



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0049626
(43) 공개일자 2009년05월19일

(51) Int. Cl.

G06F 17/30 (2006.01)

(21) 출원번호 10-2007-0115787

(22) 출원일자 2007년11월14일

심사청구일자 2007년11월14일

(71) 출원인

조선대학교산학협력단

광주광역시 동구 서석동 375, 조선대학교 산학기
획팀

(72) 발명자

정채영

광주 동구 서석동 조선대학교

강래구

광주 동구 서석동 조선대학교 375

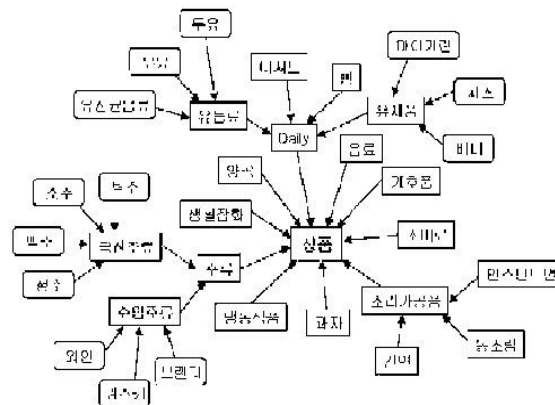
전체 청구항 수 : 총 1 항

(54) 도메인 온톨로지를 이용한 검색 시스템 설계 및 구현

(57) 요약

지금까지의 상품 검색 방법으로는 찾고자하는 정보를 검색할 때 주로 단어의 빈도수나 어휘 정보를 이용하는 키워드 기반의 검색이 주로 쓰이고 있었다. 키워드 기반의 검색에서는 사용자의 질의와 관련이 없는 문서들까지도 같은 결과로 나타내 주고 이로 인해 사용자는 제시된 결과를 한번 더 수동적으로 검색해야하는 부담을 앓게 되었다. 이러한 문제점을 해결하기 위해 온톨로지가 대두되었다. 본 논문에서는 온톨로지를 이용한 상품 검색 시스템을 직접 구축하여 분류별 검색을 통해 얼마나 정확한 검색을 하는지 실험하였다. 실험을 위해 전국적으로 On/Off 라인 할인점을 운영 중에 있는 A할인점의 상품 데이터 약 40,000여개를 데이터베이스로 구축하였고 User Interface 개발환경은 JSP와 PowerBuilder9.0을 사용하여 검색 시스템을 개발하여 실험하였다. 그 결과 본 논문에서 제안하고 설계한 상품 도메인 온톨로지를 이용한 검색 방법이 기존의 키워드 기반의 검색 방법보다 우수한 결과를 나타내고 있음을 입증하였다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 1-07-1207-001-10780-00-003

부처명 한국문화콘텐츠진흥원

연구사업명 문화콘텐츠기술연구소(CT)육성사업

연구과제명 멀티미디어 특성요소를 이용한 지능형 콘텐츠 제작기술

주관기관 조선대학교 산학협력단

연구기간 2007년 2월 1일 ~ 2008년 1월 31일

특허청구의 범위

청구항 1

키워드 기반의 검색에서는 사용자의 질의와 관련이 없는 문서들까지도 같은 결과로 나타내 주고 이로 인해 사용자는 제시된 결과를 한번 더 수동적으로 검색해야하는 부담을 안게 되는 문제점을 해결하기 위한 온톨로지를 이용하여 상품 검색 시스템을 직접 구축하여 분류별 검색을 통해 얼마나 정확한 검색을 사용자에게 제공하는 시스템;

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

- <6> 1989년 Tim Berners-Lee에 의해 처음 제안된 월드와이드웹은 HTML을 이용하여 누구나 쉽게 하이퍼텍스트 정보에 접근할 수 있게 편리성에 중점을 두으로써 결과적으로 정보의 증가를 가져왔다. 이처럼 사용하기 편리하다는 점이 웹의 주요 성장 요인이기도 하지만 반면에 많은 양의 정보가 폭발적으로 생성되는 현 상황에서 정보의 저장과 검색에 있어서 많은 문제점이 생겨나고 있다. 현재의 웹 사이트에서의 검색은 사용자가 자주 사용하는 용어와는 관계없이 웹 사이트 개발자나 데이터베이스 설계자가 일방적으로 정해 놓은 키워드를 통해서만 정보를 검색할 수 있다는 문제점을 가지고 있다. 이러한 문제점을 극복하기 위해서 1999년 W3C (World Wide Web Consortium)에서는 시맨틱 웹(Semantic Web)을 제안하게 되었고 그 중심에 온톨로지(Ontology)가 있다.[1]

발명이 이루고자 하는 기술적 과제

- <7> 지금까지의 검색은 주로 단어의 사용 빈도수나 어휘정보를 이용하여 문서의 유사도를 측정하고 순위를 부여하기 때문에 사용자가 찾고자 하는 정보와 관계가 없는 동형이의어나 동음이의어와 같은 단어가 포함된 문서들까지도 동일한 결과로 나타내었고 이로 인해 사용자는 제시된 결과를 한번 더 검색을 해야 하는 부담을 안게 되었다.

발명의 구성 및 작용

- <8> 온톨로지를 구축함에 있어서 최상위 클래스는 부모를 갖고 있지 않기 때문에 제일 상위에 위치하게 된다. 핵심 클래스란 부모를 가지고 있는 클래스로서 각 분류의 조상 클래스가 된다. 핵심 클래스는 관점에 따라 조금씩 달리 설계를 하기도 한다. 하지만 보편적으로 온톨로지를 구축함에 있어서 핵심 클래스는 2단계 혹은 3단계 수준에 위치하게 된다.
- <9> 마지막으로 행위 클래스는 트리구조에서 가장 하위에 위치해 있고 자손으로 비유된다. 최상위 클래스와 핵심 클래스를 제외한 모든 것을 행위 클래스로 간주한다.
- <10> 클래스의 계층 구조를 결정하는 방법에는 보통 상향식 방식, 하향식 방식, 조합 방식이 있다. 본 논문에서는 하향식 방식을 사용하여 클래스 계층 구조를 결정하였다. 하향식 방식은 우선 도메인의 일반적인 클래스를 정의한 후 클래스를 세부적으로 명세한다. 예를 들면 “상품” 클래스에 대한 정의를 시작한 후 “상품” 클래스의 하위 클래스인 “주류”, “Daily”, “음료” 등을 생성하고 기술한다. 그런 다음 “주류”에 대한 하위 클래스를 다시 생성한다. 이러한 과정을 통해 클래스의 계층 구조를 결정하게 된다.
- <11> 도 1은 온톨로지 클래스의 개념과 관계의 일부를 나타내고 있다.
- <12> 또한, 도 2는 구축하고자 하는 온톨로지 기반의 상품 검색 시스템의 구조이며, 사용자 인터페이스(UI : User Interface)는 일반적으로 사용하고 있는 GUI 환경에서 사용자가 원하는 검색어를 입력하고 검색 결과를 통보받을 수 있도록 시스템을 설계하였다.
- <13> 상품 검색 절차는 먼저 사용자 인터페이스를 통해 사용자가 입력한 검색어를 질의 처리기에 전송한다. 질의 처리기에 입력된 검색어는 불용어 제거를 통해 검색어를 보다 명확하게 재구성한다. 재구성된 검색어에 온톨로지를 통해 의미적인 내용을 부여함으로써 상품 검색시 사용자가 원하는 상품을 정확하게 데이터베이스에서 검색

할 수 있도록 하였다.

- <14> 또한 추론 및 확장 기능을 통해 유사 상품이나 키워드가 정확히 일치하지 않은 상품이다 하더라도 검색 조건에 적합한 상품이라면 모두 검색이 가능하게 된다.
- <15> 또한, 도 3은 찾고자 하는 상품에 대한 기본적인 지식이 있을 경우 사용하는 검색 방법으로서 상품이 해당하는 분류를 선택하고 추가적인 검색 내용을 입력하면 입력된 값들이 파라미터로 전달되고 전달된 값을 가지고 데이터베이스에서 상품을 찾게 된다.
- <16> 위 그림은 분류별 검색 방법을 나타내고 있다.
- <17> ①과 ②에 해당 하는 부분이 온톨로지로 구축된 부분으로서 사용자가 쉽게 사용할 수 있도록 재구성하여 사용자 인터페이스 화면에 디자인 하였다.
- <18> ①의 경우 각각의 상품 정보를 분석하여 구축한 온톨로지를 사용자로 하여금 검색하고자 하는 상품의 대.중.소 분류로 활용할 수 있도록 클래스를 도식화 하였다. 즉, 검색을 원하는 상품에 대한 기본적인 지식을 알고 있을 경우 찾고자 하는 상품의 대?중?소분류를 메뉴에서 선택하게 되면 선택된 각각의 분류와 구축된 온톨로지의 상?하위 클래스를 1:1로 매칭을 시켜 내부적으로는 사용자가 클래스를 직접 선택하여 상품을 검색하는 것과 같은 개념이 된다.
- <19> ②의 경우는 좀 더 구체적이고 세분화 된 검색을 하고자 할 때 추가적으로 입력이 가능한 부분으로서 주로 유의어 목록과 추론기능을 이용하게 된다. 가격, 제조사, 제품 용량, 수량뿐 만 아니라 각 상품의 특징을 추가적으로 입력함으로써 상품 검색을 보다 상세하게 할 수 있도록 해 준다.
- <20> 도 4는 위 그림은 분류별 검색 결과를 보여주고 있으며, 사용자는 검색하고자 하는 상품의 대.중.소분류를 “과자” -> “스낵과자” -> “포테이토계” 순으로 선택하였고, 이렇게 선택된 분류를 가지고 상품 도메인 온톨로지를 이용하여 상품을 검색하게 된다. 이 때 선택된 대?중?소분류와 온톨로지의 상?하위 클래스는 1:1로 서로 매칭되어 검색이 진행된다. 이러한 과정을 통해 검색된 상품들은 모두 “과자” 클래스의 하위 클래스로서 그 중 “스낵과자” 클래스에 속하고 “스낵과자” 클래스의 하위 클래스인 “포테이토계” 클래스에 속하는 상품들이다.
- <21> 이처럼 내부적으로 구축되어 있는 온톨로지의 상?하위 클래스 계층을 사용자가 인터페이스 화면에서 대.중.소분류를 이용하여 직접 선택할 수 있도록 디자인함으로써 사용자로 하여금 찾고자 하는 상품이 속해 있는 온톨로지 클래스를 직접 선택할 수 있는 효과를 주는 것이다.
- <22> 마지막으로 도 5는 검색 결과에 추가적인 검색어 입력을 통해 좀 더 세부적인 상품 검색을 실행한 결과이다. 검색 조건에 나타나 있듯이 가격대와 제조사를 추가 검색어로 입력하여 도 5에 나타난 결과 중 가격대가 “1000원 ~ 2500원” 사이에 해당하는 상품들 중 제조사가 “P&G”에 해당하는 상품들만을 검색하도록 하였다. 이 때 기존의 키워드 검색에서는 추가적으로 입력된 검색어 중 가격과 관련된 검색어는 특별한 문제가 없지만 제조사와 관련된 검색어에서는 문제가 발생하게 된다.

발명의 효과

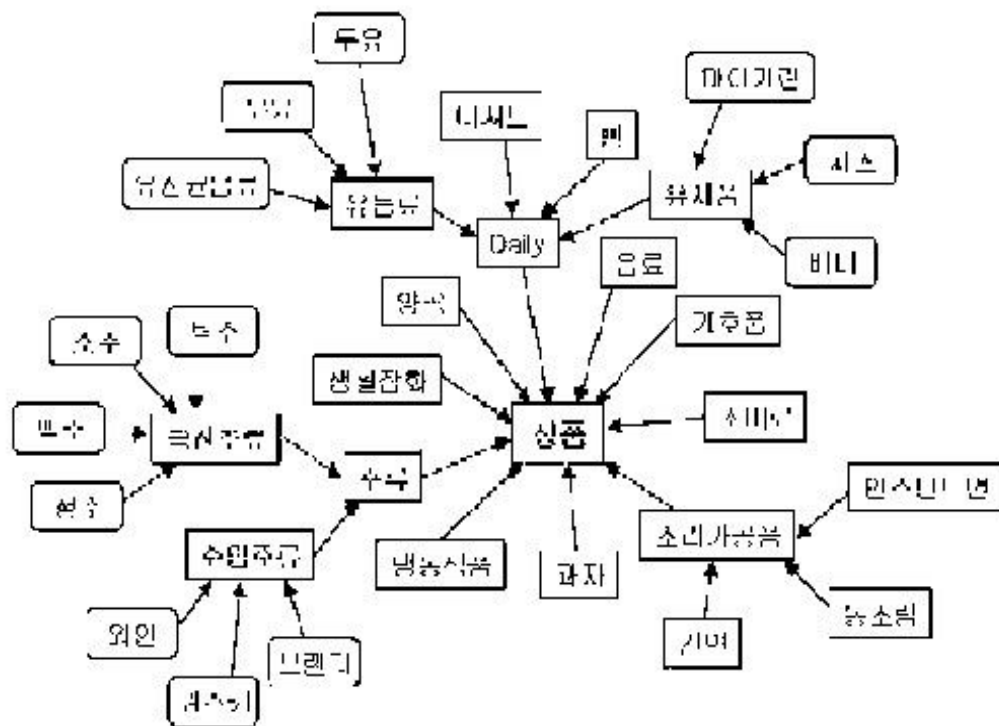
- <23> 키워드 검색에서는 검색어로 입력한 제조사의 영문 명칭인 “P&G”와 한글 명칭인 “피앤지(주)”를 같은 의미로 판단하지 못한다. 그렇기 때문에 “P&G”와 “피앤지(주)”가 동일한 제조사임에도 불구하고 데이터베이스에 저장된 상품의 제조사 키워드가 어떠한 명칭으로 저장되어 있느냐에 따라 서로 다른 검색 결과를 나타내게 된다.

도면의 간단한 설명

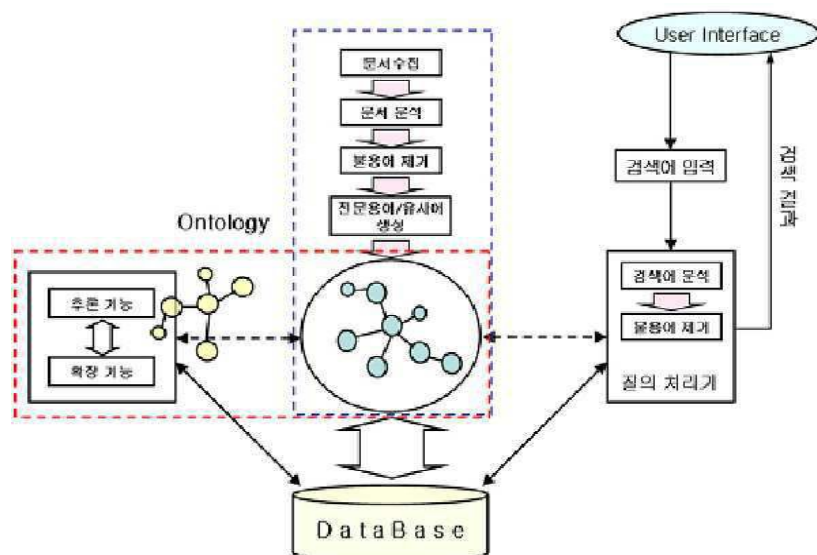
- <1> 도 1은 온톨로지 클래스의 개념과 관계 일부
- <2> 도 2는 온톨로지 기반의 상품 검색 시스템 구조
- <3> 도 3은 분류별 검색 방법
- <4> 도 4는 분류별 검색결과 1
- <5> 도 5는 분류별 검색결과 2

도면

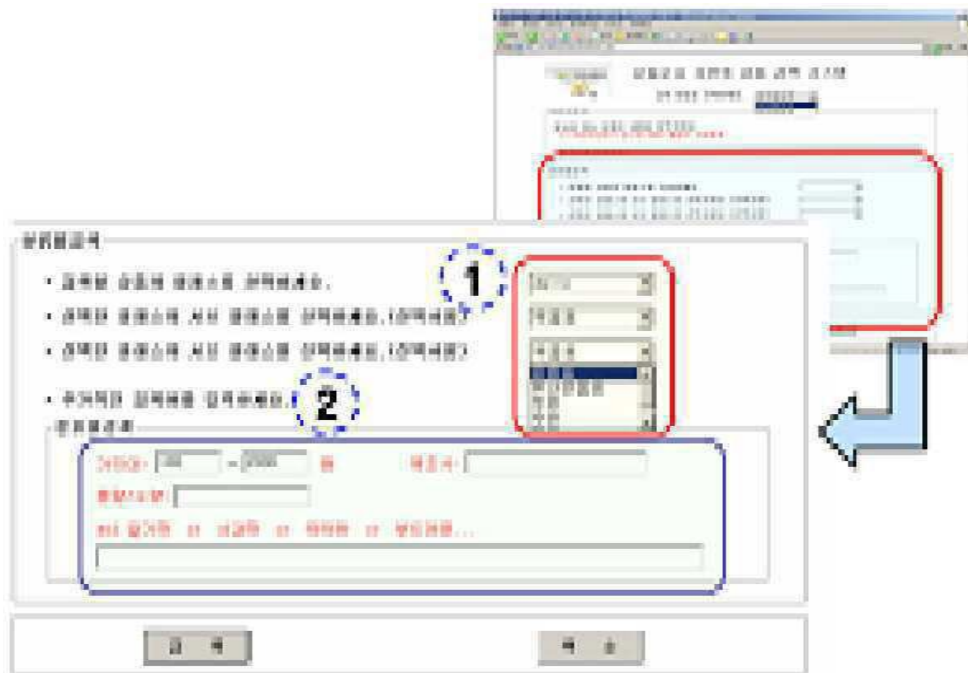
도면1



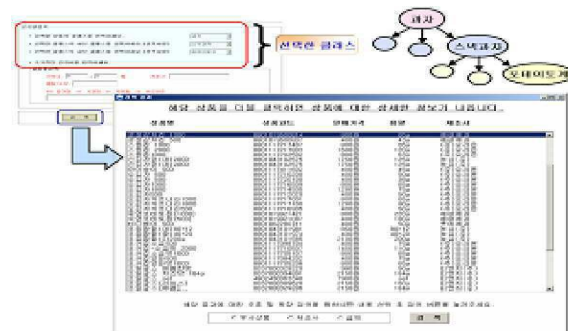
도면2



도면3



도면4



도면5

