph which selects the most relevant sentences in each cluster. They further build a word-graph model by extending that of [17] and use a ranking strategy to select the best paths in it, compressing and aggregating the selected sentences. Authors report that their approach is able to generate long sentences with little loss in grammaticality. In [20] we find another attempt to improve abstractive summarization of dialogues, this time by integrating interactive parts into the summary. They propose a sentence-gated mechanism which models the relationships between the dialogue acts and the summary. Their benchmarks with AMI meeting corpus reveal that the system outperforms the other models.

A different approach for improving abstractive meeting summaries is the one in [39] where templates are learned from human-authored summaries. A clustering sentence fusion algorithm and WordNet semantic similarities between words are used to generate templates. The meeting transcripts are segmented based on topics and the best templates for each topic are selected using the relationship between the human summaries and their sources. The evaluation shows that their system summaries are favored over human-annotated extractive ones.

4.3 Supervised and Unsupervised Summarization Methods

Data-driven machine learning models (supervised, unsupervised, both, etc.) are widespread today, even in studies about summarization of meeting dialogues. The type of machine learning approach they utilize is another way to look at these studies. It is typical to find unsupervised methods (clustering) in the initial step of a pipeline or complex system. Typical examples of this category are [39], [27], and [47]. Other frequent forms of unsupervised approaches are MMR and LSA which are based on similarity scores or the dependency graph of [3]. It is also interesting to find recent works that are unsupervised (no need for labeled data), but still produce gramatically correct summaries. One such example is [46] where they combine the strengths of various graph-based methods like the neural network sentence compression of [17], graph path reranking of [7], graph entailment of [29], etc. Authors evaluate on both AMI and ICSI datasets and report state-of-the-art results.

Supervised learning as a part of the system is even more common. SVM is clearly the most frequent algorithm followed by Naïve Bayes and maximum entropy classifier. There are even studies like [26] and [16] that perform hierarchical classification, with sub-classifiers that identify categories of different utterances and a super-classifier that produces the final summary. There are also studies like [8] where both unsupervised (directed graph and semantic similarity measures) and supervised (SVM) are combined together. Finally, among the most recent supervised approaches based on neural networks, we can mention [38] which fused verbal and non-verbal information to predict the importance of each uterance. Authors utilize MATRICS multimodal discussion corpus dataset which contains group discussions with various annotations and features (speech spectrogram, head motion spectrogram, head pose, and more). At the end, they use a multi-channel neural network architecture based on CNNs and dense layers to fuse together all types of features and predict the importance of the utterances.

5 Discussion

Let us now summarize the knowledge obtained by the survey and describe our first steps towards the creation of automatic minuting.

The types of existing meetings, minutes and datasets show a significant disproportion between the real meetings and the datasets which are available for the research. The available data are mostly in the political domain, whereas, in reality, business meetings prevail and these are also the very meetings for which the automation of minutes would bring the most benefit. For these reasons, we decided to

go beyond the political domain and focus on other types of meetings as well, first of all on international online meetings. Our goal is thus to arrange the meeting types in a way which explains the types of agendas and minutes applicable to them. From this point of view, we consider the meeting intention as the most appropriate scale in the multidimensional space of meeting types. Thus, business and decision making meetings are most structured and they require the most clear, structured and detailed agenda. The minutes are supposed to contain a list of decisions. For information sharing and status update meetings, the agenda is also extremely important. In this case, the minutes will be a refinement of ideas given in the agenda. The situation is slightly different for planning and problem solving meetings, as a number of new ideas may arise during the meeting. Innovation and idea generation meetings are creative and can include a lot of irrelevant brainstorming. During the note-taking it may not be clear what will be important in the result. The minutes do not follow the agenda consistently. Team-building meetings and other social events follow different rules, and minutes (if any) have rather different functions. Naturally, we focus on the meetings which demand for agenda and minutes. For this reason, for example, team-building meetings will not be included in our research.

Concerning the form of automatic minutes, we incline to create structured automatic notes of meetings by which the actions are fixed by the bullet system rather than detailed narrative minutes. According to the survey of meeting minutes presented in Section 2.2, for some meetings, special fields for actions or resolutions may be applied. As for linguistic form of the minutes, it will be defined by the meetings themselves, as we will primarily use extractive summarization methods.

As for the datasets, we are primarily interested in the meeting corpora which include both meeting transcripts and minutes (or other types of summarizations). Among the datasets described in Section 3, this is the AMI corpus, which will be used in our experiments as first. Other datasets with the minutes are parlamental texts in the political domain.

As most international online meetings, which we available to us, are held in English, we choose English as the main language for creating automatic minuting experiments. However, we plan to include other languages as well.[10]

6 First Steps Towards Automatic Minuting

The available meeting corpora such as those described in Section 3 can serve as a good starting point for making experiments in automatic minuting, but they do not suffice. For example, the AMI corpus (which is the most appropriate for us) is relatively small and mostly includes meetings which have been played by the actors. This fact may have significant effect on what the people say and how they act in the conversation. Therefore, we decided to extend it with our own data.

Within our project ELITR[11], we started collecting meetings of our computer science working teams. The data includes the audio recordings, ASR transcripts, pre-prepared agendas and meeting minutes created by the meeting organizers or a secretary after the meeting. For the time being, we obtained ca. 40 hours of meetings in English (of mostly non-native speakers). The corpus is under development.

7 Conclusion

In this paper, we laid out foundations for research into automatic minuting of meetings. Our main goal was to prepare the floor for automatic minuting by analyzing the sources which help to make this idea realizable. By comparing a variety of meetings and their descriptions, we tried to get a reasonable typology of meetings, summarized the types of possible minutes, described the meetings datasets and made a survey of methods of meeting summarization. We also drafted our first steps to the creation of the corpus of meetings and minutes which will be further used for developing automatic minuting.

Acknowledgement

References

[1] A.M. Attorney. *The Corporate Records Handbook: Meetings, Minutes and Resolutions*. dElta printing solutions, inc., 2007.

[2] A.M. Attorney. *Nonprofit meetings, minutes and records*. Delta printing solutions, inc., 2008.

[3] S. Banerjee, P. Mitra, and K. Sugiyama. Abstractive meeting summarization using dependency graph fusion. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, pages 5–6, New York, NY, USA, 2015. ACM.

[4] S. Banerjee, P. Mitra, and K. Sugiyama. Generating abstractive summaries from meeting transcripts. In *Proceedings of the 2015 ACM Symposium on Document Engineering*, DocEng '15, pages 51–60, New York, NY, USA, 2015. ACM.

[5] S. Banerjee, C. Rose, and A.I. Rudnicky. The necessity of a meeting recording and playback system, and the benefit of topic–level annotations to meeting browsing. In M.F.

---

[10]For example, we began to collect the meeting corpus for Czech, but, for the time being, its size is insignificant.

[11]http://elitr.eu/

Costabile and F. Paternò, editors, *Human-Computer Interaction - INTERACT 2005*, pages 643–656, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[6] F. Bargiela and S.J. Harris. *Managing Language: The discourse of corporate meetings*. John Benjamins Publishing Company, 1997.

[7] F. Boudin and E. Morin. Keyphrase extraction for n-best reranking in multi-sentence compression. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 298–305, Atlanta, Georgia, June 2013. Association for Computational Linguistics.

[8] T. Bui, M. Frampton, J. Dowding, and S. Peters. Extracting decisions from multi-party dialogue using directed graphical models and semantic similarity. In *Proceedings of the SIGDIAL 2009 Conference*, pages 235–243. Association for Computational Linguistics, 2009.

[9] Susanne Burger, Victoria MacLaren, and Hua Yu. The isl meeting corpus: The impact of meeting type on speech style. 01 2002.

[10] S.E. Butler. *Mission Critical Meetings: 81 Practical Facilitation Techniques*. Wheatmark, 2014.

[11] H. Chen, E. Benson, T. Naseem, and R. Barzilay. In-domain relation discovery with meta-constraints via posterior regularization. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 530–540, Portland, Oregon, USA, jun 2011. Association for Computational Linguistics.

[12] J. Clarke. Global inference for sentence compression: An integer linear programming approach. *J. Artif. Intell. Res.*, 31:399–429, 2008.

[13] J. Dowding, J.M. Gawron, D. Appelt, J. Bear, L. Cherny, R. Moore, and D. Moran. Gemini: A natural language system for spoken-language understanding. In *HUMAN LANGUAGE TECHNOLOGY: Proceedings of a Workshop Held at Plainsboro, New Jersey, March 21-24, 1993*, 1993.

[14] L. Espy. *Bad Meetings Happen to Good People: How to Run Meetings That Are Effective, Focused, and Produce Results*. Blue Room Press, Memphis, Tenessee, 2017.

[15] R. Fernández, M. Frampton, J. Dowding, A. Adukuzhiyil, P. Ehlen, and S. Peters. Identifying relevant phrases to summarize decisions in spoken meetings. In *Ninth Annual Conference of the International Speech Communication Association*, 2008.

[16] R. Fernández, M. Frampton, P. Ehlen, M. Purver, and S. Peters. Modelling and detecting decisions in multi-party dialogue. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, SIGdial '08, pages 156–163, Stroudsburg, PA, USA, 2008. Association for Computational Linguistics.

[17] K. Filippova. Multi-sentence compression: Finding shortest paths in word graphs. In *Proceedings of the 23rd International Conference on Computational Linguistics*, COLING '10, pages 322–330, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[18] M. Frampton, J. Huang, T. Bui, and S. Peters. Real-time decision detection in multi-party dialogue. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 1133–1141. Association for Computational Linguistics, 2009.

[19] M. Galley and K. McKeown. Lexicalized Markov grammars for sentence compression. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference*, pages 180–187, Rochester, New York, April 2007. Association for Computational Linguistics.

[20] Ch. Goo and Y. Chen. Abstractive dialogue summarization with sentence-gated modeling optimized by dialogue acts. *2018 IEEE Spoken Language Technology Workshop (SLT)*, pages 735–742, 2018.

[21] I. Gurevych and M. Strube. Semantic similarity applied to spoken dialogue summarization. In *Proceedings of the 20th International Conference on Computational Linguistics*, COLING '04, Stroudsburg, PA, USA, 2004. Association for Computational Linguistics.

[22] K.M. Hoffman. *Meeting design: For Managers, Makers and Everyone*. Louis Rosenfeld, 2018.

[23] 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. pages 364–367, 2003.

[24] R. Krishnan and S. Sarkar. Detecting Group Turn Patterns in Conversations Using Audio-Video Change Scale-Space. In *20th International Conference on Pattern Recognition, ICPR 2010, Istanbul, Turkey, 23-26 August 2010*, pages 137–140, 2010.

[25] A. Lisowska, A. Popescu-Belis, and S. Armstrong. *User Query Analysis for the Specification and Evaluation of a Dialogue Processing and Retrieval System*, pages 993–996. LREC 2004 (Fourth International Conference on Language Resources and Evaluation). ELRA - European Language Ressources Association, 2004. ID: unige:2264.

[26] F. Liu and Y. Liu. From extractive to abstractive meeting summaries: Can it be done by sentence compression? In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, ACLShort '09, pages 261–264, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.

[27] Y. Liu, S. Xie, and F. Liu. Using n-best recognition output for extractive summarization and keyword extraction in meeting speech. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5310–5313, March 2010.

[28] I. Mccowan, G. Lathoud, M. Lincoln, A. Lisowska, W. Post, D. Reidsma, and P. Wellner. The ami meeting corpus. In *In: Proceedings Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research. L.P.J.J. Noldus, F. Grieco, L.W.S. Loijens and P.H. Zimmerman (Eds.), Wageningen: Noldus Information Technology*, 2005.

[29] Y. Mehdad, G. Carenini, F. Tompa, and Raymond T. NG. Abstractive meeting summarization with entailment and fusion. In *Proceedings of the 14th European Workshop on Natural Language Generation*, pages 136–146. Association for Computational Linguistics, 2013.

[30] M. Michel, J. Ajot, and J.G. Fiscus. The NIST Meeting Room Corpus 2 Phase 1. In *Machine Learning for Mul-*

timodal Interaction, Third International Workshop, MLMI 2006, Bethesda, MD, USA, May 1-4, 2006, Revised Selected Papers, pages 13–23, 2006.

[31] M. Mieskes, Ch. Müller, and M. Strube. Improving extractive dialogue summarization by utilizing human feedback. In *Artificial Intelligence and Applications*, 2007.

[32] G. Murray and G. Carenini. Summarizing spoken and written conversations. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '08, pages 773–782, Stroudsburg, PA, USA, 2008. Association for Computational Linguistics.

[33] G. Murray, G. Carenini, and R. Ng. Generating and validating abstracts of meeting conversations: A user study. In *Proceedings of the 6th International Natural Language Generation Conference*, INLG '10, pages 105–113, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[34] G. Murray and S. Renals. Term-weighting for summarization of multi-party spoken dialogues. In *Proceedings of the 4th International Conference on Machine Learning for Multimodal Interaction*, MLMI'07, pages 156–167, Berlin, Heidelberg, 2008. Springer-Verlag.

[35] G. Murray, S. Renals, and J. Carletta. Extractive summarization of meeting recordings. In *INTERSPEECH 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology, Lisbon, Portugal, September 4-8, 2005*, pages 593–596, 2005.

[36] G. Murray and Steve Renals. Meta comments for summarizing meeting speech. In Andrei Popescu-Belis and Rainer Stiefelhagen, editors, *Machine Learning for Multimodal Interaction*, pages 236–247, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[37] R. Nallapati, B. Xiang, and B. Zhou. Sequence-to-sequence rnns for text summarization. *CoRR*, abs/1602.06023, 2016.

[38] F. Nihei, Y.I. Nakano, and Y. Takase. Fusing verbal and nonverbal information for extractive meeting summarization. In *Proceedings of the Group Interaction Frontiers in Technology*, GIFT'18, pages 9:1–9:9, New York, NY, USA, 2018. ACM.

[39] T. Oya, Y. Mehdad, G. Carenini, and R. Ng. A template-based abstractive meeting summarization: Leveraging summary and source text relationships. In *Proceedings of the 8th International Natural Language Generation Conference (INLG)*, pages 45–53, 2014.

[40] T. Rigley. *The techniques of writing minutes*. Singapore : Federal Publications, 1996.

[41] H.M. Robert. *Robert's Rules of Order: The Original Edition*. S. C. Griggs and Company, 1876.

[42] H.M. Robert, S.C. Robert, H.M. Robert III, W.J. Evans, D.H. Honemann, and T.J. Balch. *Robert's Rules of Order Newly Revised*. Da Capo Press, A Member of the Perseus Books Group, 2011.

[43] S. Rogelberg. *The Surprising Science of Meetings*. Oxford University Press, 2019.

[44] A. M. Rush, S. Chopra, and J. Weston. A neural attention model for abstractive sentence summarization. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 379–389, Lisbon, Portugal, September 2015. Association for Computational Lin-

guistics.

[45] A. See, P.J. Liu, and C. D. Manning. Get to the point: Summarization with pointer-generator networks. *CoRR*, abs/1704.04368, 2017.

[46] G. Shang, W. Ding, Z. Zhang, A. Tixier, P. Meladianos, M. Vazirgiannis, and J.-P. Lorré. Unsupervised abstractive meeting summarization with multi-sentence compression and budgeted submodular maximization. *CoRR*, abs/1805.05271, 2018.

[47] L. Wang and C. Cardie. Summarizing decisions in spoken meetings. In *Proceedings of the Workshop on Automatic Summarization for Different Genres, Media, and Languages*, WASDGML '11, pages 16–24, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.

[48] L. Wang and C. Cardie. Focused meeting summarization via unsupervised relation extraction. In *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, SIGDIAL '12, pages 304–313, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.

[49] L. Wang and C. Cardie. Domain-independent abstract generation for focused meeting summarization. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1395–1405. Association for Computational Linguistics, 2013.