

ディスカッションマイニング

－ 議論からの知識発見 － *

清水 敏之

名古屋大学 工学部 電気電子・情報工学科
shimizu@nagao.nuie.nagoya-u.ac.jp

長尾 確

名古屋大学 情報メディア教育センター
nagao@nuie.nagoya-u.ac.jp

1 はじめに

複数の人間が集まる会議、討論などの場において、スムーズに議論が進むこと、さらに、その議事内容を検索し再利用できるということは非常に有益である。しかしながら、オンラインの掲示板などによる会議とは異なり、オフラインの会議においては議論を再利用可能な形にするのは一般に困難である。

本研究で提案するシステムは、構造化された議事録データを半自動的に生成し、リアルタイムに可視化して表示する。さらに、その議事録を XML データベースとリレーショナルデータベースに登録し、検索や要約が可能な状態にする。以上のような機能により議論を支援するシステムを試作した。

2 議事録の作成

議論を再利用可能な形にするためにビデオ(映像と音)を主体として用いる手法もあるが [4]、本研究で試作したシステムにおいて、議事録は主にブラウザのフォーム上からのからのテキスト入力によって生成される。その際、入力フォームは細かく分かれており、議論の内容だけでなく、発表者、日付、議題、参加者などの情報も入力する。

本研究で試作したシステムでは、スライド情報に特に注目した。近年の会議においては資料として Microsoft PowerPoint で作成されたスライドを用意する場面がほとんどであろう。Java で COM(Component Object Model) を操作するために JACOB(A Java-COM Bridge) を利用し、Microsoft PowerPoint であらかじめ発表者がその会議のために作成しておいたスライドを指定すると、議事録サーバ側で GIF 形式に変換して、議事入力ページに表示する。同時に、スライド内の文字情報を取得して、議事録に追加する。そして、スライドと

発言を関連付けることにより議事録を構造化する。

さらに、発言の情報は「質問」「回答」「コメント」の3種類に分類して入力する。本システムにおいては発言の内容をテキストで入力するので、煩雑になってしまわないよう発言の分類を大きく3種類にとどめた。

議論は「質問」「回答」「コメント」の3種類で1セットであると考え、一つの話者からもう一つの話者になると、新たな入力フォームを出せるような仕組みを実現した。

図1に議事録の入力画面を示す。

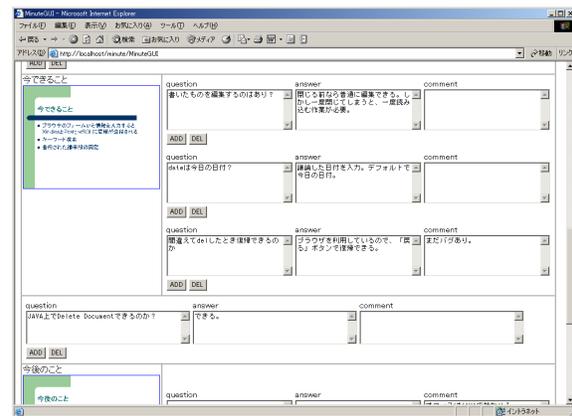


図 1: 議事録の入力作業

構造的アプローチによるグループウェアの代表的なものに gIBIS[1] があげられるが、これのように議論構造をグラフ状に表示することにより、議論内容の把握を容易にしたり、会議の場で有効な発言を促す効果があると思われる。

本システムにおいては SVG (Scalable Vector Graphics)[3] を用いた発言のグラフ表示・編集モード(図2)を設け、議事録の構造を可視化した。発言とスライドの関連情報やキーワードなどを用いて半自動で構造化して表示し、それに対してユーザが編集を行うことが可能である。

*Discussion Mining: Knowledge Discovery from Discussions by Toshiyuki Shimizu (Dept. of Information Engineering, Nagoya University) and Katashi Nagao (Center for Information Media Studies, Nagoya University)

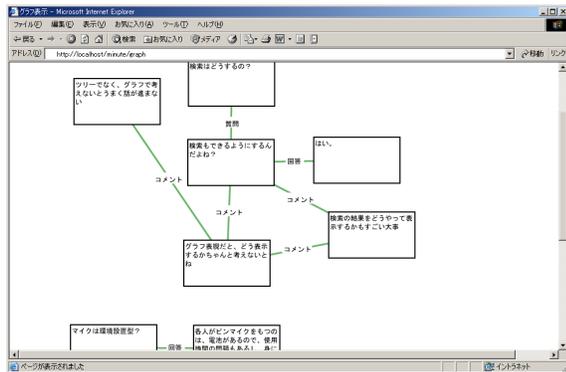


図 2: 議事録のグラフ表示

入力されたテキストとスライドから得られたテキスト情報などをあわせて議事録が作成される。作成される議事録は XML 形式であり、Java Servlet を用いてサーバマシンの XML データベースとリレーショナルデータベースに登録される。

本研究においては、XML データベースとして Xindice[2] を、リレーショナルデータベースとして PostgreSQL を用いた。

3 議事録の再利用

データベースに登録された議事録は、Xindice から取り出される DOM オブジェクトにスタイルシートをかけて表示することにより、ブラウザ上でリアルタイムで閲覧することが可能である。

さらに、過去の議事録データのうち、必要とする議事録を Web 上の検索フォーム (図 3) から検索し、取り出すことが可能である。検索する際、議題、日付、発表者など、細かく検索フィールドが分かれており、目的とする議事録を容易に探し出し、取り出して閲覧することができるようになっている。検索部分ではデータベースとして PostgreSQL を利用するので高速性も期待できる。

検索された議事録は、やはりブラウザ上で再編集することも可能であり、新たに書き加えたり、入力ミスを訂正したりすることが可能である。

議事録データを閲覧する際、要点だけを見たい場合は、発言に付けられた重要度のデータなどから要約して表示することも可能である。入力時に付けられた発言の構造、発言に含まれるキーワード、手動で付けられた発言への評価の点数を利用する。

4 おわりに

本論文では、会議における議論内容をデータベースに登録し、再利用するシステムについて述べた。

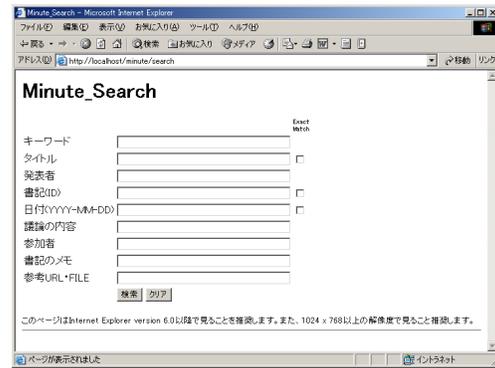


図 3: 議事録の検索

今後の課題として以下のものについて考察し、システムに反映していく予定である。

- 複数の議事録にまたがる構造化
本研究で試作したシステムにおいて作成される議事録は、特定の会議に閉じたものであり、過去の会議との関連性は明示的には入っていない。キーワードなどから半自動的に過去の議事録との関連性を調べ、複数の議事録にまたがった要約などができるようにすべきであろう。
- 音声の利用
音声を利用することにより、議事録に載せられなかった情報も、後で参照できるようになると思われる。音声はそのままでは検索しにくい情報であるが、音声認識を用いてトランスクリプトを作成するなど、アノテーションを用いることにより、検索等の問題はかなり解消される。
- 画像の利用
画像情報があると、その場の雰囲気など、他の情報からは得にくいデータも利用できるようになると思われる。

参考文献

- [1] J. Conklin and M.L. Begeman. gIBIS: A Hypertext Tool for Exploratory Policy Discussion. In *Proc. of CSCW '88*, pp. 140-152, Sep 1988.
- [2] the Apache XML Project. Apache Xindice, 2001. <http://xml.apache.org/xindice/>.
- [3] W3C. Scalable Vector Graphics (SVG) 1.0 Specification, 2001. <http://www.w3.org/TR/SVG/>.
- [4] 角康之, 間瀬健二. インタラクション・コーパス構築の試みとしてのミーティングキャプチャ. ヒューマンインタフェースシンポジウム 2002.