

協同執筆によるネットワーク上の人々が持つ情報の集積方法

門間 信行 大野 浩之

東京工業大学大学院 情報理工学研究科

概要

インターネットの多くの利用者から情報の断片を収集し、それらの断片をまとめあげた文書の作成を支援する機構を提案する。現状では、そのような情報の断片はほとんど整理されずに放置されることが多く、また、まとめあげが行なわれる場合でも少数の人が多大な労力をはらっている。

本稿で示す方法では、多人数の参加者が情報の断片として文書を書き、まとめあげに協力することにより、まとめあげ文書の作成を試みる。また、協同執筆における作業の流れのモデルを提案し、各作業の計算機支援によって従来必要であった労力を軽減する。

Method of the Information Gathering on the Internet by Collaborative Writing

Nobuyuki MONMA Hiroyuki OHNO

Graduateschool of Information Science and Engineering, Tokyo Institute of Technology.

abstract

We propose a system to gather fragments of information from Internet users about some topics into one document. Currently, such a fragment of information is rarely integrated and left alone, and a few persons summary up such information with heavy efforts.

Our system accomplish the goal by collaboration of many participants to writing documents. We propose a work flow model for writing documents, and enable to reduce efforts of collaborative workers by computer support at each work.

1 はじめに

近年のコンピュータネットワーク環境の向上にともなう、コンピュータネットワークはさまざまな人に利用されるようになり、新しいコミュニケーションの場の一つとなりつつある。さらに、複数の人々のコミュニケーション支援を目的としたさまざまなアプリケーションの普及によって、人が集まる場が生成され、その場を利用しての情報交換はさらなる活発化をみせている。人々は互いに持っている情報を出し合い、議論を重ねることにより、情報の誤りは徐々に取り除かれて行き、議論の参加者は有用な情報を得ることができる。人々が持つ情報を少しずつ集めて蓄積しておくことは有意義であると考えられる [1]。

しかし、このようにして交換される情報は有用なものも多いかかわらず、整理されずに放置される場合が非常に多い。また、それらの有用な情報は時とともに不要な情報の中に埋没し、後になって再び同じ議論が繰り返されることがままある。

このような情報の埋没を防ぐために、さまざまな工夫がなされている。特に質問と対応する回答形式でまとめたものとしてFAQ集と呼ばれる文書がある。しかし、このような情報の整理、蓄積は一人またはごく少数の作者が多大な労力をかけて行なっているのが現状である。

本稿では、このような情報の整理、蓄積を、一人ではなく多人数の参加者が協力して行ない、FAQ集のような「まとめ文書」を生成する機構を提案する。

2 コンピュータネットワーク上での 情報交換の現状

コンピュータネットワーク上に広く分散した人々が多人数で情報交換をするためのサービスには、ネットニュース、電子メールを利用したメーリングリスト、パソコン通信などにおける電子掲示板(BBS)、IRCなどに代表されるチャット、などがある。近年はWorld Wide Web(以下、WWW)の上にこれらと類似のサービスを提供するシステムもある。

これらのサービスは、情報を共有したい人の人数や関係、情報の即時性、匿名性、保存性など、それぞれ特徴がある(図1)。

ネットワーク上の人々間の情報交換は、質問をきっかけとして展開される場面がしばしば見られる。質問をきっかけとした情報交換は以下のような流れで行なわれる。

- 質問をする人は、質問の文書を作成する。
- 利用可能なサービスの中から一つを選び、質問

表 1: 各コミュニケーション手段の特性

サービス名	メーリングリスト	ネットニュース	WWW 上の BBS	チャット
サービス対象	特定	不特定	不特定	不特定
参加者の結びつき	さまざま	弱	弱	弱
情報の即時性	あり	なし	あり	あり
匿名性	低い	低い	高い	高い
情報の保存性	管理者による	高い	管理者による	低い

の文書を送出する。

- 読者へ文書が配送される。
- 読者が回答を寄せる。
- その回答に対してさらに回答が寄せられ、議論となる。

しかし、このような情報交換の方法にはいくつか欠点があり、主なものとして次の5つが挙げられる。

1. 得られる回答の内容が重複する
ネットニュースのような即時性のないサービスを利用した場合や、得られた回答を第三者が見ることができない場合に起こる。回答側は、すでに報告されている回答を知らないので、多数の回答が得られても、似通った内容であることが多い。
2. 参加者層を自由に選択できない
ある地域やある分野に極めて特化した質問の多くは、その地域の住人やその分野に興味を持つ人にだけに伝われば本来十分である。この問題は、質問を投稿するニュースグループやメーリングリストを適切にぶことによって解決できる場合もある。しかし、必ずしも最適なニュースグループやメーリングリストが存在するとは限らない。
3. 不適切な情報が寄せられる
ネットワーク上には多種多様な人々が存在するため、多くの人々の協力が期待できる反面、妨害行為を受ける可能性がある。これは特に、匿名性の高いサービスを利用した場合に生じやすく、虚偽の回答や、質問の内容と全く関係のないような回答が寄せられる場合がある。また、議論の途中で、質問者の意図とは違った、あらぬ方向へ話がそれることも多い。
4. 結論がわかりにくく、再利用性が低い
交わされた情報を後で再利用する場合、議論の過程を全て読まなければ結論がわからない。また、保存性のないサービスでは記録が残らず、情報の再利用が全くできない。
5. 情報のまとめ上げが面倒である
寄せられた回答や議論をもとにして「まとめ文書」を作成すれば、関連する全ての議論の内容を読むことなくして簡潔に結論だけを知りたい参加者の助

けになり、また、保存が容易になるために再利用性も高くなる。

しかし、「まとめ文書」の作成には多大な労力が必要なため、実際にはそのような行為が行われることは少ない。質問のためのニュースグループとして一般的に用いられている fj.questions.misc に投稿された記事を調べてみたところ、8/1~8/27 の間に投稿された 50 件の質問に対し、162 件の回答が寄せられたが、質問者が回答をまとめて記事を投稿した例は一件も確認できなかった。

1, 2, 3 の問題は、サービスの特徴に起因するものであり、複数のサービスを組み合わせることで解決が可能である。たとえば、ネットニュースのみを利用した場合には 1, 2 の問題が生ずるが、WWW を併用すれば解決できる。すなわち、質問、回答方法、および回答を保存しておく WWW ページの URL をネットニュースに投稿し、寄せられた回答を WWW ページに置くようにする。このようにすると、興味のある者だけが参加することになり、2. の問題を解決できる。また、即時性を持ったサービスを利用することで回答の重複が減少し、1. の問題を解決できる。

4. の問題はハイパーテキストの技法を用いて解決をはかることができる。たとえば、話題の流れを司るシステムを用いれば、参加者は話題の流れを追うことができる。文書の視覚化の研究の成果を用いれば、参加者は関連の深い話題を効率的に選んで読むことができる。また、5. の問題の解決は、4. の問題の解決に大きく寄与すると考えられる。

5. の問題の解決は容易ではない。これは、回答が「まとめ文書」の作成を前提にして書かれるわけではないからである。このような文書から「まとめ文書」を作成することは一般的に困難な作業である。

3 研究の目的と意義

前節で述べた問題のうち、特に 5. の問題を解決し、「まとめ文書」の作成を容易にすることがこの研究の目標である。情報交換の参加者が文書の書き手となり、「まとめ文書」の作成に協力できるような作業の枠組を提案し、かつ計算機によって途中の各作業過程を支援して「まとめ文書」を作成する人の労力を軽減する

表 2: 引用を行なっているメッセージの例

> ○○○をインストールしたのですが、
 > うまく動きません。
 > どこが悪いのでしょうか。
 ○○○はあなたの環境では
 うまく動作しないといわれています。
 私には原因は良くわからないので、
 もっと詳しい人に聞いて下さい。

> 他に似たようなソフトウェアはありませんか。
 ◎◎◎や×××がおすすです。
 動作が軽くて、設定も簡単です。

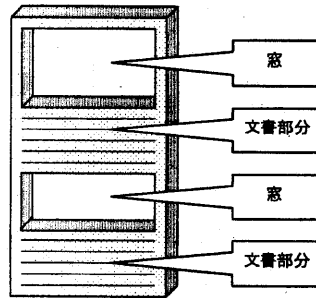


図 1: 「窓つき文書」

ことにより目標の達成をはかる。

このようにして作成された文書は複数の人が書いた部分が入り混じった文書となる。本稿では、このような文書を「協同文書」と呼び、協同文書を執筆する行為を「協同執筆」と呼ぶ。

協同執筆のような協同作業に関する従来の研究は、社内グループなどの比較的小さな、かつ結びつきが強く、お互いが信頼できるグループを対象にしていることが多い。一方、本稿で対象とするのは、コンピュータネットワークの上での比較的大きな規模の、互いが顔見知りでもないような弱い結びつきのグループである。

4 「窓つき文書」による協同文書の表現

他の人が新たに書き足すことによって文書を作成する行為である「引用」を例にとり、協同文書の表現方法について述べる。

電子メールやネットニュースに流れるメッセージに対する回答メッセージは、元のメッセージから適度に文脈を引用をすることにより、そのメッセージ単体でも、つまり元のメッセージを捜し出して読み返さなくても、意味がわかるように工夫されていることが多い。引用を行なっているメッセージの例を表 2 に示す。

数週間に到着したあるニュースグループの記事を調べたところ、明確に回答である記事のうち、元記事を全く引用していない記事は 3%程度に過ぎず、引用の行為は広く定着していると言える。

他の文書の一部を引用している文書は、図 1 のような、向こう側が見えるように「窓」がつけられた文書というモデルでとらえることができる [2]。この、「窓つき文書」を用いて引用を表現した様子を図 2 に示す。「窓と窓の対応づけ」によって、窓の向こう側を見ることができる。

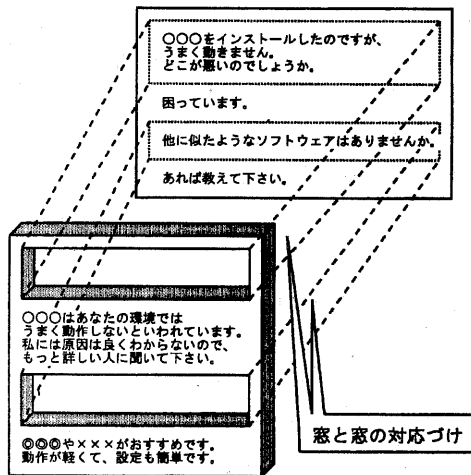


図 2: 「窓つき文書」を用いた引用の表現

ていないので「窓つき文書」の著者は一人であるといえる。「窓つき文書」を単位として、協同文書における個人の権利や責任の境界とすることができる。また、引用だけでなく、一般の校正行為、すなわち加筆、削除、訂正、複写、順序入れ換えなどは全てこの「窓つき文書」で表現することができる。

したがって、協同文書を表す構造として、「窓つき文書」を用いることは適当である。協同文書とは、「窓つき文書」をたくさん重ねて、上から見たものであると考えることができる。

5 システムの設計

3 節で述べた問題を解決するために、「まとめ文書」の作成を容易にするためのシステムについて述べる。まず作業モデルを考案し、モデル中の各作業に計算機

「窓つき文書」の中には引用元の文書は一切含まれ

支援によって連続性を持たせ、同時に作業における労力の軽減をはかる。特に、強く負担となっていた、文書の更新作業の自動化を実現する。また、協同文書の管理は「窓つき文書」を単位として行なう。

5.1 基本方針

情報を集積し、協同文書の作成を行なう人を「主催者」と呼ぶ。協同文書は、主催者によって所有され、管理される [3]。主催者は、自分の WWW ページを情報交換のための場として用意し、あらかじめネットニュースなどで協同文書作成への参加を広く呼びかけておく。他の人々は協同文書を直接ではなく、主催者に対しての働きかけを起こすことによって、間接に操作を行なう。以下、他の人々を「情報提供者」と呼ぶ。情報提供者から寄せられた文書は、主催者のもとでデータベース内に保存され、WWW ページ上で公開される。

以下では、まず作業の流れについて述べ、次にシステムの構成について述べる。

5.2 作業モデル

協同執筆作業のモデルとして、図3に示すような作業モデルを考案した。

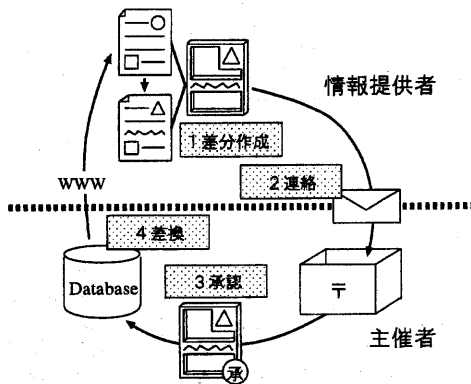


図3: 作業モデル

情報提供者はWWWを通じて協同文書を入手し、それに変更を加えて主催者に送付する。主催者は送付された内容を見て、承認/非承認を判断し、承認の場合は文書の差替を行なう。この一連の作業を繰り返すことにより、協同文書は次々と更新される。差替は自動的に行なわれるので、文書の更新にともなう主催者の労力は軽減される。

以下で各過程について述べる。

差分作成 情報提供者は、適当なエディタなどを用いて文書に変更を加え、新しい版の文書を作成する。変更内容は、変更前後の文書の相違、すなわち差分の形で保存する。この差分が「窓つき文書」である。

連絡 次に、作成した差分を主催者へ送付する。連絡手段としては電子メールを用いる。

承認 主催者は受け取った差分を見て、この変更を受け入れるかどうかを判断する。この過程は、不適切な情報が自分のホームページ上で公開されてしまうことを避けるために必要である。

差替 変更が承認された場合、差分を受け取り、差分と変更前の文書から変更後の文書を復元する。これを新しい版として保存し、文書を更新する。

5.3 システム構成

本システムの構成を以下に示す。本システムは、作業モデル中の4つの過程をそれぞれ支援するプログラムおよび、データベースを操作するプログラムの合計5つのプログラム、そして協同文書を保存するデータベースからなる。

差分作成支援部 情報提供者による変更前後の文書の差分を、差換支援部が理解する形式で生成する。

連絡支援部 情報提供者から主催者への差分の送付を行なう。

承認支援部 主催者による承認作業を支援する。差分は電子メールで送信されるので、メールリダの一部として実現される。

差換支援部 送付された差分を解析し、文書の更新に必要な、文書管理部が提供する操作の列を生成する。

文書管理部 協同文書を管理するデータベース管理プログラムである。外部に対するインターフェースとして、データに対する操作を提供する。

ハイパーテキストデータベース 協同文書の構造は、「窓つき文書」を用いて表現できることは既に述べた。本システムでは、データベース中に「窓つき文書」を保存する。「窓つき文書」はハイパーテキスト構造である。つまり、文書部分がノード、窓がアンカー、窓と窓との対応づけがリンクである。したがって、本システムで構築するデータベースはハイパーテキストデータベースである。

文書管理部は、ハイパーテキストデータベースにおけるノードやリンクなどのハイパーテキスト情報を管

理する [4]。たとえば、差分を元に差替が起こった場合、文書は実際を上書きされるのではなく、ハイパーテキスト構造を変化させることによって上書きされたように見せることになる。従って、元の版はそのまま残され、過去の履歴を参照することが可能である。

6 実装

作業モデルの妥当性、および主催者の負担が軽減されることを示すために、第一段階の実装を行なった。第一段階の実装は、協同文書に対する執筆形態を文書の末尾への追加のみに限定し、作業モデルにしたがって、最終的に情報をデータベース中に蓄積するものである。

差分作成支援部については、末尾への追加の場合、「窓つき文書」の構成は一つの文書部分と文書部分の前に現れる窓と決まっている。したがって、差分作成支援部で「窓つき文書」を生成する意味が薄く、今回は実装を行っていない。情報提供者は、文書を差分の形でなく、そのまま主催者へ送付する。

連絡支援部に関しては、既存の電子メールの配送機構をそのまま用いる。

承認支援部は、Emacs Lisp によって実装を行なった。プログラムを主催者の利用するメールリーダーに組み込み、主催者が承認/非承認の決定を行なうことができるようにした。承認された場合、連絡支援部から受け取った電子メールを差替支援部に送出する。承認の様子を図 4 に示す。図は三つの領域に分かれているが、図中の最上部がメールリーダーの機能である電子メールの一覧を表示している領域、中部が電子メールの内容を表示している領域である。最下部はミニバッファと呼ばれる領域であり、図中ではミニバッファに承認を求めるメッセージが出力されている。ここで「y」を入力すると承認が行なわれる。

```
1 Text/Plain
2 Message/Rfc822 "[hohnosec 970193] $
```

「情報通信ネットワークフォーラムの案内」

開催日：平成9年10月31日（金）
会場：京都市サテライトパーク・サイエンスホール（300名）
会費：1万円（懇親会費含む）
主催：立命館大学理工学研究所、京都市サテライトパーク（株）

```
Please answer y or n. <LI> 1997/10/31 {a href="dpsn
```

図 4: 承認の様子

差替支援部の実装には、Perl を用いた。文書管理部に対して、協同文書の末尾に追加する操作を指示し、承認支援部から受け取った電子メールの文書部分と共に送出する。

6.1 文書管理部の実装

文書管理部では、差替支援部からの命令に従い、受け取った文書をデータベースへ保存する。以下では文書管理部による協同文書のデータベースへの保存方法、および外部からの協同文書の参照に対する振る舞いについて述べる。

6.1.1 データベースへの保存

「窓つき文書」は抽象的なデータであるから、データベース中に保存する際にはそのままでは保存できない。そこで、「窓つき文書」の構成要素である文書部分、窓、窓と窓の対応づけ、にそれぞれ分けて保存する。

このとき、「窓つき文書」を取り出される際に用いる一意の文字列をキーとして与えておく。「連絡」は電子メールを介して行なわれるが、電子メールにはメッセージを一意に特定する Message-ID ヘッダが付加されているので、これを利用できる。これが「窓つき文書」を外部から参照するときに指定する識別子、すなわち URI となる。

6.1.2 外部からの参照要求

情報提供者およびその他の参加者からのデータの参照要求に対しての、文書管理部が行なう振るまいについて述べる。協同文書は WWW を通じて参照されるため、参照には HTTP プロトコルが使用される。まず、http サーバに対して協同文書を表す URI が指定される。次に、http サーバは文書管理部を CGI プログラムとして起動する。文書管理部は、データベースに対して指定された URI をキーとして検索を行なう。「窓つき文書」はデータベース中では各構成要素に分割されて保存されているが、これらをデータベース中から取り出して再構成することで、指定された「窓つき文書」がデータベースから取り出される。

「窓つき文書」は単独では意味の取れない文書となるので、参照されるときには窓を通して向こう側が見えるようにする必要がある。これは、窓と窓の対応づけをたどることで参照元の文書に到達し、窓に対応する参照部分を取り出すことで実現できる。窓の向こう側にもまた窓がある場合には、以上を再帰的に繰り返すことによって、「窓つき文書」が表す協同文書を完全に取り出すことができる。

7 試作システムの評価

現状における実装を用いた試作システムについて報告する。筆者らは試作システムを用いて研究室内で方々から寄せられる論文募集や講演参加者募集の情報を収集する機構を運用している。類似の試みとして、ネットニュースの特定のニュースグループに集められた会告記事の要約を自動作成するシステムがある [5]。

今回の試作システムでは、情報は追加されるのみであることを前提としている。結果、仕組みとしては単純になり、実装の一部でしかないが、作業モデルやデータベース登録方法の評価は可能と考えている。

情報提供者は自分が持っている論文募集や講演参加者募集の情報を主催者に電子メールで送付する。以下、作業モデルにしたがって作業が行なわれ、最終的に送付された情報はデータベースに格納され、WWWを通じて閲覧することができる。

このシステムを導入する以前、主催者は作業を手作業で行なっていたが、大量に承認の要求が寄せられて更新作業が滞ることがあった。しかし、承認支援部および差替支援部を用いて主催者に対する計算機支援を行なったところ、更新作業のために要する時間は大幅に短縮され、作業の滞りは解消された。したがって、主催者が更新作業のために必要とする労力の軽減は達成されたといえる。

現在までの半年間の運用の結果、約 90 件の承認要求が寄せられている。収集された情報を図 5 に示す。

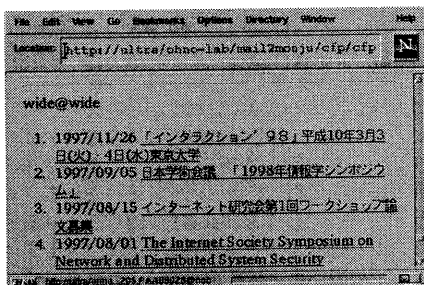


図 5: 収集された情報

文書の管理方法および文書管理部の実装に対する評価は、協同文書への追加だけでなく、さまざまな執筆形態で行なう必要がある。したがって、今回の試作システムの実装では十分な評価を与えることができない。文書管理部の実装に対する評価は、第二段階以降の実装を行なった上での、今後の課題とする。

8 今後の展望

ハイパーテキストデータベースの構築、およびこれを利用した協調作業にオブジェクト指向データベースの使用が有効であることは以前から指摘されている [6]。

本システムにおいても、「窓つき文書」のオブジェクトモデル化により、そのままの形でデータベース中に保存できる利点があるため、実装の一部としてオブジェクト指向データベースを用いることを検討している。その一例として、研究室内の備品情報をオブジェクトモデル化し、データベースとしてオブジェクト指向データベースを用いることによって、備品管理データベースシステムを運用する機構を製作中である [7]。

9 おわりに

コンピュータネットワークの上の人々が持つ情報を、協同執筆によって集積するための作業の流れを提案し、途中の過程の各作業を支援することで協同作業者の労力を軽減する仕組みを示した。また、協同文書を保存するためのデータベースの構築方法と、データベースを操作する機構を示し、設計を行なった。さらに、試作システムを実装し、運用することによって、作業の際の主催者の労力が軽減されることを示した。人は、他人からの質問に触発されてはじめて自分が持っている情報を表に出す場合が多い。したがって、人々に自分の持っている情報を積極的に文書の形に書き下すことを促す効果があると考えられる。

参考文献

- [1] 吉村 伸, 百科事典を作ろう, UNIX マガジン, Vol. 10, No. 12, pp.49-55, Dec 1995.
- [2] Theodor Holm Nelson., 竹内郁雄, 斉藤康己監訳, リテラリーマシソン, アスキー出版局, Oct 1994.
- [3] 門間信行, 大野浩之, ネットワーク上での協同執筆システム, 情報処理学会第 53 回全国大会論文集, Sep 1996.
- [4] Frank Halasz, Mayer Schwartz., The Dexter hypertext reference model. CACM, Vol. 37, No. 2, pp. 30-39, Feb 1994.
- [5] 佐藤田, 佐藤理史, 篠田陽一, 電子ニュースのダイジェスト自動生成, 情報処理学会論文誌, Vol. 35, No. 10, pp. 2371-2379, Oct 1995.
- [6] Kaj Gronbaek, Jens A. Hem, Ole L. Madsen, and Lennert Sloth., Cooperative Hypermedia Systems: A Dexter-based Architecture. CACM, Vol. 37, No. 2, pp. 64-74, Feb 1994.
- [7] 門間信行, 野田明生, 大野浩之, 小規模な組織の運営を支える情報共有機構--(2) 備品購入・管理システムの実装--, 情報処理学会第 55 回全国大会論文集, Sep 1997.