

電子メールを用いたコミュニケーションを対象とした 討議構造自動抽出エンジンの開発

山見 太郎 村越 広享 島津 明 落水 浩一郎
北陸先端科学技術大学院大学 情報科学研究科

{t-yamami,murakosi,shimazu,ochimizu}@jaist.ac.jp

概要

電子メールを介した議論の参加者にとって、議論を話題毎に整理し、議論の進捗に関する情報を獲得することは、円滑なコミュニケーションに欠かせない要素の一つである。これを支援するため、討議の道筋に関する情報を系統的に管理するシステムの開発を進めている。本報告では、討議の道筋に関する情報を電子メールに付与するためのXMLボキャブラリ『UMML+Linkbase』と、その情報をメール群に自動付与するツール、『討議構造木抽出エンジン』について述べる。また、討議の道筋を呈示する機能を有するメールクライアント『ICEMail++』の概略を述べる。

1 はじめに

近年、共同作業などにおいて、電子メールを利用した議論が盛んに行なわれている。しかしながら、電子メールを利用した場合、並列に複数の話題が議論され、討議参加者は、議論の進捗状況を即座に認識することが困難であることが予想される。討議の道筋に関する情報を系統的に整理し、円滑なコミュニケーションを促進するための情報を提供することができれば、円滑なコミュニケーションの実施に効果があるものと思われる [1]。

本報告では、話題毎に議論が整理され、議論の進捗状況に関する情報が抽出可能である討議構造モデルを述べる。また、そのモデル構造に関する情報を電子メール群に付与するためのXMLボキャブラリ『UMML+Linkbase』を定義し、その情報をメール群に自動付与するツール、『討議構造木抽出エンジン』を紹介する。さらに、円滑なコミュニケーションを支援するツールとして開発した、討議の道筋を呈示する機能を有するメールクライアント、『ICEMail++』の概略を述べる。

2 討議構造モデル

2.1 電子メールを用いた討議の特徴

電子メールを利用したコミュニケーションでは、複数の話題を同一メールに記述する傾向にあり、ある一つの

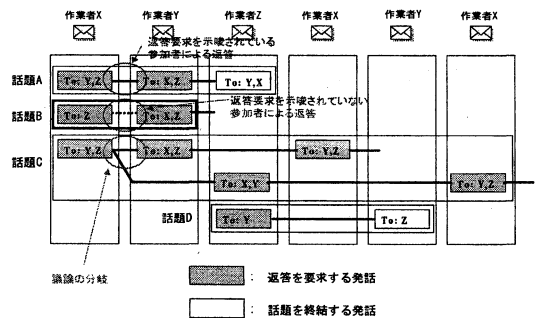


図1: 電子メールを利用した討議構造

話題を複数のメールにわたり議論する。また、以下のような特徴がある [2]。

- 次に返答すべき参加者を明示的に示唆することがある (特徴1)。
- 返答を要求されていない参加者が返答を行なう場合がある (特徴2)。
- 一つの発話に対して、複数の参加者が返答を行ない、討議の流れが分岐する場合がある (特徴3)。

図1に、電子メールを利用した討議の構造を示す。

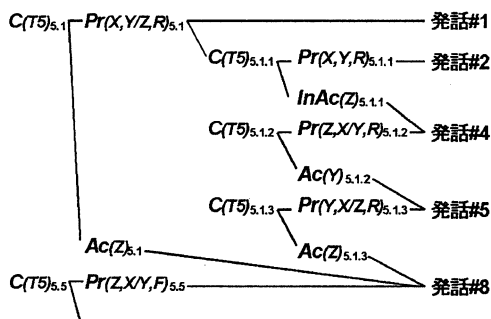


図 2: 討議構造木例

2.2 討議構造木

討議構造木は、図 1 に示される構造を形式的に表現するために、対面による 2 者の会話モデルであるコントリビューションツリー [3] を拡張したモデルである [2][4]。このモデルでは、以下のように発話、コントリビューションを定義する。

発話: 同一メール内において同じ話題に言及している文章。

宣言的な発話: 議論の流れに直接の影響を与えない発話。

返答を要求する発話: 質問や提案など、返答要求を示唆する発話。

話題を終結する発話: 同意や受諾など、話題の終結を示唆する発話。

コントリビューション (C): 話し手と聞き手からなる会話のユニット。プレゼンテーションフェーズ (Pr) とアクセプタンスフェーズ (Ac および InAc) からなる。

Pr: 作業 X が、作業グループ中のある作業員たち G に対し、話題を継続/終結する意思を伝達するための発話 u を行うフェーズ。

Ac/InAc: 作業 Y が発話 u を理解したか否かを、作業 X に伝達するための行動 e' を行うフェーズ。作業 Y が、作業 X が意図した作業員たち G に含まれるなら、Ac、含まれないなら、InAc とする。

なお、上記の各々のフェーズには、話題および発話の導入順を示すサフィックスと、意思伝達者や意思伝達を意図された参加者、話題の終結/未終結をその状態を示すための属性をもっている。図 2 に、討議構造木の例を示す。

3 討議構造木抽出エンジン

討議を構成する電子メール群から討議構造を自動抽出し、その構造を表現する UMML+Linkbase ファイル群を生成するツール、討議構造木抽出エンジンを開発した。抽出エンジンは、言語的手がかり語を用いた討議構造木の自動構築法 [2][4] に基づいて実現されている。関連研究として、討議構造を有効グラフとして捉え、引用によるメールの相関関係を抽出し、パス上にある一連のコメントをスレッドとする手法などがある [5]。

3.1 XML による討議構造の表現

電子メール群に、討議構造木に関する情報を付与するための XML ボキャブラリ、『UMML+Linkbase』を定義する。

3.1.1 UMML

UMML (Utterances-in-Mail Markup Language) は、XML 仕様に準拠したマークアップ言語で、討議構造木中に現れる発話に関する情報 (Pr フェーズの属性など) を電子メールに付与するための言語である。なお、別のメールに含まれる発話との接続関係は表現しない。

3.1.2 Linkbase

XLink とは、XML 文書のリンク機能に関する規約である [6]。異なるメールに含まれる発話間の接続関係を、XLink における Linkbase に格納したサードパーティリンクとして表現する。

3.1.3 UMML+Linkbase による討議構造木表現例

UMML+Linkbase による討議構造木表現の例を、図 3 に示す。

3.2 抽出エンジンの概要

抽出エンジンでは、電子メールを用いたコミュニケーションに現れる構造/言語的特徴を、手がかりとして利用している。

3.2.1 電子メールボディ部の特徴

引用文による手がかり 返答を含むメールのメッセージ部では、引用文が利用されるという特徴がある。引用文とは、各行の先頭に「>」、「>」といった引用符を伴い、以前の議論であらわれた文章を引用している文章のこと

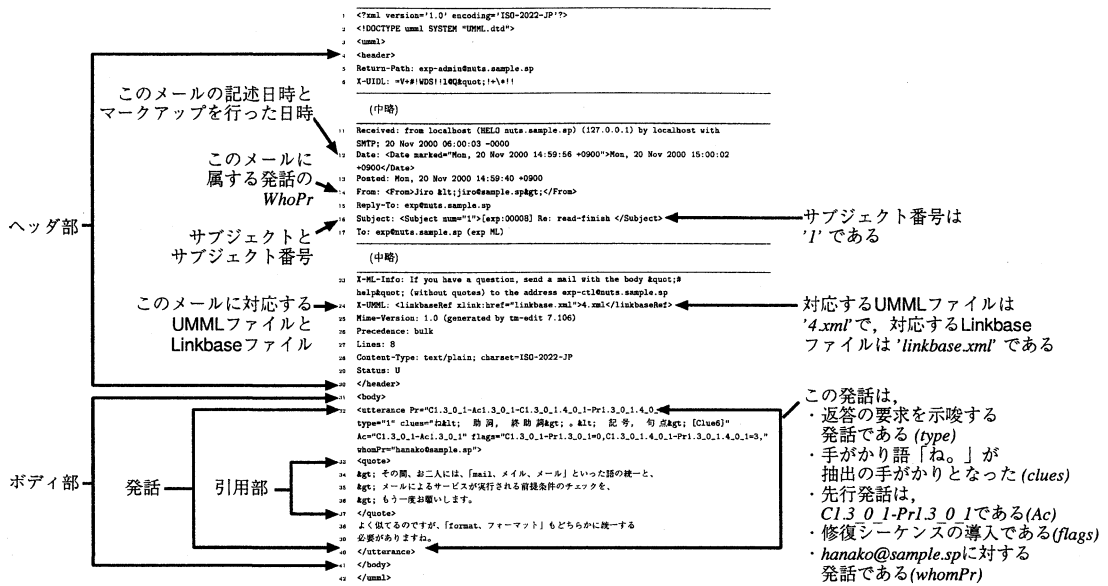


図 3: 討議構造木表現例

である。引用文は、引用文の直後にあらわれる発話の先行発話を示すことが多い。また、メッセージの最初は、自分の名前や身分を示す宣言的な発話であることが多く、メッセージの最後は、一般に「署名」と呼ばれるものであることが多い。署名は、自分の名前や身分や連絡先などを記述したもので、直前に“---”，“==--==”といった記号だけからなる行（セパレータ）を伴うことが多い。

空行や改段落による区切り 電子メールを利用したコミュニケーションでは、対面での対話と異なり、会話の一時停止などを利用した話題転換の指示ができない。その代わりに、文章中に意図的に空行や改段落を挿入することによって話題転換の指示を行なっているものと考えられる。よって、空行や改段落で区切られた文章を発話を構成する基本単位とする。ただし、発話の最低単位は一文とする。

3.2.2 文章表現の特徴

先頭文の文頭や文末などの表現に、話題の転換/終結を示唆したり、文章末に、返答の要求を示唆する表現が含まれていることがある。このような特徴を抽出手がかりとした。表 1 に、その手がかり語をまとめる。

4 討議構造参照機能を有するメールクライアント

4.1 概要

構造木抽出エンジンが生成した UMML+Linkbase ファイル群を利用し、討議の道筋を呈示するソフトウェアのプロトタイプとして、ICEMail++ 図 4 を開発した。これは、Java で記述されたメールクライアント ICEMail[7] に、討議構造呈示部を追加したものである。討議構造呈示部は、討議構造の表形式およびツリー形式呈示機能を有する。この 2 つの呈示機能は、ユーザーによる選択可能となっている。

4.2 討議構造提示機能

表形式呈示部では、抽出された発話は、導入された順序でセルに格納され、発話の種類はセルの背景色で区別する。また、引用・被引用部分を相互にリンクづけ、引用部分から被引用部分へ、あるいは、被引用部分から引用部分をたどること可能とし、討議構造木の発話を順にたどることができる。図 4 に示す討議構造呈示部は、表形式提示を行なっている。

ツリー形式呈示部では、討議構造木を呈示する。1 つのフレームは 1 つの発話に対応しており、フレーム間を

示唆する特徴	略号	例
話題の並列展開	Clue#1	列挙します, 挙げます
話題の転換	Clue#2	ところで, ~については
接続語による結束性(詳細の関係)	Clue#3	具体的には, 例えば
接続語による結束性(前提の関係)	Clue#4	それで, だから
指示語による結束性	Clue#5	この場合, これは
返答の要求	Clue#6	か, しましょう, ?
同意・了承	Clue#7	了解です, 確に

表 1: 言語的手がかりの一覧

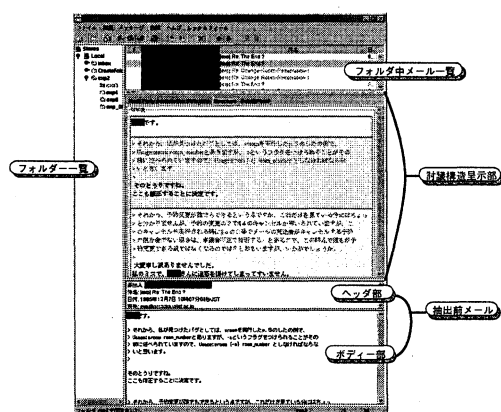


図 4: ICEMail++

矢印で結び、討議のスレッドを表現している。フレーム内の背景色は、発話の種類を示し、フレームを選択することにより、その発話が行なわれたメールを参照することが可能である。

5 おわりに

本報告では、本報告では、討議の道筋に関する情報を電子メールに付与するための XML ボキャブラリ『UMML+Linkbase』と、その情報をメール群に自動付与するツール、『討議構造木抽出エンジン』について述べた。また、討議の道筋を呈示する機能を有するメールクライアント『ICEMail++』の概略を述べた。今後の課題は、討議構造木抽出エンジンの精度を評価し、また、メールクライアントを運用することにより、円滑なコミュニケーション支援に有効であることを確かめることである。

謝辞

本研究の一部は、情報処理振興事業協会（IPA）より委託を受けた財団法人ソフトウェア工学研究財団（RISE）が実施した平成12年度「高度情報化支援ソフトウェアシーズ育成事業」での支援を受けて行なわれた。

参考文献

- [1] Koichiro Ochimizu, Chie Kadowaki, and Masa-kazu Hori. Design of an Information Repository to Support Cooperative Works over a Computer Network. In *IPSJ International Symposium on Next-Generation of Information Technologies*, pp. 79–86, 1997.
- [2] 村越広享, 島津明, 落水浩一郎. メーリングリストを利用した共同作業における討議構造の自動構築法. コンピュータソフトウェア, 2001(accepted).
- [3] Herbert H. Clark and Edward F. Schaefer. Contributing to Discourse. *Cognitive Science*, Vol. 13, No. 2, pp. 259–294, 1989.
- [4] Hiroyuki Murakoshi, Akira Shimazu, and Koichiro Ochimizu. Construction of deliberation structure in e-mail communication. *International Journal of Computational Intelligence*, Vol. 16, No. 4, pp. 570–577, 2000.
- [5] 村上明子, 長尾確. ディスカッションマイニング: 構造化されたコミュニケーションによるトピックの検索と視覚化. 言語処理学会第6回年次大会発表論文集, pp. 451–454, Mar. 2000.
- [6] XML Linking Language (XLink) Version 1.0, Jul. 2000. <http://www.w3.org/TR/2000/CR-xlink-20000703/>.
- [7] Tim Enders, Jeff Gay, and Y. Miyadate. ICEMail. <http://www.icemail.org/>.