

2015年度 修士論文

発話間関係の構造化による会議録からの
議論マップ自動生成システム

提出日: 2016年02月01日

指導教授: 山名 早人 教授

早稲田大学 基幹理工学研究科
情報理工・情報通信専攻

学籍番号: 5114F074-1

林 佑磨

概要

組織では、活動の現状を共有し、今後の方針や重要事項を決定するために会議を行う。通常、会議の時間は限られており、しばしば議題が持ち越される。そのため、参加者が過去の会議内容を事前に把握することは、会議を円滑に進める上で重要である。そのような背景から、会議録の自動要約や閲覧システムに関する研究が行われている。しかし、自動要約では発話間の関係が構造化されておらず、閲覧システムでは会議録全体が対象であるため、いずれも会議の流れや決定事項などの要旨が掴みづらいという問題がある。そこで本研究では、1)会議録からの重要発話抽出を行い、2)発話間における質疑応答の対応付けを行った後に、3)発話間の構造を保持した議論マップとして出力することで、会議要旨の素早い把握を可能にする統合システムを提案する。発話間の構造化により、重要発話抽出では既存手法のMMR-centroidと比べROUGE-1が最大0.111向上した。また、4名の被験者実験により、プレーンテキストによる出力と比較して、議論マップによる出力の有用性が示された。

目次

第1章	はじめに	1
第2章	関連研究	3
2.1	会議録の自動要約に関する研究	3
2.1.1	抽出型要約に関する研究	3
2.1.2	生成型要約に関する研究	7
2.2	テキストの構造化による可視化に関する研究	7
2.3	関連研究のまとめ	8
第3章	発話間関係の構造化による会議録からの 議論マップ自動生成システム	11
3.1	システムの概要	11
3.2	対象とする会議録データセットについて	13
3.3	システムで用いる要素技術	16
3.3.1	CW (Confidence-Weighted)	16
3.3.2	AROW (Adaptive Regularization of Weight Vectors)	17
3.4	システムの詳細	18
3.4.1	前処理	18
3.4.2	重要発話抽出	19
3.4.3	質疑応答の対応付けと重要発話の補完	20
3.4.4	議論マップによる可視化	22
第4章	実験と評価	24
4.1	会議録からの重要発話抽出に関するROUGEによる定量的評価	24
4.2	被験者実験による議論マップの有用性評価	26
第5章	おわりに	29

研究業績	31
参考文献	32
付録	35
A.1 被験者実験における質問と出力（プレーンテキストと議論マップの比較） . .	35
A.2 被験者実験における質問と出力（重要フレーズの強調と構造化の比較） . . .	43
A.3 <i>t</i> 分布表	51

図目次

図2.1.1	3種類のサンプリング ([8]のFig. 2を元に作成)	6
図3.1.1	議論マップ自動生成システムの処理フロー	11
図3.4.1	提案システムにより“ES2004c”から自動生成を行った議論マップの出力	23

表目次

表2.3.1	会議録の自動研究に関する研究のまとめ	9
表2.3.2	テキストの構造化による可視化に関する研究のまとめ	9
表3.2.1	AMI Meeting Corpusのシナリオ有り会議における参加者の役一覧	13
表3.2.2	AMI Meeting CorpusのDialogue Act一覧	15
表3.4.1	Penn TreeBankの品詞記号と意味	19
表3.4.2	AROWモデルの構築に用いる素性一覧	20
表4.1.1	比較対象の重要発話抽出手法	25
表4.1.2	ROUGE-1による重要発話抽出の評価	25
表4.2.1	被験者実験による議論マップの有用性評価	26
表4.2.2	検定で用いるデータ	27
表4.2.3	被験者実験による議論マップ内の各工夫の影響度評価	28

擬似コード目次

擬似コード3.3.1 AROWによるオンライン識別アルゴリズム ([23]より引用)	18
擬似コード3.4.2 質問発話推定	21
擬似コード3.4.3 回答発話抽出	21
擬似コード3.4.4 回答発話候補の推定	22

第1章 はじめに

企業や学術機関などの組織では、現状の共有を行い、今後の方針や重要事項を決定するために会議を行う。多くの場合、会議の時間は限られており、しばしば複数回に渡り議題が持ち越される。そのため、多くの会議では会議録の作成や会話の録音などを行い、議論内容や決定事項の共有を目的に、参加者への配布を行う。参加者が会議出席前に過去の会議内容を把握しておくことは、会議を円滑に進める上で重要である。しかし、会議中の発話を書き起こした会議録は膨大な量であり、全てを読むには多くの時間を要する。

そのような背景から、ICSI Meeting Corpus¹[1]やAMI Meeting Corpus²[2]といった、会議の円滑な進行や会議録の効率的な閲覧に関する研究の支援を目的とした、人手により様々なアノテーション情報が付与された会議録データが提供されている。人手で付与されたアノテーション情報を機械的に付与する、という研究も同時に行われているため、多くの会議録を対象とした研究では、会議録データに付属のアノテーション情報を効果的に用いている。

これらのデータセットを利用した会議録の自動要約に関する研究は、1)重要な発話を取り出す抽出型の要約手法[3][4][5][6][7][8][9][10]や、2)発話中の冗長部分削除や複数の発話をまとめる生成型の要約手法[11][12][13]など、今までに多く行われている。これらの研究では、読むべき対象となる発話の量を大幅に削減することが可能であるが、発話間の関係が構造化されておらず、質疑応答などの対応する発話の紐付けは読み手が行わなければならないという問題がある。

一方で、構文解析や意味解析を応用することで、文内および文間の構造化を行い、文章からのマインドマップ自動生成を行うような研究[14][15][16][17]が存在する。しかし、会議中の発話は自然発生的なものであるため、言い間違いなどによる不自然な単語の連続や構文的に正しくない言い回しなどが多く含まれている。このため、会議録中の発話に対しては構文解析がうまく機能せず、これらの手法の応用は困難である。また、会議内の発話を対象とした、構文解析を用いない会議内容の効率的な閲覧システムの提案[18][19][20][21]も行われているが、これらの多くは閲覧すべき発話の選定および構造化が行われておらず、会議の流れや決定事項などの要旨が掴みづらいという問題がある。

¹ICSI Meeting Corpus: <http://www1.icsi.berkeley.edu/Speech/mr/>

²AMI Meeting Corpus: <http://groups.inf.ed.ac.uk/ami/corpus/>

そこで本研究では、AMI Meeting Corpus[2]を対象に、1)発話間の関係が構造化されていない、2)確認すべき発話が限定されていない、という従来の問題点を解決し、会議の流れや決定事項などの要旨を効率良く把握することを可能にする統合システムを提案する。具体的には、以下の3つの手順により実現する。

提案システムの手順

1. 会議録からの重要発話抽出

確認対象となる議論内容を限定し、2)の問題を解決する

2. 発話間関係の構造化

重要発話内にある質問発話を特定し、元の会議録からその回答を紐付け構造化を行うことで、1)の問題を解決する

3. 議論マップによる出力

回答が補充された重要発話を、発話間の構造を保持した議論マップとして出力することで、1)と2)の問題を解決する

本稿の構成は以下の通りである。まず第2章において関連研究について説明し、第3章で提案システムに関する詳細を説明する。第4章では重要発話抽出に関する定量的評価、および被験者実験による議論マップの有用性評価とその結果を示す。第5章でまとめ、今後の課題について述べる。

第2章 関連研究

本章では、関連研究について述べる。まず2.1節で、会議録を対象とする自動要約技術について説明を行う。続く2.2節で、テキストを構造化することにより、記載内容を効率的に把握することを目指した、可視化に関する研究の説明を行う。2.3節で、関連研究のまとめを行う。

2.1 会議録の自動要約に関する研究

本節では、会議録を対象とする自動要約技術に関する先行研究の説明を行う。会議録の自動要約は、会議録の閲覧にかかる時間を短縮し会議の流れや決定事項などの重要項目を効率よく把握できるようにすることが目的である。自動要約のアプローチには大きく、1) 重要な発話を会議録中から取り出す抽出型の要約と、2) 人が要約を行う際のように、会議録に含まれないような文を新たに生成する生成型の要約がある。以下では2.1.1項で抽出型要約について、2.1.2項で生成型要約について、それぞれ研究を紹介する。

2.1.1 抽出型要約に関する研究

既存の自動要約に関する研究の大部分が抽出型要約によるものである。抽出型の要約では、会議録中から重要であると思われる発話を抜き出し、それらを並べたものを要約とする。従って、抽出型要約の目標は、議論内容における重要箇所の網羅的な抽出、および同一内容の文を含めないよう冗長性を最小化する、ということになる。この観点に基づく問題の定式化を行った研究として、以下では、会議中のどの発話が重要な発話かという事前知識を利用しないで抽出を行う「教師なし学習」による研究と、事前知識を利用して重要発話の抽出モデルを構築し、そのモデルを用いて抽出を行う「教師あり学習」による研究をそれぞれ説明する。

Gillickらの研究

Gillickら[4]は、2009年に教師なし学習の枠組みであるILP (Integer Linear Programming) を利用して、AMI meeting corpus¹[2]に含まれる会議録からの重要発話抽出を行った。Gillickらは、会議録中の発話に含まれる会議内容を特徴づけるような形態素あるいは形態素列をその会議のconceptとして定義し、なるべく広くconceptを含むように以下の式(2.1.1)としてILPの定式化を行い重要発話抽出を行う。

$$\text{Maximize : } \sum_i c_i w_i \quad (2.1.1)$$

$$\text{Subject to : } \sum_j l_j u_j < L \quad (2.1.2)$$

ここで、 c_i は*i*番目のconceptが要約結果に含まれれば1、そうでなければ0を取る2値変数であり、 w_i は*i*番目のconceptに対する重みである。また、 l_j は発話*j*の長さであり、 L は最終的な要約の長さである。なお、conceptになりうる形態素列、および重み w_i は以下として定義される。

- conceptの定義 ([4]より引用)
 1. 全ての形態素n-gram g_i を取り出す。(ただし、 $1 \leq n \leq 3$)
 2. g_i に含まれる全形態素の品詞が、1)数詞、2)外来語、3)形容詞、4)名詞のいずれかと適合する場合、concept候補とする。
 3. “manager” が “dialogue manager” に含まれるように、他のconcept候補に含まれる形態素列、および出現頻度が1のconcept候補を取り除き、残ったものをconceptとする。
- *i*番目のconceptに対する重み w_i の定義 ([4]より引用)

$$w_i = f(g_i) \cdot n \quad (2.1.3)$$

ただし、 n はそのconceptのn-gram数であり、 $f(g_i)$ は会議録における*i*番目のconceptの出現頻度である。

上記の手法により、自動要約の評価においてしばしば用いられるROUGE[22]で、それまでの一般的な文書の要約で用いられていた手法を用いた場合と比較して高いスコアを出す

¹AMI Meeting Corpus: <http://groups.inf.ed.ac.uk/ami/corpus/>

ことに成功した。しかしGillickらは、会議録中の発話構造や発話間の対応を考慮していない点や、発話の抽出による要約が人手で作成された要約とは本質的に大きく異なる点を課題として挙げている。

Xieらの研究

また、Xieら[8]は、2010年に教師あり学習の枠組みであるSVM (Support Vector Machine) およびSVR (Support Vector Regression) を利用して、ICSI meeting corpus²[1]に含まれる会議録からの重要発話抽出を行った。Xieらの研究における主な貢献は以下の2点である。

- 重要発話抽出のモデル学習時にサンプリングを適用した点
- 二値分類問題ではなく回帰問題として重要発話抽出問題を解いた点

前者に関して、重要発話としての正解ラベルが付与された発話は会議録全体に対して相対的に少ない。そこでXieらは、重要発話抽出を行うための教師あり学習モデルを構築する際に、異なる3種類のサンプリングを適用することで、重要発話を効率的に学習するという効率的な学習方法の提案を行った。具体的に適用されたサンプリングは以下である。

1. Up-sampling

正例（正解ラベルが付与された発話）と間違えてしまいやすい負例（正解ラベルが付与されていない発話）を正例へと変換することで、全体に占める正例の割合を増やす。それに伴い識別境界が変動する。（図2.1.1aは元の状態であり、Up-sampling適用後は図2.1.1bの状態となる。）

2. Down-sampling

識別境界の近くにある負例を取り除くことで、全体に占める負例の割合を減らす。それに伴い識別境界が変動する。（Down-sampling適用後は図2.1.1cの状態となる。）

3. Re-sampling

各発話と元の会議録全体との類似度を計算し、そのスコアが高いものを残すようにサンプリングを行う。

²ICSI Meeting Corpus: <http://www1.icsi.berkeley.edu/Speech/mr/>

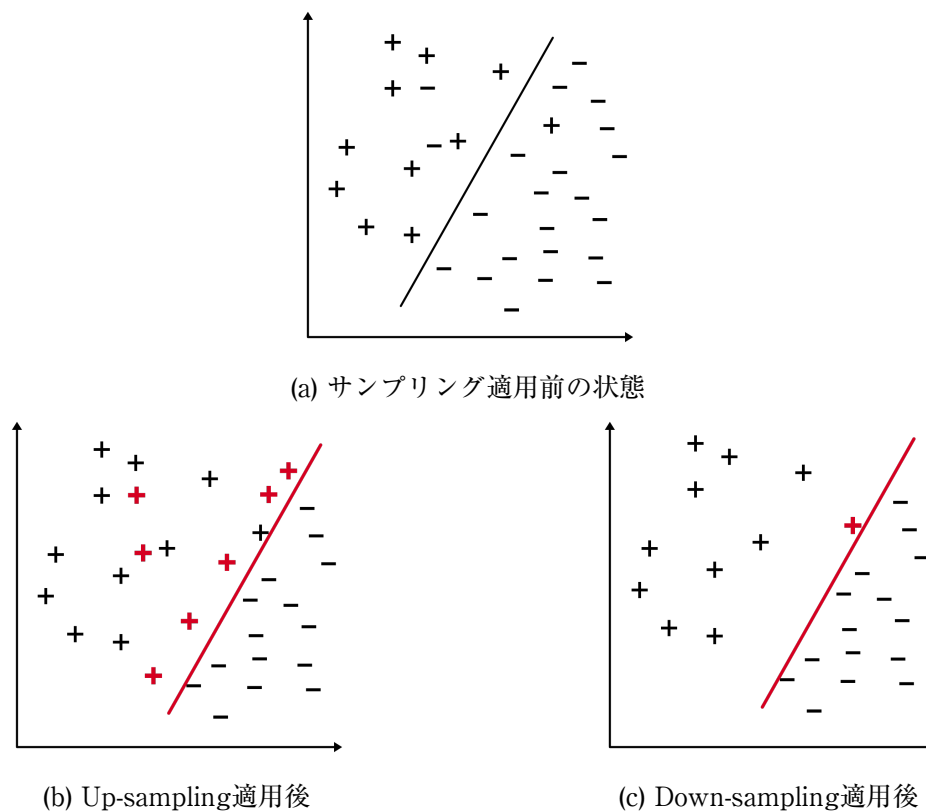


図 2.1.1: 3種類のサンプリング ([8]のFig. 2を元に作成)

なお、Xieらはモデルの構築時に用いる特徴量として、音声・音響の情報や韻律的な情報を用いておらず、テキスト上の情報のみを用いている。

また後者に関して、会議録中の重要発話は、それを選定した人により選定基準がしばしば異なる。そのため、ある人が重要発話として選定した発話が、他の人によって重要ではないとラベル付けされる問題があり、これは二値分類器の学習時に大きな問題となる。この問題に対してXieらは、重要／非重要の二値ではなく、重要合いを設けることで二値分類問題から回帰問題に帰着して解決を試みた。この際、SVMではなくSVR (Support Vector Regression) を用いて実験を行い、精度が向上するケースがあることを示した。

Xieらは、サンプリングや回帰による解き方に有効な素性の特定や、会議内容を閲覧できるブラウザの様な実際のアプリケーションへ応用した先での評価を課題として挙げている。

2.1.2 生成型要約に関する研究

近年になり，元の会議録に含まれない文を創り出すことで，抽出型要約と比べてより簡潔な要約文を生成することを目的とした生成型要約が行われるようになってきた．以下では，生成型要約による会議録の自動要約に関する研究を説明する．

Mehdadら[12]は，2013年に単語グラフを利用した複数発話の集約による文生成を行い，入力としての会議録を受け取ってから，出力としての要約を生成するまでの全過程を扱うようなend-to-endの生成型要約フレームワークを提案した．AMI Meeting Corpusを対象に実験を行い，既存の抽出型要約手法と比較して高いROUGEスコアを得た．

また，Murrayら[13]は，2015年にマルコフ決定過程（MDP: Markov Decision Process）を用いた，生成型要約手法を提案した．Murrayらの提案したMDPを用いた生成型要約手法は，特に会議録に特化した手法ではなく，一般の文書への利用を視野に入れたものである．実験においては，Mehdadらと同様にAMI Meeting Corpusの会議録データへの適応を行い，ROUGE評価の観点において有効性を示した．

しかし，いずれの生成型要約の手法においても，大きな問題として，構文的に不自然な文や意味を成さないような文を多く生成してしまう，という点が挙げられている．

2.2 テキストの構造化による可視化に関する研究

テキスト情報のみからでは，そこに含まれる文の構造が分からないという問題がある．そこで，テキストを構造化し可視化を行うことで，それらの情報を補完するという研究が存在する．

Adbeenら[14]は，2009年に純粋なテキストに対し構文解析や意味解析を適応することで，マインドマップを自動生成する最初のシステムを提案した．また2013年には，Purwariantiら[16]が，インドネシア語に特化したテキストの構造解析を用いてマインドマップの自動生成を行った．

Elhoseinyら[17]は，2015年にそれまでに提案されていたテキストからのマインドマップ生成手法を拡張し，階層的な表示が可能なマインドマップの自動生成を行うシステムを提案した．

しかし，これらの研究はいずれも，構文的な解析が可能な書き言葉のテキストを対象としている．そのため，これらの手法を構文的な解析がうまく機能しないような，会議録中

の発話を対象として適用することは不可能である。

一方で、構文的な解析を用いない、会議録を対象とする効率的な閲覧システムに関する研究が存在する。松村ら[20]は、2003年に議事録からトピックのセグメント分割を行い、セグメント単位での構造化を行うことで構造化されたマップの作成を行った。構造化においては、セグメント内の発話に含まれる単語を利用してセグメント間の類似度を取り、ある閾値以上であればエッジを張るという方法により行っている。

趙ら[19]は、2006年に議論の中で展開されるテーマは名詞の集合で表現できるという過程の元で、議事録内に含まれる名詞をノードとし、関係のあるノード間にエッジを貼った議論マップの自動生成を行った。また、森ら[21]は、2007年にリフレクションのための、発話間の関係に着目した議論構造モデルの提案と、そのモデルに基づく議事録の自動構造化手法の提案を行った。

これらの研究ではそれぞれ、1)一つのトピックセグメントに含まれる文量が多すぎる点や、2)議論の流れに沿って正しく可視化しないと余計に見づらくなる点、3)ノードが名詞であるため分かりづらい点などが問題に挙げられている。これらの問題は、議事録に含まれる発話全体が対象であるために生じている。

2.3 関連研究のまとめ

本節では、2.1節、および2.2節において説明を行った研究のまとめを行う。会議録の自動研究に関する研究のまとめを表2.3.1に、テキストの構造化による可視化に関する研究のまとめを表2.3.2にそれぞれ示す。

表 2.3.1: 会議録の自動研究に関する研究のまとめ

要約方法		関連研究	年	手法の特徴	課題点
抽出型	教師なし	Gillickら[4]	2009	なるべく広く concept を含むよう ILP を用いて定式化	発話構造や発話間の対応が考慮されていない
	教師あり	Xieら[8]	2010	モデル学習時のサンプリング適用と、二値分類から回帰問題への帰着	
生成型		Mehdadら[12]	2013	単語グラフを応用した end-to-end の生成型要約フレームワーク	構文的に不自然な文や意味を成さないような文を多く生成してしまう
		Murrayら[13]	2015	マルコフ決定過程を用いた生成型要約手法の会議録テキストへの適応	

表 2.3.2: テキストの構造化による可視化に関する研究のまとめ

対象文書	関連研究	年	手法の特徴	課題点
書き言葉	Adbeenら[14]	2009	テキストからマインドマップを自動生成する最初のシステム	構文的な解析がうまく機能しない、会議録中の発話への適用が不可能
	Purwariantiら[16]	2013	インドネシア語に特化したマインドマップ自動生成	
	Elhoseinyら[17]	2015	階層的なマインドマップの自動生成	
会議録	松村ら[20]	2003	トピックのセグメント分割とセグメント間関係の構造化による議論構造の可視化	議事録に含まれる全ての発話が確認すべき対象となっている
	趙ら[19]	2006	名詞をノードとした議論マップの自動生成	
	森ら[21]	2007	議論内容想起のための議論構造モデルと自動化手法の提案	

表2.3.1および表2.3.2より，従来の研究においては，1)発話間の関係が構造化されていない，2)確認すべき発話が限定されていない，という2つの大きな課題が存在する．本稿で提案するシステムでは，以下に示す3つのアプローチにより，これらの課題点を解決する．

提案システムのアプローチ

1. 会議録からの重要発話抽出

確認対象となる議論内容を限定し、2)の問題を解決する

2. 発話間関係の構造化

重要発話内にある質問発話を特定し、元の会議録からその回答を紐付け構造化を行うことで、1)の問題を解決する

3. 議論マップによる出力

回答が補充された重要発話を、発話間の構造を保持した議論マップとして出力することで、1)と2)の問題を解決する

第3章 発話間関係の構造化による会議録からの 議論マップ自動生成システム

本章では、発話間関係の構造化による会議録からの議論マップ自動生成システムを提案する。まず3.1節で、提案システムの概要の説明を行う。続く3.2節で、対象とする会議録のデータセットに関する説明を行い、3.3節で、提案システム内部で用いる要素技術に関する説明を行う。最後に3.4節で、提案システムの内部で行う各処理に関する詳細の説明を行う。

3.1 システムの概要

本節では、会議録からの議論マップ自動生成システムについての概要を説明する。本システムは、会議録中から重要発話を抽出し、発話間の質疑応答関係を構造化した上で、構造を保持した議論マップの自動生成を行う統合的なシステムである。ここで、議論マップとは、1)会議全体の流れ、2)会議中での重要な発話、3)会議中の重要な質問に対する回答の紐付け、をそれぞれ構造化し表現したものである。図3.1.1にシステム全体の処理フローを示す。

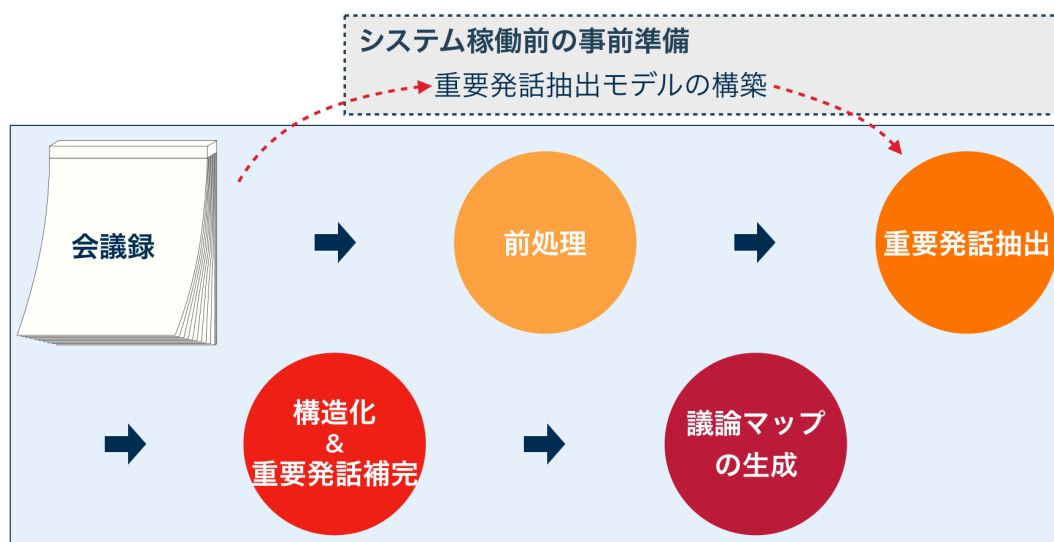


図 3.1.1: 議論マップ自動生成システムの処理フロー

本システムは、まず入力としての会議録データを受け取った後、1)前処理、2)重要発話抽出、3)質疑応答の対応付けと重要発話の補完、4)議論マップとして可視化、という大きく4つの処理を行い、出力としての議論マップを生成する。システムが内部で行う4つの主な処理に関する大まかな内容は以下である。

システム内部で行う主な処理

処理 1. 前処理

Stopwords除去や言い直しの除去を行い、発話内の重要フレーズの特定を行う。

処理 2. 重要発話抽出

事前に構築した、重要発話識別のための教師ありオンライン学習AROWのモデルを利用して、議論マップに残すべき重要発話の抽出を行う。

処理 3. 質疑応答の対応付けと重要発話の補完

取り出された重要発話中から質問であるような発話を選定し、その回答にあたるような発話の抽出と紐付けを行う。また、取り出された回答発話を重要発話として補充する。

処理 4. 議論マップによる可視化

抽出された重要発話を、会議中の主なトピックの流れや発話者、質疑応答の情報が素早く分かるような議論マップとして可視化を行う。

なお、本システムへの入力は、3.2節にて詳細を説明する、AMI Meeting Corpus¹[2]に含まれる人手により書き起こしが行われた会議録のデータである。このデータには、各発話に対する意味役割であるDialogue Actや発話開始／終了時間や話者などがアノテーションとして付与されている。

また、本システムは稼働前の事前準備として、訓練用の会議録データから、重要発話抽出のための教師あり学習モデルの構築を行っておく。事前のモデル構築に関する詳細は、3.4.2項にて述べる。

¹AMI Meeting Corpus: <http://groups.inf.ed.ac.uk/ami/corpus/>

3.2 対象とする会議録データセットについて

本研究では、会議録のデータセットとしてAMI Meeting Corpus²[2]を用いる。AMI Meeting Corpusには、1)議題が与えられているシナリオ有りの会議と、2)議題が与えられていないシナリオ無しの会議に関する会議録があり、それぞれ人手による書き起こしと機械による書き起こしのデータが存在する。本研究では、シナリオ有り会議録の人手による書き起こしデータを対象とする。なお、シナリオ有りの会議では、4人のそれぞれ異なる役をもつメンバが、架空の会社で新たなリモートコントロールの製作に関する会議を4回にわたり行う。ただし、あらかじめ定まった発話内容などは存在せず、会議中の発話は全て自然発生的なものである。

なお、シナリオ有りの会議において、4名の会議参加者に割り当てられる役は以下である³。

表 3.2.1: AMI Meeting Corpusのシナリオ有り会議における参加者の役一覧

役名	役の説明
PM	プロジェクトマネージャー (Project Manager) . 会議全体の司会進行を中心的に務める.
ME	マーケティング専門家 (Marketing Expert) . ユーザからの要求や市場のトレンド調査, プロトタイプの評価などを担当する.
UI	ユーザインタフェースデザイナー (User Interface Designer) . 開発するリモートコントロールの外見的デザインや, 提供する機能に関する責任を担当する.
ID	工業デザイナー (Industrial Designer) . 開発するリモートコントロールに搭載するボタンなどの要素決定とその機能に関する設計を担当する. また, UIと共に外見的デザインも手掛ける.

また、AMI Meeting Corpusには、人手で付与された豊富なアノテーション情報が存在する。本研究では最終的に、以下に示すアノテーション情報が付与されている、計134の会議録データを対象とした。

- 人手により記載された会議の要約文

各会議に対し、元の会議録の約6%（平均で約290単語）に相当する文量でまとめら

²AMI Meeting Corpus: <http://groups.inf.ed.ac.uk/ami/corpus/>

³各役の詳細は、AMI Meeting Corpusを提供している<http://groups.inf.ed.ac.uk/ami/corpus/scenariomeetings.shtml>に記載されている。

れた，人手による会議内容の要約である．この人手要約は，重要文抽出の評価における正解の参照要約として利用する．

- 人手により選択された会議録中の重要発話

各会議録中から，重要だと思われる発話に人手でラベルをつけたもの．この人手により取り出された重要発話の情報は，重要発話抽出のためのモデル構築時に利用する．

- Dialogue Act (DA)

会議録中の各発話に人手で付与された意味役割⁴で，全15種類が存在する．これらDAの情報は，1)重要発話抽出のためのモデル構築時と，2)質疑応答の対応付けによる構造化の際に利用する．それぞれの意味役割名とその主な機能を表3.2.2に記載する．

- 会議中のトピックセグメント

人手により付与された，会議録中の主なトピックの境界情報である⁵．トピックセグメントの情報は，1)質疑応答の対応付けによる構造化の際と，2)議論マップとして出力する際に各発話をトピック単位でまとめるために利用する．

AMI Meeting Corpusにはテスト用の会議録が用意されている．テスト用会議録の会議シリーズ（1つの会議シリーズは4回の会議からなる）は“ES2004”，“ES2014”，“IS1009”，“TS3003”，“TS3007”であり，合計20回分の会議録から構成される．

重要発話抽出のためのモデル構築時は，上記のテスト用会議録データを除く計114個の会議録データを訓練用データとして利用する．

⁴Dialogue Actの詳細および付与の基準は，AMI Meeting Corpusが提供しているアノテーションガイド (http://groups.inf.ed.ac.uk/ami/corpus/Guidelines/dialogue_acts_manual_1.0.pdf) に記載されている．

⁵トピックの境界を決定する基準の詳細は，AMI Meeting Corpusが提供しているトピックセグメンテーションに関するガイドライン (<http://groups.inf.ed.ac.uk/ami/corpus/Guidelines/TopicSegmentationGuidelines.pdf>) に記載されている．

表 3.2.2: AMI Meeting CorpusのDialogue Act一覧

DA	DAの説明
BACKCHANNEL	“uh-huh” や “yeah” , “ok” など, 第三者が会話をしている後ろで反応をとっている際に発せられた発話.
STALL	“so” や “but” , “ok” など, 話を始める前に口走った際に発せられた発話.
FRAGMENT	文頭に “I want” と述べその後に本来話したい内容を話す場合のように, BACKCHANNELやSTALL以外の特に話者が伝えたい内容ではない発話.
BE-POSITIVE	グループに対して発せられた発話で, 自分あるいは他人が好意を抱くような発話.
BE-NEGATIVE	グループに対して発せられた発話で, 自分あるいは他人の否定的な感情を表現した発話.
ELICIT-COMMENT-UNDERSTANDING	前の発話に対する理解を確認するような発話.
COMMENT-ABOUT-UNDERSTANDING	前の話者が述べた内容に対し, 理解している又はしていないことを表現している発話.
ELICIT-ASSESSMENT	“I wanted feedback” など, 他人に何らかの評価を求めるような発話.
ASSESS	“That would be great.” などのように, 何らかの評価を表現するような発話.
ELICIT-OFFER-OR-SUGGESTION	他人から提案や意思表示を引き出すことを目的に発せられた発話.
OFFER	話者が自分自身に関連する意思を表現した発話.
SUGGEST	何らかの提案を行っているような発話.
ELICIT-INFORM	他人に何か情報を求めている発話.
INFORM	何らかの情報を与える様な発話.
OTHER	その他の発話.

3.3 システムで用いる要素技術

本節では、提案システムの内部で利用する要素技術について説明する。提案システムでは、高精度かつ高速なオンライン型の線形識別器であるAROW (Adaptive Regularization of Weight Vectors) [23]を用いて重要発話抽出を行う。AROWはCW (Confidence-Weighted) [24]と呼ばれる手法の拡張であるため、以下ではまずCWについて説明し、次にCWの欠点を解決したAROWの説明を行う。なお、以下の説明は、我々が以前執筆した[25]で行った説明に基づいている。

3.3.1 CW (Confidence-Weighted)

CW[24]はAROWの元となるオンライン学習の二値分類器である。オンライン学習のため、素性ベクトル \mathbf{x}_t が与えられる度に予測ラベル \hat{c}_t を求め、正解ラベル c_t と比較することでモデルの重みベクトル \mathbf{w} を更新する。CWでは重みベクトル \mathbf{w} が平均 $\boldsymbol{\mu} \in \mathbb{R}^m$ 、分散 $\Sigma \in \mathbb{R}^{m \times m}$ の正規分布 $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ に従うと仮定されている。重みベクトルの中で分散値が大きいパラメータに関しては、まだ自信 (confidence) があまりない状態と考え、大きくパラメータを更新する。逆に、分散値が小さなパラメータに関しては、頻出な特徴のためにもう既に十分な情報が得られていると考え、小さくパラメータを更新する。実際にラベルを推定する際には、重みベクトルの期待値 $\mathbb{E}[\mathbf{w}] = \boldsymbol{\mu}$ を用いて行う。

t 回目の学習において、学習用の素性ベクトル \mathbf{x}_t および正解ラベル c_t が与えられたとする。この際、CWは以下の式(3.3.1)に示す制約付き最適化問題を解くことで、重みベクトルに関する分布を更新し、重みベクトルの新たな平均 $\boldsymbol{\mu}_t$ と分散 Σ_t を得る。

$$\begin{aligned} (\boldsymbol{\mu}_t, \Sigma_t) = \arg \min_{\boldsymbol{\mu}, \Sigma} \text{KL}(\mathcal{N}(\boldsymbol{\mu}, \Sigma) \parallel \mathcal{N}(\boldsymbol{\mu}_{t-1}, \Sigma_{t-1})) \\ \text{s.t. } \mathbb{P}_{\mathbf{w} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)}[c_t(\mathbf{w} \cdot \mathbf{x}_t) \geq 0] \geq \eta \end{aligned} \quad (3.3.1)$$

ここで、 $\mathcal{N}(\boldsymbol{\mu}_{t-1}, \Sigma_{t-1})$ は t 回目の学習による更新を行う前の重みベクトルに関する分布であり、 $\text{KL}(\mathcal{N}(\boldsymbol{\mu}, \Sigma) \parallel \mathcal{N}(\boldsymbol{\mu}_{t-1}, \Sigma_{t-1}))$ は $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ と $\mathcal{N}(\boldsymbol{\mu}_{t-1}, \Sigma_{t-1})$ 間のカルバック-ライブラーダイバージェンスである。また、 $\eta \in (0.5, 1]$ は更新の度合いを制御するハイパーパラメータである。

式(3.3.1)の制約条件において、 $\mathbb{P}_{\mathbf{w} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)}[c_t(\mathbf{w} \cdot \mathbf{x}_t) \geq 0]$ は与えられた素性ベクトル \mathbf{x}_t に対して正しくラベル予測される確率を表している。したがって、CWは $\eta \in (0.5, 1]$ 以上

の確率で正しく分類されるという条件を満たした上で、更新前の重みベクトルの正規分布に最も近い正規分布を求めることで学習を行う。この制約条件は、与えられた素性ベクトル \mathbf{x}_t を常に正しく分類出来るようにモデルを更新することを意味するため、ノイズデータに極めて弱く過学習を起こしやすいという欠点が存在する。

3.3.2 AROW (Adaptive Regularization of Weight Vectors)

AROW[23]は、上記のCWが持つ欠点を、制約条件を目的関数の一部に正則化項として持たせることにより解決した手法である。具体的には、以下の式(3.3.2)に示す最適化問題を解くことで、重みベクトルに関する分布を更新する。

$$(\boldsymbol{\mu}_t, \Sigma_t) = \arg \min_{\boldsymbol{\mu}, \Sigma} \text{KL}(\mathcal{N}(\boldsymbol{\mu}, \Sigma) \parallel \mathcal{N}(\boldsymbol{\mu}_{t-1}, \Sigma_{t-1})) + \frac{1}{2r} \ell_{h^2}(c_t, \boldsymbol{\mu} \cdot \mathbf{x}_t) + \frac{1}{2r} \mathbf{x}_t^T \Sigma \mathbf{x}_t \quad (3.3.2)$$

ここで、 $r > 0$ はモデルの更新を調節するハイパーパラメータである。式(3.3.2)は3つの項から構成され、それぞれの項は以下の様な意味を持つ。

(1) $\text{KL}(\mathcal{N}(\boldsymbol{\mu}, \Sigma) \parallel \mathcal{N}(\boldsymbol{\mu}_{t-1}, \Sigma_{t-1}))$

この項を小さくすることは、パラメータの更新を小さく抑え、更新前の重みベクトルの正規分布に最も近い正規分布を求めることを意味する。

(2) $\frac{1}{2r} \ell_{h^2}(c_t, \boldsymbol{\mu} \cdot \mathbf{x}_t)$

$\ell_{h^2}(c_t, \boldsymbol{\mu} \cdot \mathbf{x}_t) = (\max\{0, 1 - c_t(\boldsymbol{\mu} \cdot \mathbf{x}_t)\})^2$ は二乗ヒンジ損失である。この項を小さくすることは、現在与えられているデータに対する予測間違いをなるべく少なくするような重みベクトル \mathbf{w} の平均 $\boldsymbol{\mu}$ を求めることを意味する。ただし、この項には二乗ヒンジ損失関数以外の損失関数を適用することも可能である。

(3) $\frac{1}{2r} \mathbf{x}_t^T \Sigma \mathbf{x}_t$

この項を小さくすることは、重みベクトル \mathbf{w} の各素性に関する分散（自信のなさ）を、学習を進めるにつれて小さくしていくことを意味する。

以上から、AROWは1) \mathbf{w} の分布を今までの正規分布になるべく近く、2)現在の学習データを正しく分類し、3) \mathbf{w} の各素性に関する自信を少しずつ上げていくことで、CWの欠点であったノイズのあるデータに頑健なオンライン学習を実現している。

擬似コード3.3.1にAROWアルゴリズムの擬似コードを示す。

擬似コード 3.3.1 AROWによるオンライン識別アルゴリズム ([23]より引用)

入力: ハイパーパラメータ r

出力: 重みベクトル μ_T , 分散 (confidence) Σ_T

```
1:  $\mu_0 \leftarrow \mathbf{0}$ 
2:  $\Sigma_0 \leftarrow I$ 
3: for  $t \in [1, \dots, T]$  :
4:   訓練データ  $\mathbf{x}_t \in \mathbb{R}^d$  を受け取る
5:    $m_t \leftarrow \mu_{t-1}^\top \cdot \mathbf{x}_t$  ▷ マージンを計算
6:    $v_t \leftarrow \mathbf{x}_t^\top \Sigma_{t-1} \mathbf{x}_t$  ▷ 分散 (confidence) を計算
7:   正解ラベル  $c_t$  を受け取る
8:   if  $m_t c_t < 1$  : ▷ 予測を間違えた場合に更新
9:      $\beta_t \leftarrow \frac{1}{\mathbf{x}_t^\top \Sigma_{t-1} \mathbf{x}_t + r}$ 
10:     $\alpha_t \leftarrow \max(0, 1 - c_t \mathbf{x}_t^\top \mu_{t-1}) \beta_t$ 
11:     $\mu_t \leftarrow \mu_{t-1} + \alpha_t \Sigma_{t-1} c_t \mathbf{x}_t$ 
12:     $\Sigma_t \leftarrow \Sigma_{t-1} - \beta_t \Sigma_{t-1} \mathbf{x}_t \mathbf{x}_t^\top \Sigma_{t-1}$ 
```

3.4 システムの詳細

本節では、提案システムの内部で行う、1)前処理、2)重要発話抽出、3)質疑応答の対応付けと重要発話の補完、4)議論マップとしての可視化、という主な4つの処理に関する詳細をそれぞれ説明する。

3.4.1 前処理

前処理では、主に1)各発話からのStopwords除去や言い直しによる冗長性の除去と、2)発話中の重要フレーズの特定を行う。なお、英語テキストの形態素解析および品詞推定にStanford大学が提供しているCoreNLP[26]を利用する。

本研究では、Stopwordsとして、“the”や“a”などの冠詞に代表される英語テキストに多々含まれる単語に加え、“uh”や“ah”などのフィラーをはじめとする発話にしばしば含まれるが特に重要な意味を持たないような単語も含む。また、会議中の発話には、しばしば同じことを繰り返している場面が見受けられる。そこで、言い直しによる冗長性の除去として、発話中において連続する同じ単語を取り除く。

本研究において、重要フレーズとは、発話内に含まれ会議内容を特徴づけるような形態素あるいは形態素列である。重要フレーズになりうる形態素列は[5]に基づいて、以下の品詞列に適合するものとして定義する。

重要フレーズの品詞列

$$JJ^*(NN|NNS|FW|CD)^+((DT|IN)^+JJ^*(NN|NNS|FW|CD)^+)^*$$

ただし、上記の品詞記号はPenn TreeBankの品詞記号に従っており、それぞれの意味を表3.4.1に示す。

表 3.4.1: Penn TreeBankの品詞記号と意味

品詞記号	品詞記号の意味
JJ	形容詞
NN	名詞（単数形）
NNS	名詞（複数形）
FW	外来語
CD	数字
DT	限定詞
IN	前置詞，あるいは従属接続詞

例えば、上記の定義により “we are asked to make new remote control for television.” という発話から取り出される重要フレーズは，“new remote control” と “television” となる。これらの取り出した重要フレーズは、次項にて説明する、重要発話抽出のためのモデル構築の際の素性として利用する。

3.4.2 重要発話抽出

本フェーズでは、会議録からその会議における重要な発話の抽出を行う。提案システムでは、3.3.2項にて説明した、高精度かつ高速なオンライン型の線形識別器であるAROW（Adaptive Regularization of Weight Vectors）[23]を用いて重要発話抽出を行う。

AROWは教師あり学習の二値分類器であるため、モデルの構築手順は以下となる。

重要発話抽出のためのAROWモデル構築手順

1. AMI Meeting Corpusからの、学習に用いる素性の選択
2. 選定した素性を用いて、各発話の素性ベクトルを作成
3. AMI Meeting Corpusに含まれている、会議録中の各発話が人手により選択されたかどうかの情報を正解ラベルとして、AROWを学習

なお、AROWの持つハイパーパラメータ r に関しては、訓練データセットにおける10-fold cross validationを行い、グリッドサーチをかけることで適切に設定を行う。

本研究において、使用した素性の一覧を表3.4.2に示す。

表 3.4.2: AROWモデルの構築に用いる素性一覧

素性	素性の説明
話者情報	発話を行った話者の役に関する情報. PM, ME, UI, IDの4種類.
発話役割	3.2節にて記載した, Dialogue Act.
発話位置情報	会議全体を5分割したうち, 対象の発話がどの区分に属するかという, 発話の時間位置情報.
発話時間長	発話を始めてから終わるまでにかかった発話時間の長さ.
単語数	発話に含まれるユニークな単語の数
総単語数	発話に含まれる全ての単語の数
品詞数	発話に含まれるユニークな品詞の数
単語N-gram	単語のN-gram情報. ($1 \leq N \leq 3$)
品詞N-gram	品詞のN-gram情報. ($1 \leq N \leq 3$)
重要フレーズ	前処理にて特定した, 発話に含まれる重要フレーズの情報.
重要フレーズ数	発話に含まれる重要フレーズの数.

また、抽出された発話 u の重要度は、AROWが識別時に利用するマージンの値 $\mu \cdot x_u$ を用いて表現する。すなわち、マージンが大きい発話ほど重用であると定義する。

3.4.3 質疑応答の対応付けと重要発話の補完

本フェーズでは、まず重要発話抽出のフェーズにて抽出された重用発話に含まれる質問発話の選定を行い、次にその質問の発話に対する回答となるような発話の特定と紐付けを行う。最後に、特定した回答に相当する発話を重要発話として補充する。

重要発話に含まれる質問は、1) Dialogue Act (DA) の情報と、2) 文末が“?”かどうかにより、以下の擬似コード3.4.2に記載するルールに基づき推定する。

擬似コード 3.4.2 質問発話推定

入力: 対象発話 u , 質問DAの集合 DA_q

出力: 質問発話かどうかのTrue, False

```
1: procedure isQuestion( $u$ )  
2:   if  $DA(u) \in DA_q$  :  
3:     return True  
4:   else if  $u$ の末尾が “?” :  
5:     return True  
6:   else  
7:     return False
```

▷ $DA(u)$ は発話 u のDA

ただし、質問DAは、表3.2.2における “Elicit-Comment-Understanding” , “Elicit-Assessment” , “Elicit-Offer-Or-Suggestion” , “Elicit-Inform” の4種類として定義する.

また、質問に対する回答の発話は、1) Dialogue Act (DA) の情報と、2) 発話が同じトピックセグメント内にあるかどうか、3) 話者情報を用いて、以下の擬似コード3.4.3に記載するルールに基づき推定する. なお、取り出した回答発話が質問の場合は、その回答も再帰的に取り出す.

擬似コード 3.4.3 回答発話抽出

入力: 質問発話 q , 回答DAの集合 DA_a , 会議録中の全発話 U

出力: 質問 q に対する回答発話の集合 U_a

```
1: function getAnswer( $q, U, depth = 0, q_{orig} = \text{Null}$ )  
2:   if  $depth = 0$  :  
3:      $q_{orig} \leftarrow q$   
4:      $U_a \leftarrow \emptyset$   
5:     for  $u \in U$  :  
6:       if  $topic(u) \neq topic(q)$  :  
7:         break  
8:       if isAnswerCandidate( $u, q, q_{orig}$ ) :  
9:          $U_a \leftarrow U_a \cup \{u\}$   
10:        if isQuestion( $u$ ) :  
11:           $U'_a \leftarrow \text{getAnswer}(u, U, depth + 1, q_{orig})$   
12:           $U_a \leftarrow U_a \cup U'_a$   
13:        break  
14:      else  
15:        continue  
16:   return  $U_a$ 
```

▷ $depth$ は 再帰の深さ
▷ ルートとなる質問発話を設定
▷ 会議録中の全発話に対する処理
▷ 回答発話が質問の場合

擬似コード 3.4.4 回答発話候補の推定

入力: 対象発話 u , ルートとなる質問発話 q_{orig} , 直前の質問発話 q , 回答DAの集合 DA_a

出力: 発話 u が質問 q に対する回答の候補であるかどうかの True, False

```
1: function isAnswerCandidate( $u, q, q_{\text{orig}}$ )
2:   if  $DA(u) \notin DA_a$  or  $topic(u) \neq topic(q)$ :           ▷  $topic(u)$  は発話  $u$  のトピック
3:     return False
4:   if  $speaker(u) = speaker(q_{\text{orig}})$ :                     ▷  $speaker(u)$  は発話  $u$  の話者
5:     return False
6:   return True
```

上記ルールに記載したとおり，質問に対する回答発話は，1)質問発話を発した話者以外の話者による，2)同一トピックセグメント内の，3)回答になりうるDAに適合し，かつ質問発話に最も近い発話として抽出される．なお，これらの抽出された回答発話は，どの質問に対する回答かという対応を保持したまま，元の重要発話に補充する形で追加する．ただし，補充の際は，元の重要発話に含まれていない場合のみ追加を行う．

3.4.4 議論マップによる可視化

本フェーズでは，重要発話抽出および質疑応答の対応付けを行った後の，構造化された会議録情報を元に，議論マップを生成する．議論マップとは，1)会議全体の流れ，2)会議中での重要な発話，3)会議中の重要な質問に対する回答の紐付け，をそれぞれ構造化した上で素早く把握できるように可視化したものである．

具体的な議論マップの生成は，以下の手順に従って行う．

議論マップ生成の手順

1. 各発話を発した話者情報を，その発話の先頭に付与する．
2. 各発話に含まれる重要フレーズを太字で強調する．
3. 抽出された重要発話をトピックセグメントで分割．
4. トピックセグメント内で，発話を時系列順で並べる．
5. 発話が質問の場合，その質問に紐づく回答の発話を質問にぶら下げて表示する．

以上の手順により，生成される議論マップの例を図3.4.1に示す．なお，この例ではAMI Meeting Corpusのテスト用データシリーズの一つである“ES2004c”から議論マップの自動生成を行った．

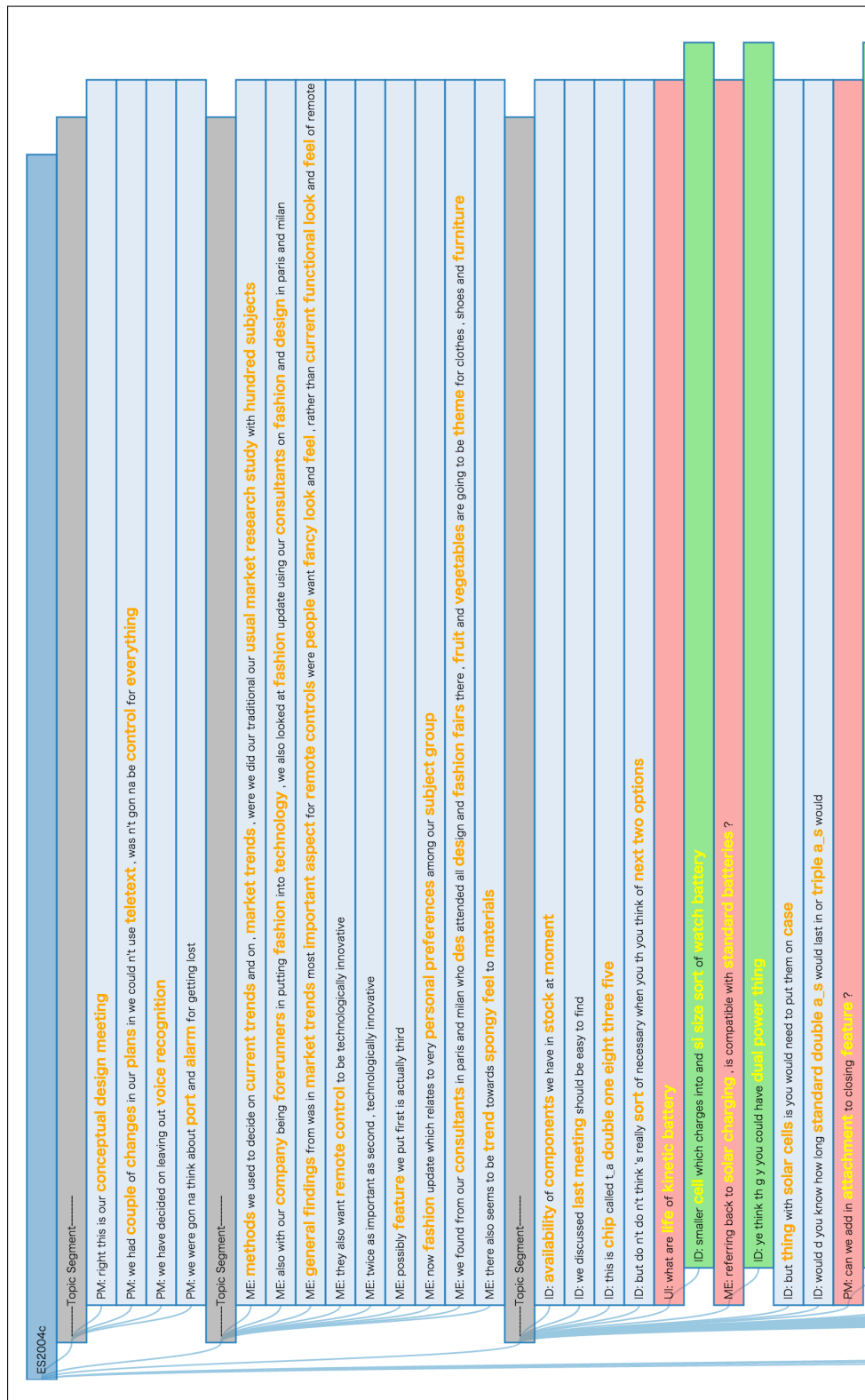


図 3.4.1: 提案システムにより “ES2004c” から自動生成を行った議論マップの出力

第4章 実験と評価

本章では、提案システムの有用性に関する評価を行う。本システムの目的は、膨大な会議録から効率的に会議の流れや会議内で話された重要事項を得ることが出来るような議論マップの自動生成である。そこで、システムの評価に関しては以下の2点に関する評価実験を行う。

システム評価実験の内容

1. 会議録からの重要発話抽出に関する評価

会議録中から重要発話をきちんと取り出せているかという点に関する評価。

2. 構造化された議論マップによる出力の有用性評価

抽出した構造化済み重要発話の出力を、i)テキストとした場合、ii)議論マップとした場合のどちらがより効率的に会議内容を把握可能か、という点に関する評価。

以下では、まず4.1節で、重要発話抽出に関する評価を行い、続く4.2節で、被験者実験による、議論マップによる出力の有用性に関する評価を行う。

4.1 会議録からの重要発話抽出に関するROUGEによる定量的評価

本節では、会議録からの重要発話抽出に関する評価を行う。評価指標には、文書要約のタスクにおいて最も広く用いられているROUGE[22]を用いた。ROUGEは、正解である参照要約と、システムが生成した要約との間で、どれだけ形態素のN-gramが一致するかの評価を行う。なお、会議中の自然発生的な発話と、人手で記載された要約とは、bi-gramよりuni-gramで重なりを評価する方が適切であるというGillickら[4]の見解に基づき、本評価実験においてはuni-gramの一致率を評価するROUGE-1におけるF-measureを評価指標として用いた。

実験のテストデータには、3.2章で説明を行った、AMI Meeting Corpusが提供する20回分のテスト用会議録データを利用した。また、参照要約にはテストデータに付属の人手により作成された会議の要約文を用いた。比較対象とする重要発話抽出手法を表4.1.1に示す。

表 4.1.1: 比較対象の重要発話抽出手法

既存／提案	手法名	
既存手法	MMR-centroid	Carbonell ら [27] に よ る MMR-centroid System
	MMR-cosine	Gillick ら [4] に よ る MMR-cosine System
	ILP-based	Gillick ら [4] に よ る ILP-based System
	TextRank	Mihalcea ら [28] に よ る TextRank System
	ClusterRank	Garg ら [6] に よ る ClusterRank System
提案手法	AROW-based	質疑応答の対応付けを行わない提案手法
	AROW-QA	質疑応答の対応付けを行った提案手法

重要発話の抽出に関する既存手法は、抽出した発話に含まれる単語数の合計が、予め与えられた単語数以下となるように重要発話の抽出を行う。そこで本実験においては既存手法との比較のため、多くの研究で採用されている、要約文に含まれる単語数が元の会議録に含まれる総単語数の6%以下となるような制約を設けた。この6%（平均で約290単語）とは、人手で作成された要約文のに含まれる単語数の平均値と概ね一致する割合である。

なお、提案手法で用いているAROWは教師あり学習の二値分類器であるため、本来は各発話が要約に含まれるか否かを識別しており、設けられた単語数制限の元で重要発話と取り出す手法ではない。そこで、本実験における単語数の制約を満たすために、AROWのマージンにより定義した発話の重要度が高いものを優先的に要約に含めた。

ROUGE-1による評価実験の結果を表4.1.2に示す。なお、表4.1.2に記載した既存手法の値は、過去に報告された同じデータセットおよび条件による実験結果の値を引用している[4][6][12]。

表 4.1.2: ROUGE-1による重要発話抽出の評価

手法名	ROUGE-1
MMR-centroid	0.182
MMR-cosine	0.21
ILP-based	0.24
TextRank	0.250
ClusterRank	0.275
AROW-based	0.291
AROW-QA	0.293

以上の結果から、本手法による重要発話抽出がより人手による要約に近い発話を取り出

していることが分かる。また、質疑応答を行ったAROW-QAのスコアが、行わなかった場合のAROW-basedと比べ高いスコアを出している。このことから、重要な質問に対する回答を要約に含めることで、有益な情報を追加出来ていることが分かる。

4.2 被験者実験による議論マップの有用性評価

本節では、重要発話抽出と構造化を行った結果の出力である議論マップの有用性に関して、被験者実験による評価を行う。実験では、以下の2通りの出力を用いて、被験者にこちらが予め用意した会議内容に関する質問に答えてもらった。

比較対象の出力

1. プレーンテキスト：構造化されていないプレーンテキストとして抽出した重要発話を並べたもの
2. 議論マップ：構造化されている議論マップとして出力したもの

ここで、予め用意した質問は全て4択であり、重要発話を全て読めば回答が可能な内容となっている。質問は全3問であり、回答時間は3分以内である。なお、AMI Meeting Corpusは英語テキストによる会議録であるため、被験者はなるべく英語力に偏りのない、20代の理工系学生4名に依頼した。また、各被験者には、プレーンテキストで会議1を、議論マップで会議2を評価という具合に、それぞれの出力で異なる1つずつの会議録に対して評価を行ってもらった。

本実験で用いる会議録データは、テストデータに含まれる20個の会議録から、無作為に取り出した“ES2004c”と“TS3003b”の2つを対象とし、被験者毎に出力との組み合わせを入れ替えた¹。実験の結果を表4.2.1に示す。

表 4.2.1: 被験者実験による議論マップの有用性評価

被験者	正答した質問数	
	議論マップ	プレーンテキスト
被験者A	2	0
被験者B	3	2
被験者C	2	1
被験者D	1	0

¹実際に出題した問題、および被験者が回答するにあたり参照した両出力は付録A.1に掲載した。

得られた表4.2.1の結果に対し、以下の条件に従い有意水準5%($\alpha = 0.05$)の t 検定を行う

条件1) 対象データ：各被験者 i の、1)議論マップ利用時の正答数 (x_i) と、2)プレーンテキスト利用時の正答数 (y_i)

表 4.2.2: 検定で用いるデータ

n	x_i	y_i	$d(= x_i - y_i)$	d^2
1	2	0	2	4
2	3	2	1	1
3	2	1	1	1
4	1	0	1	1
合計	5			7

条件2) 帰無仮説 H_0 ：両出力による正答数に差がない ($\mu_d = 0$)

条件3) 対立仮説 H_1 ：両出力による正答数に差がある ($\mu_d \neq 0$)

条件4) 検定統計量： $T = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = 5.0$

$$\text{ただし, } \begin{cases} \bar{d} = \frac{\sum d}{n} = \frac{\sum (x_i - y_i)}{n} = 1.25 \\ s_d = \sqrt{\frac{\sum d^2 - n \times (\bar{d})^2}{n - 1}} = 0.5 \end{cases}$$

条件5) 有意水準： $\alpha = 0.05$ (両側検定)

条件6) 棄却域：自由度 $f = n - 1 = 3$, $\alpha = 0.05$ であるため、付録A.3の t 分布表より、

$$t(n - 1, \frac{\alpha}{2}) = t(3, 0.025) = 3.182$$

以上の条件に従い検定を行った結果、得られた検定統計量の実測値は $T = 5.0 > 3.182$ であるため、両出力の間には正答した質問数の差がないという帰無仮説が棄却される。従って、有意水準0.05で議論マップによる出力が有意に有効であることが示された。

議論マップとしての出力では、1)重要フレーズの太字による強調と、2)トピックセグメントや質疑応答の対応付けによる構造化という大きく2つの工夫を施している。そこで、議論マップを用いて会議内容を把握する際に、どちらかの工夫がより効果的に作用しているかを検証する実験を行う。実験内容は上記の被験者実験と同様の手順で、こちらが予め用意した3つの4択問題に対して3分以内に回答してもらい、その回答の正解数を調べるものである。ただし、今回の出力はプレーンテキストと議論マップではなく、以下に示す2つである。

比較対象の出力

1. 重要フレーズの強調

プレーンテキストによる出力に含まれる、各発話中の重要フレーズを太字にして色を変えることで強調したもの

2. 構造化

重要フレーズの強調は行わずに、トピック毎に分割を行い、質疑応答の対応付けをすることで構造化を行ったもの

本実験では、被験者として、英語力に偏りのない20代の理工系学生3名に依頼した。なお、実験に用いた会議データは、テストデータに含まれる20個の会議録から、無作為に取り出した“TS3003c”と“TS3007c”の2つを対象とし、被験者毎に出力との組み合わせを入れ替えた²。実験の結果を表4.2.3に示す。

表 4.2.3: 被験者実験による議論マップ内の各工夫の影響度評価

被験者	正答した質問数	
	重要フレーズの強調	構造化
被験者A	1	3
被験者B	3	1
被験者C	2	2

表4.2.3から、被験者の正解数に両工夫間での有意差を確認することはできない。この結果から、重要フレーズの強調と構造化が相乗的に働き、上の t 検定で得られたように議論マップ自体の効率性に繋がっているものと考えられる。

²実際に出題した問題、および被験者が回答するにあたり参照した両出力は付録A.2に掲載した。

第5章 おわりに

本稿では、会議の流れや決定事項などの要旨を素早く把握することを可能とする、発話間関係の構造化による会議録からの議論マップ自動生成システムを提案した。提案システムは、1)会議録からの重要発話抽出を行い、2)発話間における質疑応答の対応付けを行った後に、3)発話間の構造を保持した議論マップとして出力を行う。これらの処理により、1)発話間の関係が構造化されていない、2)確認すべき発話が限定されていない、という従来の問題点を解決した。

重要発話抽出に関する評価では、発話間における質疑応答の対応付けを行うことで、既存手法と比べROUGE-1のスコアを最大0.111上げることに成功した。また、4名の被験者実験を行った結果、重要発話をプレーンテキストとして表示した時と比べ、議論マップとして出力を行った時の方が素早く会議内容を把握できるということが検証され、システムの有用性が示された。

今後の課題としては、システムの完全自動化に向けて、トピックセグメントの自動推定などを行い、現在用いているAMI Meeting Corpusが提供するアノテーション情報の自動付与に取り組む必要がある。また、質を向上させるため、重要発話抽出と、現在はルールベースで行っている質疑応答の対応付けの精度向上を行い、より多くの被験者による実験を行う必要がある。

謝辞

本研究を行うにあたり，数々のご指導を頂いた山名早人教授に深く感謝致します．また，研究を進める上で助言を頂いた浅井洋樹先輩やJungKyu Han先輩をはじめ，実験にも協力して頂いた研究室の同期の方々にも心より感謝致します．

研究業績

- 受賞

2014年8月：第159回情報処理学会DBS研究会「学生奨励賞受賞」

- 外部発表

1. 林佑磨, 奥野峻弥, 山名早人: “意味概念に基づいた関連論文検索システム～近傍文書からのキーフレーズ抽出を用いた自動クエリ生成～”, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 2014.
2. 林佑磨, 奥野峻弥, 山名早人: “単語の意味概念行列を用いたキーワード生成による関連論文検索システム”, 研究報告データベースシステム (DBS), Vol. 2014-DBS-159, No. 10, pp. 1-6, 2014.
3. 林佑磨, 諏訪晴士, 山名早人: “非テキスト情報のみを用いたAROWによる効率的なCTR予測モデルの構築”, 第7回データ工学と情報マネジメントに関するフォーラム (DEIM2015), 2015.
4. 林佑磨, 山名早人: “会議録から要旨可視化のための議論マップ自動生成を行う統合システム”, 第8回データ工学と情報マネジメントに関するフォーラム (DEIM2016), 2016. (発表予定)

参考文献

- [1] Adam Janin, Don Baron, Jane Edwards, Dan Ellis, David Gelbart, Nelson Morgan, Barbara Peskin, Thilo Pfau, Elizabeth Shriberg, Andreas Stolcke, and others: “The ICSI meeting corpus” , in *Proc. of ICASSP*, pp. 364–367, 2003.
- [2] Jean Carletta, Simone Ashby, and Sebastien Bourban et al: “The AMI meeting corpus: A pre-announcement” , in *Proc. of MLMI*, pp. 28–39, 2005.
- [3] Shasha Xie, Benoit Favre, Dilek Hakkani-Tür, and Yang Liu: “Leveraging sentence weights in a concept-based optimization framework for extractive meeting summarization” , in *Proc. of INTERSPEECH*, pp. 1503–1506, 2009.
- [4] Dan Gillick, Korbinian Reidhammer, Benoit Favre, and Dilek Hakkani-Tür: “A global optimization framework for meeting summarization” , in *Proc. of ICASSP*, pp. 4769–4772, 2009.
- [5] Korbinian Riedhammer, Benoit Favre, and Dilek Hakkani-Tür: “Long story short – Global unsupervised models for keyphrase based meeting summarization” , *Speech Communication*, Vol. 52, No. 10, pp. 801–815, 2010.
- [6] Nikhil Garg, Benoit Favre, Korbinian Reidhammer, and Dilek Hakkani-Tür: ” Cluster-Rank: a graph based method for meeting summarization” , in *Proc. of Annual Conf. of the Inter-Speech*, pp. 1499-1502, 2009.
- [7] Fei Liu and Yang Liu: “Using spoken utterance compression for meeting summarization: A pilot study” , in *Proc. of IEEE SLT*, pp. 37–42, 2010.
- [8] Shasha Xie and Yang Liu: “Improving supervised learning for meeting summarization using sampling and regression” , *Computer Speech & Language*, Vol. 24, No. 3, pp. 495–514, 2010.
- [9] Andreas Kathol and Gokhan Tur: “Extracting question/answer pairs in multi-party meetings” , in *Proc. of ICASSP*, pp. 5053–5056, 2008.
- [10] Michel Galley: “A skip-chain conditional random field for ranking meeting utterances by importance” , in *Proc. of EMNLP*, pp. 364–372, 2006.

- [11] Fei Liu and Yang Liu: “Towards abstractive speech summarization: Exploring unsupervised and supervised approaches for spoken utterance compression” , *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 7, pp. 1469–1480, 2013.
- [12] Yashar Mehdad, Giuseppe Carenini, Frank W. Tompa, and Raymond T. NG: “Abstractive meeting summarization with entailment and fusion” , in *Proc. of ENLG*, pp. 136–146, 2013.
- [13] Gabriel Murray: “Abstractive Meeting Summarization as a Markov Decision Process” , *Advances in Artificial Intelligence*, Springer International Publishing, pp. 212–219, 2015.
- [14] M. Abdeen and M. C.E. Yagoub: “Direct Automatic Generation of Mind Maps from text with M2 Gen” , in *Proc. of TIC-STH*, 2009 IEEE Toronto International Conference. IEEE, 2009.
- [15] Robert Kudelic, Mladen Koneck, and Mirko Malekovic: “Mind map generator software model with text mining algorithm” , in *Proc. of ICITI*, 2011.
- [16] Au Purwarianti, Atria Saelan, Irfan Aff, Filman Freudian, and Alfian Farizki Wicaksono: “Natural Language Understanding Tools with Low Language Resource in Building Automatic Indonesian Mind Map Generator” , *International Journal on Electrical Engineering and Informatics (IJEI)*, Vol.5, No.3, 2013.
- [17] Mohamed Elhoseiny, Ahmed Elgammal: “Text to multi-level MindMaps” , *Multimedia Tools and Applications*, Springer, pp. 1–28, 2015.
- [18] 松村真宏, 加藤優, 大澤幸生, 石塚満: “議論構造の可視化による論点の発見と理解” , *日本ファジィ学会誌*, Vol. 15, No. 5, pp. 554–564, 2003.
- [19] 趙雲超, 松村真宏, 谷内田正彦: “音声認識された議事録からの議論マップ自動生成” , *人工知能学会全国大会論文集*, Vol. 6, pp. 221–221, 2006.
- [20] 友部博教, 土田貴裕, 大平茂輝, 長尾確: “ディスカッションメディア: 会議コンテンツの構造化と効率的な閲覧システム” , *人工知能学会第 21 回全国大会論文集*, No. 2F3-5, 2007.
- [21] 森幹彦, 八村太輔, 喜多一: “リフレクションのための逐語議事録を用いた議論の構造化法” , *人工知能学会第 21 回全国大会論文集*, No. 2D4-1, 2007.

- [22] Chin-Yew Lin: “Rouge: A package for automatic evaluation of summaries” , in *Proc. of Workshop on Text Summarization Branches Out (WAS)*, pp25–26, 2004.
- [23] Koby Crammer, Alex Kulesza, and Mark Dredze: “Adaptive regularization of weight vectors” , in *Proc. of Advances in Neural Information Processing Systems*, NIPS’ 09, pp.414–422, 2009.
- [24] Mark Dredze, Koby Crammer, and Fernando Pereira: “Confidence-weighted linear classification” , in *Proc. of the 25th international conference on Machine learning*, ICML’ 08, pp.264–271, 2008.
- [25] 林佑磨, 諏訪晴士, 山名早人: “非テキスト情報のみを用いたAROWによる効率的なCTR予測モデルの構築” , 第7回データ工学と情報マネジメントに関するフォーラム (DEIM2015) , 2015.
- [26] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky: “The Stanford CoreNLP Natural Language Processing Toolkit” , in *Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55–60, 2014.
- [27] Jaime Carbonell and Jade Goldstein: “The use of MMR, diversity-based reranking for reordering documents and producing summaries” , in *Proc. of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 335–336, 1998.
- [28] Rada Mihalcea and Paul Tarau: “TextRank: Bringing order into texts” , in *Proc. of EMNLP*, pp. 404–411, 2004.

付録

A.1 被験者実験における質問と出力（プレーンテキストと議論マップの比較）

- “ES2004c” に関する質問と出力

- 質問

Q1. What is the theme of the meeting?

A1. kick-off-meeting.

A2. functional design.

A3. component concepts.

A4. industrial design.

Q2. How did they decided about “voice recognition” ?

A1. put a button on remote control for a voice recognition.

A2. embed a voice recognition function into new remote control.

A3. don’ t use voice recognition.

A4. they didn’ t talk about voice recognition.

Q3. What should they do before adding an attachment?

A1. change design.

A2. market research.

A3. think about budget.

A4. have nothing special to do.

- プレーンテキストによる出力

ES2004c
right this is our conceptual design meeting
we had couple of changes in our plans in we could n't use teletext , was n't gon na be control for everything
we have decided on leaving out voice recognition
we were gon na think about port and alarm for getting lost
methods we used to decide on current trends and on , market trends , were we did our traditional our usual market research study with hundred subjects
also with our company being forerunners in putting fashion into technology , we also looked at fashion update using our consultants on fashion and design in paris and milan
general findings from was in market trends most important aspect for remote controls were people want fancy look and feel , rather than current functional look and feel of remote
they also want remote control to be technologically innovative
twice as important as second , technologically innovative
possibly feature we put first is actually third
now fashion update which relates to very personal preferences among our subject group
we found from our consultants in paris and milan who des attended all design and fashion fairs there , fruit and vegetables are going to be theme for clothes , shoes and furniture
there also seems to be trend towards spongy feel to materials
availability of components we have in stock at moment
we discussed last meeting should be easy to find
this is chip called t_a double one eight three five
but do n't do n't think 's really sort of necessary when you th you think of next two options
what are life of kinetic battery
referring back to solar charging , is compatible with standard batteries ?
but thing with solar cells is you would need to put them on case
would d you know how long standard double a_s would last in or triple a_s would
can we add in attachment to closing feature ?
if we 're doing kinetic thing
should n't we do some market research on first before we add in
not something 's come up in any of our focus groups and market research
which basically has these contacts are really close together and then when you press down on rubber button , connect circuit
each switch is connected to two different legs on chip and sends different message

プレーンテキストによる “ES2004c” の出力（その1）

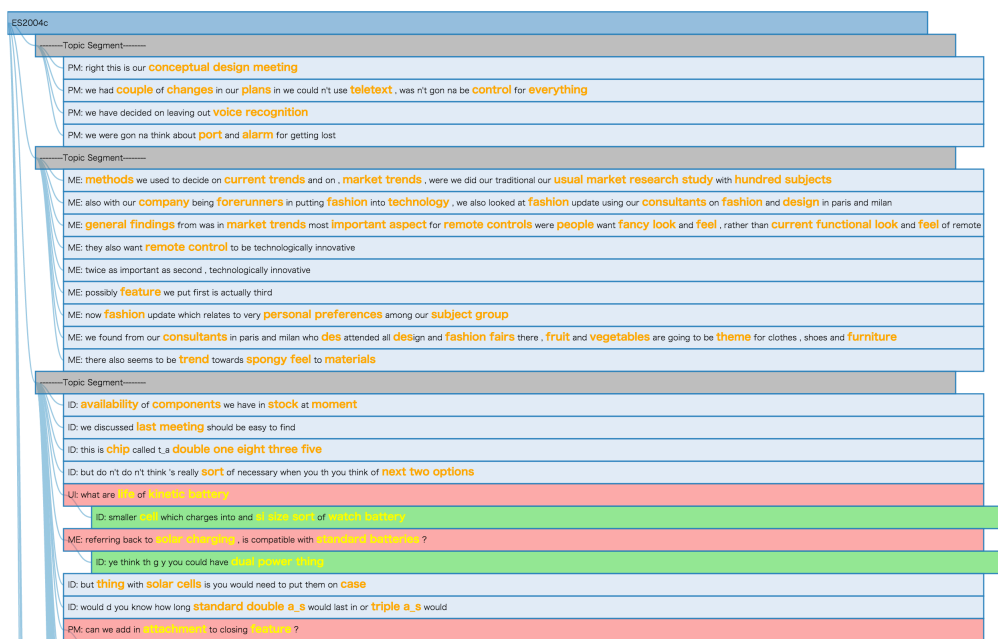
each switch is connected to two different legs on chip and sends different message
if you if you see how thin tracks are , you could you could s you could put them virtually right next to each other and have much smaller switch on each one
you 're saying people like this spongy feel this year
perhaps some something made of rubber
was just would be expensive to buy new case for , because of expense of how much titanium is
was just wondering if we make basic mould out of plastic but have like rubber cover goes over
would give spongy feel
instead of fascia comes off being plastic , fascia comes off would be rubber
just going back to interface designs with buttons
sure you could probably get in different colours
second option th they offer rubber buttons
problem with basically is 's flat and you ca n't do lots of curve curved things
or we could have rubber buttons which are made of this material which is anti-r_s_i
person who switches channels very often or does use particular function very often will find very irritating to
regarding circuit , since we are hav having flip-top , we can customise two circuits for different type of buttons
mean you 'd you 'd actually have two separate you 'd have two separate circuit boards but they 'd be joined by wires or like some cabling between them
possible when we open our flip-top shell 's little compact mirror and when you press button then goes onto phone display th remote control display thing
if you if you just put full charge through l_c_d display completely blank out
we just took input from previous meeting , especially from marketing and industrial design , to check on customer needs and feasibility
we are going to have m maybe g g_u_i interface in f in middle of flip-top
on top m repetitive buttons which are like volume or channel changing and all could be on pointing device
on s simpler board , on top we have this button , rubber buttons , to keep frequently changing channels
there was mail which mentions our division has developed new speech recognition s feature
we have to check into for financial f feasibility whether we can incorporate this at low cost
like we had g s response from customers they would like to have feature of finding
but we can take inspiration from fruit colours , like vibrant colours , red chilli
but we need to make decision about things we discussed
we could incorporate voice recognition for finding

プレーンテキストによる “ES2004c” の出力（その2）

we could incorporate voice recognition for finding
think only pitfall we would have would be how much 's going to cost if means we have to cut down somewhere else
but think pretty much we used cheap relatively cheap and simple things
to s th 's not as expensive as would be if was full colour if we just got black and white one
one of decisions we had to make was whether we had double or triple batteries
inside components is pretty standardised across board is n't ?
on bottom we were gon na have rubber , rubber ones
we decided on rubber casing for plastic shell , variety of designs
although 's rubberised and spongy , apart from we 're just going to go for sort of vegetable and fruit colours
we 're not gon na try and make actually vegetable or fruit design or anything else , like shell we discussed
this image got in my head of something kind of shaped like , maybe about size , made of plastic , fits into palm of hand , rubberised cover 's spongy
thing is rubberised covers are going to be s relatively cheap to produce
if in year 's time we get feedback from design fairs show something else is coming mean 's easy for us to just produce and can be slipped on
not actual plastic outside case , just rubber thing goes round
for decisions we made , kinetic charging , watch-type batteries
we 're gon na use fruit and vegetable colours for rubber cover , case itself is plastic
we can if financially voice recognition is not feasible we could go for whistle
took from feedback you gave in last meeting from y from y your report , 's not just case of having little r_r hidden somewhere , they do want to be obvious 's our product
have we just rejected in favour of voice recognition then ?
but if turns out 's hugely expensive we 'd have to cut down somewhere else , then 's worth thinking about

プレーンテキストによる “ES2004c” の出力（その3）

－ 議論マップによる出力



議論マップによる “ES2004c” の出力（その1）

PM: can we add in mechanism to closing switches ?
ME: should n't we do some market research on first before we add in
PM: if we're doing kinetic thing
ME: should n't we do some market research on first before we add in
ME: not something 's come up in any of our focus groups and market research
ID: which basically has these contacts are really close together and then when you press down on rubber button , connect circuit
ID: each switch is connected to two different legs on chip and sends different message
ID: if you if you see how thin tracks are, you could you could s you could put them virtually right next to each other and have much smaller switch on each one
ID: you're saying people like this spongy feel this year
ID: perhaps some something made of rubber
ID: was just would be expensive to buy new case for, because of expense of how much titanium is
ME: was just wondering if we make basic mould out of plastic but have like rubber cover goes over
ME: would give spongy feel
PM: instead of fascia comes off being plastic, fascia comes off would be rubber
ID: just going back to interface designs with buttons
ID: sure you could probably get in different colours
ID: second option th they offer rubber buttons
ID: problem with basically is 's flat and you ca n't do lots of curve curved things
ID: or we could have rubber buttons which are made of this material which is and-r-s-1
UI: person who switches channels very often or does use particular function very often will find very irritating to
UI: regarding circuit , since we are havhaving flip-top, we can customise two circuits for different type or buttons
ID: mean you 'd you 'd actually have two separate circuit boards but they 'd be joined by wires or like some cabling between them
ME: possible when we open our flip-top 's thin compact wiring and when you press button then goes onto phone display th remote control display thing
PM: we're marketing to men as much as we are to women
ID: if you if you put full charge through L_c_d display completely blank out
Topic Segment-----
UI: we just took input from previous meeting , especially from marketing and industrial design , to check on customer needs and feasibility
UI: we are going to have m maybe g g-u-1 interface in f in middle of flip-top

議論マップによる“ES2004c”の出力（その2）

UI: we are going to have m maybe g g-u-1 interface in f in middle of flip-top
UI: on top m repetitive buttons which are like volume or channel changing and all could be on pointing device
UI: on s simpler board , on top we have this button, rubber buttons, to keep frequently changing channels
UI: there was mail which mentions our division has developed new speech recognition s feature
UI: we have to check into for financial f feasibility whether we can incorporate this at low cost
UI: like we had g s response from customers they would like to have feature of finding
UI: but we can take inspiration from fruit colours , like vibrant colours , red chilli
Topic Segment-----
PM: but we need to make decision about things we discussed
PM: we could incorporate voice recognition for finding
PM: think only pitfall we would have would be how much 's going to cost if means we have to cut down somewhere else
PM: but think pretty much we used cheap relatively cheap and simple things
ID: to s th 's not as expensive as would be if was full colour if we just got black and white one
ME: one of decisions we had to make was whether we had double or triple batteries
PM: inside components is pretty standardised across boards is n't?
ID: think were we're going for L_c_d on lines , on buttons
ID: on bottom we were gon na have rubber, rubber ones
ME: we decided on rubber casing for plastic shell, variety of designs
ME: although 's rubberised and spongy, apart from we're just going to go for sort of vegetable and fruit colours
ME: we're not gon na try and make actually vegetable or fruit design or anything else, like shell we discussed
ID: no
ME: this image got in my head of something kind of shaped like, maybe about size , made of plastic , fits into palm of hand , rubberised cover 's spongy
ME: thing is rubberised covers are going to be s relatively cheap to produce
ME: if in year's time we get feedback from design fairs show something else is coming mean 's easy for us to just produce and can be slipped on
ID: not actual plastic outside case , just rubber thing goes round
Topic Segment-----
ME: for decisions we made, kinetic charging , watch-type batteries
ME: we're gon na use fruit and vegetable colours for rubber cover, case itself is plastic

議論マップによる“ES2004c”の出力（その3）

ME: we're gon na use fruit and vegetable colours for rubber cover , case itself is plastic
UL: we can if financially voice recognition is not feasible we could go for whistle
ME: took from feedback you gave in last meeting from y from y your report , 's not just case of having little r_f hidden somewhere , they do want to be obvious 's our product
ME: have we just rejected in favour of voice recognition then ?
PM: depending on how
PM: but if turns out 's hugely expensive we 'd have to cut down somewhere else , then 's worth thinking about

議論マップによる “ES2004c” の出力（その4）

- “TS3003b” に関する質問と出力

- 質問

Q1. What is the theme of the meeting?

A1. kick-off-meeting.

A2. functional design.

A3. component concepts.

A4. industrial design.

Q2. Who is the target for the new product they developing?

A1. younger people.

A2. elderly people.

A3. especially for small kids.

A4. there is no target.

Q3. Is there enough money for docking station?

A1. yes.

A2. no.

A3. depending on the quality.

A4. anyone didn' t mention.

- プレーンテキストによる出力

TS3003b
now we're going into functional design
important thing of this phase is we're going to try to get agreement about user requirements , technical function design , and working
we can have little discussion about new project requirements
we have we have been doing research in usability lab
where we observed users operating remote controls
we let them fill out questionnaire
additional eighty percent indicated they would spend more money on fancy-looking remote control
fifty percent of people indicated they only use about ten percent of buttons on remote control
audio settings , video settings , and channel settings buttons
functionality as lot of people indicated , they only use about ten percent of buttons
think we should make very few buttons
think most frequently used buttons should be emphasised
as many as lot of people indicated their remote got lost in room
but might be good idea to make docking station
this would also greatly increase production costs
but think these are just some small factors we could consider
do some little research on internet
but this remote control got lot of buttons
like on-off button
use display for specific functions of different device
you got here general functions , like on-off button sound dunno
everybody have to use of even old people young
we must keep buttons quite simple and quite large
seen something on internet not today but few weeks ago
you got quite kind of touchscreen and it's for twenty euros or less
chip controls infrared bulb and perhaps normal
think we should use default materials , simple plastics

プレーンテキストによる “TS3003b” の出力（その1）

think we should use default materials , simple plastics
think we should focus on aesthetics , design and user interface
because if you're going to use high-tech materials price is going to go sky-high
because in world of upcoming internet we think teletext is going to be thing of past
internet is also mentioned in function we can use
because we are going to develop new product which is specially designed for younger customers
suggest let's have discussion on control functions
there any discussion possible about new product requirement ?
you're you're saying teletext is gonna be old feature
but think if you're targeting people of forty plus , chance they will have t.v with internet access
in addition people indicated teletext simply was important feature for remote control
think it's pretty dumb to put no teletext feature on
besides , think market for forty plus is like pretty small
if we put lot of effort in those , we can make remote control with just two or three buttons
but don't you think if we make remote which is typically made for this market , people think people think it's device looked for although didn't realise
we do n't have to focus on design then but on functionality
new products we are gonna make , specifically design , are designed for younger people
think we have to see what requirements we need for those remote controls
think it'd definitely be bad idea not to include teletext
maybe we can make another point of advantage in our remote control , if we make a button for example for big subtitles
for elderly people they can think , wanna have subtitles
think teletext can be very useful in our advantage
they can see how much buttons there are going to be on display
but don't think if you're gonna make remote control only to operate t.v , you there's not much you can gain on having as few buttons as possible
but think there's simply not much to gain on competition when you when you're making remote control only for to operate only t.v
if you have remote control only to operate t.v there's simply not lot of buttons required
most existing remote controls simply do n't have lot of buttons either
you suggest we could better focus on for example docking station

プレーンテキストによる “TS3003b” の出力（その2）

you suggest we could better focus on for example docking station
do you think docking station will be allowed in budget we have ?
might be good idea to have just little screen on
think probably elderly people would be little bit more careful with their remote controls than youngsters
what we're going to do for next meeting is industrial designer will do components concept
just keep in mind things we said about target group , requirements , and trends which are going on
specific instructions will be sent to you by personal coach

プレーンテキストによる “TS3003b” の出力（その3）

– 議論マップによる出力

TS3003b
Topic Segment-----
PM: now we're going into functional design
PM: important thing of this phase is we're going to try to get agreement about user requirements , technical function design , and working
PM: we can have little discussion about new project requirements
Topic Segment-----
ME: we have we have been doing research in usability lab
ME: where we observed users operating remote controls
ME: we let them fill out questionnaire
ME: additional eighty percent indicated they would spend more money on fancy-looking remote control
ME: fifty percent of people indicated they only loo used about ten percent of buttons on remote control
ME: audio settings , video settings , and channel settings buttons
ME: functionality as lot of people indicated , they only use about ten percent of buttons
ME: think we should make very few buttons
ME: think most frequently used buttons should be emphasised
ME: as mo as lot of people indicated their re mo te got lost in room
ME: but might be good idea to make docking station
ME: this would also greatly increase production costs
ME: but think these are just some small factors we could consider
Topic Segment-----
UI: do done little research on internet
UI: but this re remote control got lot of buttons
UI: like on-off button
UI: use display for specific functions of different device
UI: you got here general functions , like on-off button sound dunno
UI: everybody have to use ol even old people young
UI: we must keep buttons quite s simple and quite large
UI: seen w something on internet not today but few weeks ago

議論マップによる “TS3003b” の出力（その1）

UI: seen w something on Internet not today but few weeks ago
UI: you got quite kind of touchscreen and 's for twenty euros or less
Topic Segment——
ID: chip controls infrared bulb and perf perhaps normal
ID: think we should use default materials , simple plastics
ID: think we should focus on aesthetics , design and user interface
ID: because if you 're going to use high-tech materials price is going to go sky-high
Topic Segment——
PM: because in world of upcoming Internet we think teletext is going to be thing of past
PM: Internet is also mentioned in function we can use
PM: because we are going to develop new product which is specially designed for younger customers
Topic Segment——
PM: suggest let 's have discussion on control functions
ME: there any discussion possible about new product requirement ?
PM: we can see if we can find way between functions we wan na use and market we wan na reach with our product
ME: you 're you 're saying teletext is gon na be old feature
ME: but think if you 're targeting people of forty plus , chance they will have t_v with Internet access
ME: in addition people indicated teletext simply was important feature for remote control
ME: think 's pretty dumb to put no teletext feature on
ME: besides , think market for forty plus is like pretty small
PM: if we put lot of effort in those , we can make remote control with just two or three buttons
PM: but do n't you think if we make remote which is typically made for this market , people think people think 's device looked for although did n't realise
ME: no
ID: we do n't have to focus on design then but on functionality
PM: new products we are gon na make , spec specifically design , are designed for younger people
PM: think we have to see what requirements we need for those remote controls
ME: think 'd definitely be bad idea not to include teletext
PM: maybe we can make another point of advantage in our remote control , if we make k button ex for example for big subtitles

議論マップによる “TS3003b” の出力（その2）

PM: maybe we can make another point of advantage in our remote control , if we make k button ex for example for big subtitles
PM: for elderly people they can think , wan na have subtitles
PM: think teletext can v can be very useful in our advantage
PM: they can see how much buttons there are going to be on display
ME: but do n't think if you 're gon na make remote control only to operate t_v , you there 's not much you can gain on having as few buttons as possible
ME: but think there 's simply not much to gain on competition when you when you 're making remote control only for to operate only t_v
ME: if you have remote control only to operate t_v there 's simply not lot of buttons required
ME: most existing remote controls simply do n't have lot of buttons either
PM: you suggest we could better focus on for example docking station
ME: maybe
PM: do you think docking station will be allowed in lounge we have ?
ID: should be possible yes
ME: might be good idea to have just little screen on
ME: think probably elderly people would be little bit more careful with their remote controls than youngsters
Topic Segment——
PM: what we 're going to do for next meeting is industrial designer will do components concept
PM: just keep in mind things we said about target group , requirements , and trends which are going on
PM: specific instructions will be sent to you by personal coach

議論マップによる “TS3003b” の出力（その3）

A.2 被験者実験における質問と出力（重要フレーズの強調と構造化の比較）

- “TS3003c” に関する質問と出力

- 質問

Q1. Who is the target for the new product they developing?

A1. younger people.

A2. elderly people.

A3. especially for small kids.

A4. there is no target.

Q2. Which energy source do they think more fancy?

A1. battery.

A2. solar cells.

A3. kinetic energy.

A4. not mentioned in the meeting.

Q3. Which case is ” NOT” discussed in the meeting?

A1. wood case.

A2. changeable case.

A3. metal case.

A4. curved cases.

- 重要フレーズを強調した出力

TS3003c
PM: this is our third meeting already
PM: this is second phase we're going to discuss today
PM: conceptual design meeting
PM: few points of interest in this meeting are conceptual specification of components
PM: little thing about last meeting
PM: requirements are teletext, docking station, audio signal , small screen , with some extras button information
PM: we're going to target sixty to eighty yearold customers
ME: consulted some additional trend-watch trend-watchers , after original trend-watchers return , about what best design would be
ME: about half of , half as important as innovativeness was easy to use
ME: for our group , we're focusing on people of sixty to eighty y years old, this is, these factors are slightly more equal
ME: they would prefer design where remote control just lies flat in docking station
ME: besides we would advise to bring two editions, one with wood-like colour and maybe feel, and one with grey-black
ME: grey-black colour for people with less means
UI: like speak speech recognition and special button for selecting subtitles
UI: using large buttons
UI: possible to make quite cheap system for speech recognition
UI: you can think about when you lost your remote control , you can call and gives sig signal
UI: we got buttons we have to use
UI: also we want to use little d display for displaying functions of buttons
UI: just we can put on-off button over in this corner
UI: almost e all remote controls are using on-off button on place

重要フレーズを強調した“TS3003c”の出力（その1）

UI: almost e all remote controls are using on-off button on place
UI: for our design , we have to discuss about think form of
ID: for energy source we can use basic battery
ID: or, as optional thing , kinetic energy , like in watch
ID: as case supplement , we could, thought of I later, rubber belt , like anti-slip
ID: for b buttons , we can use plastic or rubber
ID: form should follow function overall
ID: kinetic energy source is rather fancy
ID: for case, supplement and buttons , really depends on designer
PM: brings us to discussion about our concepts
PM: first think we can talk about energy source
PM: think first of all we have to see is possible to introduce kinetic energy in our budget ,
ME: how exactly does kinetic energy work ?
ME: think our docking station could be one of marketing issues with which we can
PM: do n't think they like remotes just laying everywhere in their rooms
PM: maybe docking station will help them give remote place
ME: wondering what will voice recognition mean for production price ?
ME: we also asked if w they would, if people would pay more for speech recognition in remote control
ME: on top of l.c.d screen would help in making remote control easier to use
UI: but is our research about bl large l.c.d sh display, or just small one we want to use ?
PM: personally think l.c.d screen we wan na use, with extra information , think nobody has anything against
PM: maybe we have to discard voice recognition
PM: if we do n't have voice recognition , will wo n't use lot of energy to

重要フレーズを強調した“TS3003c”の出力（その2）

PM: if we do nt have voice recognition , will we nt use lot of energy to
PM: but think just simple battery which you can reload on docking station is just as good
PM: wood is lot more expensive to produce
UI: possible to make changeable case
ME: we would have to make one w control which would fit in with wooden cover and plastic
PM: on which you can apply , remote controls , on which you can apply different case covers , for example
ME: but would have to be fancy with either wood cover or plastic one
PM: everybody s with changing covers ?
ME: heard our industrial designer talk about flat , single and double curved
ID: our d manufacturing department can also deliver single curved or double curved ca curved cases
ME: what would single curved and double curved mean ?
ME: but if we would make changing channels and changing volume button on both sides , would certainly yield great options for design of remote
ID: single curved not really sure what they 're gon na look like , but think 's something like this
ID: double curved s looks something like this guess
PM: suggest single curved
UI: but on-off button , still on top
PM: industrial designer and user interface designer are going to work together on this one
PM: but you 're going to get your instructions think send by coach
PM: make good design for all grandfathers and grandmothers , think

重要フレーズを強調した“TS3003c”の出力（その3）

– 構造化を行った出力

TS3003c
Topic Segment-----
PM: this is our third meeting already
PM: this is second phase we 're going to discuss today
PM: conceptual design meeting
PM: few points of interest in this meeting are conceptual specification of components
PM: little thing about last meeting
PM: requirements are teletext , docking station , audio signal , small screen , with some extras button information
PM: we 're going to target sixty to eighty yearold customers
Topic Segment-----
ME: consulted some additional trend-watch trend-watchers , after original trend-watchers return , about what best design would be
ME: about half of , half as important as innovativeness was easy to use
ME: for our group , we 're focusing on people of sixty to eighty y years old , this is , these factors are slightly more equal
ME: they would prefer design where remote control just lies flat in docking station
ME: besides we would advise to bring two editions , one with wood-like colour and maybe feel , and one with grey-black
ME: grey-black colour for people with less means
Topic Segment-----
UI: like speak speech recognition and special button for selecting subtitles
UI: using large buttons
UI: possible to make quite cheap system for speech recognition
UI: you can think about when you lost your remote control , you can call and gives sig signal
UI: we got buttons we have to use
UI: also we want to use little d display for displaying functions of buttons
UI: just we can put on-off button over in this corner
UI: almost e all remote controls are using on-off button on place
UI: for our design , we have to discuss about think form of
Topic Segment-----

構造化を行った“TS3003c”の出力（その1）

Topic Segment-----
ID: for energy source we can use basic battery
ID: or , as optional thing , kinetic energy , like in watch
ID: as case supplement , we could , thought of I later , rubber belt , like anti-slip
ID: for b buttons , we can use plastic or rubber
ID: form should follow function overall
ID: kinetic energy source is rather fancy
ID: for case , supplement and buttons , really depends on designer
Topic Segment-----
PM: brings us to discussion about our concepts
PM: first think we can talk about energy source
PM: think first of all we have to see is possible to introduce kinetic energy in our budget .
ID: there are four options
ME: how exactly does kinetic energy work ?
ID: you basically shake your remote , and then powers up
ME: think our docking station could be one of marketing issues with which we can
PM: do n't think they like remotes just laying everywhere in their rooms
PM: maybe docking station will help them give remote place
ME: wondering what will voice recognition mean for production price ?
ID: do n't have any information on pricing
ME: we also asked if w they would , if people would pay more for speech recognition in remote control
ME: on top of l_c_d screen would help in making remote control easier to use
UI: but is our research about bi large l_c_d sh display , or just small one we want to use ?
ME: this was for like l_c_d screen like you would have on most advanced mobile phones
PM: personally think l_c_d screen we wan na use , with extra information , think nobody has anything against
PM: maybe we have to discard voice recognition
PM: if we do n't have voice recognition , will wo n't use lot of energy to

構造化を行った“TS3003c”の出力（その2）

PM: if we do n't have voice recognition , will wo n't use lot of energy to
PM: but think just simple battery which you can reload on docking station is just as good
PM: wood is lot more expensive to produce
UI: possible to make changeable case
PM: change cases
ME: we would have to make one w control which would fit in with wooden cover and plastic
PM: on which you can apply , remote controls , on which you can apply different case covers , for example
ME: would n't design telephone
ME: but would have to be fancy with either wood cover or plastic one
PM: everybody's with changing covers ?
ID: yes
ME: heard our industrial designer talk about flat , single and double curved
ID: our d manufacturing department can also deliver single curved or double curved ca curved cases
ME: what would single curved and double curved mean ?
ID: would just only affect form , for as far as know
ME: but if we would make changing channels and changing volume button on both sides , would certainly yield great options for design of remote
ID: single curved not really sure what they're gon na look like , but think 's something like this
ID: double curved s looks something like this guess
PM: suggest single curved
UI: but on-off button , still on top
PM: still here
Topic Segment-----
PM: industrial designer and user interface designer are going to work together on this one
PM: but you're going to get your instructions think send by coach
PM: make good design for all grandfathers and grandmothers , think

構造化を行った“TS3003c”の出力（その3）

- “TS3007c” に関する質問と出力

- 質問

Q1. Which material did they decide to use?

A1. plastic.

A2. rubber.

A3. both.

A4. not decided in the meeting.

Q2. Which function is ” NOT” mentioned in the meeting?

A1. parent control.

A2. games.

A3. voice recognition.

A4. touch screen.

Q3. How did they think about ” Parent Control” ?

A1. default built in.

A2. default built in and make menu option.

A3. don’ t prepare.

A4. not mentioned.

- 重要フレーズを強調した出力

TS3007c
PM: welcome to conceptual design meeting
PM: we must reach decision on remote control concepts
ID: what did was reviewed previously used Items by twodifferent age groups
ID: shapes are curved and round
ID: but twobasic sources found were best possible were battery versus kinetic
ID: figured out , seeing you always be within i distance of t_v , and from t_v there comes whole bunch of light
ID: prefer kinetic because 's , costs are nt much higher , and , jujust bit more high-tech than normal battery
ID: we could go for single curved or double
ID: gives you younger and more high-tech look , i figured
ID: furthermore , plastic versus rubber ?
ID: you we could choose what 's better , plastic or rubber
ID: if we are going to use traditional buttons , we could go with simple chip set
ID: but if we decide to go on n l c_d screen , we would use we have to use advanced chip set
ID: what are your ideas concerning battery versus kinetic ?
UI: think you should use kinetic as back-up
ID: you use battery and w charge up with kinetic
ID: we have single-curved versus double-curved
ID: plastic versus rubber

重要フレーズを強調した“TS3007c”の出力（その1）

ID: plastic versus rubber
UI: nt possible to make combination with kind of rubber s or bendable remotes where you got
ME: but if you have square L_c_d screen , and you put case around has like bulbs or covers part of L_c_d
ME: in paris and milan , we asked different people , differing in age and in income , amount of money to spend , what they like in design and material nowadays
ME: fresh colours , green , red , strawberry red , yellow , banana
PM: but wood is not material you which you build remote control of
ME: here , think about removable covers , as seen in mobile telephon telephone market
ME: third , material being used could be stuff like rubber
ME: in my second sheet or personal preferences , we have to reconsider speech function recognition
ME: building games like snake or tetris to amuse younger users
ME: third , stick with , log-in functionality with slogan , take parental control to new dimension
UI: make menu structure changeable
UI: middle like all remotes have to be in little bit small , you can put in inside your hands
UI: but if you 're programming colour of t_v , do not display twentyother options are possible
PM: parental control , games , and voice recognition
ME: how see is you put on remote and then you have like windows log-on screen with parents , children
PM: no but think most people find much more important their children do nt watch sex or violence on television , and wait ten or fifteen seconds longer , they can finally watch because of
UI: if you w you push p children , you do nt have to log in , but you can only watch children 's channels
UI: just make option in menu , its install powerful profiles or something
UI: you 're thinking about some channels they can not see , but , when think , for f for all standard channels , and only for maybe after ten o'clock in eveningviolent films and movies will come

重要フレーズを強調した“TS3007c”の出力（その2）

UI: you're thinking about some channels they can not see, but, when think, for all standard channels, and only for maybe after ten o'clock n evening violent films and movies will come
PM: but maybe there's some option t kind of show view numbers are violent, and they are blocked out
UI: mail too from management 's very popular to use voice recon recognition
ME: like r voice recognition is too much, think
UI: some menu function, you choose parental control, and maybe four files will emerge from menu where put on
PM: like first idea from you can buy without and with parental contr control ?
UI: you can just s when you buy, you can select personal preference, parental control on, and password or something
ID: figure if you had twodifferent remotes, you could be choose one with receiver in
ID: you could actually easily build in parental control
PM: case will be double-curved and rubber, n frlflashy fruity colour with cover is removable
ME: maybe 's handy to build in expert view and simple
UI: when you push system properties, entire list pops up with e tiv as kind of t stuff you can program
UI: you could actually build in function you can program yourself, for more advanced users
PM: another thing is there were there was survey, and came out people like to buy things from single large provider, instead of those who are partnering with us
ID: if you if you have l l c, d screen, you can you can have small logo a bottom
PM: industrial development centre and user interface will work together on prototype
PM: marketing expert will be keeping busy with product evaluation
PM: but most specific instructions will be sent to you by email

重要フレーズを強調した“TS3007c”の出力（その3）

– 構造化を行った出力

TS3007c
Topic Segment-----
PM: welcome to conceptual design meeting
PM: we must reach decision on remote control concepts
Topic Segment-----
ID: what did was reviewed previously used items by twodifferent age groups
ID: shapes are curved and round
ID: but twobasic sources found were best possible were battery versus kinetic
ID: figured out, seeing you always be within l distance of t, v, and from t, v there comes whole bunch of light
ID: prefer kinetic because 's, costs are n't much higher, and, jjust bit more high-tech than normal battery
ID: we could go for single curved or double
ID: gives you younger and more high-tech look, f figured
ID: furthermore, plastic versus rubber ?
PM: soft
ID: you we could choose what 's better, plastic or rubber
ID: if we are going to use traditional buttons, we could go with simple chip set
ID: but if we decide to go on n l c, d screen, we would use we have to use advanced chip set
ID: what are your ideas concerning battery versus kinetic ?
UI: think you should use kinetic as back-up
UI: think you should use kinetic as back-up
ID: you use battery and w charge up with kinetic
ID: we have single-curved versus double-curved
UI: maybe have something in my presentation to cope with
ID: plastic versus rubber

構造化を行った“TS3007c”の出力（その1）

id: plastic versus rubber
ME: rubber
ui: n't possible to make combination with kind of rubber is or bendable remotes where you got
id: figured will be m rather than hard
ME: but if you have square l.c.d screen , and you put case around has like bulbs or covers part of l.c.d
-----Topic Segment-----
ME: in paris and milan , we asked different people , differing in age and in income , amount of money to spend , what they like in design and material nowadays
ME: fresh colours , green , red , strawberry red , yellow , banana
PM: but wood is not material you which you build remote control of
ME: here , think about removable covers , as seen in mobile telephone telephone market
ME: third , material being used could be stuff like rubber
ME: in my second sheet of personal preferences , we have to reconsider speech function recognition
ME: building games like snake or tetris to amuse younger users
ME: third , stick with , log-in functionality with slogan , take parental control to new dimension
-----Topic Segment-----
ui: make menu structure changeable
ui: middle like all remotes have to be in little bit small , you can put in inside your hands
ui: but if you're programming colour of t.v , do not display twentyother options are possible
-----Topic Segment-----
PM: parental control , games , and voice recognition
ME: how see is you put on remote and then you have like windows log-on screen with parents , children
PM: no but think most people find much more important their children do n't watch sex or violence on television , and wait ten or fifteen seconds longer , they can finally watch because of
ui: if you w you push p children , you do n't have to log in , but you can only watch children 's channels
ui: just make option in menu , ins install powerful profiles or something

構造化を行った“TS3007c”の出力（その2）

ui: just make option in menu , ins install powerful profiles or something
ui: you're thinking about some channels they can not see , but , when think , for f for all standard channels , and only for maybe after ten o'clock in evening violent films and movies will come
PM: but maybe there 's some option i kind of show view numbers are violent , and they are blocked out
ui: mail too from management 's very popular to use voice recon recognition
ME: like f voice recognition is too much , think
ui: some menu function , you choose parental control , and maybe four files will emerge from menu where put on
PM: like first idea from you can buy without and with parental contr control ?
ui: ye think best would be to put in and make menu option
ui: you can just s when you buy , you can select personal preference , parental control on , and password or something
id: figure if you had twodifferent remotes , you could bo choose one with receiver in
id: you could actually easily build in parental control
PM: case will be double-curved and rubber , in frlflashy fruity colour with cover is removable
ME: maybe 's handy to build in expert view and simple
ui: when you push system properties , entire list pops up with e ev all kind of f stuff you can program
id: you could actually build in function you can program yourself , for more advanced users
PM: another thing is there were there was survey , and came out people like to buy things from single large provider , instead of those who are partnering with us
id: if you if you have l.c.d screen , you can you can have small logo at bottom
-----Topic Segment-----
PM: industrial development centre and user interface will work together on prototype
PM: marketing expert will ke he will be keeping busy with product evaluation
PM: but most specific instructions will be sent to you by email

構造化を行った“TS3007c”の出力（その3）

A.3 t 分布表

t -分布表

$f = n - 1$	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.289	0.500	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.277	0.476	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.271	0.464	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
6	0.265	0.453	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.263	0.449	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.262	0.447	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.261	0.445	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.260	0.443	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.259	0.442	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.259	0.441	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.258	0.440	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.258	0.439	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.257	0.438	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.257	0.438	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.257	0.438	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.257	0.437	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.257	0.437	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.256	0.437	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.256	0.436	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.256	0.436	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
26	0.256	0.436	0.684	0.856	1.177	1.315	1.706	2.056	2.479	2.779	3.435
27	0.256	0.435	0.684	0.855	1.176	1.314	1.703	2.052	2.473	2.771	3.421
28	0.256	0.435	0.683	0.855	1.175	1.313	1.701	2.048	2.467	2.763	3.408
29	0.256	0.435	0.683	0.854	1.174	1.311	1.699	2.045	2.462	2.756	3.396
30	0.256	0.435	0.683	0.854	1.173	1.310	1.697	2.042	2.457	2.750	3.385
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090