



2008 年度上期未踏 IT 人材発掘・育成事業(未踏ユース)採択案件評価書

1. 担当PM

安村 通晃 PM(慶應義塾大学 環境情報学部 教授)

2. 採択者氏名

チーフクリエイター: 山本 岳洋(京都大学大学院情報学研究科 社会情報学専攻)
コクリエイター: なし

3. プロジェクト管理組織

株式会社 オープンテクノロジーズ

4. 委託金支払額

2,966,742 円

5. テーマ名

柔軟な検索結果再ランキングインタフェースの開発

6. 関連Webサイト

なし

7. テーマ概要

近年、ウェブ検索をはじめとして、論文、商品、ニュースといった様々なコンテンツを検索サービスから入手することが一般的となっている。しかし、現状の検索エンジンはユーザの検索意図に合った検索結果を返すことは困難である。例えば、異なるユーザが同じクエリを入力した場合、ユーザの求める検索結果はユーザの興味や嗜好

によって当然異なる。また、同じユーザが同じクエリを入力したとしても、その時の検索タスクによって求める検索結果は多岐にわたる。特に、検索エンジンから受け取った検索結果を閲覧しているときに、「もっとこういった結果が欲しい」とか、「こういう結果はいらない」といった検索意図が表出することが多い。このように、ユーザが検索時に入力するクエリのみを用いて、検索エンジン側がユーザの検索意図を把握することは容易ではない。

検索エンジンがユーザの検索意図を把握することが困難な原因のひとつとして、検索結果閲覧ページとユーザとのインタラクションの少なさが考えられる。現状ではユーザはクエリを入力した検索エンジンに検索意図を伝達できないが、複雑なクエリを作成することはユーザにとって敷居が高い。

そこで、本プロジェクトではユーザの検索結果閲覧中に、シンプルなインタラクションを行うだけで、検索結果をユーザの意図に合わせて再ランキングする仕組みを実現する。提案するシステムは強調・削除という二つのインタラクションを検索結果の閲覧中に導入する。ユーザは検索結果中のキーワードに対して削除操作を行うことにより「検索結果のこの部分がいない」、強調操作を行うことにより「このような検索結果をもっと上位に表示して欲しい」といった意図をシステムに伝えることができる。システムはユーザの明示的な強調・削除操作に応じて検索結果の再ランキングを行う。この仕組みによりユーザは柔軟に検索結果を再ランキングすることが可能となる。

8. 採択理由

この研究開発は、Google などの検索結果の中で、これは要らない(「削除」)とか、この属性は大事なので優先させたい(「強調」)などを指定することによって、自分の期待する検索結果に向けてチューニングし、再ランキングしようとする、非常にシンプルなアイデアに基づいている。

オッカムのカミソリではないが、物事はシンプルなアイデアこそ、その威力を発揮しやすい。このアイデアの他に、異なる検索サービス、たとえば、Amazon と価格.com の二つの検索結果の統合も考えている。こういった方法を積み重ねることによって、より利用者の希望に近い検索結果となる可能性が高まる。基本的な考えは非常に良く理解できるが、たとえば、複数の項目を共に優先させたいときにどのようにしているかなど、これから詰めるべき課題も少なくない。また、こうして得られた新しい再ランキングインタフェースをどのように評価し、アピールしていくかも課題であろう。

山本岳洋君の活躍を大いに期待したい。

9. 開発目標

本システム実現のため、本プロジェクトでは以下の開発を行なう。

1. 削除、強調による検索結果の再ランキングインタフェース
必要な検索結果の上位、不必要な検索結果を下位に再ランキングするためのアルゴリズム、および、インタフェースを開発する。
2. 異なるクエリ・サービスの検索結果の統合
複数の検索結果を一つの検索結果に統合し、ユーザーに提示するための手法を開発する。
3. ユーザーによる検索結果パーサー作成インタフェースとパーサー共有システム
さまざまな検索サービスに対してシステムを適用するためのパーサー作成を、ユーザーが自ら行なうための仕組みを開発する。

10. 進捗概要

オーデイション時には、ごく基本的な部分である、強調と削除が動いていた。その後も比較的順調であり、プロジェクトレビューも現地レビュー1回のみで済んだ。異なるクエリ検索結果統合のインタフェースを提示方法や、パーサーの部分がやや分かりづらいた改善などもあったが、総じてプロジェクトは着実に進行した。

11. 成果

本プロジェクトでは、ユーザーが検索した結果を解析し、ユーザーの指示により、再ランキングすることを基本とするシステムを開発した。システムの概要を図 1 に示す。

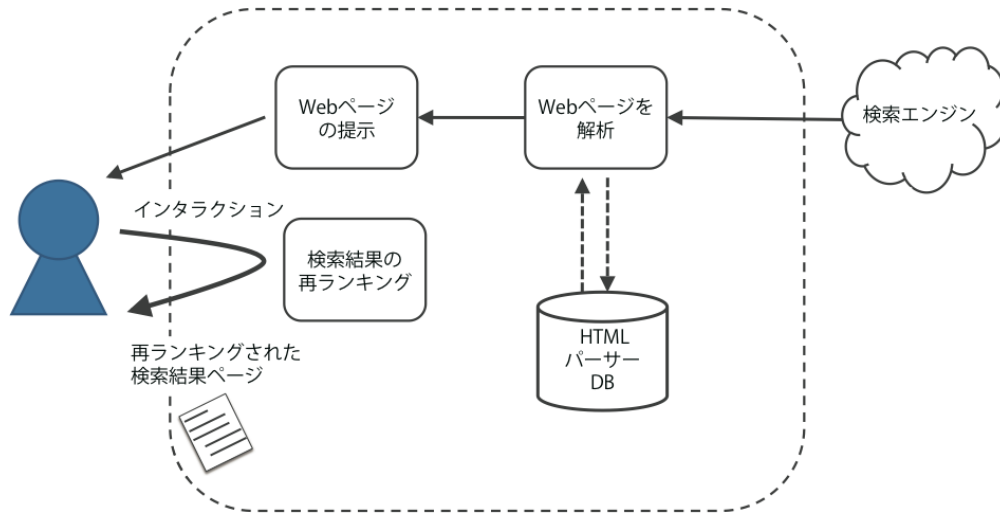
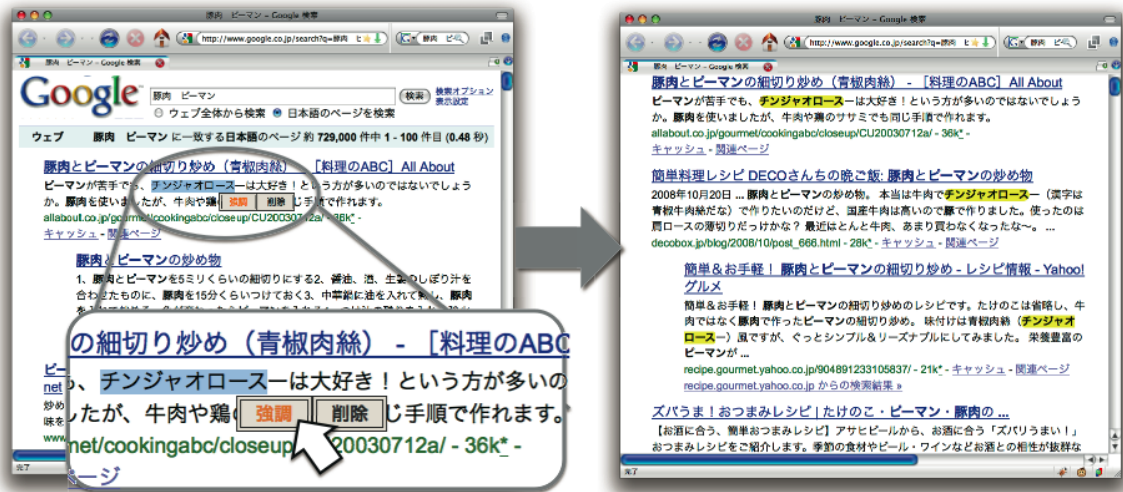


図 1 システムの概要

より具体的には、開発したシステムには、(1) 検索結果の再ランキングインタフェース、(2) ユーザーによる検索結果パーサー作成インタフェースとパーサー共有システム、(3) 異なるクエリ／検索エンジンの検索結果の統合再ランキングの 3 つの機能からなる。

まず(1)検索結果の再ランキングインタフェースは、検索結果のページ内で気になる単語を選択すると、強調と削除がセットになったメニューが表示される。この中で強調を選ぶと、その単語を含むページがより上位にランク付けされ、逆に削除を選択すると、その単語を含むページが下位にランク付けされるというものである。図 2 に強調の例を、また、図 3 に削除の例を示す。強調(削除)が行なう際に、ユーザーが選択したキーワードを含む検索結果は、直前のランキングを保ちながら上位(下位)に再ランキングするという方式となっている。つまり、直前のランキングを保持するということである。たとえば検索結果の中で単語 a を強調し、その後単語 b を強調したとする。するとランキングは b および a を含む検索結果が最上位に、続いて b のみを含む検索結果、 a のみを含む検索結果、両者を含まない検索結果と、並ぶことになる。



興味のある単語を選択し
「強調」を押すことによって…

選択したキーワードを含む
検索結果が上位に表示される

図 2 強調操作に基づく検索結果の再ランキング例

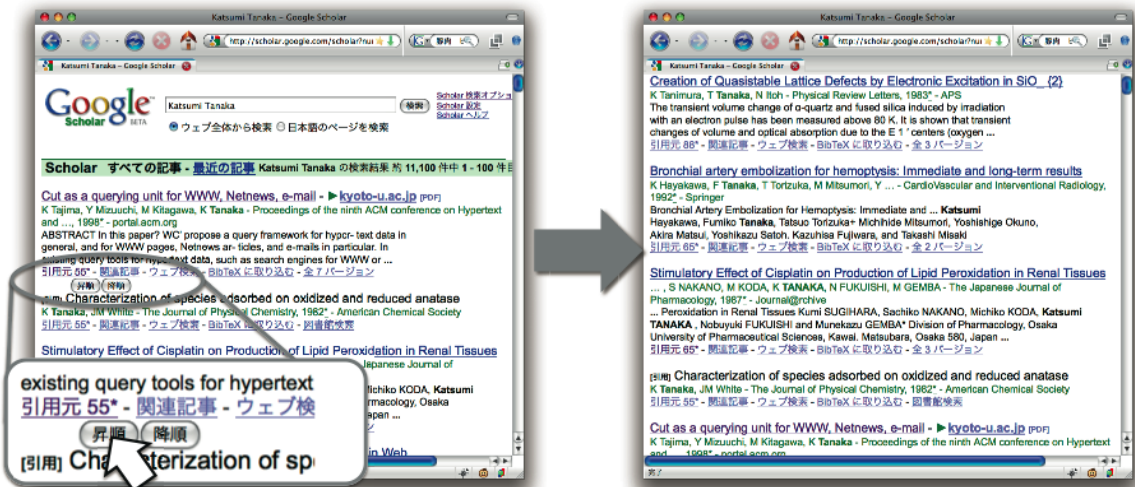


興味のない単語を選択し
「削除」を押すことによって…

選択したキーワードを含む検索結果を
下位に並びかえることができる

図 3 削除操作に基づく検索結果の再ランキング例

さらに検索した結果ページに含まれる属性値を選び、その値の昇順／降順で並べ替えることも可能である。この場合、並べ替えたい属性にマウスカーソルを合わせると、昇順／降順のメニューが出てくるので、そのどちらかだけを選べば良い。ホテルや商品の価格の順などで検索結果を再ランキングするのに用いることができる(図 4)。



並びかえたい属性に
マウスカursorを合わせることで

指定した属性に応じて検索結果を
昇順・降順に並べ替えることができる

図 4 昇順・降順に基づく属性値に応じた検索結果の再ランキング例

次に、(2)ユーザーによる検索結果パーサー作成インタフェースとパーサー共有システムであるが、これは、本システムに用意された HTML 解析のためのパーサーの基本部分をユーザーに公開し、ユーザー自らが比較的簡単に自分用のパーサーを作成できるようにしたものである。この機能を用いることにより、ユーザーは検索結果の単位を適切に調節することが可能となる。図 5 に検索結果ページの構造例を示す。

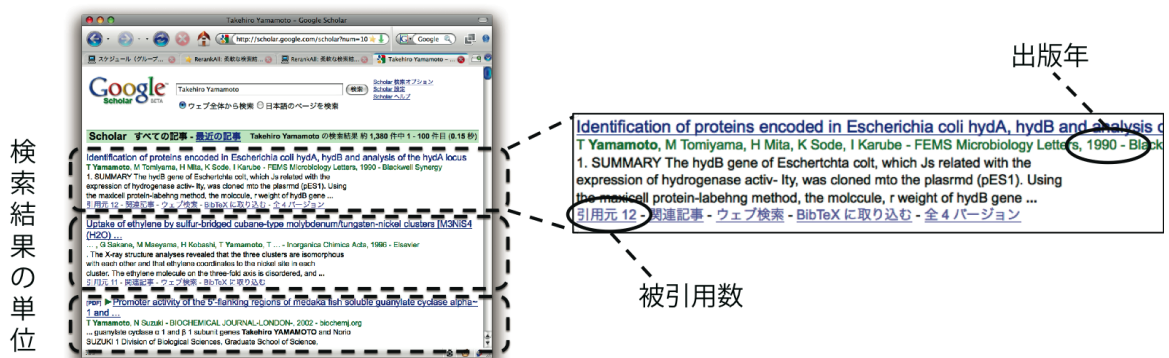


図 5 検索結果ページの構造例

図 6 に検索結果 HTML のパーサー作成例を示す。また、一度作成したパーサーを登録しておき、登録したパーサーを別のユーザーが利用する、パーサー共有機能も開発した。

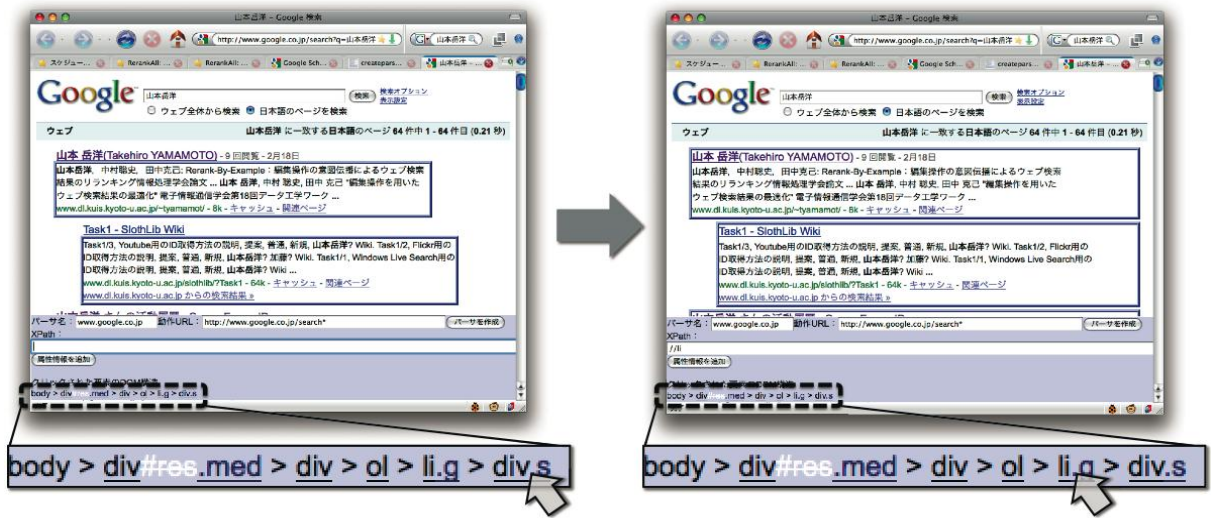


図 6 検索結果の HTML パーサー作成例

さらに、(3)異なるクエリ、検索エンジンの検索結果の統合再ランキングでは、たとえば、Amazonで検索した結果と、価格.comや楽天での検索結果を比較照合したいときに、これらの複数の検索結果を統合する機能である。このとき、ある尺度でランキングした結果は、別の検索結果にも反映するようになっている。たとえば、図7に示すのは、じゃらんとYahoo!トラベルの検索結果を価格について同時再ランキングした結果である。



図 7 じゃらんとYahoo!トラベルの検索結果を価格で同時再ランキングした結果

12. プロジェクト評価

従来の Web 検索では、一度検索した結果に、さらに条件を追加して再検索したり、あるいは、複数のブラウザ上で異なる検索結果を開いて見比べる、などが必要であった。山本岳洋君の今回の柔軟な再ランキングインタフェースの開発により、Web ユーザーは、ほんの簡単な操作だけで、特定の視点のものを強調したり、不要なものを削除したりできるし、あるいは、価格や数量などでランキングを変えたりも容易にできるようになった。また、複数の検索を連動させながら行なうことも可能になり、まさに、Web 検索ランキングを非常に柔軟かつ容易に扱えるものとなった。まさに、未踏に相応しいチャレンジングなプロジェクトであった。

13. 今後の課題

今後は開発したシステムを広く公開し、ユーザーからの意見も貰って、改善・改良を加えて行って欲しい。また、とくに、パーサーの作成と共有の部分は、特に実際に使ってもらった上での有効性の検証が必要であろう。