

## キーワード・サーチ機能を備えた視覚的検索インタフェースの提案

野崎陽平† 平川正人‡ 吉高淳夫‡ 市川忠男‡

†広島大学大学院工学研究科

‡広島大学工学部

〒739-8527 広島県東広島市鏡山 1-4-1

{nozaki, hirakawa, yoshi, ichikawa}@isl.hiroshima-u.ac.jp

複数の検索キーワードを入力してそれらの検索キーワードを持つ情報を探し出す一般の検索システムでは、ユーザの望まない情報も含む大量の検索結果が得られることが多々あり、それに対して適切な絞り込みキーワードを入力することは難しい。このことはユーザが好ましい情報を獲得することを困難にしている。

本研究では、検索結果の各検索キーワードに対する関連性を表わす機能と適切な絞り込みキーワードの表示機能を備えた視覚的インタフェースを提案する。本インタフェースは視覚的に検索結果を絞り込める操作を提供している。これにより絞り込みキーワードの表示にユーザの好みも反映されて、検索結果の絞り込みをより効率よく行える。

## A Visual Interface for Information Retrieval with Keyword Search Facility

Youhei Nozaki† Masahito Hirakawa‡ Atsuo Yoshitaka‡ Tadao Ichikawa‡

† Graduate School of Hiroshima University

‡ Faculty of Engineering,

The University of Hiroshima

1-4-1, Kagamiyama, Higashihiroshima, Hiroshima, 739-8527

Generally, keywords are used in retrieving items from a huge information source. However, the user most likely gets too many items including those the user doesn't want, and it is difficult to find appropriate next keywords to refine the search. This makes it difficult to obtain meaningful results.

We propose a visual interface with the functions capable of representing the relation between the results and each keyword and displaying appropriate next keywords. The interface supports extraction of desired items visually and enables next keywords be reflected by user's intention. This makes the retrieval efficient.

## 1 はじめに

WWW の普及に代表されるように、現在、大量で多種多様な情報が氾濫しているが、そのような状況下においての情報獲得の手法としては、複数の単語（検索キーワード）を入力して、それらの単語をキーワードとして持つ情報を検索結果として探し出すというのが一般的である。

この手法は単語を入力するだけで良いという手軽さがある。しかしながら大量の検索結果が得られた場合、その中から目的とするものを抽出するための適切な追加検索キーワードをどの様に選んでいけば良いか明示的ではなく、得られた検索結果の表示を閲覧してもどれが所望のものかを判別することが難しいという欠点を持つ。これにより、絞り込み検索時に入力する検索キーワードを模索したり、得られた結果の内容を実際に確認しながら欲しい情報を探す、といったことが多々生じ、所望の情報を簡単に得ることができるとは言い難い。

本稿ではキーワードが与えられているデータ集合からの検索の際に、これらの問題を解決するためのインタラクティブな検索システムを提案し、それに関して評価、考察をする。

## 2 本研究の手法

ドキュメントのようにキーワードを所有しているデータ集合から検索を行う際に、ユーザが手際良く検索を行うためには、

- 各検索結果の内容の違いを認識しやすいこと。
- より良い検索キーワードを発見しやすいこと。
- 連続的かつ動的なインタラクションをサポートしていること。

などの要素が重要であると考えられる。

上の条件の中で特定の条件を満たすものについては、既に様々な手法が提案されている。

例えば、

- 各検索キーワードに対しての関連性に基づいて検索結果を二次元空間に分類し、各検索結果の内容の違いをユーザに認識させる手法 ([1], [2])。
- 各情報の内容とそれらに含まれるキーワードの関連の度合いを一空間内でグラフにより表現し、直接操作、動的更新を適用した手法 ([3])。
- 検索結果からのキーワード抽出とそれらの概念の構造化を行い、これを表示してユーザの検索キーワード選択を支援する手法 ([4])。

などである。

しかしながら上の条件をすべて満たした際のインタラクションについて考えられているものは少ない。DocSpace [3]ではキーワードと情報間の関係を視覚化して動的なインタラクションを提供しているが、ブラウジングの要素が強く、我々が本稿で議論の対象としている検索とは言えない。

これに対して我々は、VIBE [1]のように検索結果を二次元空間に視覚的な分類を行う手法と、DualNavi [4]のように検索結果からキーワードを抽出して、それらの概念の構造化を行いユーザの絞り込みキーワードの選択支援を行う手法を一つに統合したシステムを提案する。併せて、検索結果の視覚的な分類については、VIBE [1]やGUIDO [2]などの手法では触れられなかった検索結果の配置位置の重なりに対する対処や、一部分に検索結果が密集し認識し難い場合の対策としてバンドル作成という手法を提案する。

我々の提案するシステムにおいては、二次元空間に視覚的に分類された検索結果に対して、その空間の一部分を指定して視覚的に絞り込みを行える機能を実装している。この機能を絞り込みキーワードの追加と組み合わせると効率よく検索が行えるだけではなく、視覚的な絞り込み範囲の指定が絞り込みキーワードの抽出と表示に影響を与える。

## 3 インタフェースの概要

本章では構築したシステムのインタフェースの概要を述べる。システムのインタフェースを図1

に示す。

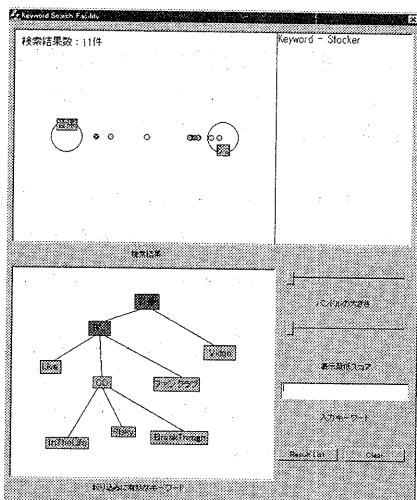


図 1：システムのユーザインタフェース

図 1 の上部のウィンドウは検索結果を視覚的に表示する部分であり、下部が現在の検索結果から抽出されたキーワードの概念的関係を表現している部分である。

検索結果を表示するウィンドウでは、検索キーワードは白い円、検索結果が色のついた円として表示される。また各検索結果がどの程度信頼できる情報かという指標が色で表現されている。

検索結果から抽出されたキーワードの概念的関係が表現されているウィンドウでは、それらの概念的関係が分析されて木構造に視覚化されている。

上で述べた二つのウィンドウの片方に指示を与えるとそれに応じて動的にもう片方も更新される。

本システムで検索を行い十分に検索結果が絞りこまれた後は、検索結果のタイトル、評価値、そして要約などが表示されたリストをみて最終的にユーザの望む情報を得ることができる。

#### 4 検索結果の表示部分

本システムでは入力された検索キーワードをすべて持っている情報を検索結果とする。

検索結果の内容が特定検索キーワードの意味する話題について触れられている度合いをそのキーワードに対する関連度と呼ぶことにする。

検索結果の各検索キーワードに対する関連度の総和をその検索結果に対する評価値とし、これを各検索結果がどれほど信頼できるかを表す指標とする。

#### 4.1 検索結果の表示

基本的に VIBE [1] で提案されたものと同様に、検索キーワードの配置位置に検索キーワードとの関連度に応じた重みを置き、その重み付き重心を求めることにより各検索結果の配置位置を決定する。

我々のシステムでは、特定領域を指定することによって、視覚的に二次元空間に配置された検索結果を絞り込むことができる。ここで、視覚的な絞り込み機能を実装するにあたっては、次の条件を満たす必要がある。

- ① 検索結果の配置位置は曖昧性を持ってはならない。
- ② どれ程の数の検索結果がどの位置にあるかを認識しやすくなければならない。

① に対しての理由であるが、検索結果の配置位置に曖昧性があると、視覚的絞り込みにより検索結果を絞り込む際にユーザの望む情報が欠落したり、逆にユーザの望まない情報が得られたりするためである。

ここで二次元平面を仮定する限りでは、最大三つまでの検索キーワードへの正確な関連度の比率を表わすことができず、検索キーワードが四つ以上になるとその配置に曖昧さが生じる。この問題に対して、VIBE [1] では、ユーザの配置された検索キーワードの直接操作に従って検索結果の配置位置を更新するという方法を取っている。しかしながら、本システムでは、上の理由によって一度に関連度を観察できる最大検索キーワード数を三つとする。

次に② に対しての理由であるが、検索結果の分布傾向をしっかりと把握できれば、視覚的絞り込みを行う際に、より適切な検索結果を適切な量絞り込むことができる。そこで、ユーザに検索結果の配置位置を把握しやすくする手法として検索結果の重なりに対する対処、そしてバンドル作成という手法を提案する。

**検索結果の配置位置：**

本システムでの検索結果との関連度を観察する検索キーワードが一つ、二つ、三つの場合の検索結果配置位置をそれぞれ図 2 から図 4 に示す。なお、各検索結果は円で表わされ、その評価値にしたがって色付けされる。

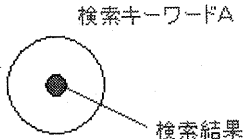
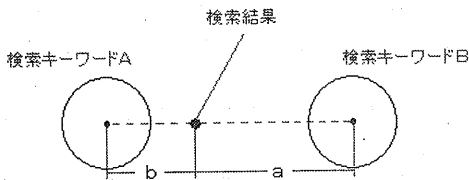
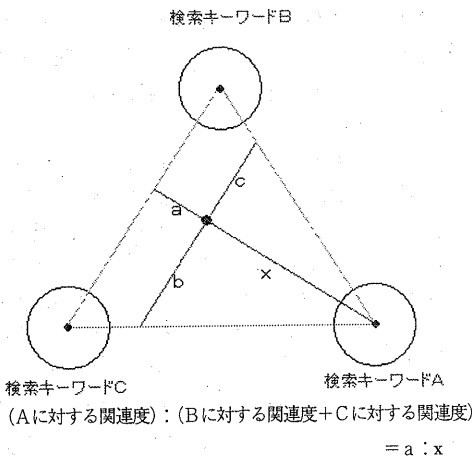


図 2：キーワードが一つの場合の検索結果の配置位置



※ (Aに対する関連度) : (Bに対する関連度) = a : b

図 3：キーワードが二つの場合の検索結果の配置位置



※ (Aに対する関連度) : (Bに対する関連度 + Cに対する関連度) = a : x

※ (Bに対する関連度) : (Cに対する関連度) = b : c

図 4：キーワードが三つの場合の検索結果の配置位置

**検索結果の重なりに対する対処：**

検索結果の配置位置が重なった場合の対処法として我々は次のような手法を提案する。

- ①各位置において、検索結果を評価値に応じて幾つかのグループに振り分ける。
- ②各位置において、検索結果の数の多いグループから順に(数が同じ場合には評価値の低いグループから順に)、各グループに属する検索結果の数を円の大きさと表わし描く。

この手法により検索結果が重なった場合でも、ユーザはどの程度の検索結果があり、その評価値が如何ほどか認識することができる。

図 5 には本処理を行わない場合の検索結果の表示結果が、そして図 6 には、評価値に従って三つのグループに振り分け、本処理を行った過程と処理結果が示されている。

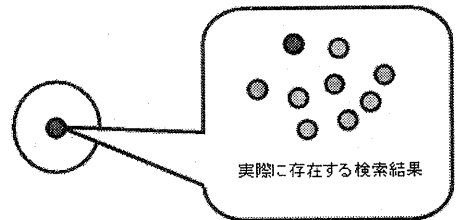


図 5：重なりに対する処理を行わなかった場合

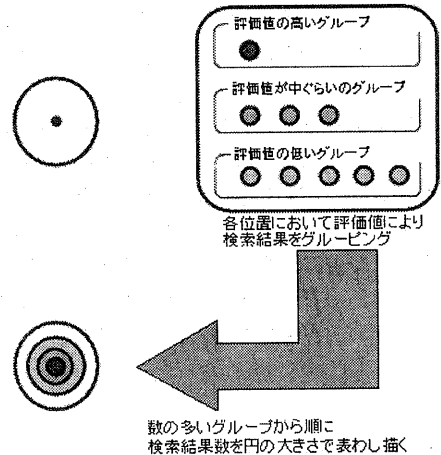


図 6：重なりに対する処理過程と処理結果

図 6 に見える通り、本手法を用いることで検索結果の位置が重なっているにも関わらず、どの程度の評価値の検索結果がどの程度の量あるかが認識できることが分かる。

### バンドル作成:

検索結果の配置の傾向をより認識しやすいものとする手法として、隣接座標にある検索結果をひとまとめにして、それを連続的に変化させるという手法を提案する。

具体的には検索結果を表わす二次元空間のx軸、y軸を共に一定間隔で区切り、個々の区切られた範囲内にある検索結果をその中心座標にあるとみなすという機能である。分割領域内の中心座標にあるとみなされた検索結果に対して、検索結果が重なった場合の処理が行われる。図7に処理の過程を図示する。

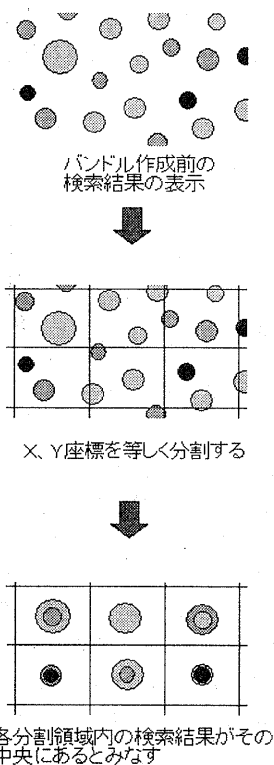


図7:バンドル作成の処理過程

この機能においてユーザは作成する分割領域のサイズを大きくすれば検索結果の分布傾向がだまかに認識でき、また分割するサイズを小さくしていけばより詳細な分布傾向が認識できる(図8)。

具体的なシステムの実装レベルではスライダにより作成する分割領域のサイズを変更することができ、その動きに対応して検索結果の視覚分類は

動的に更新される。よって、スライダを動かすことによって検索結果を表わす円が結合して大きくなったり分裂して小さくなったりする。このような動的な検索結果表示の更新をユーザが観察することにより、検索結果の分布傾向がより認識しやすくなる。

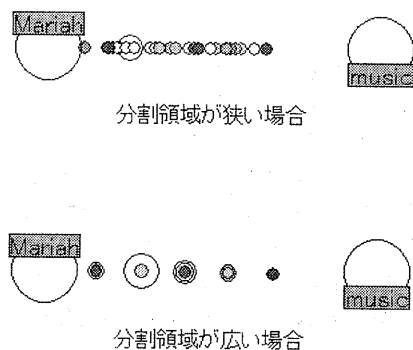


図8:バンドル領域の違いによる表示の違い

### 4.2 検索結果表示部分とのインタラクション

本節では検索結果表示部分とのインタラクション機能である、視覚的絞り込み、検索キーワードのグルーピング機能、そして、キーワード・ストッカーについて述べる。

#### 視覚的絞り込み:

視覚的絞り込みを行う際の具体的な操作は、ラバーバンドと呼ばれる円を描くことにより、その範囲内にある検索結果のみを選択することができる(図9)。

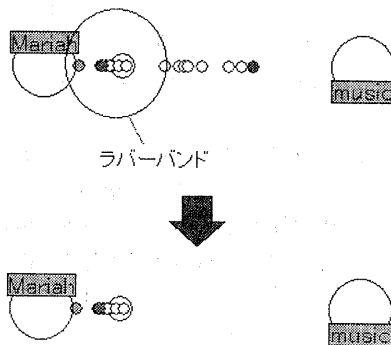


図9:ラバーバンドによる絞り込み

### 検索キーワードのグルーピング機能:

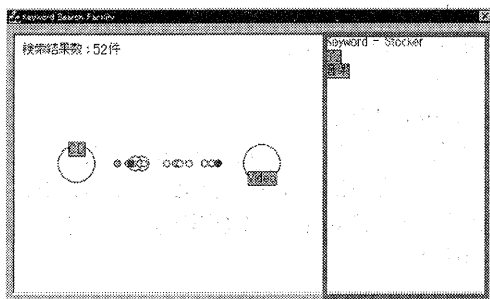
ユーザは任意の検索キーワードをグループ化し、一つのキーワードと同等に扱うことができる。これによりユーザの意図に沿ったキーワードの合成が可能となる。

合成されたキーワードに対する関連度は、そのグループに属するキーワードに対する関連度の総和とする。

### キーワード・ストッカー:

ユーザはすべての検索キーワードに対して関連度の比率を見たいとは限らない。そこで特に関連度の比率を見なくてもよいキーワードについては、それらをキーワード・ストッカーと呼ばれる領域に置くことができる。キーワード・ストッカーにある検索キーワードは検索結果の評価値には影響するが、検索結果の配置自体には関与しない。

例えば、図 10 では、“Bz”と“音楽”という検索キーワードをキーワード・ストッカーにおくことで、“Video”と“CD”という検索キーワードに対してのみ検索結果の関連度の比率を見ることが可能になっている。



キーワード・ストッカー

図 10: キーワード・ストッカー

ユーザはこれらのキーワードのグルーピング機能、キーワード・ストッカーを用いてユーザの意図に沿った検索結果と検索キーワードの関連度を観察することができ、視覚的絞り込みの際、適切な絞り込みを行うことができるようになる。

### 5 絞り込みキーワードの表示

ユーザは、大量の検索結果に対して絞り込みキーワードが思い付かない際は、ここで述べる表示の上で直接キーワードを指定し、それを絞り込みキーワードとすることができる。ユーザの絞り込み検索時のキーワード選択支援を行うこの表示部分は、検索キーワードの追加、ラバーバンドによる検索結果の一部選択など、検索結果が変化する度に、更新される。

検索結果から抽出されたキーワードの概念的相互関係を表わす手法として、本システムでは DualNavi [4]と同様の手法を用いている。この手法においては、木構造でキーワード間の関係が表わされ、画面上部に位置するものほど全検索結果内で使用頻度の多いキーワードとなっている (図 11)。

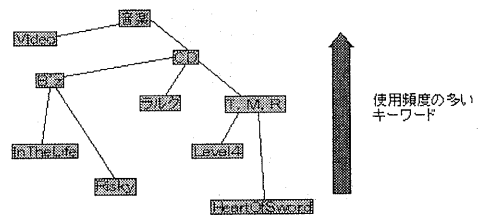
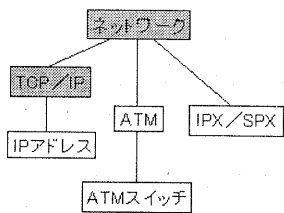


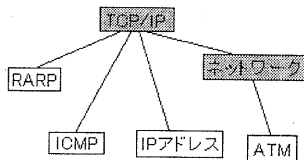
図 11: 絞り込みのためのキーワード表示例

絞り込みキーワードの表示にあたって、本システムではユーザの好みも反映されたものになる。その理由は、例をあげると、“ネットワーク”と“TCP/IP”というキーワードで検索を行う際、“ネットワーク”寄りの検索結果を円で囲めば、その付近は“ネットワーク全般について触れられていて TCP/IP にも触れられている情報”と考えられる。よって、それらから絞り込みのキーワードの抽出が行われるのでネットワーク全般に関するキーワードが画面に出現する (図 12. (a))。

逆に“TCP/IP”寄りを囲めば、“TCP/IP についての詳細な情報”からキーワードの抽出が行われるため、TCP/IP についてより詳細な機能などを表わすキーワードが出現する (図 12. (b))。



(a) ネットワーク寄りを選んだ場合



(b) TCP/IP寄りを選んだ場合

図 12：絞り込み範囲の違いによる表示の違い

この様に、ユーザのラバーバンドによる検索結果の抽出が絞り込みキーワードの表示に影響し、ユーザにとってより認識しやすい表示となり、適切な絞り込みキーワードを発見し易くなる。

## 6 評価実験

本章では、本システムの評価実験について説明する。

実験は、次の項目について検証することを主な目的としている。

- ① 重なりに対する対処やバンドル作成をすることによって認識性は向上しているか。
- ② [4]と比べてユーザの認識のしやすい絞り込みキーワードの表示となったか。
- ③ ユーザが効率よく検索を行えるか。

### 6.1 実験手法

評価にあたっては、検索対象となる情報として、音楽全般についての触れられている Web サイト 200 件程度の URL を用意した。

本実験の被験者数は 7 人であり、事前に全被験者に本システムの操作方法を説明した。

実験方法は次の通りである。

実験 1：①を検証するために、(1)バンドル作成、重なりに対する対処という機能を実装したシステム、(2)それらを実装していないシステム、をそれぞれ使用してもらい、それぞれの認識性を 5 段階で評価してもらった。

実験 2：②を検証するために、ラバーバンドによって絞り込みを行う際、(1)絞り込み後の絞り込みキーワードの表示、(2)絞り込み前の絞り込みキーワードの表示、のそれぞれの認識性を 5 段階で評価してもらった。

実験 3：③に関しては各被験者に、(1)本システム、(2)本システムから絞り込みキーワード表示部分を除いたもの、(3)本システムから検索結果の視覚的分類部分を除いたもの、(4)本システムから検索結果の視覚的分類部分と絞り込みキーワードの表示部分を除いたもの、の四つを使用してもらい、それぞれどの程度手際良く検索が行えたか被験者に 5 段階で評価してもらった。

### 6.2 実験結果と考察

実験 1 の結果を図 13 に示す。

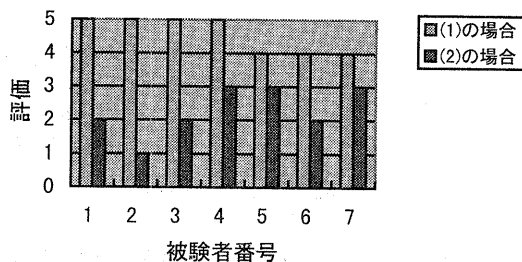


図 13：実験 1 の結果

図 13 が示す通り、全ての被験者はバンドル作成、重なりに対する対処により、認識性が向上しているとみなしている。

次に実験2の結果を図14に示す。

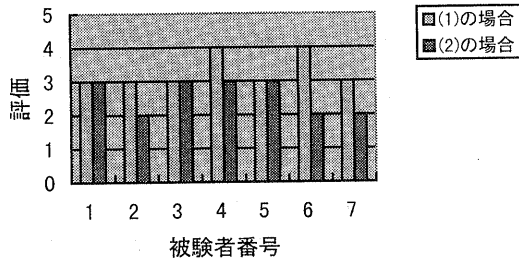


図14：実験2の結果

実験2では全体的に余り良い結果が得られているとは言えない。理由として、検索対象として用意したデータが200件程度と少なかったために絞り込みキーワードの抽出と概念の構造化が効果的に行えなかったことが考えられる。しかしながら、過半数の被験者はラバーバンドによる絞込み後に、絞り込みキーワードの認識性が向上したと答えている。

実験3の結果を図15に示す。

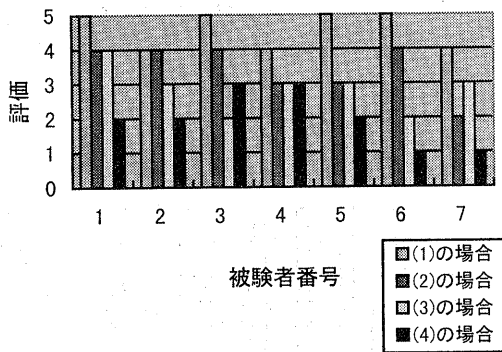


図15：実験3の結果

実験3では、実験2の結果が影響しているのが原因であろうが、絞り込みキーワードの表示のみを実装した場合の評価が全体的に低い。しかしながら、本システムのように絞り込みキーワードの表示と検索結果の視覚分類、そして視覚的絞り込みを組み合わせさせた効果はこの実験結果に表れている。実際の実験においても各被験者は、本稿で提案した様々なインタラクションを連続的に行い手際良く検索を行っている。

## 7 まとめ

本稿では、検索結果の視覚的分類と絞り込みキーワードの表示によるユーザ支援を統合したシステムを提案した。

実際のシステムでは、最初のキーワードを入力した後はマウスでの連続的操作により検索を行うことができ、ユーザは手際良く所望のデータを獲得することができる。

## 参考文献

- [1] Kai A. Olsen, and Robert R. Korfhage. Desktop Visualization. Proc. of 1994 IEEE Symposium on Visual Languages, pp.239-244 (1994).
- [2] Assadaporn Nuchprayoon & Robert R. Korfhage. Guido, a Visual Tool for Retrieving Documents, IEEE Symposium on Visual Languages, St. Louis, Missouri, pp.64-71 (1994).
- [3] 館村 純一. DocSpace:文献空間のインタラクティブ視覚化. インタラクティブシステムとソフトウェアIV, 日本ソフトウェア科学会 WISS '96, pp.11-20 (1996).
- [4] 西岡 真吾, 丹波 芳樹, 岩崎 真, 高野 明彦. 文献検索支援インタフェース DualNAVI, 日本ソフトウェア科学会 WISS '97, pp.43-48 (1997).