

知的支援機能を備えた電子会議システム

ELECTRONIC DISCUSSION SYSTEM WITH INTELLIGENT SUPPORTING FUNCTIONS

古田一雄¹・前原基芳²・高島亮祐³・中田圭一⁴

¹Ph.D. (工学) 東京大学大学院 新領域創成科学研究科環境学専攻 (E-mail: furuta@k.u-tokyo.ac.jp)

²M.A. (環境学) 東京大学大学院 新領域創成科学研究科環境学専攻

³M.A. (工学) 東京大学大学院 工学系研究科システム量子工学専攻

⁴Ph.D. (人工知能) 東京大学大学院 新領域創成科学研究科環境学専攻 (E-mail: keiichi.nakata@i-u.de)

社会的決定においては公衆の参加による社会的合意形成への要望が高まる一方、インターネットの急速な展開にともない、電子掲示板などを用いた電子会議が社会的合意形成に一定の役割を果たすようになった。本研究では、こうした電子会議を用いた合意形成における参加者の発言内容や会議の進行に対する理解を深め、円滑な合意形成を支援するために、会議発言録から話題を抽出し、要約を作成して参加者に提示する手法を開発した。さらに開発手法を組み込んだ電子会議システム TSS (Transcript Summarization System)を開発し、既存の会議発言録を用いて機能確認を行った。

キーワード：社会的合意形成，電子会議，話題抽出，自動要約，統計的自然言語処理

1. はじめに

これまで迷惑施設立地，公共事業，先端科学技術の受容などに関して社会的決定が行われる際には，専門家の意見に基づいて行政が決定し，公衆には結果だけが報告されることが多かった。しかし近年，社会的決定の場に公衆が参加し，合意に基づく決定が求められる機会が増えている。しかし社会的合意形成においては，多様な人々が参加することにより論点が広がり，議論の齟齬や誤解が生じて合理的決定を円滑に行うことが困難であることなどの問題が予想される。こうした問題を改善するには，発言内容や会議の進行に対する各参加者の理解を深めるための何らかの方策が必要であると考えられる。

一方，近年のインターネットの急速な展開により電子情報ネットワークを介したコミュニケーションが職場から家庭までのあらゆる局面で広がっている。¹⁾例えば掲示板 (BBS) と呼ばれるシステムを通じ，不特定多数の参加者による特定テーマに関する意見・情報の交換が行われている。これらは通常，匿名性を前提とした自由な発言の場としての性質が強いが，マルチメディアなどさまざまな形式の情報の利用や，時間や空間に束縛されず多数の利害関係者の参加を可能にする電子会議としての側面が注目できる。

しかし電子会議を社会的合意形成の手段として適用するにはまだ多くの問題点がある。電子会議の目的を参加者間の合意形成とし，その参加者による協調作業と捉え

た場合，その協同作業の対象は問題認識・問題解決・社会的意思決定といった会議そのものである。このように作業対象が主観的・非明示的であり，見る人や立場により解釈が異なるような場合，協調作業を進めることは困難である。そこで議論を構造化したり，発言間の関係を明示するための方法がいくつか提案されている。なかでも非同期分散型の議論支援グループウェアの代表的な例として挙げられる IBIS (Issue-Based Information System) では意見に対する賛否を示しながら議論を進め，これらの意見をハイパーテキストで結びつけて構造化していきながら議論の進行および記録を支援することを目的としている。²⁾

しかし IBIS では発言における二項関係のみしか捉えることができず，発言の内容についての分析はされないため大局的な意見や立場の把握は困難である。また，テキスト情報の内容を理解することに対しては十分な考慮がされていない。会議の内容を理解するためには全発言録を読まなければならない，相当な時間と労力が必要であり，これは特に途中からの参加者にとって多大な負担となる。そこで，会議記録からの話題抽出や自動要約の機能を提供することは，参加者が会議の内容を簡単に理解する上で効果が大きいと考えられる。

以上の背景をふまえ，本研究では会議の発言録に対する話題抽出，自動要約の手法を開発し，要約の結果を明瞭な形で参加者に提示することで会議内容の理解を支援することを目的とする。

2. 話題抽出

2.1. 特徴語抽出のための尺度

会議の発言録に特化した話題抽出や自動要約の研究はこれまで報告された例はないが、既存の要約手法、情報検索、統計的自然言語処理における手法が応用できると考えた。しかし発言録特有の特徴のため、一般的な文献に対する検索や要約の手法をそのまま適用したのでは、なかなか意味のある結果が得られない。発言録の処理においては発言録の特徴を捉え適切な対応を考える必要があるが、発言録の特徴としては、

- a. 話題が推移する。
- b. 話し言葉である。
- c. 発言の長さが極端に異なる場合がある。

という3点があげられる。

話題が推移する問題を解決するために、発言録を話題ごとに自動分割することを考えた。もし同じ話題についての発言であれば、抽出される特徴語は同じものが多いはずであるが、話題が変わると、違う特徴語が抽出されると考えられる。この特性を利用して、話題ごとに発言録を分割する。

特徴語の選択のための統計的尺度としては、 tf (term frequency), idf (inverse document frequency), $tf \cdot idf$, 相互情報量, 対数尤度比などが利用されている。³⁾ それぞれ一長一短あるが、頻度と情報量とをうまく組み合わせ、低頻度、高頻度語それぞれに対する重みのバランスをとることが必要である。そこで、語の網羅性と特定性を考慮した $tf \cdot idf$ を用いた。自動構文解析は構文的不完全性の大きい話し言葉にはうまく適用できないため、このような統計的手法により話し言葉である問題にも対応する。

語 w_i , 文書単位 d_j に対する $tf \cdot idf$ の定義は以下の式で与えられる。

$$tf \cdot idf(w_i, d_j) = \frac{F_{ij}}{F_i} \log \frac{N}{N_i} \quad (1)$$

ここで F_{ij} は語 w_i の文書単位 d_j における出現数, F_i は全文書中の出現数である。また, N_i は語 w_i を含む文書単位数, N は文書単位の総数である。文書単位としては、発言者ごとの全発言, ある時間継続して行われた1回の発言機会, その他特定長の発言などを、分析目的に応じて使い分けることとする。

形態素解析には「茶筌」⁴⁾ を利用し、分析対象語には名詞のみを用いた。また、さらに名詞のうち、形態素解析の際の明らかなエラー、1文字の語、代名詞などを不要語リストに登録して除外した。

2.2. 発言者ごとの話題の抽出

文書単位を発言者ごとの全発言とし、 $tf \cdot idf$ の高い語を抽出すれば、各発言者が興味を持っている話題を抽出できると考えられる。一例として、1996年に開催された「原子力政策円卓会議」⁵⁾ の第1回発言録を対象に、 $tf \cdot idf$ の高い特徴語を参加者ごとに上位4位まで示すとTable 1 のようになる。

Table 1 発言者を特徴づける語
(「原子力政策円卓会議」第1回)

発言者	特徴語	$tf \cdot idf$
A	モデレーター	0.0347
	休憩	0.0235
	事務局	0.0235
	反応	0.0189
B	立地地域	0.0328
	プルサーマル計画	0.0246
	合意形成	0.0226
	地元	0.0198
C	広報活動	0.0424
	地域住民	0.0318
	地方自治体	0.0239
	放射線	0.0239
D	振動	0.0538
	設計不良	0.0269
	核燃料リサイクル技術	0.0269
	平和利用	0.0269

これを見ると各参加者が興味を持つ話題の特徴がある程度抽出できている。たとえば、A は会議そのものの進行についての語を多用しているが、これはAがこの会議のモデレータであるためである。またBとCは原子力発電の建設に関わる地域問題、Dは原子力の技術的問題にそれぞれ関心があることがうかがわれる。BとCは原子力産業界の代表者、Dは原子力工学を専門とする学識経験者であることを考えると、上記の関心は納得できる。ただし、各発言者の興味は抽出できているとしても、それぞれの発言者が会議の中でどのような立場にあるのかは、これだけでは読みとれない。

2.3. 語のクラスタリングによる話題抽出

会議における話題を抽出するもう一つの方法として、発言録に出現する語をクラスタリングすることを試みた。クラスタが得られれば、発言者がどのクラスタ(話題)の語を用いているか見ることによって、各発言者が興味

を持つ話題を比較することができる。語のクラスタリング方法としてさまざまな可能性があるが、同一話題に関する語は発言中で共起することが多いと考え、語同士の共起関係を表す共起グラフを分割してクラスタを作った。共起強さを示す共起度は、次式で計算する。⁶⁾

$$C(w_i, w_j) = P(w_i, w_j) \log \frac{P(w_i, w_j)}{P(w_i)P(w_j)} \quad (2)$$

ただし、 $P(w_i)$ は語 w_i の出現頻度、 $P(w_i, w_j)$ は語 w_i と w_j の共起頻度である。この共起度を用いて、発言録から共起関係を取り出し、共起グラフを作成する。

つぎに、この共起グラフ G に対して推移律に基づくクラスタリング⁷⁾を行う。アルゴリズムは以下の通りである (Fig. 1)。

- 1) $G_0 = \phi, i = 0$
- 2) G_0, \dots, G_i に含まれない互いにリンクした 3 ノードによって構成される任意の三角形を G から選び G_i に入れる。
- 3) G_i のノードによって作られる任意の三角形について、 G_0, \dots, G_i に含まれないノード v と 3 ノードとの間にリンクがあるとき、 v を G_i に入れる。さらに、 v と G_i のノードとの間のリンクをすべて G_i に入れる。
- 4) 3) を G_i が拡張できなくなるまで繰り返す。
- 5) $G_{i+1} = \phi, i = i+1$ 、として 2) に戻る。

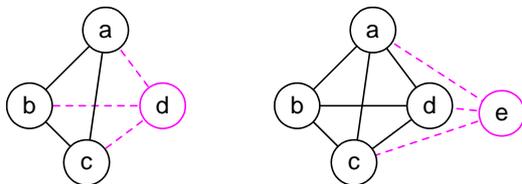


Fig. 1 推移率に基づくグラフのクラスタリング

どこまで強い共起関係を取り出すかによって結果は異なり、共起度の閾値が低いと大きな一つのクラスタ、高いと多数の小さなクラスタができていく。ここでは語数 10 以上のクラスタが最も多く得られるように閾値を調整することにした。ここで用いた閾値が一般的に適切であるか否かは、さまざまな会議を対象とするテストによって確認する必要があるが、人間の理解に適したクラスタ数には直接記憶範囲から常識的な上限があると考えられる。

前出の「原子力政策円卓会議」第 1 回発言録に本手法を適用した例を示す。延べ 777 語から 104 個のクラスタが作られたが、10 語以上を含む大きいクラスタは以下に示す 11 個であった。

- 依存 選択 システム ピーク エネルギー 日本 平準 議論 夏場 高校野球 …
- 持続 技術レベル 中断 蓄積 技術発展 阻害 散逸 維持 モラトリアム提案 …
- 火力発電 水準 堅持 太陽光発電 化石燃料 地球 エネルギー源 石炭 石油 …
- エネルギー供給構造 導入 二酸化炭素問題 増大 エネルギー分野 規模 国内 選択肢 …
- 福井 少数 福島 総合エネルギー調査会原子力 改定 報道 実施 中間報告 弾力 …
- 行政機関 審議 ブラックボックス 政策議論 大学研究機関 ヒューマンスケール …
- ソフト ハード先行 肥大 手続き 法律 …
- 原子炉設置許可 上程 調整審議会 申請 …
- 総理 具体 指示 地域フォーラム 試み …
- 原子力連絡調整 機能結合 原子力公報専門 …
- 原発反対 保守党 資格 議員 クーラー …

以上の結果を見ると、各クラスタが何らかの共通概念を表していることが示唆されるが、このクラスタリング手法の性能については第 4 章で評価する。

なお、本手法を複数回の会議の発言録や、規模の大きな発言録に適用した場合、有意なクラスタが得られない場合がある。これは、会議全体を対象とすると共起関係が広範囲に及ぶようになるために、共起度の閾値を高くしてもほとんどの語が巨大なクラスタに包含されてしまうためである。このような問題に対処するためには、次節に述べるような話題による分割を行ってからクラスタリングすることが必要となる。

2.4. 話題による発言録の分割

発言録における話題の推移に対処するため、類似の話題が議論されている範囲で発言録を複数ブロックに分割することを試みた。その手法としては、まず 10 発言単位ずつを文書単位とし、移動平均をとるような形で文書単位的位置を 1 発言ずつずらしながら $f \cdot idf$ を計算し、各部の特徴語を抽出する。つぎに、2 つ先の文書単位との上位特徴語 50 語の重複率を類似度として求め、類似度が大きく低下する位置を話題の転換点とみなす。10 発言を文書単位としたのは、文書単位が小さすぎても大きすぎても特徴がつかめないことから、何通りかの試行の結果、10 発言が適当であると判断したためである。また、いくつ先の文書単位との間で類似度をとるかについては 1~5 単位の間でサーベイを行った結果、2 以上で類似度プロフィールの形に大きな変化がなくなることから、2 つ先の文書単位との間で類似度をとることとした。

一例として、1998 年に開催された「地球環境と夏時間を考える国民会議」⁸⁾ の第 3 回発言録を対象として、上

位特徴語の類似度の推移を計算した結果を Fig. 2 に示す。文書全体での類似度の平均値0.80を分割の基準値に用いると、この発言録は10の話題に分割できる。

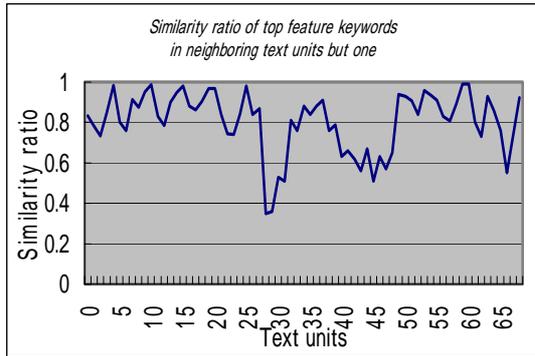


Fig. 2 発言録中での特徴語類似度の推移(「地球環境と夏時間を考える国民会議」第3回)

3. 自動要約

3.1. 展望台システム

テキストの自動要約は、一般的にテキスト中の文あるいは形式段落を文書単位とし、何らかの情報に基づいてそれらに重要度を付与し、重要度の高い文書単位を選択、抜き出すことで要約文を生成するというものである。従来手法における重要度の評価基準は、以下の5項目に分類できる。⁹⁾

- a. テキスト中の語の頻度を利用する。
- b. タイトル中の語を利用する。
- c. テキスト中、段落中での文の位置を利用する。
- d. 手がかり語を利用する。
- e. 文の間の関係を利用する。

従来手法の要約対象としては新聞記事や論文などが中心であるが、本研究で対象とする会議発言録は新聞記事や論文ほどよく構造化されていないので、b, c, eを用いることはできない。そこで、a, dを基準に自動要約手法を考えることとする。

語の出現頻度と手がかり語に基づく比較的新しいテキスト要約手法として、展望台システム¹⁰⁾が提案されている。この方法では、以下の式(3)~(5)で名詞 w の3種類の重要度指標 $\lambda_1, \lambda_2, \lambda_3$ を評価する。

$$\lambda_1(w) = \frac{F(w)}{F} \tag{3}$$

$$\lambda_2(w) = \sum_{s \in S} \frac{n(w \cap s)^2}{n(w)n(s)} \tag{4}$$

$$\lambda_3(w) = \sum_{s \in S} \frac{n(w \cap s)}{n(w)} \tag{5}$$

ここで、 $F(w)$ は語 w の出現回数、 F は文書中の総語数、 S はユーザが主観的に選択した要約の観点を表す観点語の集合、 $n(w)$ は語 w を含む文数、 $n(w \cap s)$ は語 w と s を両方含む文数を表す。もしユーザから観点語が与えられた場合には $\lambda_1, \lambda_2, \lambda_3$ の全てを考慮して要約するが、観点語が与えられない場合には λ_1 のみを考慮する。

つぎに、文 T に含まれる全語 ($w \in T$) についてその重要度 λ_i ($i=1,2,3$) の総和を計算する。

$$\Lambda_i(T) = \sum_{w \in T} \lambda_i(w) \tag{6}$$

文の重要度は、 $\Lambda_1, \Lambda_2, \Lambda_3$ の重みつき合計として評価し、重要度の高い文だけを抜き出して要約文を作成する。

3.2. 会議発言録への適用

展望台システムの会議発言録への適用を試みる。

まず、発言録に適用する場合には話題が推移する中で重要文が偏って抽出されないようにするため、2.3節に説明した方法で話題ごとに分割し、各話題ブロックを単位として語の重要度を $\lambda_1, \lambda_2, \lambda_3$ で評価し、ブロックごとに要約する。ここで一般語によるノイズを減らすため、式(3)の代わりに $tf \cdot idf$ を λ_1 として用いる。

ところで、展望台システムでは重要文抽出のための重要度評価を文単位で行うので、長文ほど抽出されやすく短文が無視されるという傾向があり、発言の長さが極端に異なる場合がある発言録に対しては必ずしも有効でない。そこで、文ごとに重要度評価せずに、50字を単位として1字ずつずらしながら移動平均をとるようなやり方で抽出部分の重要度評価を行い、最も重要度が高い部分を含む文を抜き出すことにした。最重要部分が複数の文にまたがる場合には、当該部分を含む全ての文を重要文と見なした。なお、50字を単位とした理由は、名詞が4~5個含まれるのが50字であり、重要度を評価するのに適当な長さだと判断したためである。

各話題ブロックから同じ数だけ重要文を抜き出すとすると、長いブロックと短いブロックとで要約の粒度が異なってしまう。そこで、話題ブロックの長さ按比例する数の重要文を抜き出すことにしたが、要約という性格上、どんなに短いブロックでも最低1文は抜き出すようにした。

4. 手法の性能評価

4.1. 評価尺度

本研究では、文献検索や重要文抽出などの分野で一般

的に用いられている評価尺度である適合率，再現率を用いて手法の性能評価を行うこととした。

いま，評価用の標準解集合と評価対象となる解集合とが Fig. 3 に示す関係にあるとする．このとき，適合率 P と再現率 R は以下のように定義される．

$$P = \frac{w}{w+x} \quad (6)$$

$$R = \frac{w}{w+y} \quad (7)$$

すなわち，適合率は手法の選択性を，再現率は手法の網羅性を示している．また，両者を考慮した総合的評価尺度として，次式で定義される van Rijsbergen の E 尺度¹¹⁾を用いる．

$$E = \frac{2}{1/P+1/R} = \frac{2w}{x+y+2w} \quad (8)$$

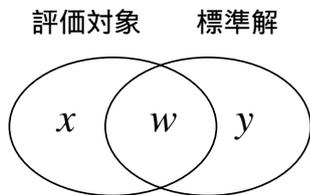


Fig. 3 適合率と再現率

4.2. 語のクラスタリングによる話題抽出の評価

2.3.節の手法で抽出された話題がどれほど妥当であるか評価するために，例題として用いた会議の議事概要を利用して話題を決定した．議事概要において，各発言者の意見は12の項目に分類されており，さらに項目ごとに発言の要旨がまとめられている．そこで項目ごとに発言要旨中の名詞を取り出し，複数項目にわたって登場するような一般語を削除することによって，語のクラスタを作り，評価用クラスタとした。

つぎに，評価用クラスタとどれだけ近いクラスタが得られたか評価するため，推移率に基づくクラスタ結果と評価用クラスタのすべてのクラスタ対に対して E 尺度を計算した．その結果，104 クラスタ中 E 尺度が 0.1 を超えるクラスタ数は 35 個，0.2 を超えるクラスタは 8 個であり，あまり一致はよくない。

一致がよくない原因としては，評価用クラスタ作成に議事概要を用いたことが考えられる．議事概要では，内容を簡潔にまとめるために，一般的な語が用いられる．たとえば，「石炭」「石油」「天然ガス」と書く代りに，「化石燃料」という語で代表させてしまう．このため，特定の語が十分にカバーされていない可能性がある．

4.3. 一般的な自動要約の評価

自動要約に本要約手法の妥当性を確認するため「地球環境と夏時間を考える国民会議」の発言録を対象に1人の被験者に評価用重要文を抜き出してもらい，自動要約の結果と比較した．参考まで，マイクロソフト社の Word とジャストシステム社の一太郎に附属する要約機能と，本手法との性能比較を行った結果を Table 2 に示す．

Table 2 被験者の抜き出した重要文との比較(「地球環境と夏時間を考える国民会議」第3回)

判定基準	本手法	Word	一太郎
再現率 文単位完全一致	7/21	0/21	6/21
再現率 文単位前後一文	11/21	0/21	9/21
適合率 文節単位	24/75	-	20/88
再現率 文節単位	24/54	-	20/54
再現率 文単位被験者割合	7.9/21	-	6.2/21

Word の要約は評価用重要文と全く関係のない部分が抽出されており，芳しい結果が得られない．一太郎では被験者が抜き出した部分が抽出されているので正解と見なすことができるが，抽出される部分が長くなる傾向がある．

次に一太郎と比べてどの程度，本手法によって抜き出された文が評価用重要文に長さが近く，かつ，同様の内容が抜き出されているかを評価するために，抽出された文を文節に分けて，適合率，再現率を調べることにした．再現率としてはやや本システムのほうが優れているが，適合率は一太郎の方がより多い意味単位を含んでいるために，本システムの方が短い文が抽出され，かつ適合する割合が多いことがわかる．

ところで，発言録に対する評価は被験者によって大きく異なる．そこで，5人の被験者に重要文を抜き出してもらい，その文を抜き出した被験者の割合を基準に再現率を評価した．その結果でも，本手法は一太郎に比較して若干ではあるが優れた成績をあげることがわかった．

なお，複数の被験者が抜き出した重要文の一致は極めて悪く，被験者5人が抜き出した評価用重要文のうち5人全員が重要と見なした部分は1文，4人が重要と見なした部分は4文，3人が重要と見なした部分は10文にすぎなかった．人間が要約を作成する場合，一般的にその

人の興味や目的によって重要と見なす部分は大きく変動する。このことから、要約の目的や観点を指定しない一般的な自動要約の評価が非常に難しいことが示唆される。

4.4. 観点キーワードを指定した要約の評価

観点語が指定された場合における自動要約機能の評価のため、会議録中ある話題ブロックを一つ選び、そこから50個の語を選んだ。その中から4名の被験者に興味のある語を選んでもらい、その観点に基づいて評価用重要文集合を作成してもらった。そして、被験者に選んでもらった語を観点語として自動要約を行い、それと被験者による評価用重要文と比較した。

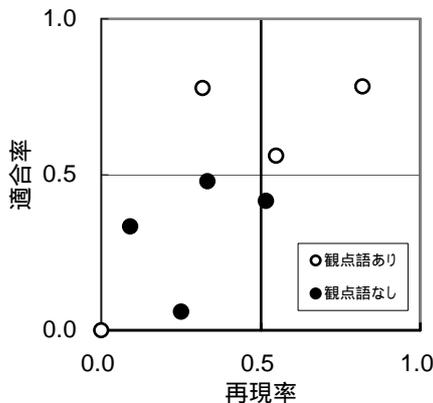


Fig. 4 観点語ありとなしの場合の再現率、適合率の比較 (「地球環境と夏時間を考える国民会議」第3回)

Fig. 4 は観点語ありの場合となしの場合とで、同一の発言録に対する重要文の再現率、適合率を比較したものである。観点語なしの一般的な要約をしたときの再現率、適合率に比べ、観点語を入れたときは両方とも改善され、被験者の興味を要約に反映させることができるという結果が得られた。

一事例において一般的な要約よりも劣った結果となったが、これは被験者が観点語を含まないが興味に関連した文を選び、さらに短く簡潔な部分を抜き出した結果、自動要約の結果に適合しなかったと考えられ、自動要約の性能評価の難しさをうかがわせる。

5. 知的支援機能を備えた電子会議システム

以上に述べた発言録に対する話題抽出や自動要約の手法を適用し、インターネット上で行う電子会議参加者に会議の概要をわかりやすく提示する電子会議システム TSS (Transcript Summarization System) を作成した。

このシステムでは、発言録名、観点語、文数を指定し

て実行ボタンをクリックすると、サーバ側に情報が渡され、話題別の重要文が表示されるとともに発言録中の重要文が色づけされる。Fig. 5 に TSS のインタフェース画面を示す。おのこの、左上は発言録名と観点語の入力、左下は話題項目の一覧表示、右上は重要文の表示、右下は重要文が強調された発言録全体の表示を行うためのペインである。各発言者がどの話題に興味を持っているかを表示するためには、Fig. 6 に示すような発言者×話題行列を提示する。これは発言者を行、話題を列とした行列であり、どの発言者がどの話題で発言しているかの概略が把握できる。



Fig. 5 TSS のインタフェース画面

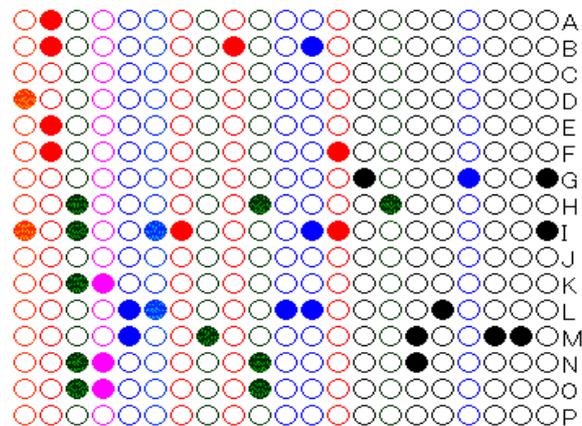


Fig. 6 発言者と話題の関係表示の例

このシステムにより、テキストベースの発言録だけを提供する場合に比較して、参加者による会議概要の理解を支援する情報が提供できるものと考えられる。なぜならば、話題の特徴語やそれに基づいて抽出された重要文は、参加者が会議の主課題に対して態度決定する場合の価値基準を反映していると考えられるからである。ただし、議論の流れを理解するためには、話題や重要文以外に話題に対する各参加者の立場や、それらの間の論理構造などを把握する必要があり、こうした高次情報をどう提示して

行くかが今後の課題である。また、本システムによって会議参加者の理解が進み、合意形成が促進されるかどうかについては、実際の会議の場面におけるフィールドテストの結果を待たなければならない。

6. 結論

本研究では、会議発言録に対する話題抽出、自動要約の手法を開発した。発言録の特徴から、発言録の処理では構文解析的な手法よりも統計的自然言語処理が有効であり、また話題の推移や発言の長さのばらつきを考慮した処理が必要となる。そこで、特徴語の変化に基づいて話題による発言録の分割を行い、それぞれの話題ブロックに対する要約を行う手法を開発した。また、特徴語から各発言者が興味を持つ話題を抽出する手法、語の共起グラフのクラスタリングにより話題を抽出する手法を提案した。さらに、会議参加者に会議の話題と要約を提示する機能を有する電子会議システム TSS を作成した。

本システムは、他の既存の要約システムに比べて被験者が作成した評価用重要文により近い結果を与え、会議発言録の自動要約手法として有効であることが確認された。しかしながら、既存システムに対する優位は大きなものではなく、再現率、適合率ともまだ十分とは言えない。話題抽出や自動要約の性能評価には一般的にさまざまな問題があるため、本研究で行った評価実験には限界があり、さらに検討が必要である。

参考文献

- 1) 木村忠正, 土屋大洋(1998)『ネットワーク時代の合意形成』NTT 出版.
- 2) Conklin, J., and Begeman, M.L. (1988). gIBIS: A Hypertext

Tool for Exploratory Policy Discussion. *ACM TOOLS*, 6(4), 303-331.

- 3) 徳永健伸(1999)『情報検索と言語処理』東京大学出版会.
- 4) 松本裕治, 北内 啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原正幸(2000)「日本語形態素解析システム『茶筌』 version 2.2.1 使用説明書」奈良先端科学技術大学院大学自然言語処理学講座.
- 5) 原子力委員会(1996)『原子力政策円卓会議』
<http://aec.jst.go.jp/jicst/NC/iinkai/entaku/> [2003, May 26].
- 6) 相澤彰子(2000)「語と文書の共起に基づく『特徴量』の定義と適用」『情報処理学会研究報告』F157-4, NL136-4, 25-32.
- 7) 田中久美子, 岩崎英哉(1996)「推移律に基づく共起グラフのクラスタ分け」『情報処理学会研究報告』NL115-15, FI43-7, 107-113.
- 8) 「地球環境と夏時間を考える国民会議」事務局(1998)『地球環境と夏時間を考える国民会議』
<http://www.kokuminkaigi.gr.jp/> [2001, September 1].
- 9) 奥村 学, 難波 英嗣(1999)「テキスト自動要約に関する研究動向」『自然言語処理』6(6).
- 10) 砂山 渡, 谷内田正彦(2002)「観点に基づいて重要文を抽出する展望台システムとそのサーチエンジンへの実装」『人工知能学会誌』17(1), 14-22.
- 11) van Rijsbergen, C.J. (1979). *Information Retrieval* (2nd ed.). Butterworths

本研究は、社会技術研究システム ミッション・プログラム「安全性に係わる社会問題解決のための知識体系の構築（平成13～14年度は日本原子力研究所の事業，平成15年度からは科学技術振興事業団の事業）の研究として行われた。

ELECTRONIC DISCUSSION SYSTEM
WITH INTELLIGENT SUPPORTING FUNCTIONS

Kazuo FURUTA¹, Motoyoshi MAEHARA²,
Ryosuke TAKASHIMA³, and Keiichi NAKATA⁴

¹Ph.D. (Eng.) The University of Tokyo, Institute of Environmental Studies (E-mail: furuta@k.u-tokyo.ac.jp)

²M.A. (Environmental Studies) The University of Tokyo, Institute of Environmental Studies

³M.A. (Eng.) The University of Tokyo, Dept. of Quantum Engineering and Systems Science

⁴Ph.D. (A.I.) The University of Tokyo, Institute of Environmental Studies (E-mail: keiichi.nakata@i-u.de)

While social demands for consensus development through public involvement are uprising in social decision-making, electronic discussion has come to play a particular role in consensus development along with rapid propagation of the internet. Methods of topic extraction and automated summarization from public discussion records have been proposed in this work so that participants' mutual understanding on progress of discussion and dominant trends of opinions can be promoted. In addition, we developed an electronic discussion system called TSS (Transcript Summarization System), which is equipped with these supporting functions, and then tested the functionality of the system using some actual meeting records.

Key Words: *Consensus development, electronic discussion, topic extraction, automated summarization, statistical natural language processing*