

## テキストファイルの更新状態の可視化と編集ツール

村野 紀代      粕川 正充

お茶の水女子大学 理学部 情報科学科

murano@is.ocha.ac.jp, kasukawa@is.ocha.ac.jp

diff コマンドを使用して、テキストファイルの version 更新によって生じた変更箇所を示し、旧 version の参照が可能で、かつ複数 version の結合を容易にするツール Windiffs を作成した。Windiffs は、1つの window の中に、テキストの遷移を知ることのできる疑似3次元的な情報画面を提供している。

本研究では、典型的なテキストファイルの結合方法について考察を述べる。評価実験から複数ファイルの結合について、通常のエディタを使用しての作業と比較して有効であることが明らかになった。

## 'Windiffs', a window oriented diff program

Kiyo Murano      Masaatsu Kasukawa

Ochanomizu University  
2-1-1 Otsuka Bunkyo-ku, Tokyo 112, JAPAN.

We developed a program named 'Windiffs', a new window oriented diff program. It could show a visual history of differences between some versions of texts and merge those versions at ease. This program only uses one window to show those texts' transitions and displays pseudo-3D view on it.

In this report, we studied about a typical work on the text merging. Also we reported an experimental result; some text merge task compare with the normal text editors. It concluded our program tasks were about twice faster than text editor tasks.

## 1 はじめに

現代はパソコンが普及してきており、パソコン通信やインターネットなどを利用した電子メールによる文書のやりとりが企業や大学以外でも盛んに行なわれている。筆者らの一人も同窓会名簿の作成などを電子メールの共同作業で行なっているが、以下のようなことができないかと日頃から不満を持っている。

- 自分が送付したファイルに対して、どのような更新があったのかを知りたい。
- 作業の進行した自分のファイルとの結合を簡単に行ないたい。

一般にテキストファイルを編集する際には、入力を完了して推敲を始める前のファイルを version1 と考え、修正を加えた後でもこの時点のファイルと比較可能なように、また加えた編集作業を取り消す場合には、この時点に戻れるように version1 ファイルを別扱いして保存する傾向がある。

テキストファイルがドキュメントファイルならば、他者からのコメントや自らの読み直しの結果、推敲が加えられてファイルは更新されていく。また、ソースプログラムファイルでは、機能追加やバグの除去のためにファイルを頻繁に更新するため、RCS [1] や SCCS を用いて管理されることが多い。こうしたツールを用いた場合に自動的に挿入される履歴や更新情報などは変更の概要に留まることが多く、作業者はもっと詳しい情報を得たいが、詳細な情報を記述することは、負担でもある。

筆者たちは、複数のファイルを簡単に比較することのできる計算機上のツールが非力なために、日頃以下のような不便を感じている。

- バグが生じた場合には、いつの version から潜んでいたものなのかをチェックする必要がある。現在テキストファイルの中にある文字列がいつの version からあったのかを即座に知ることができない。
- ある関数を追跡したいような場合に、特定の箇所がどのような変遷をしているのかを即座に得ることができない。
- version の派生が生じた場合の結合や、ある関数のみ数 version 前に戻したいといった場合など、単純に結合するのではなく、内容を検討しながらの複雑な作業となる。
- いつの version の変更が最も大きいのか、テキストのどの箇所に変更が多く起こって

いるのか、といった全体を通しての雰囲気を読むことができない。

このような問題点を解決するために、テキストファイルのバージョンの更新状態を色分けして示し、ファイル間の結合を簡単な操作で行なうことのできるツール Windiffs を作成した。

## 2 diff の視覚化における考察

人間が複数のテキストの間で比較を行なう場合には、おおむね以下のような場合が考えられる。

1. 変更の有無を知りたい。
2. 変更の量や位置を知りたい。
3. 変更箇所と変更後の内容を知りたい。または変更前の内容を知りたい。
4. 変更前と変更後の内容を比較したい。

発生頻度は、1から4の順番に少なくなると考えられる。

1のケースでは、diff コマンド [2] の出力のみで十分である。

2のケースでは、diff コマンドの出力のみでは情報が不十分であって、テキスト全体に対する比重や位置を視覚的に把握したい。

3のケースでは、挿入・削除・変更が生じたことを察知しながら、テキストを自然に読み進むことが可能でなければならない。diff コマンドの出力のみでは、前後の文脈が得られにくい。

4のケースでは、変更の前後をできるだけ近い位置に並べて比較できるようにする必要がある。例えば、2行を比較する際に上下に並べられている場合が最も比較しやすい。離れている場合には、頻繁な視線の動きが必要であり、疲労も大きい。

Display 上で2つのテキストの比較する場合には、通常2つの window を上下又は左右に並べるといった方法をとる。しかし、小さな Display 上で2つの window を表示すると、空間が分割され、与えられたスペースが非常に小さくなり、窮屈に感じる。そして、位置合わせのために、常に2つの window をスクロールしなければならない。また、4のようなケースでは、window 上で作業している限り変更の箇所を見比べるのに目を動かさなければならない。1文字の空白の挿入や打鍵誤りがバグの原因となることが、よく生じるが、このような違いを見つけるためには、別の window では、発見しづらい。

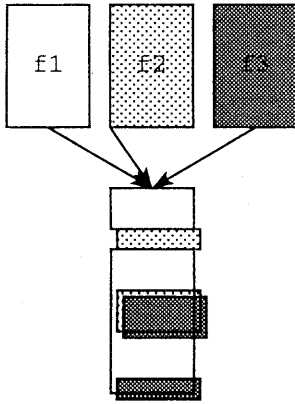


図 1: Windiffs の重ね合わせ方法

紙に印刷したものとを比べるには、重ね合わせて透かしてみることが出来る。同様の操作を計算機上で実現するために、筆者らは、更新状態を一つの window にだけ反映させ、古い版を必要に応じて色分けして示すことを考えた。

### 3 Windiffs の概要

#### 3.1 実現の方法

Windiffs は、Tcl/Tk で作成したツールである [3]。Tcl/Tk のバージョンは、Tcl7.3 Tk3.6 で作成した。また、差分情報の抽出は UNIX の diff を起動して行なっている。Windiffs は、2 節で述べた diff の視覚化についてのすべての要求を叶えるために、テキストを重ね合わせるという手法をとった。

Windiffs は、Tcl/Tk の GUI 部品である canvas widget の中で人間がよく行なう紙での切り貼り作業を行なうものである (図 1)。テキストを行単位に分解し、異なる version を示すためにテキストの背景色を変えて表示する。diff コマンドを使用して、挿入が生じると、以降の行を切って下げて新しい部分を挿入する。変更が生じると、変更箇所を紙を張り付けるように、テキストを重ね合わせる。削除が生じると、黒く塗りつぶすように全面黒のテキストを重ね合わせる。そして、canvas widget 上で最上位に表示するものを制御することにより、挿入したり重ね合わせた紙をめくったり、塗りつぶした黒をはがすように、変更前の内容を見ることが出来る。また変更前に戻すことを可能にしている。また、重なった紙だけを引き出すことによ

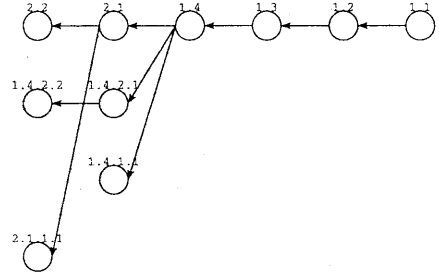


図 2: Windiffs の RCS のノード表示

り、必要な場合には特定場所の遍歴を並べて見ることも可能である。そして、重ね合わさった紙の組合せをユーザが指定して、ファイル間の結合を行なうことができる。

このようにして、単一の window の中に、version の更新情報をつめこんだ疑似 3 次元的な情報画面を実現している。

#### 3.2 機能

##### 1. RCS の revision 番号のノード表示

Windiffs は通常のファイルと同様に、RCS ファイルからの入力を指定できる。RCS ファイルを指定する場合は、ポップアップ window が開き、revision の状態を以下のようなノードで表示する (図 2)。マウスでこのノードを指定したり、ノード群を囲んだりすることで RCS からの入力を容易に指定できる。

##### 2. 表示 Window

Windiffs は 4 つの window を持つ (図 3)。

- メイン window  
メニューボタンと version 更新の情報を持つ window で、各 version ごとに固有の背景色を持つテキストの重ね合わせの状態を表示する。これにより、テキストの内容を読みながら、変更や挿入を知ることができる。また、重ね合わせた紙を横から見るように、変更回数の厚みを左端に表示する。最新の变更箇所は赤枠で囲まれる。
- エディタ window  
最新の追加箇所は赤文字で表示される。最新の变更箇所は、赤文字で且つ背景色が淡い黄色で表示される。指定バージョンの中で、変更が発生した箇所はすべて背景色が淡い黄色で表示

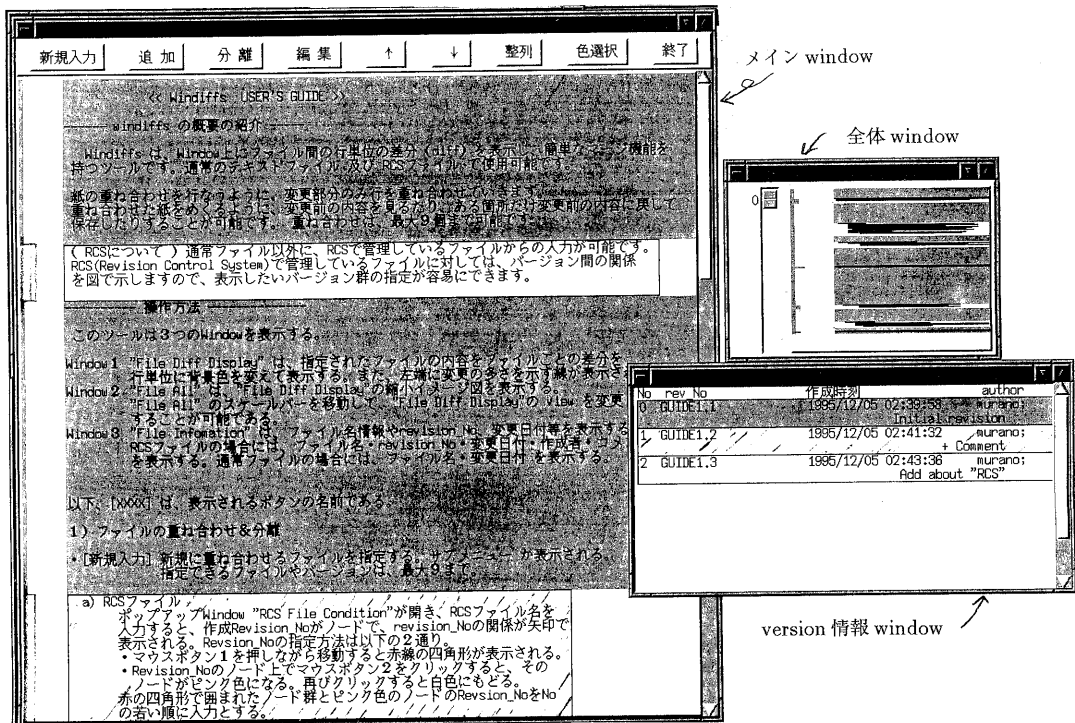


図 3: Windiffs の表示 window

される。なお現在この画面は、Tk の text widget の編集機能プラス矢印キーによるカーソル移動のみのため、編集能力はかなり低い。Tk version 4.0 に移行することによって、emacs と同様のキー操作による編集が可能になる。

• 全体 window

メイン window の全体の縮小図を表示する。最新の変更箇所を赤ラインで表示し、変更の量や位置を即座に把握できる。メイン window のスクロールと全体 window のスケールバーは連動しており、常に全体の中の位置を把握することができる。

• version 情報 window

各 version またはファイルの更新時刻、作成者などの情報を固有の背景色とともに表示する。

3. 追加と分離

[追加] ボタンにより、現在表示されているメイン window にさらに重ね合わせが可能であり、重ねられるファイルの数は最大 10 までである。[分離] ボタンにより、重ねた最も上の紙をはがすように、最後に重ねた

ファイルを分離する。[追加][分離] により、最新の変更箇所は変化する。

4. 最新変更箇所の検索

メイン window の中で、最新の変更箇所は赤枠で囲まれて表示される。[↑][↓] ボタンにより、上方向又は下方向の最新変更箇所へジャンプする。

5. 旧 version の参照

メイン window の中で、挿入・変更・削除が生じている箇所、マウス左ボタンをクリックすると、1version 前のテキストの状態を覗くことができる。複数 version で変更が重なっている場合には、クリックするごとに、1version 前の状態になる。マウス中央ボタンをクリックすると元に戻る。[整理] ボタンで、全変更箇所をファイルの重ね順の通りに戻す。

6. 特定箇所の変遷の表示

ある関数の数 version にわたる経過を表示したい場合、また変更前と変更後をじっくりと比較したい場合が生じる。このような場合には、マウス右ボタンのドラッグ操作により、青斜線の四角形が表示されるので、欲する領域を指定する。マウス左ボタンクリックでポップアップ window が現れ、重

なっていたものが引き出されるように、古い version から順にその領域のテキストが表示される。

#### 7. 複数ファイルの結合 以下の手順で行なう。

- 旧 version の参照機能を用いて、最適なテキストを表示する。
- 変更が overlap しているが、どちらの内容も必要な場合には、マウス右ボタンをドラッグして、青斜線の四角形で囲んでおく。
- [編集] ボタンにより、画面に表示されているものの組合せでエディタ Window に読み込むことが可能である。

このように複数 version を組み合わせたファイルの結合を「WYSIWYG」で行なうことができる。

#### 8. 背景色のカスタマイズ

マシンごとに使える色の数は様々である。[色選択] ボタンにより、テキストの背景色は、各ユーザがカスタマイズ可能である。ある version 番号に特徴的な色合いを持たせたり、共同作業者がいる場合には、作成者ごとに色系統を分けることができる。

## 4 ファイルの結合機能について評価

### 4.1 通常のテキストエディタとの比較 実験

果たして本当に Windiffs が優れているかを検証するため、以下のような実験を行なった。

- 被験者  
お茶の水女子大学理学部情報科学科の2年生20名
- 実験内容  
ファイル1と、ファイル1に3箇所挿入・変更を行なったファイル2、同じく3箇所の挿入・変更を行なったファイル3の3つのファイルを指示書どおりに編集し、ファイル1にファイル2から2箇所の変更・挿入、ファイル3から2箇所の変更・挿入を反映させて、ファイル4を作成してもらう。被験者群を先に Windiffs を使用するグループと、先に通常のエディタを使用するグループに分け、この課題を同一の被験者によって実施し、おのおのの課題が完了するまでの時間を UNIX の time コマンドを呼び出す Tcl/Tk のプログラムで計測した。

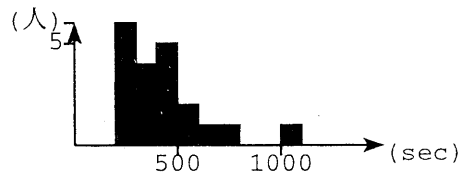


図 4: Windiffs を使用の作業時間分布図

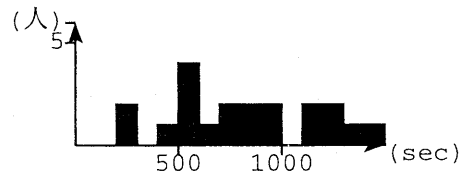


図 5: 通常の Editor を使用の作業時間分布図

- 実験環境  
通常のエディタを使用した計算機は、SUN4-IPC, SunOS Release 4.1.3-JLE1.1.3, 主記憶 24MB で、Windiffs を使用した計算機はネットワーク接続された IBM-PC 互換機で、AMD 社 DX4-120MHz, Linux1.2.8, 主記憶 12MB である。  
このような構成で実験を行なった理由は、Windiffs が SUN 純正の window システムである OpenWindows の下では動作が不安定になることがあるためである。
- 実施手順  
先に実験内容を説明した後、Windiffs の使用方法を約 10 分間説明し、10 分ほど使用させた後、作業時間を計測し、編集後のファイル4を収集した。なお、被験者のほぼ全員がエディタとして、Sun ワークステーションの OpenWindows 上の TextEdit を使用した。
- 実験結果  
実験の結果、通常エディタ使用の作業平均時間は 824sec, Windiffs 使用の作業平均時間は 439sec であった。エディタ使用では作業時間のバラツキが大きいのにに対し、Windiffs 使用では、80% 以上が 200sec ~ 600sec の間に作業を完了することができた(図4 図5)。t 検定の結果 95% の信頼度で、Windiffs 使用の方が優位であった。

## 4.2 実験についての議論

収集した編集後のファイルを見ると、正確さも Windiffs 使用の場合の方が高かった。実験前の説明で、3枚の色紙を用いて Windiffs での表示方法を説明したところ、被験者はみなよく理解してくれたようである。2年生ということもあり、通常のエディタもあまり習熟していなかったようであるが、おおむね好意的な感想が寄せられており、Windiffs が初心者にもなじみやすいことが実証できた。

Windiffs を使用することにより、ある特定範囲の変遷の情報を得ることができる。また、変更が overlap する箇所でも、内容を比較しての結合が簡単に行なえる。筆者らは version の更新ということを念頭においていたが、上記のことから、例えば、定型のフォーマットに回答の入力されたテキストファイルが複数枚存在する場合など、Windiffs ならば、10枚のファイルを重ねあわせて、回答の箇所のみをまとめて比較する、または結合するといった処理がやりやすい。また現段階では、メイン window の文字の変更は不可であるが、テキストの変更を許可すれば、複数のテキストの特定箇所の変更を単一の window 内で1回で行なうことが可能になる。

現在はネットワークの発達により、特にテキストファイルは電子メールなどを媒体として、様々なプラットフォーム上で、やりとりされている。Tcl/Tk の動作環境は、MS-Windows および Windows-NT、また Mac-OS 上にも移植されている。そこで Windiffs も Windows 上で動作可能なように対応した。操作性は、Unix 上と比べると格段に劣るが、RCS からの入力を除く上記の機能はすべて Windows 上でも動作可能となっている。

## 5 問題点

### • 処理速度

テキストの行数が長くなるほど、またテキスト間の差分が大きいくほど、表示に要する処理時間が長くなり、マウス操作の応答速度も落ちる。現状では、1000行のテキストがほぼ限界である。これは Windiffs のソースが全て Tcl/Tk で書かれているためである。Tcl は "Everything is string" というルールを持っており、そのため for 文などでは変数のカウントのたびごとに string から integer への変換が必要となる。この点を改善するだけでも処理速度は、5~10倍になると報告されている [4]。従って、C

言語に書き直すことにより処理速度の向上が見込まれる。現在は、C 言語へのコンパイラとして、ICEM 社から "Tcl C Compiler" が発売されている [5]。

### • Windiffs の限界

Windiffs では、UNIX コマンドの diff を使用して、行単位での変更を抽出している。しかし、行の移動や段落の組み替えなどには対応できないため、削除と新規挿入として扱っている。この問題にとり組むためには、構文の解析を行なう独自の差分抽出機構を組みこむ必要があるだろう。

## 6 結論

テキストファイルの変更点を視覚化し、ファイルの結合を容易にするツール Windiffs を作成した。Windiffs は、単一の window の中に、version の更新情報を格納した疑似 3 次元的な情報画面を実現している。

評価実験から複数ファイルの結合について、通常のエディタを使用している作業と比較して有効であることが明らかになり、テキストを重ね合わせるという手法が人間と親和性があることがわかった。

## 参考文献

- [1] Walter.F.Tichy. *RCS A system for Version Control SOFTWARE.PRACTICE AND EXPERIENCE, Vol.15(7),637-654(JULY 1985)*
- [2] UNIX online manual diff
- [3] John K.Ousterhout. *Tcl and the Tk Toolkit. ADDISON-WESLEY PUBLISHING COMPANY(1994)*
- [4] Adam Sah,Jon Blow,and Brian Dennis. *An Introduction to the Rush Language (May 29,1994)*
- [5] Forest R.Rouse,Wayne Christopher. *A Tcl to C Compiler(<ftp://ftp.dnai.com/i/icemcfd/tclc>).*
- [6] 岡田謙一, 松下温: *本メディアを越えて: Book Window, 情報処理学会論文誌 Mar.1994, Vol.35 No.3*
- [7] 岡野 裕之: *時系列情報を用いるテキストエディタ「雲水」とそのエディタカーネル, ヒューマンインタフェイス 37-4(1994.7.4)*