

개념 검색어 확장을 통해 질의 형식화를 도와주는 “개념 마법사”의 설계 및 구현

강 현 규[†]

요 약

정보 검색 시스템이나 웹(Web)이 방대해지고 초보적 수준의 사용자들이 늘어남에 따라 간단한 조작만으로 원하는 정보를 얻어낼 수 있는 도구의 개발이 점점 중요해지고 있다. 일반적으로 정보 검색 시스템이나 검색 엔진을 통한 질의 입력 방법이나 연산자들이 매우 다양하며 일반 사용자들은 질의를 형식화 하는 것이 쉽지 않다. 본 논문은 일반 사용자가 정보 검색 시스템이나 검색 엔진을 통하여 정보를 검색하기 위한 검색어 형식화를 도와주는 **개념 마법사**를 제안한다. 본 논문에서 제시한 **개념 마법사**는 실세계 지식의 부족을 시소러스를 이용하여 인터랙티브(inter-active) 하게 제시하고 웹을 기반으로 하는 플러그인(plug-in)으로 제공함으로써 유용성 및 확장성이 크다.

Design and Implementation of “Concept Wizard” Supporting Query Formulation with Concept Term Expansion

Hyun-Kyu Kang[†]

ABSTRACT

There are some important that development of tools to retrieve information by simple operation in large of nave users in the world wide web. In general, query formulation method and operators are variety, not easy to formulate query in information retrieval system or web based retrieval engine. In this paper, we propose “Concept Wizard” to support query formulation with concept term expansion in natural language query information retrieval system. The Concept Wizard are interactively supporting query formulation using thesaurus and providing plug-in on the web.

키워드 : 정보 검색(Information Retrieval), 질의 형식화(Query Formulation), 웹(Web), 자연어 질의(Natural Language Query), 개념 검색어(Concept Term), 개념 마법사(Concept Wizard)

1. 서 론

주요한 정보원으로서의 인터넷의 잠재적인 능력은 인터넷 상에서 사람들이 어떻게 원하는 정보를 빠르고 정확하게 찾아 낼 수 있느냐에 달려있다. 웹(Web)이 방대해지고 초보적 수준의 사용자들이 늘어남에 따라 간단한 조작만으로 원하는 정보를 얻어낼 수 있는 도구의 개발이 점점 중요해지고 있다. 사용자들은 컴퓨터에게 원하는 분야를 말하고 어떤 결과를 얻는다는 개념에는 익숙치 않다. 예를 들어 사용자가 도서관에 가서 사서에게 정보를 요청한다는 것은 이해할 수 있지만 인터넷에서 원하는 정보를 찾아내기 위해 검색엔진을 사용하는 것은 비록 두 절차가 논리적으로 동일하다고 해도 사용자들은 쉽게 이해할 수 없다[1].

일반적으로 정보검색 시스템(검색엔진 포함)의 질의 입력 방법이나 연산자들이 매우 다양하며 일반 사용자들은 질의를 형식화 하는 것이 쉽지 않다. 따라서 사용자가 검색 하고자 하는 검색어의 형식화를 쉽게하고 문서를 검색하지 못하는 경우 이를 해결할 수 있는 방법이 필요하다.

본 논문에서는 일반 사용자가 처음에 정보를 어떻게 검색해야 할지 모를 경우 개념을 나타내는 시소러스를 이용하여 사용자가 검색하고자 하는 의도를 나타낼 수 있는 검색어를 형식화할 수 있도록 도와주는 **개념 마법사(Concept Wizard : CW)**를 제안한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 관련연구에 대하여 기술하였다. 제 3장에서는 개념 마법사의 요구사항 및 설계 고려사항에 대해서 설명하고, 제 4장에서는 개념 마법사의 구성에 대하여 설명하였다. 제 5장에서는 개념 마법사의 구현 기능을 화면을 통해 자세하게 설명하였다. 마지막으로

[†] 중신회원 : 건국대학교 컴퓨터과학과 교수
논문접수 : 2001년 10월 16일, 심사완료 : 2002년 8월 16일

제 6장에서 결론을 맺는다.

2. 관련 연구

질의 용어와 검색 문서 사이의 관계를 이해하는 시스템으로 VIBE[4]는 여러 용어를 명시하고 이들과 가장 유사한 문서의 클러스터링을 이용하여 문서와의 연관성을 나타낸다. 이는 단지 검색어로 제시한 용어들을 나열하고 그와 연관성있는 문서들을 클러스터링한다. InfoCrystal[5]은 질의 용어 표현 수를 나타내기 위하여 기하학적인 모양을 연결하는 라인으로 구성되는 그래픽컬 질의언어를 사용한다.

Scatter/Gather[6]는 용어들과 국소적으로 밀접한 클러스터 그룹으로 문서들을 모으고 사용자에게 보여준다. 각 클러스터를 특징짓는 클러스터 용어들과 클러스터의 간단한 내용인 제목들로 구성되어 있다. 또한 검색 문서의 연관성을 판단하는데 도움을 주는 시스템으로 TileBars[7]는 질의 용어 분포와 검색 문서에서의 일치 부분을 시각적으로 보여준다. Scatter/Gather나 TileBars는 주로 검색 용어들의 그룹들이 문서상에서 어떻게 존재하는가에 관심을 가지고 있다. Envision[8]은 연관순서와 질의에 대한 각 문서의 특성을 추정된 연관성의 둘다를 이용하여 검색 결과를 시각화(visualization) 한다. 검색 문서의 X-Y축의 속성을 달리하고 다양한 아이콘 모양이나 색을 이용하여 검색 정보를 시각화 하고 있다.

그러나 위의 시스템들은 검색어를 확장하거나 검색어들에 대하여 가중치등을 부여하지는 않는다. 아울러 위의 시스템들은 단독(stand-alone) 시스템으로 동작하는 시스템으로 웹을 고려한다거나 질의 형식화를 쉽게 도와주는 그러한 시스템들은 아니다.

3. 개념 마법사의 요구사항 및 설계 고려사항

3.1 질의 형식화(formulation)

일반 사용자들은 질의 형식화하는 것을 어려워한다. 실생활에서는 도서관 등에 가서 길거나 부정확한 용어를 사용하여 요구사항을 표현하거나 메뉴를 통해 항목(item) 들을 선택한다. 사용자들은 요구사항을 표현하기 위해 인위적인 문자열을 사용하는데 익숙하지 못하다. 일반 사용자들은 다음과 같은 시도나 오류를 범한다.

- 컴퓨터가 자연어로 표현된 요구를 이해할 수 있는 기대
- 동시에 여러 검색어들을 표현하려는 시도
- 너무 적은 수의 단어나 과도하게 많은 수의 단어를 검색

어로 사용함

따라서 일반 사용자들이 원하는 정보검색을 표현하기 위해 쉽게 질의를 형식화하는 방법이 있어야 한다.

3.2 검색어 선정

질의 형식은 사용자가 자연스럽게 표현할 수 있는 자연어 질의를 원하게 된다. 내재적으로 사용자는 동음이의어 등의 모호성 처리가 되는 것으로 기대하고 있다. 따라서 시스템이 질의를 해석하기 위해서 피드백(feedback) 정보가 필요하다. 또한 파라미터 등을 조절할 수 있는 사용자 제어(user control)를 제공해야 한다.

3.3 검색 실행

명시적으로 검색을 실행하는 경우에 검색을 하며 내재적으로 질의에 대하여 검색어가 변할때마다 계속하여 보여 주고 수정할 수 있어야한다. 또한 네트워크의 사정이나 검색 시간이 지연되는 경우 이를 중지시킬 수 있는 실행 정지가 필요하다.

3.4 실세계 지식 요구

검색에 대해 필요한 실세계 지식의 부족으로 인해 쉽게 얻을 수 있는 정보를 얻는데 어려움을 겪는다. 따라서 실세계의 지식을 표현하는데 도움을 받을 수 있는 개념적으로 질의를 확장할 수 있는 방법이 있어야 한다.

3.5 질의 형식화 내지는 재형식화의 다양성이 해결되어야 한다. 사용 연산자의 표준화가 필요하다.

- AND, 그리고, \$, *, 와, 과, 및
- OR, 또는, +, |, 혹은
- NOT, -, ~, !
- 절단검색 : *, %
- 유의어 : #
- 위치검색 : NEAR, ^, { }
- 구단위검색 : "x x x"
- 기타 : !&, &!, (x)

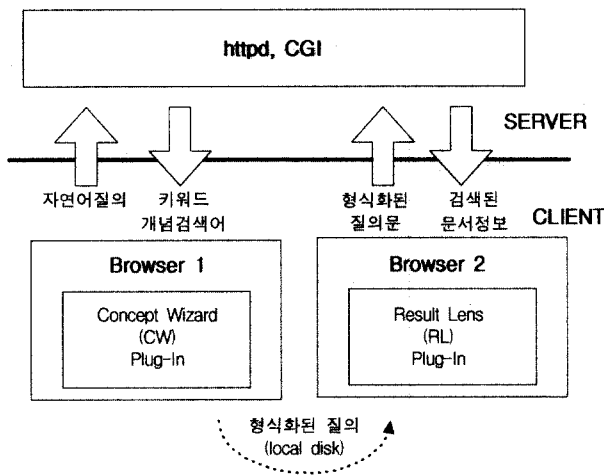
3.6 기타 고려사항

- 질의 발자취(history)를 유지해야 한다.
- 가능한 한 색(color)나 크기(size)를 다양하게 유지하여 한눈에 알아볼 수 있도록 도움을 주어야 한다.
- 보다 미세한 검색어의 가중치(중요도)를 조정할 수 있어야 한다.
- 검색 문서의 갯수를 지정할 수 있어야 한다.
- 사용자 인터페이스는 단순 명료하고 편리해야 한다.

4. 개념마법사의 구성

4.1 전체 시스템 구성

본 시스템은 (그림 1)과 같이 클라이언트(client)/서버(server)로 구성되어 있다. 클라이언트는 사용자의 자연어 질의를 형식화하고 정제(refine) 하기위한 플러그인 프로그램인 개념 마법사(Concept Wizard : CW)와 추출된 대량의 문서를 순위화하여 시각적으로 보여주는 플러그인 프로그램인 결과렌즈(Result Lens : RL) [11]로 구성되어 있다. CW와 RL은 사용자의 요구에 따라 브라우저를 통해 서버측의 httpd나 CGI (Common Gate Interface)를 호출하여 서버와 데이터를 교환한다.



(그림 1) 전체 시스템 구성도

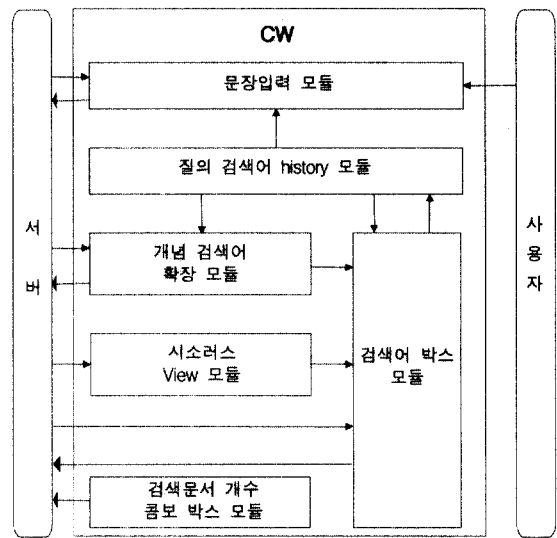
사용자의 검색요구는 CW를 통하여 이루어지며 RL은 검색되어 나온 추출된 문서 결과를 다양한 알고리즘을 적용하여 여러 관점으로 볼 수 있는 환경을 제공한다. 사용자의 검색요구에 의해 서버는 자연어 질의로부터 키워드(원 질의어)들을 추출하고 시소러스를 통해 각 키워드에 해당하는 개념 검색어들과 함께 CW로 전송한다. 사용자는 CW를 이용하여 서버로부터 전송받은 검색어들 중에서 원하는 검색어들을 선택하여 세세한 가중치 조정도 할 수 있으며 간단한 마우스 조작에 의해 질의의 형식화를 손쉽게 할 수 있다.

서버에게 문서에 대한 추출을 요구하기 위해 사용자는 ‘검색 시작’ 버튼을 선택하여 로컬 디스크(local disk)를 통해 형식화된 질의문을 RL에게 넘겨주게 된다. RL은 자신이 실행되자마자 로컬 디스크로부터 질의문을 읽어들이어 서버에게 문서들을 요구한다. 질의문을 CW가 직접 서버에게 전달하지 않고 RL이 로컬 디스크에서 질의문을 받아서 서버에게 요구하는 것은 각 클라이언트들이 요구한 질의문과 문서 검색 결과를 서버가 관리할 필요를 없애주기 때문에 많은 사

용자가 서버에게 검색을 요구할 때 서버의 부담을 줄일 수 있게 된다.

4.2 개념 마법사 시스템 구성

CW는 여러 구성요소(component) 들이 각각의 기능을 수행하면서 사용자의 행위(action)에 대해 서로 정보를 교환한다. CW는 6개의 모듈로서 ‘문장입력 모듈’, ‘질의검색어 리스트 모듈’, ‘개념검색어 확장 모듈’, ‘시소러스 View 모듈’, ‘검색어 박스 모듈’, ‘검색문서개수 콤보박스 모듈’로 구성되어 있으며 (그림 2)와 같다.



(그림 2) 개념 마법사 시스템 구성도

‘문장입력 모듈’은 사용자의 질의를 받는 모듈이고 ‘질의검색어 history 모듈’은 사용자가 이전에 형식화 했던 질의 정보들을 그대로 저장해 놓았다가 사용자에게 의해 다시 그 전 상태로 복원시키기 위한 모듈이다. ‘개념검색어 확장 모듈’과 ‘시소러스 뷰(View) 모듈’은 사용자가 질의한 원 질의어 외에 원 질의어와 관련이 있는 개념 검색어들을 기하학적인 모델을 통해 사용자들에게 적절한 검색어들을 연상시키는 역할을 한다.

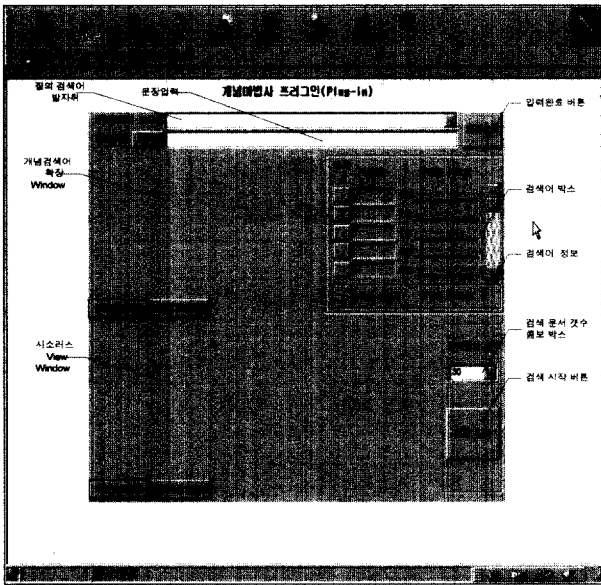
사용자에 의해 선택된 모든 검색어들은 ‘검색어 박스 모듈’이 처리하게 되며 단순히 검색어외에 검색어의 중요도에 따라 가중치를 조정할 수 있으며 ‘검색 X’와 같은 검색어 단위의 동작(operation)을 정의할 수 있다.

5. 개념 마법사의 구현 기능

개념마법사(CW)는 사용자 질의의 형식화를 도와주는 정보검색 툴(tool)로서, PC, Windows NT상에서 개발되었으며 개발 툴로는 Visual C++ 6.0을 사용하였다.

5.1 개념 마법사의 초기 화면

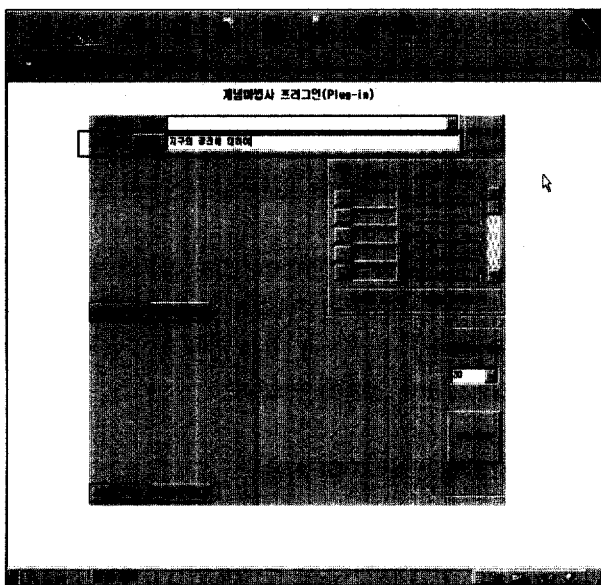
개념마법사(CW)는 (그림 3)에서와 같이 '문장입력', '질의 검색어 발자취', '입력완료 버튼', '검색어 박스', '검색어 정보', '개념검색어 확장 Window', '시소러스 View Window', '검색 문서 갯수 콤보 박스', '검색 시작 버튼'으로 구성되어 있다.



(그림 3) 개념 마법사의 초기 화면

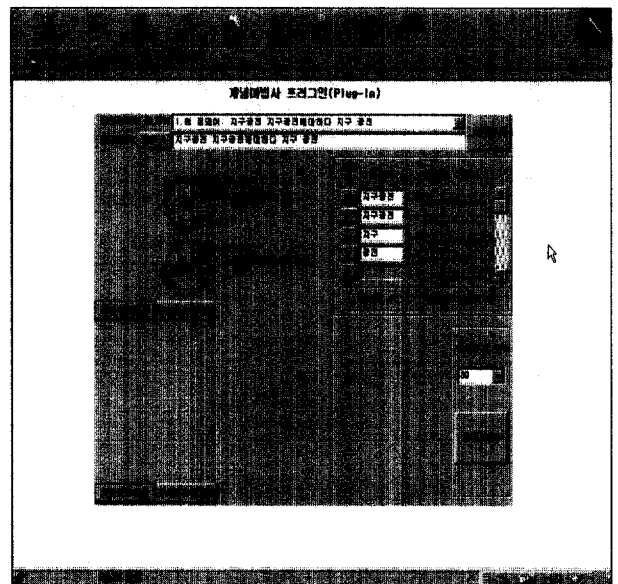
5.2 사용자의 자연어 질의문 입력

사용자는 (그림 4)에서와 같이 구성요소(component) '문장 입력'을 이용하여 자연어 질의를 할 수 있다.



(그림 4) 자연어 질의문 입력

자연어 질의를 하고 나서 검색어의 추출을 위해 '입력완료 버튼'을 누른다. 버튼을 누르면 개념마법사는 '문장입력' 모듈로부터 자연어 질의를 추출하여 서버에게 전달한다. 서버가 개념 마법사로부터 자연어 질의를 받으면 질의문으로부터 키워드를 추출한다. 그리고 각 키워드의 동음이의어와 동의어를 시소러스로부터 추출하여 개념마법사에게 추출된 검색어들을 전달한다. 개념마법사는 서버로부터 전달받은 검색어들을 '문장입력', '질의 검색어 발자취', '개념검색어 확장 Window', '검색어 박스'로 보낸다. 이 때, '문장입력'과 '검색어 박스'에는 원 질의어만 보내고, '질의 검색어 발자취'와 '개념검색어 확장 Window'에는 원 질의어, 동음이의어, 동의어를 보낸다. 이에 따른 실행결과는 아래 (그림 5)와 같다.



(그림 5) 검색어 추출 및 개념 검색어 확장 화면

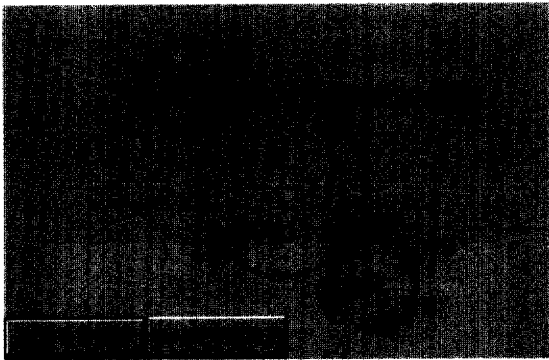
5.3 개념검색어 확장 window에서의 개념검색어

(그림 6)은 원 질의어와 그에 해당하는 동음이의어, 동의어를 원형모델로 나타낸 '개념 검색어 확장 window'이다.



(그림 6) 개념 검색어 확장 window

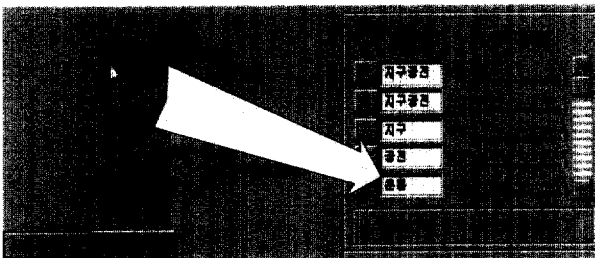
중앙의 가장 큰 붉은색의 궤도(orbit)는 사용자가 입력한 자연어 질의로부터 추출한 원 질의어를 나타내는 궤도이고 붉은색의 노드를 중심으로 하는 푸른색의 작은 궤도는 붉은색 노드의 동음이의어들을 나타낸 궤도이다. 동음이의어를 나타내는 궤도는 위의 (그림 6)과 같이 노드를 중심으로 하여 표현하지만, 동의어는 (그림 7)과 같이 노드를 외접하는 형태로 표현한다. ((그림 7)은 자연어 질의로 ‘말갈과 고구려 발해에 대하여’라고 질의를 내렸을 때의 결과이다.)



(그림 7) 동의어를 갖는 개념 검색어 확장 window

(그림 7)에서 붉은색의 궤도는 활성화된 궤도를 나타내고 푸른색의 궤도는 비활성화된 궤도를 나타낸다. 활성화된 궤도란, 화면 상의 여러 궤도들에 대해 사용자가 조작할 수 있는 활성화된 궤도를 말한다.

활성화된 궤도상의 노드는 마우스로 클릭하여 선택할 수 있으며, 활성화된 궤도상의 노드를 선택하려면 먼저 궤도를 선택한 후에 노드를 선택해야만 한다. 궤도를 선택하기 위해서는 궤도를 클릭하거나 궤도의 내부를 클릭하면 활성화가 된다. 또한 궤도의 내부를 클릭했을 때 클릭한 위치가 여러 궤도에 속한다면 클릭할 때마다 해당하는 궤도들을 차례로 활성화한다. 그리고 선택된 노드의 해당 검색어는 검색어 박스에 삽입이 된다.(아래 (그림 8)은 ‘공전 3(운동)’을 선택했을 때 검색어 박스로 검색어 ‘운동’이 삽입되는 그림이다.)



(그림 8) 개념 검색어 확장 window에서의 검색어 삽입

하나의 궤도는 최대 6개까지의 검색어들을 나타낼 수 있기 때문에 해당 궤도에 대해 검색어의 갯수가 최대 갯수를 넘었을 경우에 ‘시계방향회전’이나 ‘반시계방향회전’ 버튼을 클릭하여 화면 상에 표시되지 않은 노드들을 볼 수 있다.

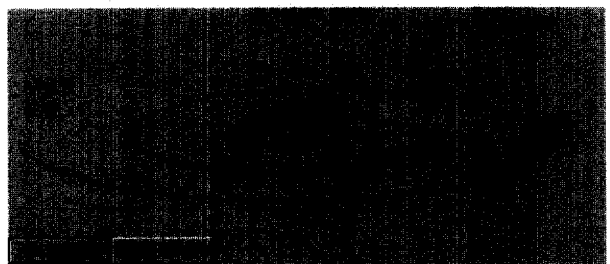
5.4 시소러스 View Window에서의 개념검색어

‘시소러스 View Window’를 이용하면 동음이의어나 동의어에 ‘상위개념어’, ‘하위개념어’ 그리고 ‘관련개념어’를 볼 수 있다. (그림 9)처럼 검색어 ‘공전 3’의 ‘관련개념어’를 보기 위해 노드 상에서 마우스의 오른쪽 버튼을 클릭하면 팝업메뉴가 화면 상에 디스플레이(display) 된다. 사용자는 팝업메뉴로부터 원하는 검색어를 클릭하여 선택할 수 있고 해당 검색어(예에서는 ‘공전 3’)와 메뉴를 통해 선택한 검색어 종류(예에서는 ‘관련개념어’)가 서버로 전송된다.



(그림 9) 시소러스 개념어 보기

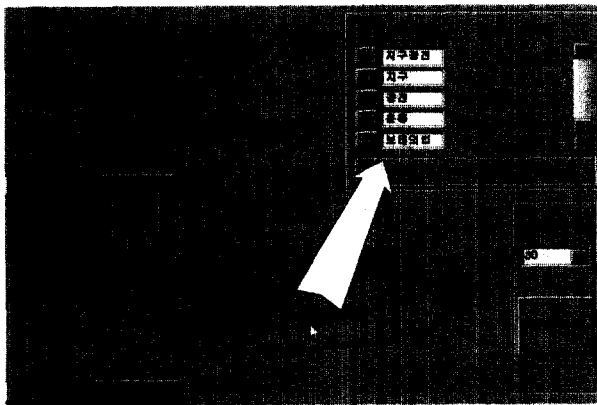
서버는 개념마법사로부터 받은 검색어와 원하는 검색어 종류를 시소러스를 이용하여 찾아서 개념마법사로 다시 전송한다. 전송받은 검색어들(예에서는 ‘관련개념어’)을 화면상에 표시한 그림은 (그림 10)과 같다.



(그림 10) 관련 개념어 보기

전송받은 검색어들은 나선의 바깥쪽에서 안쪽으로 차례로 디스플레이(display)되며 검색어들이 많아서 한 화면에 디스플레이될 수 없는 검색어들은 왼쪽 하단의 회전버튼을 클릭

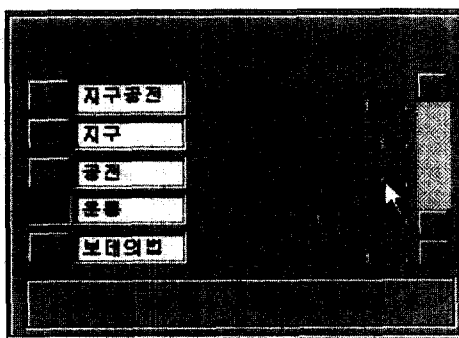
하여 보이지 않는 검색어들도 볼 수 있도록 하였다. '시소러스 View window'도 '개념검색어 확장 window'와 같이 노드를 선택하여 검색어 박스로 검색어를 삽입할 수 있다. (그림 11)은 노드 '보태의법칙'을 클릭하여 검색어 박스로 검색어를 삽입하는 그림이다.



(그림 11) 시소러스 view에서의 검색어 삽입

5.5 검색어 박스에서의 질의 형식화

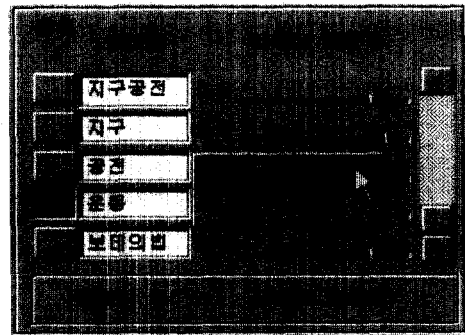
사용자가 입력한 자연어 질의문으로부터 추출한 원 질의어뿐만 아니라 '개념검색어 확장 window'나 '시소러스 View window'로부터 선택한 개념 검색어들은 모두 검색어 박스로 모아지게 된다. 검색어 박스는 사용자가 직접 질의했거나 선택한 검색어들을 질의의 형식화하는 기능을 담당한다(그림 12). 따라서 검색어들에 대한 연산을 제공하기 위해 '검색 X' 체크 박스와 '검색어 가중치' 슬라이더 바가 각 검색어마다 지정할 수 있도록 되어 있다. '검색 X' 체크 박스는 해당 검색어가 문서 상에 포함되어 있다면 그 문서는 검색대상에서 제외하기 위한 것이다. 또한 '검색어 가중치' 슬라이더 바는 각 검색어



(그림 12) 검색어 조정

단위로 가중치를 부여할 수 있도록 하여 사용자가 검색어의 중요도를 지정할 수 있도록 하였다.

사용자는 '개념검색어 확장 window'나 '시소러스 View window'를 통해 검색어들을 검색어 박스로 삽입할 수 있지만 직접 검색어 박스에 검색어를 입력할 수도 있다. 새로운 검색어를 검색어 박스에 삽입하기 위해 (그림 13)과 같이 검색어 박스 위에서 마우스의 오른쪽 버튼을 클릭하면 팝업메뉴가 디스플레이된다.



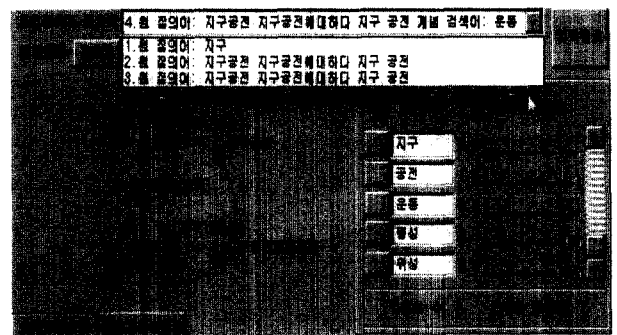
(그림 13) 검색어 삽입

메뉴의 '검색어 삽입'을 선택하면 마우스 커서가 위치한 다음 라인에 빈(empty) 라인이 삽입이 되고 여기에 새로운 검색어를 입력할 수 있다.

또한, 원하지 않는 검색어는 팝업메뉴의 '검색어 삭제'를 선택하여 검색어 박스로부터 검색어를 삭제할 수 있다. 그리고 검색어 박스의 총 검색어 수와 '검색 X'가 선택된 검색어 수는 검색어 정보란에 표시된다.

5.6 질의검색어 발자취를 이용한 상태 복원

개념마법사는 사용자가 질의를 내릴 때마다 검색어 박스에 있는 검색어와 검색어에 대한 정보들을 '질의검색어 발자취'에 저장한다. 이와 같이 하는 것은 언제라도 개념마법사가 이전의 상태로 돌아갈 수 있도록 하기 위한 것이다. 개념마법사가 검색어들을 저장하는 시기는 '사용자가 자연어 질의를 한 후, 입력완료 버튼을 누를 때'와 '검색시작 버튼을 누를 때'이

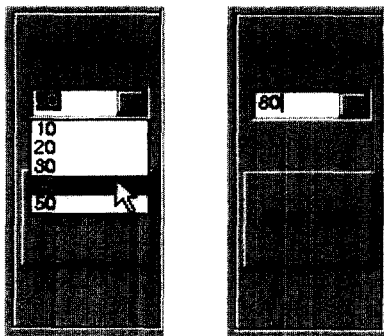


(그림 14) 질의 검색어 발자취

며 검색어 박스에 있는 검색어들을 원 질의어와 개념검색어로 나누어 리스트 박스에 저장한다. 이 때 사용자가 직접 입력한 자연어 질의로부터 추출한 검색어들은 원 질의어로 분류하고 시소러스로부터 얻은 검색어들은 개념검색어로 분류하여 저장한다. (그림 14)는 검색어들이 원 질의어와 개념검색어로 분류되어 '질의검색어 발자취'에 저장되어 있는 것을 보여준다. 사용자는 이 중에서 원하는 리스트를 선택하여 이전의 상태로 돌아갈 수 있다.

5.7 검색 문서 갯수 지정과 검색 시작

사용자는 검색어 박스에 있는 검색어들로 서버에게 문서 검색을 위해 질의를 내리게 된다. 문서 검색을 위해 질의를 내릴 때 사용자는 추출할 문서의 갯수를 지정할 수 있는데 문서 검색 리스트 박스를 선택하여 지정하거나 직접 입력하여 문서 갯수를 지정할 수 있다(그림 15).



(그림 15) 검색 문서 개수 지정

문서 검색의 갯수를 지정한 후에 '검색시작' 버튼을 클릭하여 서버에게 검색을 요구할 수 있고 검색되어 나온 문서들은 RL을 통해 디스플레이된다.

6. 결 론

본 논문에서는 현재의 정보검색 시스템 및 웹(web)상의 검색엔진들에 대한 개선되어야 할 요구사항 및 고려사항들을 정의하였다. 또한 정보검색 시스템의 문제점을 지적하고 일반 사용자가 처음에 정보를 어떻게 검색해야 할지 모를 경우 개념을 나타내는 시소러스를 이용하여 사용자가 검색하고자 하는 의도를 나타낼 수 있는 검색어를 형식화할 수 있도록 도와주는 개념 마법사(Concept Wizard : CW)를 제안하였다. 구체적으로 화면을 중심으로 구현된 기능들을 설명하였다.

사용자 선택에 의해 질의, 개념이 확장이 가능하고 동음이의어 및 동의어 처리가 가능하며 실세계 지식의 부족을 인터

랙티브(interactive) 하게 제시함으로써 확장가능 하도록 하였다. 현재 웹을 기반으로 하는 검색엔진 및 정보검색 시스템에 적용할 수 있도록 하기 위하여 플러그 인(Plug-in)으로 구현 하였고 서버(server) 측에서는 클라이언트(client)의 검색어 추출을 위해 CGI방식으로 구현되었다. 클라이언트와 서버 사이의 검색을 의뢰하고 검색결과를 주고 받기 위해서 기본적으로 서버는 IIS(Internet Information Server)를 이용하고 그 결과를 웹브라우저 클라이언트의 플러그인에서 받아 기능을 수행한다.

참 고 문 헌

- [1] A. Pollock and A. Hockley, "What's Wrong with Internet Searching," <http://www.dlib.org/dlib/march97/bt/03pollock.html>, D-Lib Magazine, March, 1997.
- [2] M. Hearst, G. Kopec, and D. Brotsky, "Research in Support of Digital Libraries at Xerox PARC," <http://www.dlib.org/dlib/june96/hearst/06hearst.html>, D-Lib Magazine, June, 1996.
- [3] G. Salton, Automatic Text Processing, Addison-Wesley Publishing Co., Reading, MA, 1989.
- [4] K. A. Olsen, R.R. Korfhage, K. M. Sochats, M. B. Spring, and J. G. Williams, "Visualization of a Document Collection : The VIBE system, Information Processing and Management," 29(1) : 69-81, 1993.
- [5] M. Beaulieu, "Experiments on Interfaces to support Query Expansion," Journal of Documentation, Vol.53, No.1, January, 1997.
- [6] D. R. Cutting, D. Karger, and J. Pederson, "Constant Interaction-Time Scatter/Gather Browsing of Very Large Document Collections," SIGIR '93, Pittsburgh, PA, pp.126-134, 1993.
- [7] M. A. Hearst, "TileBars : Visualization of Term Distribution Information in Full Information Access," CHI '95, Denver, Colorado, pp.67-73, 1995.
- [8] L. T. Nowell, R. K. France, D. Hix, L. S. Heath, and E. A. Fox, "Visualizing Search Results : Some Alternatives to Query- Document Similarity," Proceedings of SIGIR '96, Zurich, Switzerland, pp.67-75, 1996.
- [9] Kang, H. K., Choi, K. S., "Two-level Document Ranking Using Mutual Information in Natural Language Information Retrieval," Information Processing & Management, Vol.33, No.3, pp.289-306, 1997.

- [10] 강현규, 박세영, “정보 검색”, 정보처리학회지, 정보처리학회, 제5권 제5호, pp.37-47, 1999.
- [11] 강현규, 박세영, “정보 검색 결과의 효과적 제시를 위한 ‘결과 렌즈’의 설계 및 구현”, 정보과학회논문지(C), 제5권 제3호, 한국정보과학회, pp.375-384, 1999.
- [12] 알타비스타, <http://www.altavista.co.kr/>.
- [13] 라이코스, <http://www.lycos.co.kr/>.
- [14] 네이버, <http://www.naver.com/>.
- [15] 한미르, <http://www.infocop.com/>.
- [16] 심마니, <http://www.simmani.com/>.
- [17] 엠파스, <http://www.empas.com/>.



강 현 규

e-mail : hkkang@kku.ac.kr

1985년 홍익대학교 전자계산학과 졸업(학사)

1987년 한국과학기술원 전산학과(석사)

1992년 정보처리 기술사 자격 취득

1997년 한국과학기술원 전산학과(박사)

1987년~2000년 한국전자통신연구원 책임

연구원

2001년~현재 건국대학교 컴퓨터과학과 교수

관심분야 : 정보검색, 자연어처리, 한국어정보처리, XML, 디지털 라이브러리 등