

電子メールをベースにした討議型共同文書作成支援システムの設計

堀雅和[†], 加藤康記[†], 村越広享[‡], 山見太郎[‡], 島津明[‡], 落水浩一郎[‡]

[†]インテック・ウェブ・アンド・ゲノム・インフォマティクス(株)

〒930-0804 富山県富山市下新町3番23号

[‡]北陸先端科学技術大学院大学情報科学研究科

〒923-1292 石川県能美郡辰口町旭台1-1

あらまし

現在我々は、「複数の人がメーリングリストを用いて議論しながら共有文書を作成(修正)する」という状況で、議論した結果に基づき発生する共有文書の修正作業を支援することを目的とした、以下のような特徴を持つシステムを研究している。

- ① メーリングリスト上に発生した会話を構造化し、わかりやすい形式にて表示する機能
- ② 会話情報から共有文書の修正に関する ToDo リストを作成し、対象文書の修正状態に基づき ToDo リストを更新する機能

本稿では、本システムの設計方針、機能の概要、実装後のシステムの評価対象について述べる。

Design of System to Support Cooperative Writing in E-Mail Discussion

Masakazu HORI[†], Yasunori KATO[†], Hiroyuki MURAKOSI[‡], Taro YAMAMI[‡], Akira SHIMAZU[‡],

Koichiro OCHIMIZU[‡]

[†]INTEC Web and Genome Informatics Corporation

3-23, Shimoshin-machi, Toyama, 930-0804 Japan

[‡]School of Information Science

Japan Advanced Institute of Science and Technology

1-1, Asahidai, Tatsunokuchi-machi, Nomi-gun, Ishikawa, 923-1292 Japan

Abstract

We currently investigate a system to support a modification process of a shared document based on a discussion result under the situation such that several people cooperatively create a document discussing in a mailing list. The system has the following features:

1. Structure the conversation in a mailing list and visualize the information; and
2. Create a ToDo list about modification of a shared document from a mail data and update a modification state in the list after the document is modified.

In this paper, we describe the design policy, the overview of the features, and the evaluation points after the completion of the implementation.

1 はじめに

ネットワークを介してソフトウェア仕様書などの文書を共同作成する場合、関係者の中で議論した結果に基づいて実施される共有文書の修正プロセスを支援することが重要である[1]。共有文書の修正要求が発生した場合、修正要求の発生理由、修正要求が承認されたのか否か、承認された場合は修正作業に必要な情報や要求通りに修正されたか等に関する情報が必要となる。すなわち、交わされた会話の内容に応じてコミュニケーションに付随して発生するプロセスを支援できるシステムが要求される。(半)構造化メッセージを用いたシステムを用いると、単に情報交換するだけでなく、コミュニケーションに付随して発生するプロセスの支援も可能である。しかし、ユーザの使い勝手に関する制約が強くなるためか、実際にはほとんど普及していないのが現状である。一方、一般に用いられている電子メールを使用した場合、制約が少ない反面、コミュニケーションに付随して発生するプロセスを支援する機能が弱い。そのため、ネットワーク上での共同文書作成に電子メールを用いると、修正内容に関する情報だけでなく、修正プロセスを制御するための情報も交換し、手動で制御する必要がある。このように、使い勝手に関する制約を少なくする一方で、会話の情報から修正プロセスに有効な情報を抽出し、作業プロセスに反映することが重要になってくる。

我々は、以上のような状況を改善する目的で、以下のような特徴を持つシステムを研究している。

1. メーリングリスト上に発生した会話を、内容の点から構造化し、わかりやすい形式にて表示する機能
2. 会話情報から共有文書の修正に関するToDo リストを作成し、対象文書の修正状態に基づきToDo リストを更新する機能

本稿では、現在研究開発中のシステムの設計方針、機能の概要、実装後のシステムの評価対象について報告する。本稿の構成は、以下のとおりである。第2節では、本システムを設計するうえでの基本的な考え方および目標とする機能要件につ

いて述べる。第3節では、実現する機能の内容とその範囲について述べる。第4節では、システムを構成する代表的なモジュールの構成について述べる。第5節では、本システムの実装終了後の評価対象に関する考察結果について述べる。

2 設計方針

本研究では、「複数の人がメーリングリストを用いて議論しながら共有文書を作成(修正)する」という状況で、議論した結果に基づき発生する共有文書の修正作業を支援するシステムを実現することを狙いとしている。本研究における基本的な考え方は、以下のとおりである。

- ・使用するツールやコミュニケーションの形態など、既存の作業プロセスに極力制約を与えずに実現する必要がある。
- ・可能な限りシステムによってプロセスを自動化する。但し、明らかに人間が手動で対処すべき処理は、人間が介入できるようにする。

以上のような考え方に基づき本研究では、以下の2つの機能の実現を目標とする(図1参照)。

1. 討議ストリームの自動抽出機能
メーリングリスト中に発生した会話情報を解析し、以下のような観点から会話情報を構造化してデータ管理する機能を実現する。
 - ・一通のメールよりも細かい粒度による、同一話題に関する会話の履歴情報
 - ・会話に関係すると想定される人の情報
 - ・討議の状態(終結/未終結)の情報
2. 共有文書の修正に関するToDo リストの生成および修正監視機能
共有文書の修正プロセスを、以下のように支援する機能を実現する。
 - ・会話中で共有文書の変更に関して言及された場合、その情報を共有文書の修正に関するToDo リストに登録する。
 - ・また、ToDo リスト中に登録された共有文書の修正状態を監視し、修正が完了したらToDo リストから対応する項目を削除する。

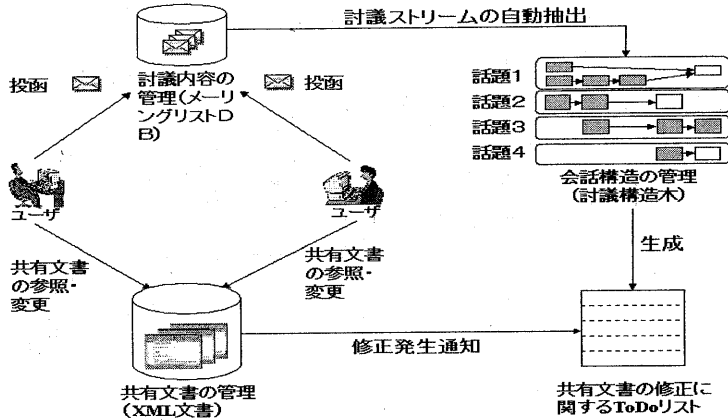


図 1 システムイメージ

3 機能の概要

3. 1 討議ストリームの自動抽出機能

本機能は、電子メールのデータを解析して討議構造木と呼ぶ会話間の関係を管理するデータ構造を生成する機能である。

討議ストリームとは、同一話題に言及する一つ以上の返答を要求する発話と話題を終結する発話による連鎖と定義する。また発話とは、同一メール内で同じ話題に言及している文章と定義する。電子メールのデータから討議ストリームを自動抽出する方法の概要は、下図のとおりである(理論の詳細は文献[2,3]、実装の詳細は文献[4]参照)。

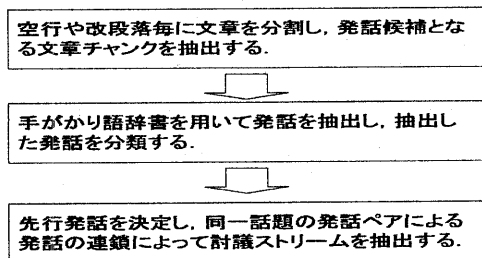


図 2 討議ストリーム抽出アルゴリズム

図 2 のアルゴリズムにより、メーリングリスト中で交わされた同一の話題に関する会話の時系列関係を、発話という単位で取得できる(図 3 参照)。

話題転換した箇所の自動抽出が可能である知的

ニュースリーダー「HISHO」[5]や、メールの引用情報に基づきスレッドを生成し視覚化する「ディスカッションマイニング」[6]に比べ、抽出された討議ストリームは以下のような特徴を持つ。

- ・ 同一メール中に異なる話題が並列に記載されている場合でも、それぞれを異なる話題として扱うことが可能である。
- ・ 討議ストリームで扱っている話題に関係している人が誰かわかる。
- ・ 返答を要求している発話に返答したかという観点から、討議の状態(終結/未終結)がわかる。

表 1 に、今回設計した討議構造木 DB のマスターテーブルを構成する主要なデータ要素を示す。

表 1 討議構造木のマスターテーブル

名称	概要
討議ストリーム ID	討議ストリームの識別子
サブジェクト ID	所属するサブジェクトの識別子
先頭発話 ID	討議ストリームに属する発話のうち先頭の発話の識別子
開始日付	討議ストリームが生成された日付
終了日付	討議ストリームが終結した日付
修正項目 ID	修正対象となる文書領域情報を格納したデータの識別子
討議状態	「終結」あるいは「未終結」

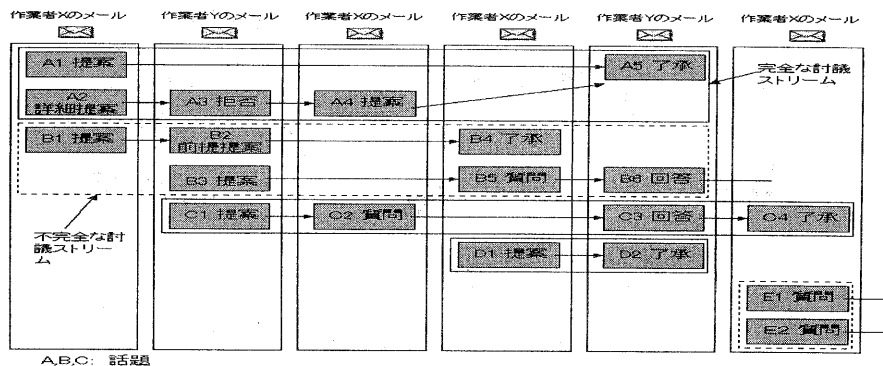


図3 電子メールを利用した議論の流れ

3.2 共有文書の修正に関する ToDo リストの生成および修正監視機能

本機能は、電子メール中で共有文書の修正に関する議論がなされていた場合に、共有文書の修正に関する ToDo リストを自動生成し、ToDo リスト中に表れた修正対象文書が議論完了後に修正されたかチェックする機能である。

本研究では、修正領域の粒度がファイルより小さい領域に対して修正対象の正確な特定を可能にする目的で、対象とする共有文書は XML 形式の文書とする。ToDo リストを構成するデータ項目の属性は、(ファイル名、タグ名、修正日、修正状態)の4つである。ファイル名とタグ名は、電子メールのデータから自動抽出する。修正日は、指定された文書の指定されたタグが修正された日、修正状態は当初「未完了」状態が設定されているが文書名とタグ名で指定された領域が修正されると「完了」状態に遷移する。

電子メールのデータからファイル名とタグ名を自動抽出する方法は以下のとおりである。ダブルクォートで囲まれた文字列は、ファイル名あるいはタグ値の可能性のあるものとする。その文字列の近辺に、「ファイル名」という文字列が発見された場合はファイル名、XML 文書で使用されているタグ名と同一の文字列が発見された場合はタグ名が記載されていると判断する。例えば、以下のような文章からファイル名(モジュール.xml)、タグ名(関数名)、タグ値(init)を自動抽出する。

例文：

ファイル名"モジュール.xml"の関数名"init"の引数を以下のように修正願います。

但し、同一の発話中で複数の指摘はできないという制約があるものとする。

4 システム構成

本システムは、サーバはサーブレット、クライアントはアプレットを用いて実装する(図4参照)。以下にそれぞれの構成について説明する。

4.1 サーバシステム

サーバを設計するにあたり、多くの人に簡単に使用してもらう目的で、以下のような条件を満たすことが重要であると考える。

- ・ インターネット環境からアクセス可能
 - ・ 既存システムにアドオンする方法で使用可能
- 以上の条件を満たすために、Web サーバのバックエンドであるサーブレットを用いて実装する。サーバと外部システムとの関係および、サーバを構成する主要な2つのモジュール (a) 討議構造木抽出エンジンと (b) 共有文書修正監視モジュールと DB との関係は以下に説明する(図4参照)。

サーバは、POP プロトコルを用いて定期的にメールサーバから新着メールを取得する。取得した新着メールは、討議構造木抽出エンジンを用いて解析する。解析した結果生成された討議構造木は、討議構造木 DB に、また ToDo リストは ToDo リ

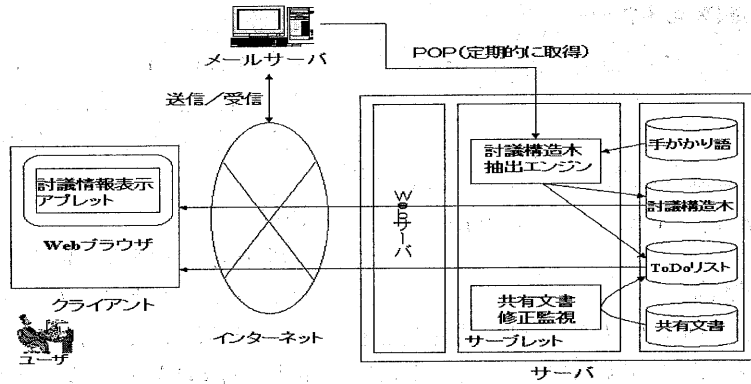


図4 本システムの構成

ストDBに登録する。討論構造木DBおよびToDoリストDBはともに、関係データベースを用いて管理する。

共有文書DBで管理しているデータのうち文書の内容は、ユーザが任意の編集ツールを使用して修正できるようにするため、サーバ機のファイルシステムを用いて管理する。共有文書修正管理モジュールは、共有文書の修正状態を監視しており、ToDoリスト中に登録されている共有文書が修正されたらToDoリスト中の修正状態を変更する。

4. 2 クライアントシステム

サーバ同様クライアントも、以下の点から手軽に使用できることを主眼として設計する。

- ・ 使用するメールリーダに制約を与えずに使用可能
- ・ Webブラウザから参照可能

以上の条件を満たす目的で、Java アプレットを使用することとする。以下に、クライアントを構成する主要なダイアログの設計について説明する。

図5は、最初に表示されるメインの画面例である。画面上部のボタン類により、表示件数、メールの送受信対象者名、開始/終了日付等の検索条件を設定することができる。画面中央の表にて、指定された検索条件に合致したメールに付随する件名等の属性情報の一覧を表示する。また画面下部の表にて、選択されたメール中に含まれる発話の一覧情報を表示する。

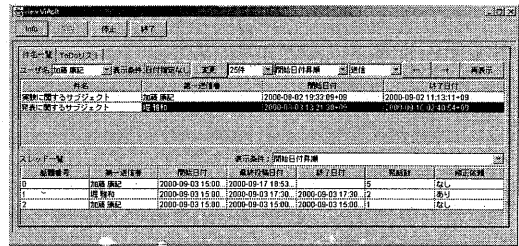


図5 件名一覧表示ダイアログ

図6は、発話間の履歴を表示した画面例である。討論構造木表示ダイアログで扱う発話の履歴は、直列的な関係だけでなく、分岐や階層が深くなる場合も存在する。図6は、分岐の場合だけでなく、パネルの上にパネルを配置することで階層が深くなる場合に対処していることもわかる。

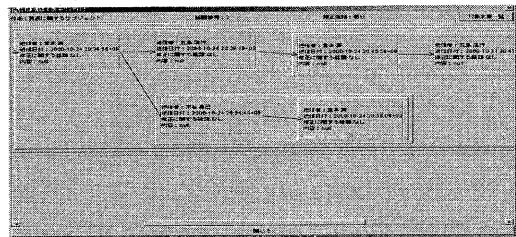


図6 討論構造木表示ダイアログ

図7は、共有文書の修正に関するToDoリスト一覧の表示例を表している。修正状態はシステムによって自動的に設定されるが、必ずしも実状にあっていない場合が想定されるので、ユーザによって手動で変更することも可能とする。

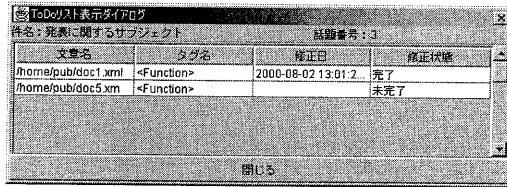


図7 ToDo リスト表示ダイアログ

5 今後の課題

本節では、今回設計したシステム実装後に実施予定の本システムの評価の対象について考察する。

5.1 自動抽出における間違い発生率の測定

本システムを用いて討議構造木およびToDo リストを自動生成した場合の抽出間違いの発生率を測定する。また、間違い発生率を測定するだけでなく、間違いが発生する理由(事例)も調査する必要がある。これらの情報を用いて、システムを使用する人が異なることで、どのような要素が間違い発生に影響を与えているかを考察することが重要である。

5.2 修正作業プロセスの変化の調査

本システムを使用することでユーザは、修正作業のコストが減るなどの恩恵を得ることができる。と予想しているが、我々の予想が正しいか検証する必要がある。そのために、本システムを使用した場合と使用しない場合の作業プロセスを比較し、観察される違いについて調査する必要がある。

5.3 有効に機能する作業形態の調査

本システムが有効に機能する作業形態を明らかにする目的で、異なる作業形態に本システムを適用し、ユーザが本システムの機能をどの程度活用したかを調査する。その時に考慮すべきパラメタとして例えば、以下の3つが考えられる。

- ・ 文書作成作業に関わるコミュニケーションの比重の違い
- ・ 共同作業での意思決定プロセスの違い
- ・ 変更作業の役割分担の違い

6 まとめ

本稿では、電子メールによる議論の結果に基づき、共有文書を作成するプロセスを支援するシステムの設計結果について報告した。

謝辞

本研究は、情報処理振興事業協会(IPA)の平成12年度「高度情報化支援ソフトウェアシーズ育成事業」に採択された「電子メールをベースにした討議型共同文書作成支援システムの研究開発」の一環として実施している[7]。

参考文献

- [1] 堀, 落水: ソフトウェア分散共同開発におけるコミュニケーション支援法に関する一提案. FOSE '99, レクチャーノート/ソフトウェア学22, (1999), pp. 116-123, 近代科学社.
- [2] Murakosi, H., Shimazu, A., Ochimizu, K.: A Method of Computing Measure for Evaluating Conversational Coherency in Email Communication. In IWCMC '99, pp. 66-73, 1999.
- [3] Murakosi, H., Shimazu, A., Ochimizu, K.: Construction of Deliberation Structure in E-Mail Communication. Computational Intelligence, Vol. 16, No. 4, pp.570-577, 2000.
- [4] 山見, 村越, 島津, 落水: 電子メールを利用したコミュニケーションにおける討議スレッド自動抽出法の実装と評価 情報処理学会 2000-NL-137, pp.69-76, June 2000.
- [5] 井佐原, 小作, 内元: 討論型ニュースグループを対象とする知的ニュースリーダの開発. 情報処理学会研究報告, Vol. 97, No. 53, pp. 13-18, May 1997.自然言語処理.
- [6] 村上, 長尾: 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.
- [7] (財)ソフトウェア工学研究財団. <http://www.rise.or.jp/>.