



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0010556
(43) 공개일자 2009년01월30일

(51) Int. Cl.⁹

G06F 17/40 (2006.01)

(21) 출원번호 10-2007-0073755

(22) 출원일자 2007년07월23일

심사청구일자 2007년07월23일

(71) 출원인

고려대학교 산학협력단

서울 성북구 안암동5가1 고려대학교 내

(72) 발명자

정인정

경기 성남시 분당구 서현동 시범단지우성아파트
212동 502호

양진혁

서울 중구 신당4동 약수하이츠 동아아파트 116동
307호

(뒷면에 계속)

(74) 대리인

현종철

전체 청구항 수 : 총 10 항

(54) 데이터베이스로부터 온톨로지를 생성하는 방법 및 그 장치

(57) 요약

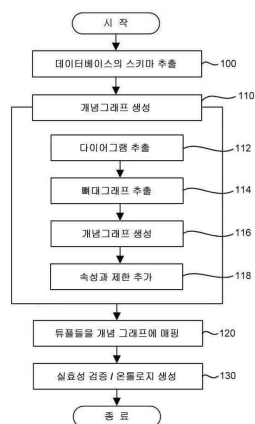
데이터베이스로부터 온톨로지를 생성하는 방법 및 그 장치가 개시된다.

본 발명에 따른 데이터베이스로부터 온톨로지를 생성하는 방법은, 데이터베이스로부터 온톨로지를 생성하는 방법에 있어서, 입력된 데이터베이스의 스키마를 추출하는 단계, 상기 데이터베이스 스키마의 연관 관계를 이용하여 개념그래프(conceptual graph)를 생성하는 단계, 및 상기 데이터베이스의 튜플들을 상기 개념 그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 단계를 포함하며,

본 발명에 따른 데이터베이스로부터 온톨로지를 생성하는 장치는, 데이터베이스를 온톨로지 형태로 생성하는 장치에 있어서, 데이터베이스를 입력으로 하여 그 데이터베이스의 스키마를 추출하는 스키마추출부, 상기 추출된 데이터베이스 스키마에서 각 테이블들 간의 상호 연관 관계에 따라 개념그래프를 생성하는 개념그래프생성부, 및 상기 데이터베이스의 튜플들을 상기 생성된 개념그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 형태결정부를 포함한다.

본 발명에 의하면, 데이터베이스의 튜플들을 상기 개념 그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하여, 상기과 같은 데이터베이스로부터 온톨로지의 생성은 시간과 비용의 소비를 줄이는 많은 연구의 단점인 분석 및 설계과정을 축소하고 키워드(keyword) 색출, 데이터 선별 등과 같이 시간과 비용을 많이 소비하는 부분을 최소화함으로써 시간과 비용의 소비를 감축시키는 효과가 있다.

대 표 도 - 도1



(72) 발명자

손종수

충남 연기군 조치원읍 서창리 170-2

이지홍

서울 동작구 신대방1동 600-44호 1/6

김태석

서울 영등포구 신길6동 우성3차아파트 303동 1302호

특허청구의 범위

청구항 1

입력된 데이터베이스의 스키마를 추출하는 단계;

상기 데이터베이스 스키마의 연관 관계를 이용하여 개념그래프(conceptual graph)를 생성하는 단계; 및

상기 데이터베이스의 튜플들을 상기 개념 그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 단계;를 포함하는 데이터베이스로부터 온톨로지를 생성하는 방법.

청구항 2

제1항에 있어서,

상기 개념그래프를 생성하는 단계는,

상기 데이터베이스 스키마에서 각 테이블 간의 관계인 다이어그램(diagram)을 추출하는 단계;

상기 다이어그램을 기반으로 각 테이블들 간의 상호 연관 관계에 따라 상기 데이터베이스 스키마에 대응하는 뼈대그래프를 추출하는 단계;

상기 추출된 뼈대그래프의 각 테이블들 간의 상하위 및 상호 연관 관계를 추가하여 개념그래프를 생성하는 단계; 및

상기 개념 그래프의 각 노드에, 그 노드에 대응하는 데이터베이스 테이블의 속성(attribute)과 제한(restriction)을 추가하는 단계를 포함하는 것을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 방법.

청구항 3

제1항 또는 제2항에 있어서,

상기 개념그래프를 생성하는 단계는,

상기 데이터베이스 스키마의 테이블의 이름을 상기 개념그래프의 노드 이름으로, 상기 데이터베이스 스키마의 테이블 간의 관계를 상기 개념 그래프의 노드 간의 관계로 설정하는 것을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 방법.

청구항 4

제3항에 있어서,

상기 속성은 상기 데이터베이스 테이블의 속성이 상기 개념그래프의 노드의 속성이며, 상기 제한은, 상기 데이터베이스 테이블의 제한은 상기 개념그래프 노드의 제한이며, 상기 데이터베이스의 튜플은 상기 개념그래프 각 노드 하위의 객체임을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 방법.

청구항 5

제3항에 있어서,

상기 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 단계는,

상기 튜플들을 상기 개념 그래프의 대응되는 각 노드에 연결하여 매핑하는 단계를 포함하는 것을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 방법.

청구항 6

제1항 또는 제2항에 있어서,

상기 개념그래프를 생성하는 단계는,

상기 개념그래프에 매핑된 각 튜플들이 적합하게 매핑되었는지의 여부에 따라 실효성 검증을 하는 단계; 및

상기 실효성 검증이 된 상기 개념 그래프를 온톨로지 기술 언어를 통해 온톨로지 생성하는 단계;를 더 포함하

는 것을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 방법.

청구항 7

데이터베이스를 입력으로 하여 상기 데이터베이스의 스키마를 추출하는 스키마 추출부;

상기 추출된 데이터베이스 스키마에서 각 테이블들 간의 상호 연관 관계에 따라 개념그래프를 생성하는 개념그래프 생성부; 및

상기 데이터베이스의 튜플들을 상기 생성된 개념그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 형태 결정부;를 포함하는 데이터베이스로부터 온톨로지를 생성하는 장치.

청구항 8

제7항에 있어서,

상기 개념그래프 생성부는,

상기 데이터베이스 스키마에서 각 테이블 간의 관계를 추출하는 수단;

상기 테이블 간의 관계를 기반으로 하여 각 테이블을 노드로, 테이블 간의 관계는 노드 간의 관계로 하여 각 노드간의 상하위 및 상호 연관 관계를 생성하여 개념그래프를 생성하는 수단; 및

상기 생성된 개념그래프의 각 노드에 그 노드에 대응하는 데이터베이스 테이블의 속성(attribute)과 제한(restriction)을 추가하는 수단;을 포함하는 것을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 장치.

청구항 9

제8항에 있어서,

상기 형태 결정부는 상기 데이터베이스의 튜플들을 상기 개념 그래프의 대응되는 각 노드에 연결하여 매핑하는 것을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 장치.

청구항 10

제7항 내지 제9항 중 어느 한 항에 있어서,

상기 개념그래프에 매핑된 각 튜플들이 적합하게 매핑되었는지 확인하여 실효성 검증을 하고, 상기 실효성 검증된 그래프를 온톨로지 기술 언어를 이용하여 온톨로지를 생성하는 검증 생성부를 더 포함하는 것을 특징으로 하는 데이터베이스로부터 온톨로지를 생성하는 장치.

명세서

발명의 상세한 설명

기술 분야

- <1> 본 발명은 데이터베이스 분야에 관한 것으로서, 설계과정을 축소하고 키워드(keyword) 색출, 데이터 선별 등과 같은 시간과 비용 소비를 최소화하여 데이터베이스로부터 온톨로지를 생성하는 방법 및 그 장치에 관한 것이다.

배경 기술

- <2> 월드와이드웹(WWW) 기술은 단순 키워드 검색을 통한 데이터의 선별을 기본으로 하고 있다. 이로 인하여 사용자가 원하는 정보 외에 사용자가 원하는 키워드를 내포하고 있는 많은 정보를 같이 선별해 줌으로 선별된 정보를 사용자가 다시 선별해야 하는 문제점을 지니고 있다. 이러한 문제는 단순 키워드(keyword) 검색에서 비롯된 것으로 키워드의 의미를 기계가 이해 할 수 있다면 사용자가 원하는 정보만을 검색 할 수 있다. 이를 가능하게 하기 위하여 정보의 관계가 정의된 지식저장소인 온톨로지를 기반으로 한 시맨틱웹(semantic web)이 대두되게 되었다.
- <3> 온톨로지는 전통적인 생성방법으로 인지, 생성, 평가, 기록의 과정을 거쳐서 생성된다. 전통적인 방법은 많은 부분을 휴리스틱(heuristic)에 의존한다. 휴리스틱을 통한 온톨로지의 생성은 풍부한 어휘와 정형적 의미에 충

실한 온톨로지를 생성 할 수 있는 장점이 있다.

<4> 그러나 많은 부분을 휴리스틱에 의존하기 때문에 많은 시간과 비용을 소비 하는 단점이 있다. 휴리스틱에 의존하는 부분은 시간과 비용의 감축에 저해 요소가 된다. 이런 단점을 극복하기 위하여 제시된 방법으로는 UML을 이용한 도메인 온톨로지 생성방법, 형식적 개념 분석(Formal Concept Analysis)을 이용한 도메인 온톨로지 생성방법, 데이터마이닝 기법(ID3, AOI, association rules mining, clustering)을 이용한 온톨로지 생성방법, 도구들(Protege 2000, etc)을 이용한 온톨로지 생성방법들이 있다.

<5> 그러나, UML을 이용한 도메인 온톨로지 생성방법은 분석, 설계의 과정에 많은 부분을 휴리스틱에 의존하여 많은 시간과 비용을 소비한다. 형식적 개념 분석을 이용한 방법은 문서(document)들에서 유효한 키워드를 색출하는 과정에서 많은 부분을 휴리스틱에 의존하며 데이터마이닝 기법 역시 데이터마이닝 기법을 적용하기 위한 데이터들을 선별하는 과정에 많은 시간과 비용을 소비하는 단점을 지니고 있다. 도구를 사용하는 방법은 휴리스틱에 의존하는 일부분을 도구를 통하여 시간과 비용의 소비를 감축시키나 근본적으로 전통적인 생성방법에 도움을 줄 뿐 휴리스틱에 많은 부분을 의존함으로 시간과 비용을 많이 소비하며, 구성된 온톨로지의 실효성 검증이 되지 않는 문제점이 있었다.

발명의 내용

해결 하고자하는 과제

<6> 본 발명이 해결하고자 하는 첫 번째 과제는, 데이터베이스로부터 온톨로지를 효율적으로 생성할 수 있으며, 실효성 검증도 가능하게 할 수 있는 데이터베이스로부터 온톨로지를 생성하는 방법을 제공하는 것이다.

<7> 본 발명이 해결하고자 하는 두 번째 과제는, 상기 데이터베이스로부터 온톨로지를 생성하는 방법을 이용한 데이터베이스로부터 온톨로지를 생성하는 장치를 제공하는 것이다.

과제 해결수단

<8> 상기 첫 번째 과제를 해결하기 위하여 본 발명은,

<9> 데이터베이스로부터 온톨로지를 생성하는 방법에 있어서, 입력된 데이터베이스의 스키마를 추출하는 단계, 상기 데이터베이스 스키마의 연관 관계를 이용하여 개념그래프(conceptual graph)를 생성하는 단계, 및 상기 데이터베이스의 튜플들을 상기 개념 그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 단계를 포함하는 데이터베이스로부터 온톨로지를 생성하는 방법을 제공한다.

<10> 한편, 상기 개념그래프를 생성하는 단계는, 상기 데이터베이스 스키마에서 각 테이블 간의 관계인 다이어그램(diagram)을 추출하는 단계, 상기 다이어그램을 기반으로 각 테이블들 간의 상호 연관 관계에 따라 상기 데이터베이스 스키마에 대응하는 뼈대그래프를 추출하는 단계, 상기 추출된 뼈대그래프의 각 테이블들 간의 상하위 및 상호 연관 관계를 추가하여 개념그래프를 생성하는 단계, 및 상기 개념 그래프의 각 노드에, 그 노드에 대응하는 데이터베이스 테이블의 속성(attribute)과 제한(restriction)을 추가하는 단계를 포함할 수 있다.

<11> 그리고, 상기 개념그래프를 생성하는 단계는, 상기 데이터베이스 스키마의 테이블의 이름을 상기 개념그래프의 노드 이름으로, 상기 데이터베이스 스키마의 테이블 간의 관계를 상기 개념 그래프의 노드 간의 관계로 설정하는 것이 바람직하다.

<12> 그리고, 상기 속성은 상기 데이터베이스 테이블의 속성이 상기 개념그래프의 노드의 속성이며, 상기 제한은, 상기 데이터베이스 테이블의 제한은 상기 개념그래프 노드의 제한이며, 상기 데이터베이스의 튜플은 상기 개념그래프 각 노드 하위의 객체임을 특징으로 할 수 있다.

<13> 한편, 상기 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 단계는, 상기 튜플들을 상기 개념 그래프의 대응되는 각 노드에 연결하여 매핑하는 단계를 포함할 수 있다.

<14> 그리고, 상기 개념 그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 단계는 상기 개념 그래프에 매핑된 각 튜플들이 적합하게 매핑되었는지 확인하며 실효성 검증을 하는 단계, 및 상기 실효성 검증이 된 상기 개념 그래프를 온톨로지 기술 언어를 통해 온톨로지 생성하는 단계를 더 포함할 수 있다.

<15> 상기 두 번째 과제를 해결하기 위하여 본 발명은,

<16> 데이터베이스를 온톨로지 형태로 생성하는 장치에 있어서, 데이터베이스를 입력으로 하여 그 데이터베이스의 스

키마를 추출하는 스키마추출부, 상기 추출된 데이터베이스 스키마에서 각 테이블들 간의 상호 연관 관계에 따라 개념그래프를 생성하는 개념그래프생성부, 및 상기 데이터베이스의 튜플들을 상기 생성된 개념그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 형태결정부를 포함하는 데이터베이스로부터 온톨로지를 생성하는 장치를 제공한다.

<17> 그리고, 상기 개념그래프 생성부는, 상기 데이터베이스 스키마에서 각 테이블 간의 관계를 추출하는 수단, 상기 테이블 간의 관계를 기반으로 하여 각 테이블을 노드로, 테이블 간의 관계는 노드 간의 관계로 하여 각 노드간의 상하위 및 상호 연관 관계를 생성하여 개념그래프를 생성하는 수단, 및 상기 생성된 개념그래프의 각 노드에 그 노드에 대응하는 데이터베이스 테이블의 속성(attribute)과 제한(restriction)을 추가하는 수단;을 포함할 수 있다.

<18> 상기 형태결정부는 상기 데이터베이스의 튜플들을 상기 개념 그래프의 대응되는 각 노드에 연결하여 매핑하는 것을 특징으로 할 수 있다. 여기서, 상기 개념그래프에 매핑된 각 튜플들이 적합하게 매핑되었는지 확인하며 실효성검증을 하고, 검증된 개념그래프를 온톨로지 기술 언어를 통해 온톨로지로 생성하는 검증/생성부를 더 포함할 수 있다.

효 과

<19> 본 발명에 의하면, 데이터베이스로부터 온톨로지를 생성하는 방법에 있어서, 입력된 데이터베이스의 스키마를 추출하고, 데이터베이스 스키마의 연관 관계를 이용하여 개념그래프(conceptual graph)를 생성하며, 데이터베이스의 튜플들을 상기 개념 그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하여, 상기과 같은 데이터베이스로부터 온톨로지의 생성은 시간과 비용의 소비를 줄이는 많은 연구의 단점인 분석 및 설계과정을 축소하고 키워드(keyword) 색출, 데이터 선별 등과 같이 시간과 비용을 많이 소비하는 부분을 최소화함으로써 시간과 비용의 소비를 감축시키는 효과가 있다.

<20> 본 발명에 따른 방법에서도 휴리스틱에 의존하는 부분이 일부 있으며, 그에 따른 시간과 비용을 소비한다. 그러나 본 발명에 따른 온톨로지의 생성방법은 분석, 설계, 키워드 색출과 같은 사전 작업을 배제함으로써 그 시간과 비용은 종래 온톨로지 생성방법들보다 적게 소비되며, 일부 휴리스틱을 이용하는 것은 상기와 같은 장점에 비해서는 아주 작은 부분이다. 따라서 전체적으로 본 발명과 같이 데이터베이스로부터의 온톨로지의 생성은 종래의 온톨로지 생성방법보다 많은 시간과 비용을 감축하는 효과가 있다.

발명의 실시를 위한 구체적인 내용

<21> 이하, 본 발명의 바람직한 실시예를 첨부도면에 의거하여 상세히 설명하기로 한다.

<22> 그러나, 다음에 예시하는 본 발명의 실시예는 여러 가지 다른 형태로 변형할 수 있으며, 본 발명의 범위가 다음에 상술하는 실시예에 한정되는 것은 아니다. 본 발명의 실시예는 당 업계에서 평균적인 지식을 가진 자에게 본 발명을 더욱 완전하게 설명하기 위하여 제공된다.

<23> 도 1은 본 발명에 따라 데이터베이스로부터 온톨로지를 생성하는 방법의 흐름을 도시한 것이다.

<24> 입력된 데이터베이스의 스키마를 추출하고(100 단계), 데이터베이스 스키마의 연관 관계를 이용하여 개념그래프(conceptual graph)를 생성한다(110 단계).

<25> 데이터베이스 스키마에서 각 테이블 간의 관계인 다이어그램(diagram)을 추출하며(112 단계), 상기 다이어그램을 기반으로 각 테이블들 간의 상호 연관 관계에 따라 상기 데이터베이스 스키마에 대응하는 뼈대그래프를 추출하고(114 단계), 상기 추출된 뼈대그래프의 각 테이블들 간의 상하위 및 상호 연관 관계를 추가하여 개념그래프를 생성하며(116 단계), 개념 그래프의 각 노드에, 그 노드에 대응하는 데이터베이스 테이블의 속성(attribute)과 제한(restriction)을 추가하여(118 단계) 개념 그래프를 완성한다.

<26> 그리고 상기 데이터베이스의 튜플들을 상기 개념 그래프에 매핑하여 온톨로지 기술 언어(Ontology Web Language; OWL)로 기술될 수 있는 형태로 결정한다(120 단계).

<27> 상기과 같이 결정된 개념그래프에 매핑된 각 튜플들이 적합하게 매핑되었는지 확인하며 실효성 검증을 하며, 상기 실효성 검증이 된 개념 그래프를 온톨로지 기술 언어를 통해 온톨로지로 생성한다(130 단계).

<28> 도 2는 본 발명에 따라 데이터베이스로부터 온톨로지를 생성하는 장치의 블록도이다.

<29> 도 2를 참조하면, 본 발명에 따른 데이터베이스로부터 온톨로지를 생성하는 장치는, 데이터베이스(200)를 입력

으로 하여 그 데이터베이스의 스키마를 추출하는 스키마 추출부(210), 상기 추출된 데이터베이스 스키마에서 각 테이블들 간의 상호 연관 관계에 따라 개념그래프를 생성하는 개념그래프 생성부(220), 상기 데이터베이스의 튜플들을 상기 생성된 개념그래프에 매핑하여 온톨로지 기술 언어로 기술될 수 있는 형태로 결정하는 형태 결정부(230) 및 상기 개념그래프에 매핑된 각 튜플들이 적합하게 매핑되었는지 확인하며 실효성검증을 하고, 검증된 개념그래프를 온톨로지 기술 언어를 통해 온톨로지를 생성하는 검증 생성부(240)를 포함한다.

- <30> 한편, 개념그래프 생성부(220)는, 상기 데이터베이스 스키마에서 각 테이블 간의 관계를 추출하는 수단, 상기 테이블 간의 관계를 기반으로 하여 각 테이블을 노드로, 테이블 간의 관계는 노드 간의 관계로 하여 각 노드간의 상하위 및 상호 연관 관계를 생성하여 개념그래프를 생성하는 수단 및 상기 생성된 개념그래프의 각 노드에 그 노드에 대응하는 데이터베이스 테이블의 속성(attribute)과 제한(restriction)을 추가하는 수단을 포함한다.
- <31> 데이터베이스는 분석, 설계 단계를 거쳐 데이터 모델을 생성하고 데이터 모델을 기반으로 구현된다. 따라서 데이터 모델로부터 구현된 데이터베이스는 그 데이터 모델들의 집합인, 예를 들면 도메인을 잘 나타내는 데이터를 저장하는 저장소로 그 스키마는 도메인의 추상적 구성을 연관 관계를 통하여 잘 나타내고 있다.
- <32> 스키마는 데이터베이스 개체의 특성을 나타내는 속성과 속성들의 집합으로 이루어진 개체, 개체들 사이에 존재하는 관계에 대한 정의와 이들이 유지해야 할 제약조건들을 기술한 것으로 데이터베이스에 어떤 구조로 데이터가 저장되는가를 나타낸다. 다시 말해 데이터베이스 구조를 스키마라고 할 수 있다.
- <33> 또한 데이터베이스 튜플(tuple)들은 스키마에 따르는 정보들로 도메인을 잘 나타내는 데이터들로 이루어져 있다. 이러한 구성은 온톨로지를 생성하기 위한 좋은 기초 자료가 된다. 튜플은 릴레이션을 구성하는 각각의 행 속성의 모임으로 구성되는, 파일 구조에서는 레코드와 같은 의미로 사용된다.
- <34> 예를 들어, 어떤 도메인에 대한 데이터 모델을 기반으로 구현된 데이터베이스를 가정한다. 그 데이터베이스는 그 도메인을 잘 나타내는 정보를 포함하고 있으며, 그 정보들은 어떤 구체적인 규칙 하에서 데이터베이스로 구성된 것이며, 상기에 설명된 것과 같이 그 도메인의 추상적인 구성은 그 데이터베이스의 스키마의 연관 관계의 의해 나타내진다.
- <35> 스키마 추출부(210)는 이와 같은 데이터베이스(200)를 입력으로 하여, 그 데이터베이스의 스키마를 추출한다. 그와 같이 추출된 스키마의 예가 도 3에 표시되어 있다. 데이터베이스를 입력으로 하여 도 3과 같은 스키마를 추출하는 것은 본 발명이 속한 기술 분야의 통상의 지식을 가진 자에게는 자명한 것이므로 별도의 설명은 생략한다.
- <36> 개념그래프 생성부(220)는 데이터베이스 스키마를 이용하여 개념그래프의 모체가 되는 뼈대그래프를 생성한다. 뼈대그래프는 개념그래프의 기본골격을 나타내는 그래프로 개념그래프의 관계 정의가 없는 그래프이다. 즉, 본 발명에서 뼈대그래프라고 칭하고 있는 그래프는 각 노드 간의 관계가 명확하지 않은 트리(tree) 또는 그래프(graph)와 같은 단계의 그래프이다.
- <37> 개념그래프 생성부(220)는 생성된 뼈대그래프를 기반으로 데이터베이스 스키마의 연관 관계를 이용하여 개념그래프로 확장한다.
- <38> 이를 위해 데이터베이스 스키마에서 각 테이블 간의 관계인 다이어그램(diagram)을 추출하며(112 단계), 상기 다이어그램을 기반으로 각 테이블들 간의 상호 연관 관계에 따라 상기 데이터베이스 스키마에 대응하는 뼈대그래프를 추출한다(114 단계). 도 3에서 하이클래스(HighClass)(380)부터 딜리버리(Delivery)(380)의 테이블들이 도 4의 개념 그래프에서 참조번호 400 내지 480 까지의 노드로 표시될 수 있다. 이와 같이 노드만이 표시된(즉, 도 4에서 노드간의 연결 관계가 없는 모습) 모습이 뼈대 그래프라고 할 수 있다.
- <39> 그리고, 데이터베이스 스키마의 연관 관계 즉, 추출된 뼈대그래프의 각 테이블들 간의 상하위 및 상호 연관 관계를 추가하여 개념그래프를 생성한다(116 단계). 도 3에서의 각 테이블들 간의 상하위의 관계와 연관 관계에 따라 도 4에서 노드들간의 관계 혹은 연관 관계가 실선을 통해 표시되어 있다.
- <40> 그리고 개념그래프 생성부(220)는 개념 그래프의 각 노드에, 그 노드에 대응하는 데이터베이스 테이블의 속성(attribute)과 제한(restriction)을 추가한다(118 단계).
- <41> 결국 뼈대그래프는 데이터베이스의 테이블들의 상호연관을 이용하여 완성하는 것이며, 완성된 뼈대그래프에 관계 정의를 하고 제한과 속성을 부가함으로써 개념그래프로 확장되는 것이다.
- <42> 형태 결정부(230)는 상기 데이터베이스의 튜플들을 110 단계에서 생성된 개념 그래프에 매핑하여 온톨로지 기술

언어로 기술될 수 있는 형태로 결정한다(120 단계).

- <43> 표 1은 상기에 설명된 데이터베이스 스키마와 온톨로지 매핑 관계를, 그리고 표 2는 데이터베이스 각 구성과 온톨로지 매핑 관계를 대비한 것이다.

표 1

<44>

데이터베이스 스키마	온톨로지
테이블 이름	개념그래프의 노드 이름
테이블 관계	개념그래프의 노드 관계

표 2

<45>

데이터베이스 구성	온톨로지
테이블 속성	각 노드의 속성
테이블 각 속성의 제한	각 노드의 제한
튜플	각 노드 하위의 개체

- <46> 즉, 개념그래프를 생성할 때에, 데이터베이스 스키마의 테이블의 이름을 상기 개념그래프의 노드 이름으로, 데이터베이스 스키마의 테이블 간의 관계를 상기 개념 그래프의 노드 간의 관계로 설정하며, 상기 데이터베이스 테이블의 속성이 상기 개념그래프의 노드의 속성이며, 상기 데이터베이스 테이블의 제한은 상기 개념그래프 노드의 제한이며, 상기 데이터베이스의 튜플은 상기 개념그래프 각 노드 하위의 객체가 된다.

- <47> 그 결과 도 3의 데이터베이스 스키마는 도 4와 같은 개념 그래프의 형태로 변환된다.

- <48> 검증 생성부(240)는 개념그래프에 매핑된 각 튜플들이 적합하게 매핑되었는지 확인하며 실효성 검증을 한다(130 단계).

- <49> 매핑이 적합하게 매핑되었는지의 확인은 휴리스틱에 의존하여 수행하며, OWL validator 프로그램을 이용하여 실효성 검증을 할 수 있다. 도 5가 온톨로지 검증 결과의 예를 도시한 것이다.

- <50> 그리고 도 4와 같은 온톨로지 기술 언어로 기술될 수 있는 형태인 개념 그래프를 온톨로지 기술언어인 OWL을 이용하여 코딩하여 100 단계에서의 데이터베이스에 대응하는 온톨로지를 생성한다(130 단계). 표 3이 도 3의 데이터베이스 스키마를 거쳐서 생성된 도 4의 개념 그래프에 대응하여 코딩된 혹은 생성된 온톨로지이다.

표 3

<51>

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
  <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
]>
<rdf:RDF
  xmlns      = "http://localhost/Goods#"
  xmlns:owl  = "http://www.w3.org/2002/07/owl#"
  xmlns:rdf  = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd  = "http://www.w3.org/2001/XMLSchema#">
  <owl:Class rdf:ID="Goods">
  <owl:Class rdf:ID="Goods">
  <rdfs:subClassOf rdf:resource="#Low class"/>
  </owl:Class>
  <owl:ObjectProperty rdf:ID="Part of">
  <rdfs:domain rdf:resource="#Buyinfo"/>
  <rdfs:range rdf:resource="#members"/>
  </owl:ObjectProperty>
  .....
  <owl:Class rdf:ID="Middle class">
  <rdfs:subClassOf rdf:resource="#High class"/>
  <owl:Restriction>
  <owl:onProperty rdf:resource="#ClassOf">
  </owl:Restriction>
  </owl:Class>
  .....
  <owl:ObjectProperty rdf:resource="Account">
  <owl:Restriction>
  .....
  </owl:Restriction>
  </owl:ObjectProperty>
  .....
```

<52>

이와 같은 코딩은 도 4와 같은 그래프에서 각 노드 간의 연결 관계 및 제한/속성 그리고 각 노드의 하위 개체인 튜플에 대한 정보를 바탕으로 OWL을 이용하여 자동적으로 생성되도록 적절한 프로그램을 통해 실시될 수 있을 것이다.

<53>

결국 도 3의 스키마로 나타내지는 데이터베이스에 대응하는 도메인이 표 3과 같은 오톨로지로 표시된 것이다.

<54>

상기와 같은 데이터베이스로부터 온톨로지의 생성은 시간과 비용의 소비를 줄이는 많은 연구의 단점인 분석 및 설계과정을 축소하고 키워드(keyword) 색출, 데이터 선별 등과 같이 시간과 비용을 많이 소비하는 부분을 최소화함으로써 시간과 비용의 소비를 감축시키는 효과가 있다.

<55>

본 발명에 따른 방법에서도 휴리스틱에 의존하는 부분이 일부 있으며, 그에 따른 시간과 비용을 소비한다. 그러나 본 발명에 따른 온톨로지의 생성방법은 분석, 설계, 키워드 색출과 같은 사전 작업을 배제함으로써 그 시간과 비용은 종래 온톨로지 생성방법들보다 적게 소비되며, 일부 휴리스틱을 이용하는 것은 상기와 같은 장점에 비해서는 아주 작은 부분이다. 따라서 전체적으로 본 발명과 같이 데이터베이스로부터의 온톨로지의 생성은 종래의 온톨로지 생성방법보다 많은 시간과 비용을 감축하는 효과를 갖는다.

<56>

본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 본 개시된 실시예 들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 상기의 설명에 포함된 예들은 본 발명에 대한 이해를 위

해 도입된 것이며, 이 예들은 본 발명의 사상과 범위를 한정하지 않는다. 상기의 예들 외에도 본 발명에 따른 다양한 실시 태양이 가능하다는 것은, 본 발명이 속한 기술 분야에 통상의 지식을 가진 사람에게는 자명할 것이다. 본 발명의 범위는 전술한 설명이 아니라 청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

<57> 또한 본 발명에 따른 상기의 각 단계는 일반적인 프로그래밍 기법을 이용하여 소프트웨어적으로 또는 하드웨어적으로 다양하게 구현할 수 있다는 것은 이 분야에 통상의 기술을 가진 자라면 용이하게 알 수 있는 것이다.

<58> 그리고 본 발명의 일부 단계들은, 또한, 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, CD-RW, 자기 테이프, 플로피디스크, HDD, 광 디스크, 광자기 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드로 저장되고 실행될 수 있다.

도면의 간단한 설명

<59> 도 1은 본 발명에 따른 데이터베이스로부터 온톨로지를 생성하는 방법의 흐름도를 도시한 것이다.

<60> 도 2는 본 발명에 따른 데이터베이스로부터 온톨로지를 생성하는 장치의 블록도를 도시한 것이다.

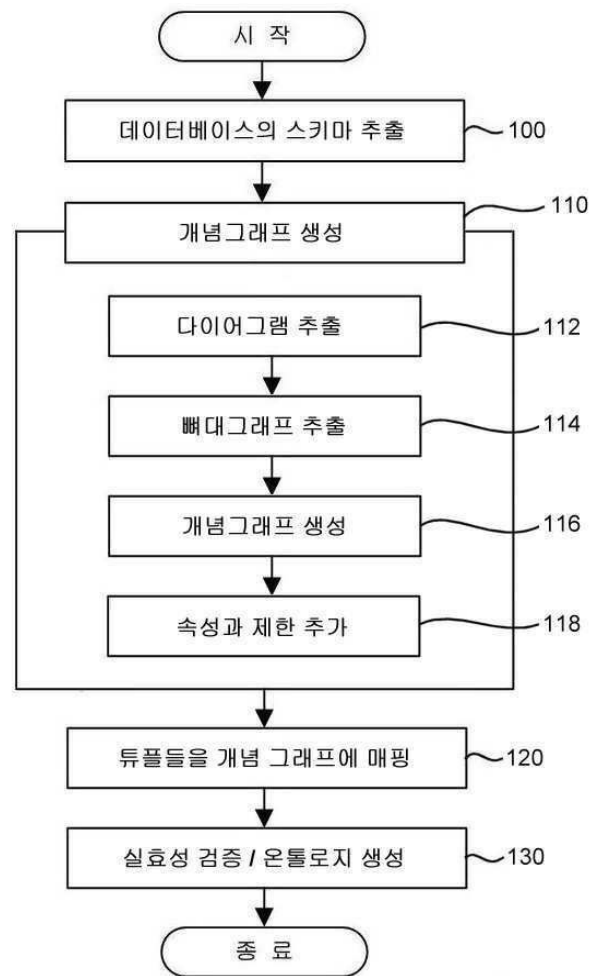
<61> 도 3은 본 발명에 따른 데이터베이스로부터 온톨로지를 생성하기 위한 데이터베이스 스키마의 예를 도시한 것이다.

<62> 도 4는 도 3의 스키마로부터 생성된 개념 그래프의 일 예를 도시한 것이다.

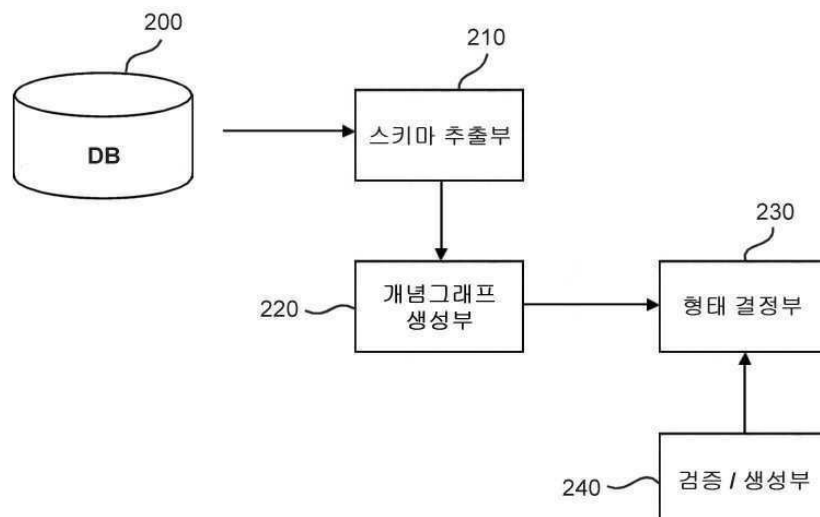
<63> 도 5는 온톨로지 검증 결과의 예를 도시한 것이다.

도면

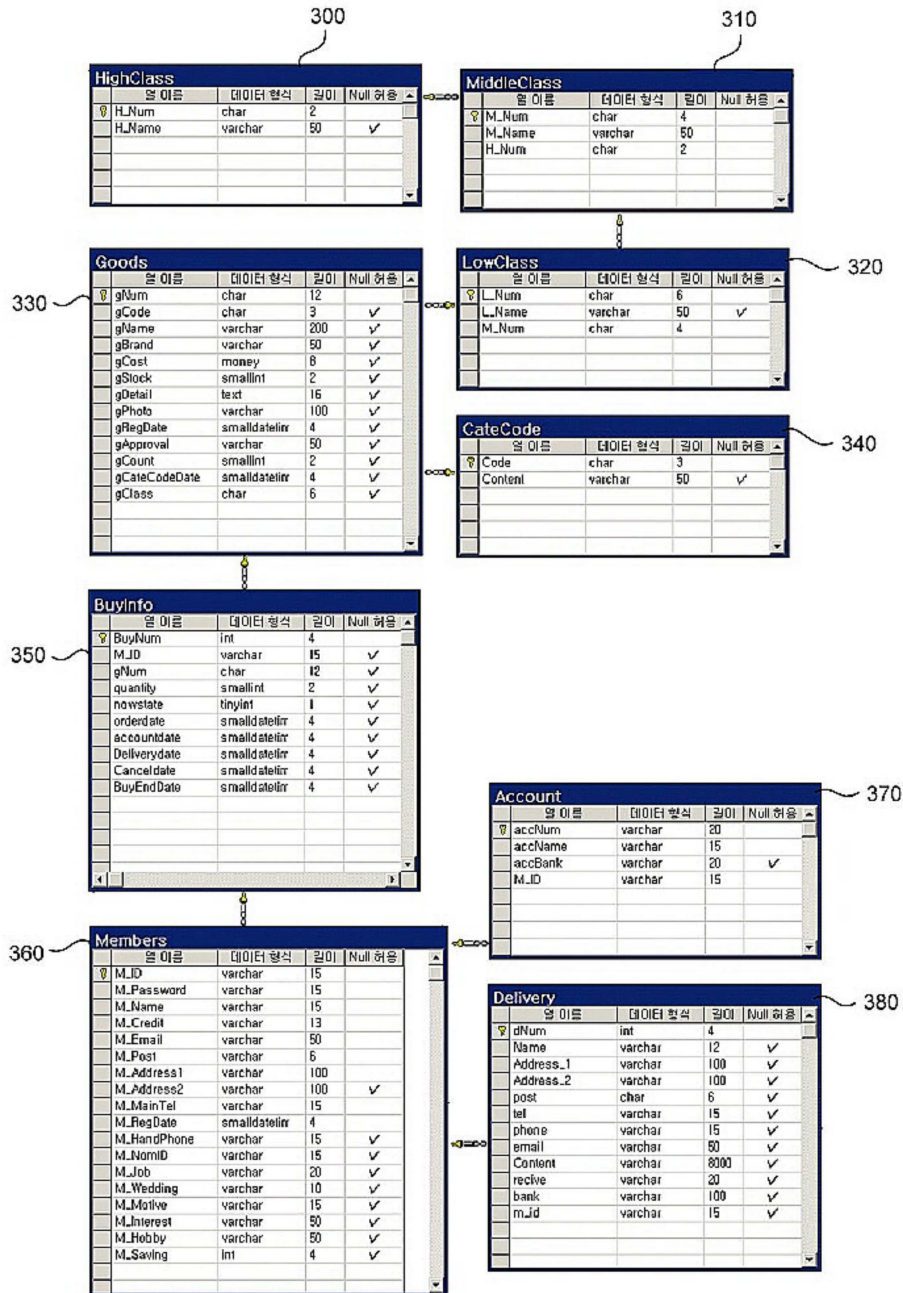
도면1



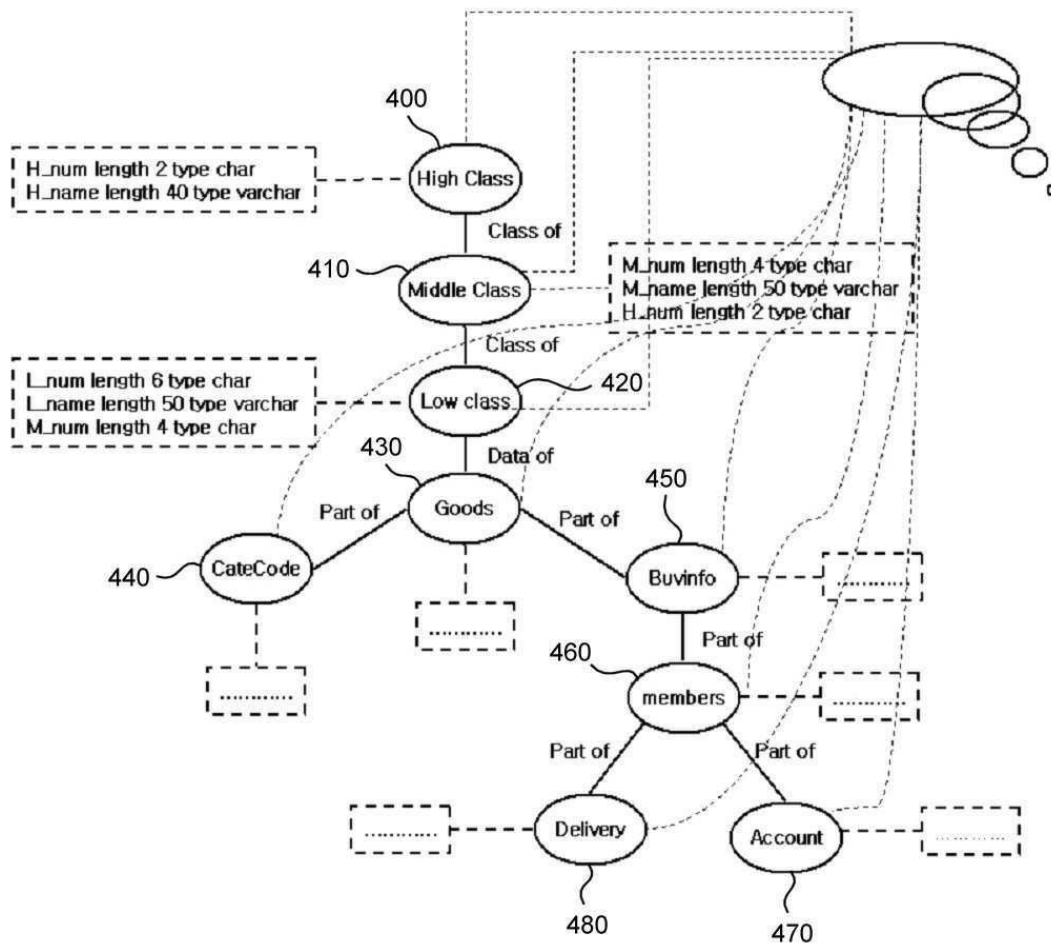
도면2



도면3



도면4



도면5

