

# 멀티에이전트를 이용한 WBI 학습 지원 시스템

## WBI Support System using Multiagent

노은영\* · 허철회\*\* · 정환묵\*

\* 대구가톨릭대학교 컴퓨터 정보통신공학부

\*\* 성덕대학 컴퓨터정보계열

Eun-Young Roh and Chul-Hoi Her and Hwan-Mook Chung

Faculty of Computer & Information Communication Engineering,

Catholic University of Daegu

E-mail : eyroh@naver.com

### 요 약

웹을 이용한 교수-학습모형인 WBI 학습을 촉진시키기 위해서 교육전문가에 의해 검증된 학습자료를 효율적으로 관리하는 시스템을 구축하고, 사용자가 원하는 문서를 추정하여 제시한다. 사용자가 참조한 문서들에 대하여 에이전트가 추가, 삭제 등을 자율적으로 하며, 참조 문서간의 유사도를 측정한다. 이 유사도를 이용하여 적합한 문서를 추정하고 제시함으로써 사용자의 불명확한 정보요구에도 적합한 문서를 제공할 수 있다. 따라서, 사용자의 정보수집을 지원하고 웹을 통한 학습환경을 개선하여 WBI 학습을 촉진시킬 수 있다.

### 1. 서론

교육 인적자원부의 제 7차 교육과정에 의해 구성주의 학습과 개별화 학습이 강조되면서 Web을 이용한 교수-학습 모형인 WBI(Web based instruction)가 부각되고 있다. WBI란 특정하고 미리 계획된 방법으로 학습자의 지식이나 능력을 육성하기 위한 의도적인 상호 작용을 Web을 통해 전달하는 활동이다[1]. WBI에 사용되는 자료는 대부분 교사가 검색엔진에 의해 검색한 것이며, 자료들은 한번 사용한 후 버려지고 있다.

그러나 교사가 사용하였거나 참조한 자료는 그 자체로도 충분한 학습자료로서의 가치를 지닌다[2]. 이러한 자료를 체계적으로 구축하여 재사용(교수 또는 학습)할 수 있도록 함으로써 검증된 학습자료의 활용도를 높이고, 웹을 통한 상호작용을 증진시킴으로 WBI 학습을 촉진시키고 지원할 필요가 있다.

대부분의 사용자들은 자료를 검색하는데 있어서 자신의 정보요구를 명확히 하지 못하고 있으며, 정보요구를 명확히 하더라도 시스템 설계자가 아닌 이상 시스템 내의 자료를 충분히 활용하기에는 많은 시간과 노력을 필요로 한다.

본 논문에서는 사용자가 참조하는 문서의 유사도 측정을 통하여 필요로 하는 정보내용을 추론하고, 적합한 문서를 검색하여 제시한다. 이것은 사용자의 정보수집을 지원하고, 웹을 통한 학습환경을 개선하여 WBI 학습을 지원한다.

### 2. 정보수집과정

#### 2.1 정보 수집 과정 지원

Taylor는 정보 수집 과정에서 사용자의 정보 요구 상태를 4단계로 구분하고 있다[3].

[1단계] 직관적 요구: 현재 상태에 만족하지 못하고, 구체적인 표현이 언어로는 설명이 불가능한 상태.

[2단계] 의식된 요구: 문제를 인식할 수 있으나, 효과적인 언어로 표현할 수 없는 상태.

[3단계] 형식화된 요구: 문제에 대한 요구사항의 언어표현이 가능한 상태

[4단계] 조정 해결의 요구: 문제 해결에 필요한 정보를 활용할 수 있을 만큼 문제가 구체화 되어 있는 상태

정보 요구의 명확화 과정은 [1단계]에서 [4년

계]의 과정으로 이루어지며, [1단계]와 [2단계]에 서는 사용자가 막연한 정보요구 밖에 가지지 못 한다. 그리고, 검색 과정을 지원하기 위한 방법으 로 검색 질문 확장(query expansion)방법과 적합 성에 대한 피드백(relevance feedback)방법이 있 다.

검색 질문 확장방법은 사용자가 검색 시스템에 제시하는 검색 요구에 대해서 검색 키워드를 보 충함으로써 검색 요구를 자동적으로 확장하는 방 법이다.

적합성에 대한 피드백 방법은 사용자가 자신의 검색요구에 대한 문서의 적합성 판정을 행하고, 검색 시스템은 결과를 기본으로 다음 검색 식을 재구성하는 것을 반복해서 단계적으로 적합문서 후보를 좁히는 것으로 정보 수집 과정을 지원하 는 것이다. 따라서, 적합성에 대한 피드백에서는 사용자의 단계적인 정보요구를 명확하게 하고 있 다. 그러나, 적합성에 대한 피드백방법에서 사용 자는 검색 시스템이 제시하는 적합한 문서가 자 신의 정보요구에 대해서 적합성 판정이 가능한 것 을 전제로 하고 있기 때문에 참조 정보에 의한 사용자의 정보 요구가 영향을 받게되어 변할 수 있다. Taylor의 분류에 의하면 검색 질문 확장방 법은 [3단계]에 있는 이용자를 [4단계]로 지원하 고 있으나 [1,2단계]는 지원하지 않으며, 정보 요 구를 위하여 1,2단계에 있는 사용자에게 적합성 에 대한 피드백 방법을 이용하면 일관된 적합성 판정을 행하지 못하는 문제점이 있다.

## 2.2 적합 문서 평가 방법

검색 결과에 대한 문서 평가 방법인 순위 결 정 방법에는 TF-IDF, PageRank 등이 있다. TF-IDF를 이용한 순위 방법은 검색 키워드가 문서에서 출현하는 빈도에 의해 평가되어 순서화 되며, 문서가 가지는 정보를 생각하지 않는다. 따 라서 검색 결과에 나타난 적합 문서의 정보 내용 에 관해서는 표현하지 못하므로 중립적인 순서 기법이다. Google 등의 검색 시스템에서 사용하 고 있는 PageRank와 HITS 등 문서간에 참조성 에 기초한 적합도 평가에 의한 순위의 경우에는 적합문서에 관심도가 높은 문서가 포함되어 있는 특정의 정보에 대하여 공유하는 다수의 정보제공 자들의 순서에 의한 특정의 내용을 갖는 문서가 상위 에 있으며 그 외의 정보를 가진 문서는 하위 에 있게된다. 이것은 상위에 출현하는 적합문서 의 정보 내용에 편중된다. 그러므로 정보 요구를 명확하게 하지 못하는 일반 사용자에게 대한 순위 기법으로 적합하지 못하다. 또한 검색 결과에 나 타난 적합문서가 가지는 정보 내용간의 관련성은 적합도 평가가 반영되지 않아 적합문서의 다양성 이 적합도 평가에 반영되지 않는다.

## 3. WBI지원 시스템

### 3.1 시스템구성

검색질문 확장 방법과 적합성에 대한 피드백 방법의 단점을 보완하고 적합문서를 평가하여 일 반 사용자들이 문서를 검색하고 검색된 문서들을 참조하여 사용자들이 원하는 정보를 쉽게 찾을 수 있도록 그림 1 과 같이 시스템을 구성한다.

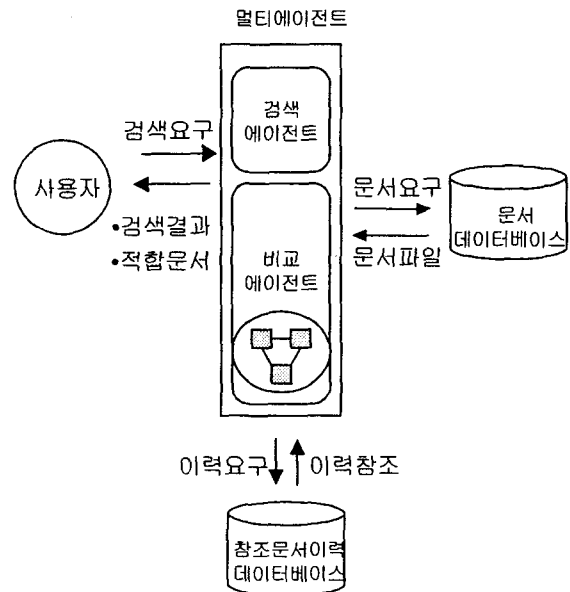


그림 1 WBI 지원 시스템

#### (1) 사용자

교육 및 학습을 위하여 키워드 형태로 정보를 요구하고 멀티에이전트에 의하여 적합한 문서를 제공 받는다.

#### (2) 검색에이전트

벡터공간모델을 이용하여 문서데이터베이스의 자료를 검색하고, 사용자에게 제시한다.

#### (3) 비교에이전트

참조문서간의 유사도를 측정하고 문서에 평가 값을 부여한다. 평가값을 통하여 사용자가 어떤 내용의 문서를 적합해 하는지 추정한다.

#### (4) 데이터베이스

문서데이터베이스는 검증된 학습자료를 보관하 는 데이터베이스이며, 참조문서이력데이터베이스 는 이용자가 검색결과 문서 중에서 참조하는 문 서의 이력을 저장하는 데이터베이스이다.

### 3.2 정보 요구의 모델화

사용자의 정보 요구를 추정하기 위해서 사용자 의 정보 수집 과정에서 참조하는 문서들을 문서 집합으로 하여 이용한다. 정보 수집 과정에서 참 조 문서들은 사용자의 정보요구에 유익한 정보를 가지는 문서이다. 이러한 참조 문서 사이에서 현 재의 참조 문서는 정보 요구에 대해서 유익한 정

보를 가질 가능성이 높으며, 비유사한 문서는 가능성이 낮다. 따라서, 문서집합에 포함된 정보의 유사성을 네트워크로 형성하고 이를 참조문서 네트워크라 한다.

참조문서 네트워크를 구성하는 문서들은 정보 요구에 대한 유용성의 정도를 평가값로 나타낸다. 네트워크에서 참조문서를 추가할 때 유사한 참조문서의 평가값은 올리고, 또 비유사한 참조문서에 관해서는 평가값을 내린다. 이러한 사용자의 정보 요구로 모델화 한 네트워크인 참조문서 네트워크는 사용자가 어떤 내용의 문서를 유용하다고 판단하고 있는가를 추정한다.

### 3.3 참조문서네트워크 모델

비교에이전트의 참조문서네트워크에서 참조문서 평가값의 계산순서는 그림 2 와 같다.

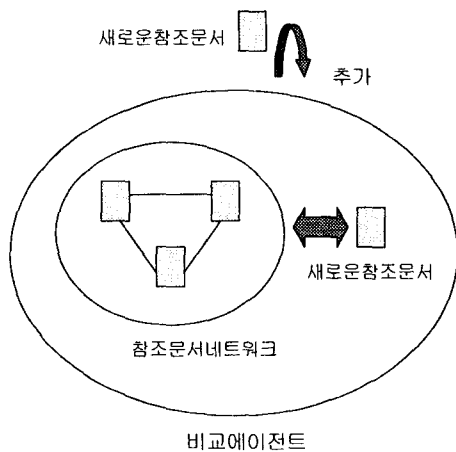


그림 2 참조문서 네트워크

참조문서의 네트워크는 다음과 같이 실행된다.

- [1단계] 비교에이전트에 참조문서가 3개가 될 때까지 추가한다.
- [2단계] 각 문서를 문서 벡터로 표현한다. 문서벡터간에 유사도를 측정하여 평균 유사도를 구하고 문서에 평가값을 부여한다.
- [3단계] 새로운 참조문서가 추가되면 문서벡터로 표현하여 평균 유사도와 유사정도를 비교하고 평균 유사도를 수정한다.
- [4단계] 평균유사도 보다 높은 유사도를 가지면 새로운 참조문서와 유사한 문서의 평가값은 상향 조정하고 비유사한 문서의 평가값은 하향 조정하여 순위를 정하고 임계치 값 이하의 문서는 제거한다.

### 3.4 다양성 관련 문서 추출

참조문서는 네트워크의 검색 결과에 적합 문서로 포함되어 추가된다. 이런 변경된 참조문서 네트워크에서 적합문서의 평가값은 검색 요구에

대한 적합도이며 그림 3와 같이 나타낸다.

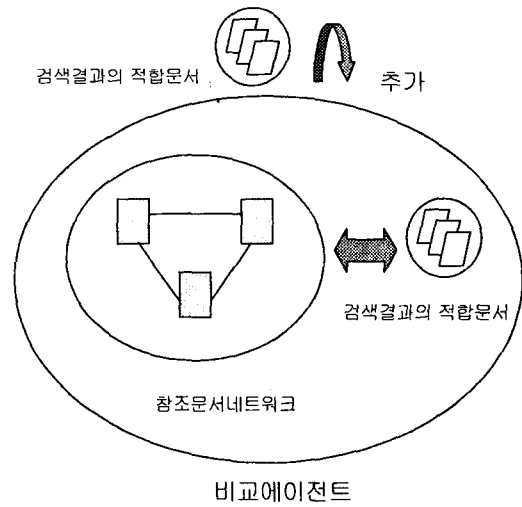


그림 3 적합문서 평가

다양성 있는 문서를 추출하기 위해 다음과 같이 실행한다.

- [1단계]이용자의 요구가 검색이면 2단계, 문서 참조이면 7단계로 이동한다.
- [2단계]사용자의 검색요구에 따라 검색 실행한다.
- [3단계]사용자가 참조한 문서를 수집한다.
- [4단계]참조문서 네트워크를 구성한다.
- [5단계]검색 결과의 적합문서를 참조문서 네트워크에 추가하고 문서 평가값을 계산하여 참조문서 네트워크를 재구성한다.
- [6단계]참조문서 네트워크의 문서 순위를 수정하고 결과를 이용자에게 제시하고 스텝 1로 돌아간다.
- [7단계]사용자의 문서 참조요구에 따라 문서를 수집하고 이용자에게 제시한다.
- [8단계]참조문서 네트워크를 참조한다.
- [9단계]사용자의 참조문서를 참조문서 네트워크에 추가하고 문서 평가치를 계산하여 참조문서 네트워크를 재구성한 후, 1단계로 돌아간다.

### 3.5 유사도 측정

벡터공간 모델에서는 질의나 문서의 키워드에 적절한 가중치를 할당할 수 있으며 질의와 문서의 유사한 정도에 따라 순위를 매길 수 있다는 장점을 갖는다[5]. 또한, 용어 빈도는 문서에 단어가 얼마나 자주 나타나는지를 측정하는 척도를 식 1과 같이 나타낼 수 있다.

$$f_{i,j} = \frac{freq_{i,j}}{\max_l freq_{l,j}} \quad (1)$$

$freq_{i,j}$  : 문서  $d_j$ 에서 용어  $k_i$ 의 출현 빈도수  
 $\max_l$  : 문헌  $d_j$ 에서 빈도수가 가장 큰 용어

역문서빈도는 문서빈도와 반비례하는 값으로 서 문서빈도는 특정 단어가 나타난 문서들의 개수를 의미한다. 역문서빈도가 높을수록 그 단어는 몇몇 문서에만 나타났다는 말이 되므로, 그 단어는 문서들을 구별지어주는 특징이 있다고 생각할 수 있으며 식 2와 같이 나타낸다.

$$idf_i = \log \frac{N}{n_i} \quad (2)$$

$n_i$  : 색인어  $k_i$ 가 출현한 문헌 수

$N$  : 시스템 내의 총 문헌의 수

용어빈도와 역문서빈도를 사용하여 문서  $i$ 에서 단어  $j$ 의 가중치를 구하면 식 3과 같다.

$$w_{i,j} = f_{i,j} \times \log \frac{N}{n_i} \quad (3)$$

따라서, 질의에 나타난 키워드의 가중치는 식 4와 같이 계산된다.

$$w_{i,q} = \left(0.5 + \frac{0.5 \text{freq}_{i,q}}{\max_l \text{freq}_{l,q}}\right) \times \log \frac{N}{n_i} \quad (4)$$

$\text{freq}_{i,q}$  정보요구  $q$  텍스트에서의 용어  $k_i$ 의 빈도수

문서와 질의에 나타난 키워드들의 가중치가 계산되면, 아래와 같은 방법으로 질의와 문서의 유사도를 계산할 수 있다. 또한, 이 유사도 값을 사용하여 질의와 유사한 문서 순위를 결정할 수 있게 된다.

$$\vec{q} = (w_{1,q}, w_{2,q}, w_{3,q}, \dots, w_{i,q}) \quad w_{i,q} \geq 0$$

$$\vec{d}_j = (w_{1,j}, w_{2,j}, w_{3,j}, \dots, w_{i,j}) \quad w_{i,j} \geq 0$$

$$\text{sim}(d_j, q) = \frac{\vec{d}_j \times \vec{q}}{|\vec{d}_j| \times |\vec{q}|} = \frac{\sum_{i=1}^l w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^l w_{i,j}^2} \times \sqrt{\sum_{i=1}^l w_{i,q}^2}}$$

#### 4. 시스템 구축

시스템 구축을 위한 운영체제는 윈도우즈2000 서버, 데이터베이스로 MS-SQL과 IIS 5.0 웹서버를 이용하였으며, 언어는 ASP를 이용하였다. 자료는 국민학교 3학년 국어 교육용 자료 13가지를 이용하였다.

##### 4.1 사용자인터페이스

사용자는 키워드를 입력하고 검색결과를 마우스로 선택할 수 있도록 하였다

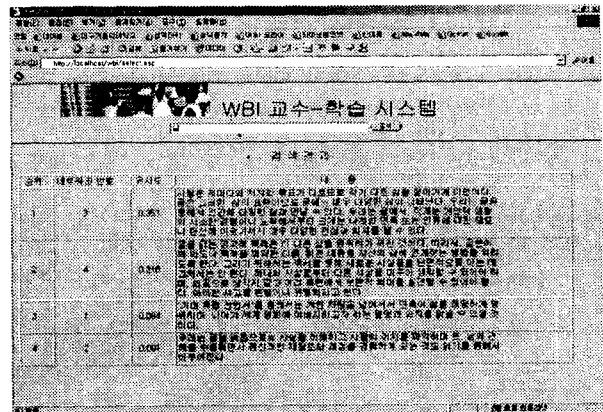


그림 4 사용자인터페이스

#### 4.2 고찰

기존의 정보검색 시스템은 키워드를 이용하여 검색하고, 검색된 결과들의 정보를 순위를 부여하여 사용자에게 제공하고 있으며, 사용자가 원하는 정보를 얻기 위해서는 키워드를 확장하거나 정보의 적합성에 대한 피드백으로 원하는 정보를 검색하고 있다. 또한, 정보 유익성의 판정 기준에 애매한 경우 지원 시스템이 제시하는 정보에 의해 이용자 자신의 정보 요구의 방향성이 달라질 수 있다. 그러나 본 논문에서는 사용자가 원하는 다양한 정보요구에 적절하게 대응할 수 있도록 참조 문서 네트워크를 만들고, 유사도를 이용하여 사용자가 원하는 다양성 있는 문서 추출로 정보를 찾을 수 있도록 하였다.

#### 5. 결론

정보 요구가 막연한 일반 사용자에게 정보 수집을 지원할 목적으로 참조한 문서의 문장을 이용하였다. 문서의 다양성을 유지하면서 환경에 적용할 수 있도록 에이전트를 이용하여 정보 수집 과정을 지원하였으며 정보수집 과정을 재사용할 수 있다. 또한, 사용자의 정보수집의 경험을 공유하여 학습적인 검색을 할 수 있다.

#### 6. 참고문헌

- [1]khan. B.H. Web-based instruction, New Jersey: Educational Technology Publications, 1997
- [2] 이건민, 웹 에이전트를 이용한 웹 기반 교수-학습 시스템, 서울교육대학교 교육대학원 석사학위논문, 2000
- [3] Stefano Mizzaro, On the Foundations of Information Retrieval, 1996.
- [4] 이상학, Meta-Search Engine에 기반한 인터넷 검색에이전트 개발에 관한 연구, 경희대학교 석사학위논문, 1997
- [5] 김명철, 최신정보검색론, 홍릉과학출판사, 2001