

VCoCoA : VCCML を用いた非同期型協同添削支援システム

緒方 広明[†] 葉田 善章[†] 矢野 米雄[†]

近年のインターネットの普及により、電子メールを用いて文章を交換しながら、複数の人物と協同で文章を執筆することが容易になった。しかし、従来の紙メディアでの添削と比較して、電子メールでは、修正が分かりにくいという問題がある。そこで我々は、紙面上での添削記号を用いた添削と同様の環境をコンピュータ上に実現することを目指した CoCoA を構築した。CoCoA は、挿入、削除などの添削記号を SGML に基づく通信添削用マークアップ言語 CCML で表現し、電子メールでの送受信による非同期の添削を可能とする。本論文では、CCML を拡張した VCCML によりバージョン管理を可能とした VCoCoA を提案する。これは、(1) 複数の添削者による複数回の添削で生成されたバージョンを管理する、(2) 確信度を導入して複数の添削結果のマー지를支援する、という特徴を持つ。評価実験の結果、紙メディアでの添削に比べて、VCoCoA が添削に役立つことが分かった。

VCoCoA: Markup Based Asynchronous Collaborative Correction System Using VCCML

HIROAKI OGATA,[†] YOSHIAKI HADA[†] and YONEO YANO[†]

Recently, users can write documents with other users using e-mail. We developed a computer supported communicative correction environment called *CoCoA* that supports editors and authors to exchange marked-up documents via e-mail. *CoCoA* provides a very similar environment to a real world one in which people use pen and paper. In order to record and exchange corrected compositions with marks and comments, we also proposed *CCML* based on SGML. This paper proposes how to manage the versions of corrected documents in an online markup environment. In particular, it is very hard to merge some corrected documents by different editors. This paper describes *VCCML* for versioning marked-up documents and a semi-automatical merge method with a confidence degree to help users. We have developed a prototype system called *VCoCoA* and the experimental results of its use showed that it could facilitate document correction, even better than those done with pen and paper.

1. はじめに

近年のインターネットの普及により、離れた場所に存在する複数の利用者が、電子メールを用いて協同で文書を作成することも行われるようになった。たとえば、企業では技術系文書や会議の議事録などを作成したり、大学では、共同研究報告書の作成などが行われている。しかし、電子メールでの文章の変更は、従来の紙面上の添削に比べて添削が行いにくく、その結果も分かりにくいという問題がある。

CSCW (Computer Supported Cooperative Work) の研究分野では同期 (リアルタイム) 型、非同期型の協同執筆システムが数多く提案されてきた¹⁾。たとえば、同期型として GROVE²⁾、ShrEdit³⁾ などがあり、

非同期型として Quilt⁴⁾、PREP⁵⁾ などがある。これらは、ともに 1 つの文章を効率的に協同作成することを主な目的としており、文章中にコメントなどを付加して、文章を訂正するものがほとんどであり、高度な添削機能は提供されてない。

一方、文章を書くには、内容のプランニングや書きながら文章を推敲していくモニタリング技能が必要となる^{6),7)}。執筆者自身が文章を推敲するだけでなく、文章を他者に添削してもらうことは、執筆者自身のセルフモニタリング技能を高める意味で重要である。そこで、我々は協同執筆作業で行われている紙面上での添削をコンピュータ上で可能とすることを目的として、ネットワーク型添削支援システム CoCoA (Communicative Correction Assisting System) を構築した^{8),9)}。CoCoA は執筆者と添削者が電子メールを用いて文書を交換し、文章の添削により協同で文章を作成する作業を支援する。CoCoA

[†] 徳島大学工学部知能情報工学科

Department of Information Science and Intelligent Systems, Faculty of Engineering, Tokushima University

では添削に用いる添削記号を統一し、執筆者と添削者には同様の添削結果を表示する。また、交換する添削文書の記述には SGML (Standard Generalized Markup Language)¹⁰⁾に基づく通信添削用マークアップ言語 CCML (Communicative Correction Markup Language)¹¹⁾を用いる。

我々は、外国人留学生を対象とした日本語の作文教育で CoCoA を使用し、その有効性を確かめた¹²⁾。作文教育の授業では、学生が書いた文章を教師が電子メールで受け取り、CoCoA を用いてそれを添削して返信する作業が行われた。CoCoA では、このように執筆者と添削者の間の 1 回だけの添削しか考慮されていない。しかし、一般の協同執筆作業では、最終的に文章を作成するまでの添削は 1 度だけで十分とは限らず、複数の添削者によって複数回添削が行われる場合が多い。

そこで本研究では、非同期での一般的な協同添削作業を支援するシステムとして CoCoA を拡張することを目的とする。本論文では、この場合に問題となるバージョン管理に関して、以下の手法を新たに提案する。

- (1) 最初の文章から最終的な文章になるまでの間、複数の添削者により複数回の添削が行われる過程を自動的にバージョン管理するため、CCML を拡張した VCCML (Versioned CCML) を提案する。これにより、執筆者は誰がどのように文章を修正したかを知ることができる。また、各添削者も文章を添削するにあたり、文章が変更された経緯を参照できる。
- (2) 複数の添削者が複数の紙面上で添削を行った場合、それらを 1 つにまとめる作業は困難となる。本論文では、複数の添削結果をマージする支援として、採用する添削記号を決定する指標となる確信度を導入し、半自動マージ機能を提案する。

これらの提案をもとに、我々は CoCoA を拡張し、非同期型協同添削支援システム VCoCoA (Versioned CoCoA) を構築した。

以下、2 章で紙メディアと電子メディアによる文章の添削手法について考察を行い、3 章でネットワーク型添削支援システム CoCoA の概要と特徴について述べる。また、電子メールによる非同期添削におけるバージョン管理手法とマージ手法を 4、5 章でそれぞれ提案する。さらに、CoCoA にバージョン管理に付加した VCoCoA の実装を 6 章で述べ、最後に、7 章でシステムを評価実験し、その考察を行う。

2. 紙メディアと電子メディアによる文章添削

2.1 紙メディアを用いた文章添削の特徴

文章の添削は従来、紙面上で行われてきた。その特徴を以下に示す。

- (1) 添削記号を用いることで添削の位置、添削内容が分かりやすい。
- (2) 文章の行間など紙面の余白を添削に用いる。余白が広いほど、複雑な添削が可能である。
- (3) 原文と添削後の文章の対応がとりやすく、一覧性が良い。
- (4) 訂正後の文章を作るには、添削を参照しながら文章を書き直す必要がある。
- (5) 最終的に文章を作成するまでに繰り返された添削結果を管理することは困難である。
- (6) 複数の人物に添削を依頼した場合、その異なる添削結果を 1 つにまとめることは困難である。

本研究では、電子メディアを用いた添削により、紙面上での添削の (1)~(3) の特徴を保持しつつ、(4)~(6) の改善を目指す。

2.2 電子メディアによる文章添削手法の分類

これまで、コンピュータを用いて様々な方法で文章の訂正・添削が行われてきた。これは、以下のように分類される¹³⁾ (図 1 参照)。

- (I) **Silent editing model**: 原文に直接上書きする方法である。これは最も単純な方法であり、特別な処理を必要としない。しかし、訂正文から原文との相違点を探すことは困難である。電子メールを用いた文章の添削では、多くの場合このモデルが用いられる。
- (II) **Comment model**: 原文中に特別な注釈記号などを挿入し、変更内容をコメントとして付加する方法である。このモデルは特定のグループ内でのアドホックな利用に向いている。Quilt⁴⁾、PREP⁵⁾はこのモデルを採用している。
- (III) **Edit trace model**: 添削者が原文を変更していく過程を記憶しておき、修正終了後にその変更点を示す方法である。このモデルでは、添削者は Silent editing model と同じように簡単に編集を行うことができる。その反面、原文を無視して過度に変更してしまう問題がある。また、変更は基本的に挿入と削除の 2 種類の添削記号で表されるため、複雑な修正が行われるとその結果が分かりにくくなる。このモデルは、Microsoft Word などの多くのワードプロセッサに採用されている。

Original text

CCML (Communicative Correction Markup Language) は、テキストで書かれた日本語添削文章の交換のために用いるマークアップ言語であり、SGML (Standard Generalized Markup Language) を基にしている。

SGML は、文章交換を目的としたもので1986年に、ISOで定められた企画である。

(I) Silent editing model

CCML (Communicative Correction Markup Language) は、テキストベースの日本語添削文書の交換に用いるマークアップ言語であり、SGML (Standard Generalized Markup Language) を基にしている。SGML は、文書交換を目的として1986年に、ISOで定められた規格である。

(II) Comment model

CCML (Communicative Correction Markup Language) は、テキストで書かれた日本語添削文章*1の交換のために用いるマークアップ言語であり、SGML (Standard Generalized Markup Language) を基にしている。*2 SGML は、文章交換を目的としたもので1986年に、ISOで定められた企画*3 である。	*1 文章より文書の方が適切です。 *2 新しい段落にする必要はありません。 *3 「規格」ですね？
--	--

(III) Edit trace model

CCML (Communicative Correction Markup Language) は、テキストベースで書かれた日本語添削文章文書の交換のために用いるマークアップ言語であり、SGML (Standard Generalized Markup Language) を基にしている。SGML は、文章文書交換を目的としたもので1986年に、ISOで定められた企画規格である。

(IV) Traditional markup model

CCML (Communicative Correction Markup Language) は、テキストベースで書かれた日本語添削文章の交換のために用いるマークアップ言語であり、SGML (Standard Generalized Markup Language) を基にしている。SGML は、文章交換を目的としたもので1986年に、ISOで定められた企画規格である。

図1 コンピュータを用いた添削モデル

Fig. 1 Computer based correction model.

(IV) **Traditional markup model**: 紙面上での添削と同様に、原文中に添削記号(削除・挿入・移動など)を挿入して添削を行う方法である。添削に用いる記号は執筆者や添削者にとって直感的に分かりやすく、親しみやすい。たとえば、MATE¹⁴⁾はこのモデルを実装している。MATEではユーザがタブレットを用いてテキスト上に行った添削をシステムが認識して文章を訂正する。しかし、使用する添削記号は統一されておらず、バージョン管理も考慮されていない。

この分類では、(I)~(IV)の順にユーザにとって添削結果が分かりやすくなる。そのため、本研究ではTraditional markup modelモデルに着目する。

3. CoCoAの概要

3.1 CoCoAの特徴

CoCoAの特徴を以下に示す。

- (1) 電子メールを利用して文書を交換することにより、執筆者と添削者が非同期に添削作業を行えるため、時間的・場所的な制約を受けない。また、添削文書の電子化により、その管理や分析が容易になる。
- (2) WYSIWIS (What You See Is What I See)¹⁵⁾の概念に基づき、原文執筆者と添削者には同様の添削記号を用いた添削結果を表示する。
- (3) CoCoAは文書を作成するエディタや電子メールのツールとは独立であり、ユーザは使い慣れた環境をそのまま利用することができる。
- (4) 通信添削用マークアップ言語CCML形式で添削文書を表現するため、電子メールを用いて添削文書を交換できる。
- (5) CCML形式の添削文書から原文や清書文を自動的に作成できる。
- (6) 添削者は各添削記号に重要度を付けることができ、執筆者は重要な添削から順に見ることができ、具体的には、重要度は(1)訂正しなくてもよい、(2)訂正したほうがよい、(3)必ず訂正しなければならない、の3段階である。

3.2 CCMLの特徴

通信添削用マークアップ言語CCMLは、文書添削に用いる添削記号と添削内容を構造化して表現する。これは以下の特徴を持つ。

- (1) JIS規格の添削記号¹⁶⁾から抽出した電子文章の添削に必要となる6つの添削記号と注釈を表1に示すタグで表現する。
- (2) CCMLタグをとると原文になり、タグを適用すると訂正後の文章になる。
- (3) テキスト形式であるため、コンピュータの機種やソフトウェアに依存せず、電子メールで送受信できる。

3.3 CoCoAの動作

CoCoAは添削者が利用するCoCoA-Editorと、添削結果を表示するCoCoA-Viewerからなる。CoCoA-Editorのインタフェース例を図2に示す。以下、CoCoAの動作を示す。

- (1) 執筆者は添削者に原文を電子メールで送る。
- (2) 添削者は電子メールで送られた原文をCoCoA-Editorに読み込む。CoCoA-Editorは、添削が行いやすいように行間を空けて表示する。

表1 CCMLで用いる添削用タグ
Table 1 CCML tags for correction.

添削	タグ表現
挿入	<Insert string="挿入文字列">
訂正	<Change string="変更後文字列"> 変更前文字列 </Change>
削除	<Delete> 削除文字列 </Delete>
改行	<Separate>
追込	<Join>
移動	<MoveFrom refid="番号"> <MoveTo id="番号"> 移動文字列 </MoveTo>
注釈	<Annotate type="質問 説明" string="文字列">

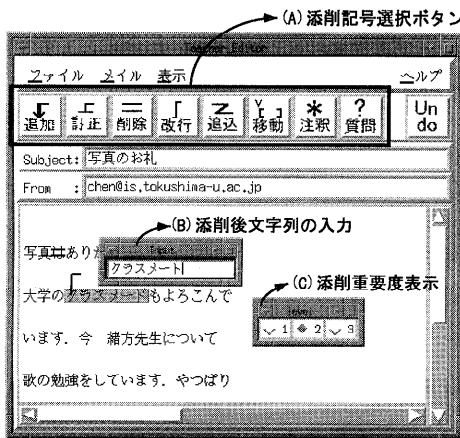


図2 CoCoA-Editorのインタフェース例
Fig. 2 Interface of CoCoA-Editor.

- (3) 添削者は図2(A)のボタンを用いて添削記号を選択し、テキスト中に添削を行う。(B)は訂正後の文字列を入力する画面である。また、(C)により、添削記号ごとに重要度を指定できる。
- (4) CoCoA-Editorは添削結果をCCML形式で保存し、添削者はそれを電子メールで原文の執筆者に送る。
- (5) 執筆者はCoCoA-Viewerを用いてCCML文書を読み込み、添削結果を参照する。

このようにCoCoAは原文執筆者と添削者の間での文章交換は可能であるが、同時に複数の人物に添削を依頼したり、添削を繰り返し行うことを考慮していない。本論文では、CCMLを拡張して、添削文章のバージョン管理を行う手法を提案する。

4. バージョン管理手法

4.1 添削文書の流れ

1つの文章を協同で作成するとき、複数回の変更を行ったり、複数の添削者に添削を行ってもらう場合がある。CoCoAを用いた場合、図3(A)のように執筆者

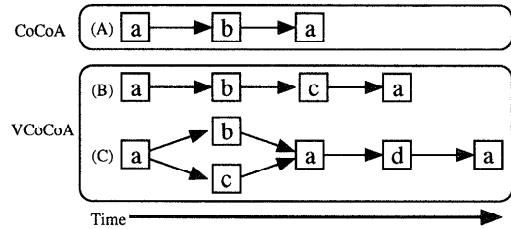


図3 CoCoAとVCoCoAの添削文章の流れ
Fig. 3 Document flow of CoCoA vs. VCoCoA.

者aと添削者bの間の1回だけの添削しかCCMLで管理できない。添削者bの添削後、さらに他者に添削してもらう場合、別のCCML文書が作成されるため、CoCoAでは前のCCML文書と関連付けできない。

そこでVCoCoAでは、図3(B)、(C)に示すように複数の添削者による複数回の添削過程で生成されたバージョンの管理を目指す。ここで、図3(B)のように添削者が逐次添削していく場合を直列添削型と呼び、(C)のように複数の添削者が並行して添削する場合を並列添削型と呼ぶ。並列添削型の場合、複数の添削結果のマージを行う必要がある。

4.2 バージョン管理の方針

VCoCoAでのバージョン管理の方針を以下に示す。

- (1) 複数回の添削・複数の添削者に対応：複数の添削者に添削を同時に依頼したり、添削を繰り返し行う場合のバージョンを管理可能とする。
- (2) 添削されるごとにバージョンを更新：文章に対する添削が終了するたびにバージョンを更新する。
- (3) 文章ごとにバージョン管理：バージョン管理は、1つの文章が最終的な文章になるまでを対象とする。
- (4) 全バージョンを電子メールで送受信：添削者がそれまでの添削の経緯を理解できるようにするため、最初の文章からの全バージョンを添削者に送信する必要がある。

4.3 VCCMLによるバージョン表現

上述の方針に基づき、添削文書のバージョンを表現するVCCMLを以下のように設計する。

- (1) 添削が終了するたびに作成されるCCML文書をバージョンの1単位とし、それらを双方向リンクで関連付ける。そのため、前バージョンの清書文が次バージョンの原文となる。
- (2) 同一ファイル内に各バージョンを記録し、そのファイルを電子メールで交換する。

CCMLに追加されたVCCMLタグを表2に示す。従来のCCML文書の内容をバージョンを表すVersionタグで囲み、prev, nextの2つの属性でバージョン

表2 VCCMLで用いるバージョンタグ
Table 2 VCCML tags for version management.

タグ	属性	属性の意味
Version	curr	現在のバージョンの ID
	prev	前のバージョンの ID
	next	次のバージョンの ID

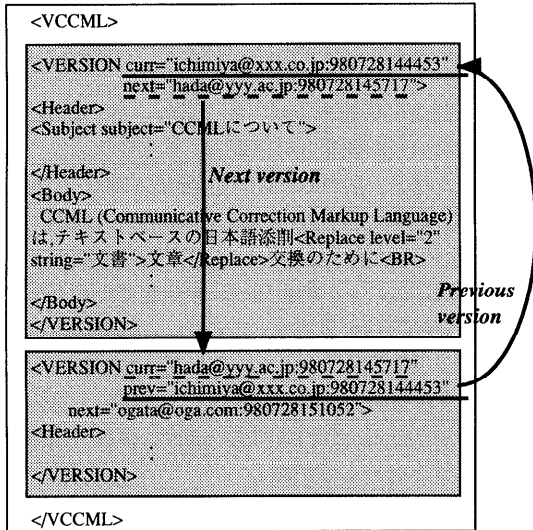


図4 VCCML 文書の例

Fig. 4 Example of VCCML document.

間をリンクすることで、バージョン管理を実現する。VCCML 文書の例を図4に示す。このようにタグ内の prev, next の2つの属性は、前のバージョンと次のバージョンを双方向リンクで関連付けを行う。ユニークなバージョン ID の生成には、添削者のメールアドレスと添削時刻を結合した文字列を用いる。

4.4 関連研究

ソフトウェアの協同開発においてソースコードのバージョン管理を行うツールとして、SCCS (Source Code Control System)¹⁹⁾, RCS (Revision Control System)²⁰⁾, CVS (Concurrent Version System)²¹⁾がある。これらはともに、ソフトウェアの開発段階において各ファイルをコンパイルする必要があるため、バージョン管理の対象となるファイルを1カ所で集中管理する。これに対して、協同添削作業では様々な添削者に個別に添削をしてもらい、最終的に文章を作成できればよく、SCCSのように全バージョンを集中管理する必要はない。そこで VCoCoA では、協同添削の過程で電子メールで各バージョンを送受信することにより、それらをユーザごとに分散して管理するアプローチをとる。また、SCCS などではシステム側がバージョン間の差異を抽出して、各バージョンの更新

内容を表現する。したがって、ユーザが行った変更とは表現が異なる場合がある。一方、文章の協同添削を考えた場合、添削者が文章のどこをどのように変更したのか、を添削依頼者側に正確に提示することが重要となる。VCoCoA では添削者が挿入した添削記号に従って文章の変更内容を記録するため、添削者が行った文章の変更を正確に表現できる。

また、文書の協同編集において、そのバージョン情報を記録するマークアップ言語 VTML (Versioned Text Markup Language)²²⁾が提案されている。これは、主に World Wide Web 上の HTML 文書をバージョン管理の対象とする。この方式は、挿入、削除のタグを用いて同一文書上に全バージョンを合成して表現するため、文字列の移動などの複雑な添削や、複数の添削者による複数回の添削を正確に表現することは困難である。これに対して、本論文で提案する VCCML は、バージョンごとに分割して同一ファイル内に記録するため、Traditional markup model で用いる複雑な添削記号が表現可能であると同時に、複数の添削者による複数回の添削という複雑なバージョンの管理も可能である点で優れている。

5. 添削文章のマージ手法

複数の添削者によるバージョンのマージは、システムが完全に自動的に行うことは困難である。そこで、本論文ではマージ対象となる文章中の添削記号が一意に適用できるかどうかを表した確信度を導入し、半自動マージを提案する。半自動マージでは、添削記号の確信度を参考にして、システムがある程度マージ処理を行うが、その後、ユーザが最終的に添削記号の決定を行う。

5.1 半自動マージの手法

半自動マージに関するシステムの処理の流れを以下に示す。

- (1) 添削記号の標準化: マージする複数の添削文章内の添削記号を標準化し、添削方法の比較処理をしやすいとする。
- (2) 添削記号の比較: 添削者ごとに添削内容を比較し、他の添削者による添削記号との重複 (Overlap) や衝突 (Conflict) を調べる。
- (3) 確信度の算出: 全添削記号に対してマージ確信度を算出する。
- (4) 採用する添削記号の決定: 算出した確信度やユーザの設定をもとに、マージした結果に採用する添削記号と重要度を決定する。不採用の添削記号は、ユーザに判断を促す。

5.2 添削方法の標準化

添削者により、添削後の文章は同一になるが、添削の方法が異なる場合があり、これは添削記号を比較する際に障害となる。たとえば、“クラスメード”を“クラスメート”と訂正する場合を考える。この場合、単語全体の“クラスメード”を“クラスメート”と修正する添削者もいれば、訂正部分である「ド」を「ト」と修正する添削者もいる。この場合、システムは異なった添削と見なしてしまう。そこでシステムが原文と訂正文の差異の抽出を行い、添削方法を統一することで、添削記号の標準化処理を行う。これにより、システムが類似した添削をまとめることができるため、マージにかかるユーザの負担を軽減できる。

5.3 マージ確信度

マージ確信度 (Confidence degree of marge: CDM) はシステムがその添削記号を一意に決定できるかどうかを表した数値である。この値は1から100の整数値をとり、マージしたすべての添削記号に対し個々に求める。以下に確信度を求める方針を示す。

- (1) 添削記号が一意に決定できるほど値を高くする。
 - (2) 添削者全員が同一の添削箇所に対して、まったく同じ添削を行った場合、値を100とする。
 - (3) 同一の添削箇所に対して異なる添削がある場合、数値を低くする。
 - (4) 重要度にばらつきがある場合、数値を低くする。
- これらの点を考慮した算出式を以下に示す。

$$CDM = \left(1 - \sqrt{\frac{1}{C} \sum_{i=1}^C (w_i^2 - \bar{w}^2)} \right) \frac{O}{P} \times 100 \quad (1)$$

ここで、

P : 全添削者数

C : その添削箇所での添削数

O : その添削箇所での添削の重複数

\bar{w} : その添削箇所での添削記号の重要度の平均

w_i : その添削箇所での i 番目の添削記号の重要度である。上式は添削記号の重要度のばらつきと、添削記号の重複度を乗じたものである。

たとえば、ひらがなを漢字に直したり、漢字のミスのような比較的簡単な誤りを考える。このような誤りは、同時に複数の人が添削する確率が高く、確信度は高くなる。一方、文章の内容の添削の場合、添削記号の重複が少なくなり、確信度が低くなる。これは、内容に関しては添削者それぞれに捉え方が異なり、表現方法が異なるためである。この場合、ユーザはこの部分の添削結果を再確認する必要がある。このように、確信度はユーザがマージを行う際、再確認が必要かど

うかを表す指標となる。

5.4 添削記号の決定

マージを行う場合、同じ箇所異なる添削が行われる場合がある。この場合、以下の優先順位で添削を採用する。

- (1) 確信度の最も高い添削を優先する。この場合、その箇所における添削のうち、最も頻度が高いものが採用される。
- (2) 特定の添削者の添削を優先する。
- (3) 重要度が高い添削を優先する。

ここで初期値は(1)であるが、マージを行う前にユーザが選択可能とする。

5.5 マージ後の重要度の決定

添削記号の重要度は添削者によりまったく同一添削でも異なる場合がある。同一の場合は問題なく処理できるが、異なる場合はマージ後の重要度をシステムで決定する必要がある。そこでマージ後の重要度の決定方法として、以下を用意する。

- (1) 全重要度の平均値を用いる。
- (2) 多数決により最も多い重要度を用いる。
- (3) 基準となる添削者がいる場合、その添削者の重要度を用いる。

これらは、添削記号の決定方法と同様に、マージを行う前にユーザが選択する。

6. VCoCoAの実装

我々はCoCoA-EditorとCoCoA-Viewerを利用し、VCoCoAを試作した。開発はSun Workstation上でを行い、Tcl/Tk²⁴⁾を用いた。詳細は文献(17)、(18)を参照されたい。以下、VCoCoAで実装した機能を述べる。

6.1 バージョン管理の支援機能

VCoCoAのバージョン管理の機能を以下に示す。

- (1) バージョンネットワークの視覚化: 各バージョンのつながりを図5(A)のように視覚的に表示する。画面中の矢印は文章が添削された流れを表し、名前は添削者を意味する。この例は、“一宮”が最初に文章を作成し、次に“葉田”と“森”が添削した後、“一宮”がマージしたことを表す。
- (2) バージョンの表示: これは、ユーザが指定したバージョンの添削結果を表示する機能である。原文表示、訂正文表示、重要度別添削表示など、CoCoA-Viewerと同様の機能を持つ。また、複数の添削を重ねて表示することが可能である。この場合、添削者ごとに色を変えて添削記号が

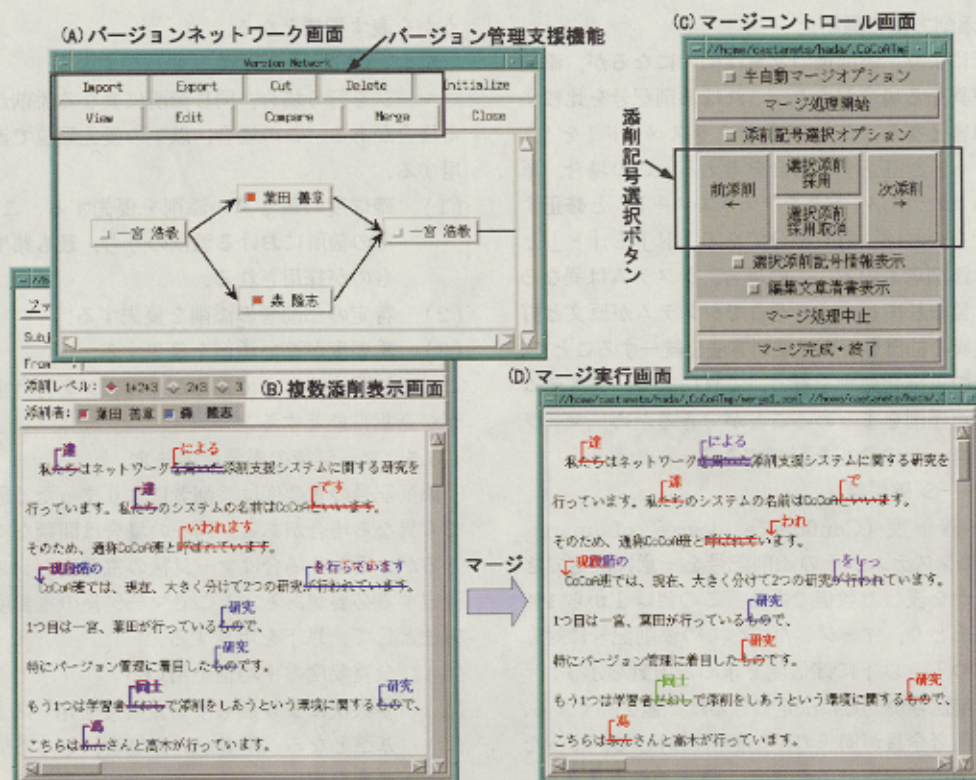


図5 VCoCoAのインタフェース
Fig. 5 Interface of VCoCoA.

表示されるため、添削が同じ箇所や異なる箇所の確認が容易に行える。

- (3) バージョンの追加・編集：ユーザが指定したバージョンの訂正文を CoCoA-Editor で添削し、その後、新たにバージョンを追加する機能である。
- (4) バージョンの結合：複数の添削者に添削を依頼した場合、添削されて執筆者に戻る CCML 文書は複数のファイルに分かれる。そこで、この機能は分割された CCML 文書を1つにまとめる。
- (5) バージョンの削除：バージョン管理機構では、1つのファイルに複数のバージョンを含むため、添削が多く繰り返されると、ファイルサイズが大きくなる。そこで、この機能により不要となったバージョンを削除する。
- (6) バージョン間の差異表示：複数の人物に添削を依頼したとき、それぞれの結果を比較する場合がある。この機能は、原文が同じ任意の2つのバージョンにおいて、それらの添削の差異を表示する。返信された添削結果を比べる場合に利用できる。この訂正文の比較には diff²³⁾を用い、2つの添削結果の差異は添削記号を用いて表現する。

6.2 半自動マージ機能

システムの流れに沿ってマージ機能を説明する。

- (1) ユーザがマージを行うバージョンを画面 (A) で選択する。
- (2) ユーザはマージに関して以下のオプションを設定できる。

確信度の基準値：この基準値以上の確信度の添削記号をマージ後の文章に採用する。初期値は 50 である。

添削記号の採用優先方法：デフォルトでは、確信度に従って添削記号の採用を決定するが、採用を優先する添削者が存在する場合、あるいは、重要度を優先する場合は、オプションを指定する。

重要度の決定方法：重要度の決定方法を指定する。

- (3) ユーザは画面 (D) に示すマージ結果を参照しながら添削文書を修正する。修正する添削記号を画面 (B) で選択し、採用または不採用を決定する。(D) 上の各添削記号は次のように色分けされる。

緑 ユーザが選択している添削記号

赤 マージ後の文章に採用される添削記号
 青 確信度は高いが、ユーザにより採用を取り消された添削記号

紫 確信度が低く、採用されていない添削記号
 ユーザが望むマージ結果が得られない場合は、オプションを再設定し、再びマージ処理を行う。

7. 評価実験と考察

我々は試作したVCoCoAを用いて評価実験を行った。

7.1 実験方法

本実験での評価項目を以下に示す。

- (1) 様々な添削作業でシステムを利用してもらい、バージョンの作成状況、添削状況などを調査する。
- (2) 半自動マージ機能の有効性を検討する。
- (3) 利用者にアンケートに答えてもらい、システムを全体的に評価する。

我々は研究室内の6つのグループ、計20人を対象に評価実験を行った。1グループは2~4人で構成されており、1グループあたりの平均人数は3.3人である。各グループに1週間VCoCoAを利用してもらい、以下の3つの課題を与えた。ここで、並列型協同添削作業は、文章全体を添削する全体添削型と、文章を複数の執筆者で分担して添削する分担添削型に分けた。

作業1: 直列添削型 具体的には、グループ内の最年少の人物が自分の研究について紹介文を作成し、自分より年上の人物に添削を依頼していき、最終的に最初の人物に返してもった。この場合、マージは起こらない。

作業2: 並列・全体添削型 グループ内の1人が研究グループ全体の紹介文を作成し、他のメンバに同時に添削してもらい、結果をマージして文章を作成した。

作業3: 並列・分担添削型 グループ内で1つの住所録を作成するために、グループ内の1人が研究グループ全員の名前を書いた原文を作成し、他の全員に送信した。受信者は原文に個人情報を添削形式で記入し、最終的に原文作成者がまとめて住所録を完成した。

評価中はシステム内で各添削者の行動を保存し、評価後にこのログを解析した。作業の終了後、5段階形式と記述形式でアンケートを行った。

7.2 実験結果と考察

以下、実験結果を示し、考察を行う。

7.2.1 システムの利用状況

各作業ごとの添削の状況を表3に示す。1グループ

表3 各作業における添削状況

Table 3 The state of correction in each group.

1グループの平均	作業1	作業2	作業3
添削前の文字数	327.2	529.5	242.0
添削後の文字数	386.2	552.5	382.8
添削回数	3.2	3.7	2.5
添削記号数	24.6	20.3	8.8
バージョン参照回数	15.4	15.3	3.3

が作成した文章の平均文字数は、添削後に増加した。これは、情報を付加するための添削が主に行われたためである。全グループが行った各作業の添削回数の合計は56であった。各作業における内訳は作業1から順にそれぞれ19, 22, 15である。表3に示すように、各作業に対して1グループ内で平均3回添削が行われた。グループごとの添削記号数を見ると、作業3が最も少ない。これは、添削対象が文章の一部だけであったためである。

一方、ユーザが添削を行う際、過去のバージョンを参照した回数を表すバージョン参照回数は、作業1, 2ともに平均15回以上であった。このように、自分で添削する前に他者の添削履歴を参照しており、バージョン管理機能は有効であることが分かった。

7.2.2 マージ機能に関する評価

実験中に全ユーザ実行したマージの合計回数は6回であり、その内訳は作業2が2回、作業3が4回であった。我々は、以下に示す指標を用いて、バージョンのマージ機能を評価した。

$$\text{再現率} = \frac{x \text{のうちシステムが採用した添削数}}{\text{ユーザが採用した添削数}(x)} \quad (2)$$

$$\text{適合率} = \frac{y \text{のうちユーザが採用した添削数}}{\text{システムが採用した添削数}(y)} \quad (3)$$

$$\text{重複率} = \frac{\text{重複のある添削数}}{\text{マージ対象の全文書中の添削数}} \quad (4)$$

$$\text{衝突率} = \frac{\text{衝突のある添削数}}{\text{マージ対象の全文書中の添削数}} \quad (5)$$

それぞれの算出結果を表4に示す。作業2はグループ内のメンバ全員が文章全体を添削したため、添削の重複および衝突があった。再現率が86.4%となったのは、システムによるマージ処理の後、ユーザが添削を追加したためである。しかし、システムが提案したマージ結果は適切であったため、適合率は100%となった。このようにマージ文章中に重複や衝突がある場合でも、VCoCoAはユーザのマージ作業に役立つ試案を生成する意味で有効であった。

一方、作業3の住所録作成では個人の担当部分のみを修正するため、添削記号の重複および衝突はまったくなかった。マージ処理では、システムはすべての添削記号を採用し、ユーザはそのマージ結果を変更しな

表4 マージ結果の適合率, 再現率, 重複率, 衝突率 (%)
Table 4 The ratio of relevance, recall, overlap and conflict in the marge results.

作業	再現率	適合率	重複率	衝突率
作業2	86.4	100.0	33.3	2.9
作業3	100.0	100.0	0.0	0.0

表5 アンケートの回答結果 (5段階形式)
Table 5 The results of questionnaire.

質問内容	結果	人数
(i) 紙面上での添削に比べて役立つと思いますか?	4.38	20
(ii) 過去の添削結果は参考になりましたか?	3.63	20
(iii) バージョンネットワークの表示は分かりやすかったですか?	4.38	20
(iv) 各添削記号のマージ確信度は適切でしたか?	3.80	5
(v) マージした結果は適切でしたか?	4.80	5

かったため、適合率・再現率ともに100%であった。したがって、このように重複率が高く衝突率が低い文章をマージする場合、ほぼ自動的に行えるため、本システムは有効である。

7.2.3 アンケート結果

評価実験終了後に行ったアンケート結果を表5に示す。値は質問に最も同意する場合を5とし、1から5をとる。表の人数欄は回答を行った人数を示す。

(i) に関して、VCoCoAは文章の添削過程を自動的にバージョン管理できるため、ユーザの負担を軽減でき便利である、という意見が多く、グループ内で通常行われている紙面上での添削に比べて、VCoCoAによる添削は高く評価された。また、(ii)の結果から、ユーザは過去の添削結果をある程度参考にしながら添削を行っており、最初の文章から現在までの全バージョンを電子メールで送ることに意味があることが分かる。さらに、(iii)の結果から、バージョンネットワークの表示は分かりやすかったといえる。半自動マージ処理の評価に関して、(iv)の結果から、各添削記号に付与された確信度はほぼ適切と評価された。また、(v)の結果から、本実験においては、システムはユーザが十分満足するマージ結果を提示できた。

システムの全体の評価に関して、ユーザがすべての添削を行うことは負担である、という意見もあり、文法チェックやスペルチェックなどの文章誤りの訂正技術²⁵⁾を導入した自動添削機能を今後検討する必要がある。

8. おわりに

本論文では、従来、紙面上で行われてきた添削作業をネットワークを用いてコンピュータ上に実現することを目指し、電子メールを用いた非同期の協同添削作

業支援システム VCoCoA を提案した。本システムの特徴は、VCCML を用いて添削過程を記録することにより、バージョン管理を可能とし、複数の添削者による添削結果を1つにまとめるマージ作業を支援することである。我々は VCoCoA を実装し、評価実験を行った結果、紙面上の添削作業に比べて本システムが有効であり、バージョン管理、マージ機能にも高い評価を得た。なお、本論文では研究室の学生を対象とした3種類の協同添削作業をもとに本システムを実験的に評価した。そのゆえ、本実験は被験者などの点で限界があり、その結果も限定的なものである。今後は、VCoCoA を利用するユーザグループや、文章内容、また文章量などを変更して、さらに実験データを増やし、VCoCoA の有効性を検証する必要がある。

近年、オフィスではネットワーク化やペーパーレス化が展開されており、本研究はネットワークを用いて文章を交換し、文章の添削・推敲プロセスを記録するという、次世代のワードプロセッサを目指す意味で重要であると考えられる。今後は、テキスト情報だけでなく、図表、音声、動画などを含めたマルチメディアデータを扱えるようにシステムを拡張していく予定である。

謝辞 本研究の一部は、文部省科研費、基盤研究(B)(2) No.09558017, No.09480036, 萌芽的研究 No.11878032, および電気通信普及財団の援助を受けた。また、本システムの開発にあたり、一宮浩教氏(現在、株式会社 APEX 勤務)にご尽力いただいた。ここに記して謝意を表する。

参考文献

- 1) 松下 温, 岡田謙一: コラボレーションとコミュニケーション, 共立出版(1995).
- 2) Ellis, C.A., Gibbs, S.J. and Rein, G.L.: Groupware Some Issues and Experiences, *Comm. ACM*, Vol.14, No.1, pp.33-58 (1991).
- 3) McLaughlin, H.C. and Olson, G.M.: Unblock-ing Brainstorming Through the Use of a Simple Group Editor, *Proc. CSCW '92*, pp.99-106 (1992).
- 4) Leland, M.D.P., Fish, R.S. and Kraut, R.E.: Collaborative Document Production Using Quilt, *Proc. CSCW '88*, pp.206-215 (1988).
- 5) Neuwirth, C.M., Kaufer, D.S., Chandhok, R. and Morris, J.: Issues in the design of computer support for co-authoring and commenting, *Proc. CSCW '90*, pp.183-195 (1990).
- 6) Scadamalia, M. and Bereiter, C.: Research on Written Composition, *Handbook of research on teaching*, Wittrock, M. (Ed), pp.778-801, Macmillian (1986).

- 7) 波多野 諄余夫 (編): 学習と発達, 認知心理学, 5, 東京大学出版会 (1996).
- 8) 矢野米雄, 緒方広明, 榊原理恵, 脇田里子: 日本語作文教育のためのネットワーク型添削支援システム CoCoA の構築, 教育システム情報学会論文誌, Vol.14, No.3, pp.21-28 (1997).
- 9) Ogata, H. and Yano, Y.: CoCoA: Communicative Correction Assisting System for Composition Studies, *Proc. ICCE 97*, pp.461-468 (1997).
- 10) Herwijnen, E.: *Practical SGML*, Kluwer Academic Publishers (1990). SGML 懇談会実用化 WG (監訳): 実践 SGML, 日本規格協会 (1992).
- 11) Ogata, H., Yano, Y. and Wakita, R.: CCML: Exchanging Marked-up Documents in a Networked Writing Classroom, *Computer Assisted Language Learning: An International Journal*, Vol.11, No.2, pp.201-214 (1998).
- 12) 脇田里子, 緒方広明, 矢野米雄: 作文教育のためのネットワーク型添削支援システム CoCoA の実践と評価, 教育システム情報学会論文誌, Vol.15, No.4, pp.270-275 (1999).
- 13) Farkas, K.D. and Poltrock, E.S.: Online editing mark-up models, and the workspace lives of editors and writers, *IEEE Trans. Professional Communication*, Vol.38, No.2, pp.110-117 (1995).
- 14) Hardock, G., Kurtecbach, G. and Buxton, W.: A marking based interface for collaborative writing, *Proc. ACM Symposium on User Interface Software Technology*, pp.259-266 (1993).
- 15) Stefik, M., Bobrow, D., Foster, G., Lanning, S. and Tatar, D.: WYSIWIS Revised: Early Experiences with Multiuser Interfaces, *ACM Trans. Office Information Systems*, Vol.5, No.2, pp.147-167 (1987).
- 16) 島野 一: 校正実務ハンドブック, みき書房 (1986).
- 17) 一宮浩教, 葉田善章, 緒方広明, 矢野米雄: ネットワーク型添削支援システム CoCoA におけるバージョン管理手法の提案, 教育システム情報学会研究報告, Vol.98, No.3, pp.46-53 (1998).
- 18) 葉田善章, 一宮浩教, 緒方広明, 矢野米雄: 協同添削支援システムにおけるバージョン管理機構の構築, 信学技報, 教育工学, pp.95-102 (1999).
- 19) Rochkind, M.J.: The Source Code Control System, *IEEE Trans. Softw. Eng.*, Vol.SE-1, No.4, pp.364-370 (1975).
- 20) Tichy, W.F.: RCS - A System for Version Control, *Software - Practice and Experience*, Vol.15, No.7, pp.637-654 (1985).
- 21) Berliner, B.: CVS II: Parallelizing Software Development, *Proc. USENIX Conf. Winter*, pp.341-352 (1990).
- 22) Vitali, F. and Durand, D.: Using versioning to support collaboration on the WWW, *Proc. International World Wide Web Conference* (1995). available at <http://www.cs.unibo.it/~fabio/bio/papers/1995/WWW95/version.html>
- 23) Neuwirth, C.M., Chandhok, R., Kaufer, D.S., Erion, P., Morris, J. and Miller, D.: Flexible Diff-ing in a Collaborative Writing System, *Proc. CSCW '92*, pp.147-154 (1992).
- 24) Ousterhout, J.: *Tcl and Tk Toolkit*, Addison-Wesley (1994).
- 25) Webster, R., 中川正樹: 英語と日本語を対象にした文法誤り検出・訂正の共通点と相違, 情報処理, Vol.37, No.9, pp.865-871 (1996).

(平成 11 年 4 月 7 日受付)

(平成 11 年 10 月 7 日採録)



緒方 広明 (正会員)

1992 年徳島大学工学部卒業。1994 年同大学院博士前期課程修了。1995 年同博士後期課程退学。同年, 同大工学部助手。現在, 同講師。博士 (工学)。CSCW, CSCL に興味を持つ。教育システム情報学会論文賞受賞。WebNet 99 Top Paper Award 受賞。電子情報通信学会, 教育システム情報学会, ACM, IEEE, AAI, AIED, AACE 各会員。



葉田 善章 (学生会員)

1997 年徳島大学工学部知能情報工学科卒業。現在, 同研究科博士前期課程在学中。ヒューマンコンピュータインタラクションに興味を持つ。



矢野 米雄 (正会員)

1969 年大阪大学工学部卒業。1974 年同大学院博士課程修了。工学博士。同年徳島大学工学部助手。1990 年同教授。1979~1980 年米国イリノイ大学 Computer-based Education Research Laboratory 客員研究員。知的教育システム, 柔軟なデータベースの研究に従事。教育システム情報学会副会長・編集委員長。日本教育工学協会理事。電子情報通信学会和文誌編集委員。日本教育工学会, IEEE, AACE 各会員。