

## 遠隔教育システムにおける効果的な教材検索および表示

4 X - 5

## 機能の設計\*

井上 勝隆† 大和田 勇人† 溝口 文雄†

東京理科大学 理工学部 経営工学科†

## 1 はじめに

教材をHTMLで記述し、それをブラウザで見ることはネットワークを介した遠隔教育システム[1][3]において、非常に有用である。しかしながら、教育システムを用いて経験のない分野についての学習を行なう場合、理解が不十分な単語が出現する機会が必然的に多くなる。そのため、任意の単語について、その単語が記されている教材を検索するシステムを実装することにより、ユーザの学習に対する補助や、学習能率の向上を促進することができると考えられる。本稿では、これに加えて、教材構造[2]を考慮した検索や、教材とWWWのシームレスな検索、検索語句に関わる演習問題の自動生成、学習の進捗情報を考慮した検索などの特徴を持った検索システムを設計する。そして、本システムによる学習者における学習効率の変化についても検討する。

## 2 設計方針および特徴

本稿で述べる検索システムは、以下のような設計方針および特徴を備えている。

- 検索語がはっきりと定まっている場合には、検索語を入力することにより語句検索を行なうシステムが、WWWやファイルの検索などにおいて一般的である。それに対して、遠隔教育システムにおいては、検索対象となる教材は、章、節、項という項目にしたがって構成されているため、同一の項目に属するページについて、内容的に関連があると考えられる。そこで、検索機能として、内容的に近接するページを集合体とみなすことにより、ページの再評価を行い、検索結果をまとめる。
- 教材を検索する場合に、検索結果が少数しか存在しないような場合には、検索結果の信頼性に疑問が生

じる。そこで、教材中のみを検索結果が2つ未満の場合には、WWWについてのシームレスな検索を行うことで、検索語に関連するページを探し出す。

- 検索語句について理解が不十分であると判断される場合は、検索語句に関連したページを表示するだけでは、再び十分な理解が得られるとは判断しにくい。そこで、検索語句として入力されたキーワードについての演習問題における理解度を参照し、もしも理解が不十分であると判断される場合は、検索語句に最も関連があると判断されるページについて、検索語句を含めた重要語を抽出する。そして、抽出した重要語に対する演習問題を自動生成する。
- WWWにおける検索では、検索語句に対する重みが大いページが検索結果として優れていると判断される。しかし、遠隔教育システムの場合、教材には教材参照順序があり、同時に進捗情報が存在する。この進捗情報よりも先のページ、すなわち、まだ学習が行なわれていないページについての内容は、基本的に学習者にとって理解しにくいものであると考えられる。そのため、教材についての索引および検索結果を表示する際に、この学習者の進捗状況を考慮し、まだ学習していないページについては、検索結果に含めないものとする。

## 3 教材構造を考慮した検索

学習者からの検索語の入力を受けて、次のように検索を行なう。

1. 学習者から入力された検索語について、全文検索を行なう。検索語が含まれるページ数および出現数を計算し、検索語についての重みをそれぞれTF・IDF法によって計算する。TF・IDF法とは、全文書中において、キーワードが出現する頻度から、そのキーワードの重要度を計算する方法である。
2. 検索語が出現するページのうち、同一の章、節に属するものをまとめて、TF・IDF値を再計算する。

\*Design of effective educational search and indication system for remote education system.

†Katsutaka INOUE, Hayato OHWADA, Fumio MIZOGUCHI

†Dept. of Industrial Admin. Faculty of Sci. and Tech. Science University of Tokyo

### 3. 得られた TF・IDF 値の順にリンクを表示する。

このような教材構造を考慮した検索を行った場合、結果として得られるのはページおよび、ページの集合体であるため、ページ単位のみで検索を行なう時と違って、結果表示方法について検討しなければならない。教材では、Web ページと違い、教材ページ間の上下関係や順序関係がはっきりしているの、これを利用して、集合体におけるページの位置関係がわかりやすいように、ツリー形式により、検索結果を表示する。

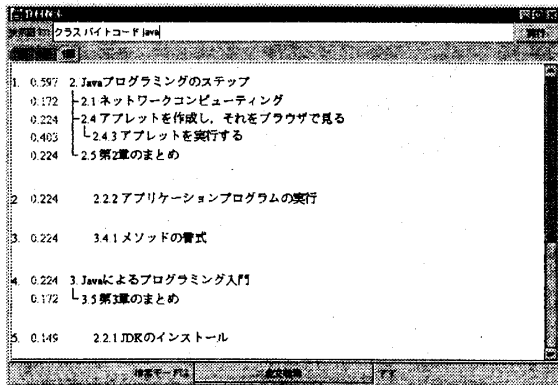


図 1: 検索結果表示

## 4 シームレスな WWW 検索

検索実行時において、検索結果が2つ未満の場合に、WWW コンテンツについても検索を行なう。WWW の検索方法としては、メタサーチを利用する。

1. 検索語句を代表的な複数のサーチエンジン (Goo, AltaVista, e.t.c.) に、GET リクエストを通じて渡し、それぞれ検索結果を得る。
2. 得られた検索結果をひとつにまとめ、検索結果に対する TF・IDF を用いた再評価を行ない、ランク付けを更新する。この過程で、検索結果として不適切なページは排除される。
3. 更新された情報を基に、リンクを表示する。

これにより、教材中においてあいまいにしか定義されないような単語についても、検索結果に有意なページを含めることを可能にする。

## 5 検索語の演習問題自動作成

1. 対象となるページからの重要語の抽出に際して、ページ内の文書を形態素に分割し、自立語についてそれぞれ TF・IDF 値を計算する。

2. 算出された各語の重みから、閾値を越えるものについて、重要語と定義する。

3. 重要語についての穴埋め形式の演習問題を解説文より自動的に作成する。穴埋め問題は、検索語句および重要語句のリストから、適正な語句を選択する形式で出題される。

このような演習問題を出題することにより、学習者の語句についての理解度を向上させることが期待できる。

## 6 進捗情報の利用

教育システム側から現在学習を行なっているファイル名を取得し、その情報をもとにした検索の制御を行なう。検索システムは、あらかじめ教材構造やキーワードを記述したキーワードデータファイルを保持している。このキーワードデータファイル中から教材構造を取得し、現在学習を行なっているファイル名を基に、既に学習が行なわれたページを教材構造中から特定する。既に学習が行なわれたページとして取得したファイルについてのみ、検索を実施することにより、ユーザが学習を行っていないページを表示することによる混乱を避け、あまり学習が進んでいない段階において、検索結果の数が膨れ上がることを防止する効果が期待できる。

## 7 結論

遠隔教育システムにおける検索として、検索語についての検索を実施するだけではなく、演習問題や WWW などを考慮することにより、教材に記述されている解説の枠にとらわれない、より学習効果の高い検索システムを設計した。とくに、教育システムにおける検索は、学習の効率を向上させる可能性を大いに秘めていると考えられる。

## 参考文献

- [1] 矢口 敬正, “インターネットを利用した学習支援システムの設計”, *AIST-TR-98-031, SUT Infomation Media Center*, 1997.
- [2] 山中 崇史, 矢口 敬正, 大和田 勇人, 溝口 文雄, “遠隔学習に適した教材記述言語の開発”, *情報処理学会第48回全国大会論文集*, 1999
- [3] 山中 崇史, 大和田 勇人, 溝口 文雄, “遠隔学習のための適応型教育システムの設計”, *人工知能学会第19回全国大会論文集*, pp.88-91, 1999