



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년01월20일
 (11) 등록번호 10-1009924
 (24) 등록일자 2011년01월14일

(51) Int. Cl.
G06F 17/30 (2006.01) *G06F 17/20* (2006.01)
 (21) 출원번호 10-2010-0060280
 (22) 출원일자 2010년06월25일
 심사청구일자 2010년06월25일
 (56) 선행기술조사문헌
 KR100820746 B1*
 KR1020080019439 A*
 KR100525072 B1*
 KR1020050046973 A*
 *는 심사관에 의하여 인용된 문헌

(73) 특허권자
한국과학기술정보연구원
 대전광역시 유성구 어은동 52-11
 (72) 발명자
정한민
 대전광역시 유성구 어은동 99 한빛아파트 129동 305호
김평
 대전광역시 유성구 반석동 양지마을 204동 1304호
 (뒷면에 계속)
 (74) 대리인
특허법인화우

전체 청구항 수 : 총 25 항

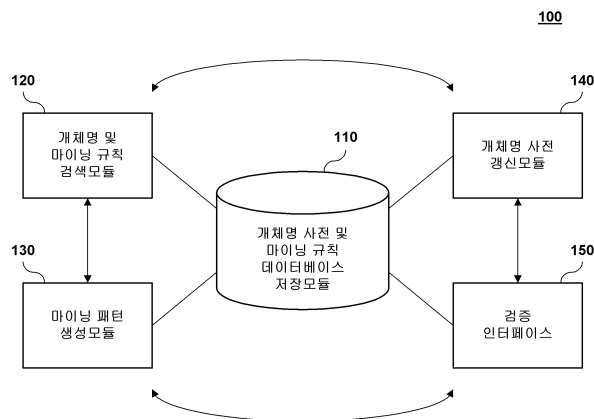
심사관 : 박지은

(54) 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 또는 마이닝 규칙 데이터베이스 갱신 장치 및 방법

(57) 요약

본 발명은 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 또는 마이닝 규칙 데이터베이스 갱신 장치 및 방법에 관한 것으로, 온톨로지 스키마와 결합되고 용어의 개체명이 정의된 개체명 사전, 상기 온톨로지 스키마와 결합되고 관계명에 대응되는 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈, 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 마이닝 패턴 생성모듈, 상기 인식된 용어와 상기 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 상기 개체명 사전과 상기 마이닝 규칙 데이터베이스에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈, 및 상기 인식된 용어에 대응되는 개체명이 상기 개체명 사전으로부터 검색되지 않고 상기 마이닝 패턴에 대응되는 마이닝 규칙이 상기 마이닝 규칙 데이터베이스로부터 검색되는 경우, 상기 마이닝 규칙을 이용하여 용어의 개체명을 추정하고, 상기 추정된 용어의 개체명을 사용자의 선택에 따라 개체명 사전에 저장하는 개체명 사전 갱신모듈을 포함한다.

대표도 - 도1



(72) 발명자

이승우

대전광역시 서구 내동 맑은아침아파트 119동 1001호

이미경

대전광역시 유성구 도룡동 383-2 과기원교수아파트 2동 101호

서동민

충청북도 청주시 흥덕구 복대동 주은아파트 101-1509호

성원경

대전광역시 유성구 전민동 엑스포아파트 207-704

특허청구의 범위

청구항 1

온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 상기 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈;

텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 마이닝 패턴 생성모듈;

상기 인식된 용어와 상기 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 상기 개체명 사전과 상기 마이닝 규칙 데이터베이스에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈; 및

상기 인식된 용어에 대응되는 개체명이 상기 개체명 사전으로부터 검색되지 않고 상기 마이닝 패턴에 대응되는 마이닝 규칙이 상기 마이닝 규칙 데이터베이스로부터 검색되는 경우, 상기 마이닝 규칙을 이용하여 용어의 개체명을 추정하고, 상기 추정된 용어의 개체명을 사용자의 선택에 따라 개체명 사전에 저장하는 개체명 사전 갱신모듈;을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 2

온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 상기 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈;

입력된 마이닝 패턴과 상기 마이닝 패턴에 포함되는 용어를 이용하여 대응되는 마이닝 규칙과 개체명을 상기 마이닝 규칙 데이터베이스와 상기 개체명 사전에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈; 및

상기 용어에 대응되는 개체명이 상기 개체명 사전으로부터 검색되지 않고 상기 마이닝 패턴에 대응되는 마이닝 규칙이 상기 마이닝 규칙 데이터베이스로부터 검색되는 경우, 상기 마이닝 규칙을 이용하여 용어의 개체명을 추정하고, 상기 추정된 용어의 개체명을 사용자의 선택에 따라 개체명 사전에 저장하는 개체명 사전 갱신모듈;을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 3

제1항 또는 제2항에 있어서,

상기 개체명 사전 갱신모듈은, 상기 용어와 대응되는 개체명의 특성을 표시하는 검증 인터페이스와 연결되며,

상기 검증 인터페이스를 통해 사용자로부터 상기 용어와 대응되는 개체명의 특성을 확인한 후, 상기 확인 결과를 상기 개체명 사전에 저장하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 4

제1항 또는 제2항에 있어서,

상기 개체명 및 마이닝 규칙 검색모듈은,

상기 용어의 개체명을 상기 개체명 사전으로부터 검색하는 개체명 검색부; 및

상기 마이닝 패턴에 대응되는 마이닝 규칙을 상기 마이닝 규칙 데이터베이스로부터 검색하는 마이닝 규칙 검색부;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 5

제1항 또는 제2항에 있어서,

상기 개체명 사전 갱신모듈은,

검색 결과, 상기 용어가 상기 개체명 사전으로부터 검색되었는지 판단하는 검색결과 판단부;

판단 결과, 상기 용어가 상기 개체명 사전으로부터 검색되지 않고 상기 마이닝 패턴이 상기 마이닝 규칙 데이터베이스로부터 검색되는 경우 상기 마이닝 규칙을 이용하여 상기 용어의 개체명을 추정하고 사용자의 선택에 따라 개체명 사전에 저장하는 개체명 사전 갱신부;

상기 검색결과 판단부의 판단 결과에 따라 상기 개체명 사전 갱신부에 개체명 특성을 정의하여 저장되도록 제어하는 제어부;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 6

제5항에 있어서,

상기 개체명 사전 갱신부는, 상기 개체명의 특성을 상기 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 상기 개체명으로 분류되는 용어, 상기 용어의 식별자, 대표 용어와 상기 대표 용어의 식별자로 구성된 전거 데이터를 포함하여 정의하고 하나의 포맷으로 연결하여 저장하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 7

제1항 또는 제2항에 있어서,

상기 개체명 사전은, 상기 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 상기 개체명으로 분류되는 용어, 상기 용어의 식별자, 대표 용어와 상기 대표 용어의 식별자로 구성된 전거 데이터를 하나의 포맷으로 연결하여 저장하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 8

제1항에 있어서,

상기 마이닝 패턴 생성모듈은,

상기 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 상기 인식된 용어를 상기 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 개체명 인식부; 및

상기 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 패턴 변환부;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치.

청구항 9

온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 상기 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈;

텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 마이닝 패턴 생성모듈;

상기 인식된 용어와 상기 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 상기 개체명 사전과 상기 마이닝 규칙 데이터베이스에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈; 및

상기 마이닝 패턴에 대응되는 마이닝 규칙이 상기 마이닝 규칙 데이터베이스로부터 검색되지 않고 상기 인식된 용어에 대응되는 개체명이 상기 개체명 사전으로부터 검색되는 경우, 상기 인식된 용어의 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 대응되는 마이닝 규칙을 생성한 후, 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장하는 마이닝 규칙 데이터베이스 갱신모듈;을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치.

청구항 10

온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 상기 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈;

입력된 마이닝 패턴과 상기 마이닝 패턴에 포함되는 용어를 이용하여 대응되는 마이닝 규칙과 개체명을 상기 마이닝 규칙 데이터베이스와 상기 개체명 사전에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈; 및

상기 마이닝 패턴에 대응되는 마이닝 규칙이 상기 마이닝 규칙 데이터베이스로부터 검색되지 않고 상기 용어에 대응되는 개체명이 상기 개체명 사전으로부터 검색되는 경우, 상기 용어의 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 대응되는 마이닝 규칙을 생성한 후, 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장하는 마이닝 규칙 데이터베이스 갱신모듈;을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치.

청구항 11

제9항 또는 제10항에 있어서,

상기 마이닝 규칙 데이터베이스 갱신모듈은, 상기 마이닝 패턴에 대응되는 마이닝 규칙을 표시하는 검증 인터페이스와 연결되며,

상기 검증 인터페이스를 통해 사용자로부터 상기 마이닝 패턴과 대응되는 마이닝 규칙을 확인한 후, 상기 확인 결과를 상기 마이닝 규칙 데이터베이스에 저장하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치.

청구항 12

제9항 또는 제10항에 있어서,

상기 개체명 및 마이닝 규칙 검색모듈은,

상기 용어의 개체명을 상기 개체명 사전으로부터 검색하는 개체명 검색부; 및

상기 마이닝 패턴에 대응되는 마이닝 규칙을 상기 마이닝 규칙 데이터베이스로부터 검색하는 마이닝 규칙 검색부;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치.

청구항 13

제9항 또는 제10항에 있어서,

상기 마이닝 규칙 데이터베이스 갱신모듈은,

검색 결과, 상기 마이닝 패턴이 상기 마이닝 규칙 데이터베이스로부터 검색되었는지 판단하는 검색결과 판단부;
 판단 결과, 상기 마이닝 패턴이 검색되지 않고 상기 용어가 상기 개체명 사전으로부터 검색되는 경우 검색된 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 상기 관계명에 대응되는 마이닝 규칙을 생성한 후 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장하는 마이닝 규칙 데이터베이스 갱신부; 및
 상기 검색결과 판단부의 판단 결과에 따라 상기 마이닝 규칙 데이터베이스 갱신부에 마이닝 규칙을 정의하여 저장되도록 제어하는 제어부;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치.

청구항 14

제13항에 있어서,
 상기 마이닝 규칙 데이터베이스 갱신부는, 상기 마이닝 규칙을 상기 온톨로지 스키마의 관계명(Property), 상기 관계명에 대응되는 하나의 마이닝 패턴, 상기 마이닝 패턴과 대응되는 하나 이상의 RDF 트리플을 포함하여 정의하고 하나의 포맷으로 연결하여 저장하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치.

청구항 15

제9항에 있어서,
 상기 마이닝 패턴 생성모듈은,
 상기 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 상기 인식된 용어를 상기 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 개체명 인식부; 및
 상기 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 마이닝 패턴 변환부;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치.

청구항 16

미리 저장된 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 상기 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 저장모듈, 마이닝패턴 생성모듈, 개체명 및 마이닝규칙 검색모듈, 개체명사전 갱신모듈을 포함하는 장치에서의 개체명 사전 갱신 방법으로서,
 상기 마이닝패턴 생성모듈이 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계;
 상기 개체명 및 마이닝규칙 검색모듈이 상기 인식된 용어와 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 상기 개체명 사전과 마이닝 규칙 데이터베이스로부터 각각 검색하는 단계;
 상기 개체명 사전 갱신모듈이 상기 검색결과로부터 상기 인식된 용어 또는 마이닝 패턴과 일치하는 용어 또는 마이닝 패턴이 검색되었는지 판단하는 단계;
 상기 판단 결과, 일치하는 마이닝 패턴이 상기 마이닝 규칙 데이터베이스로부터 검색되었으나 일치하는 용어가 상기 개체명 사전으로부터 검색되지 않은 경우 상기 개체명 사전 갱신모듈이 상기 마이닝 패턴에 대응되는 마이닝 규칙과 온톨로지 스키마를 이용하여 용어의 개체명을 추정하는 단계; 및
 상기 개체명 사전 갱신모듈이 상기 추정된 개체명을 사용자의 선택에 따라 상기 개체명 사전에 저장하는 단계;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법.

청구항 17

제16항에 있어서,

상기 추정된 개체명을 사용자의 선택에 따라 상기 개체명 사전에 저장하는 단계는,

상기 인식된 용어와 상기 추정된 개체명을 검증 인터페이스에 표시하여 사용자로부터 검증을 받는 것을 특징으로 하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법.

청구항 18

제16항에 있어서,

상기 개체명 사전은, 상기 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 상기 개체명으로 분류되는 용어, 상기 용어의 식별자, 대표 용어와 상기 대표 용어의 식별자로 구성된 전거 데이터를 하나의 포맷으로 연결하여 저장하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법.

청구항 19

제16항에 있어서,

텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계는,

상기 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 상기 인식된 용어를 상기 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 단계; 및

상기 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 단계;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법.

청구항 20

제16항 내지 제19항에 있어서,

상기 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법을 수행하는 프로그램이 기록되고 전자 장치에서 판독 가능한 기록 매체.

청구항 21

저장된 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 상기 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 저장모듈, 마이닝패턴 생성모듈, 개체명 및 마이닝규칙 검색모듈, 마이닝규칙 데이터베이스 갱신모듈을 포함하는 장치에서의 마이닝규칙 데이터베이스 갱신 방법으로서,

상기 마이닝패턴 생성모듈이 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계;

상기 개체명 및 마이닝규칙 검색모듈이 상기 인식된 용어와 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 상기 개체명 사전과 마이닝 규칙 데이터베이스로부터 각각 검색하는 단계;

상기 마이닝규칙 데이터베이스 갱신모듈이 상기 검색 결과로부터 상기 인식된 용어 또는 마이닝 패턴과 일치하는 용어 또는 마이닝 패턴이 검색되었는지 판단하는 단계;

상기 판단 결과, 일치하는 용어가 상기 개체명 사전에서 검색되었으나 일치하는 마이닝 패턴이 상기 마이닝 규칙 데이터베이스로부터 검색되지 않은 경우 상기 마이닝규칙 데이터베이스 갱신모듈이 상기 용어의 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 대응되는 마이닝 규칙을 생성하는 단계; 및

상기 마이닝규칙 데이터베이스 갱신모듈이 상기 추정된 관계명과 마이닝 규칙을 사용자의 선택에 따라 상기 마이닝 규칙 데이터베이스에 저장하는 단계;

를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법.

청구항 22

제21항에 있어서,

상기 추정된 관계명과 마이닝 규칙을 사용자의 선택에 따라 상기 마이닝 규칙 데이터베이스에 저장하는 단계는,

상기 추정된 관계명, 마이닝 패턴 및 마이닝 규칙을 검증 인터페이스에 표시하여 사용자로부터 검증을 받는 것을 특징으로 하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법.

청구항 23

제21항에 있어서,

상기 마이닝 규칙 데이터베이스는, 상기 마이닝 규칙을 상기 온톨로지 스키마의 관계명(Property), 상기 관계명에 대응되는 하나의 마이닝 패턴, 상기 마이닝 패턴과 대응되는 하나 이상의 RDF 트리플을 포함하여 정의하고 하나의 포맷으로 연결하여 저장하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법.

청구항 24

제21항에 있어서,

텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계는,

상기 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 상기 인식된 용어를 상기 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 단계; 및

상기 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 단계;를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법.

청구항 25

제21항 내지 제24항에 있어서,

상기 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법을 수행하는 프로그램이 기록되고 전자 장치에서 판독 가능한 기록 매체.

명세서

기술분야

[0001] 본 발명은 미리 저장된 개체명 사전 및 마이닝 규칙 데이터베이스를 이용하여 온톨로지 스키마와 결합된 개체명 또는 마이닝 규칙을 검색하고 일치하는 개체명이 검색되지 않는 경우 마이닝 규칙을 이용하여 개체명을 추정하고, 마이닝 규칙이 검색되지 않는 경우 개체명을 이용하여 마이닝 규칙을 추정하여 개체명 사전 또는 마이닝 규칙 데이터베이스를 갱신하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 또는

마이닝 규칙 데이터베이스 갱신 장치 및 방법에 관한 것이다.

배경 기술

- [0002] 최근 시멘틱 웹(Semantic Web)의 연구와 관련하여 온톨로지(Ontology)를 이용한 데이터 모델이 주목받고 있다. 온톨로지는 특정한 영역 내에 있는 실체 및 상호 작용의 작업 모델을 의미한다. 즉, 온톨로지는 특정한 도메인(Domain) 내의 지식을 개념화하고 이를 명세화한 것으로 그 도메인 내에 사용되는 개념들의 관계성을 가진 네트워크 또는 그래프로 정의할 수 있다.
- [0003] 온톨로지는 특정 분야의 관련된 문서로부터 온톨로지에 구축할 온톨로지의 구성 요소 즉, 개념, 개념의 속성, 개념 간 연결 관계 등에 관한 지식을 얻어 개념과 속성을 정의하고 개념과 개념의 관계를 설정하여 구축된다.
- [0004] 또한, 웹 자원에 대한 의미성을 부여하기 위하여 자원과 속성, 속성값 등의 3차원 구조인 리소스 기술 프레임워크(RDF: Resource Description Framework)로 나타낼 수 있다.
- [0005] RDF는 리소스(Resource), 특성(Property), 서술문(Statement)의 개념으로 구성되고, 입력된 문서의 모든 사물(Thing)은 리소스로 표현되고, 각 리소스의 특성이나 다른 리소스의 관계 등의 특성을 나타낼 수 있다.
- [0006] 입력된 문서로부터 개체명을 검출하고 마이닝 규칙(Mining Rule)을 이용하여 대응되는 RDF 트리플로 변환하여 RDF 네트워크를 구축할 수 있다.
- [0007] 개체명(Named Entity)은 사람(Person), 기관(Organization) 등의 이름, 곡명, 방송명 또는 지명과 같이 분류될 수 있는 단어 또는 일련의 단어들의 집합을 의미하고, 마이닝 규칙은 마이닝 패턴과 그로부터 획득할 수 있는 RDF 트리플로 구성된 규칙을 의미한다.
- [0008] 개체명 검출은 특히, 음성 언어와 관련된 지식 학습 분야에 있어서 매우 중요한 연구 분야 중 하나이며, 개체명 검출을 위한 많은 알고리즘이 제안되고 있다.
- [0009] 개체명 검출의 가장 기본적인 방법으로 사전(Dictionary)을 기반으로 하는 방법으로, 입력되는 문장으로부터 개체명이라고 판단되는 단어 또는 단어의 집합을 추출하여 미리 저장된 개체명 사전에 저장된 내용과 비교하는 방법으로 개체명을 검출하는 것이다.
- [0010] 또한, 마이닝 규칙을 이용하여 RDF 트리플로 변환하는 방법은 미리 저장된 마이닝 규칙 데이터베이스로부터 일치하는 마이닝 패턴을 검색하여 대응되는 RDF 트리플로 변환하는 것이다.
- [0011] 그러나, 개체명 또는 마이닝 규칙은 시간에 따라 고정적인 것이 아니라 사회, 문화 등에 따라 새롭게 생성되고 시간이 지나면 사라지는 특성이 있으므로 개체명 또는 마이닝 규칙의 업데이트가 필요하다.

발명의 내용

해결하려는 과제

- [0012] 본 발명은 개체명 사전 및 마이닝 규칙 데이터베이스를 입력된 문서로부터 인식된 용어 및 마이닝 패턴을 이용하여 갱신하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 또는 마이닝 규칙 데이터베이스 갱신 장치, 그 방법 및 이를 기록한 전자 장치에서 판독 가능한 기록매체에 관한 것이다.

과제의 해결 수단

- [0013] 본 발명의 한 측면에 따르면, 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈, 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 마이닝 패턴 생성모듈, 인식된 용어와 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 개체명 사전과 마이닝 규칙 데이터베이스에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈, 및 인식된 용어에 대응되는 개체명이 개체명 사전으로부터 검색되지 않고 마이닝 패턴에 대응되는 마이닝 규칙이 마이닝 규칙 데이터베이스로부터 검색되는 경우, 마이닝 규칙을 이용하여 용어의 개체명을 추정하고, 추정된 용어의 개체명을 사용자의 선택에 따라 개체명 사전에 저장하는 개체명 사전 갱신모듈을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치를 제공한다.

- [0014] 본 발명의 다른 측면에 따르면, 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈, 입력된 마이닝 패턴과 마이닝 패턴에 포함되는 용어를 이용하여 대응되는 마이닝 규칙과 개체명을 마이닝 규칙 데이터베이스와 개체명 사전에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈, 및 용어에 대응되는 개체명이 개체명 사전으로부터 검색되지 않고 마이닝 패턴에 대응되는 마이닝 규칙이 마이닝 규칙 데이터베이스로부터 검색되는 경우, 마이닝 규칙을 이용하여 용어의 개체명을 추정하고, 추정된 용어의 개체명을 사용자의 선택에 따라 개체명 사전에 저장하는 개체명 사전 갱신모듈을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치를 제공한다.
- [0015] 개체명 사전 갱신모듈은, 용어와 대응되는 개체명의 특성을 표시하는 검증 인터페이스와 연결되며, 검증 인터페이스를 통해 사용자로부터 용어와 대응되는 개체명의 특성을 확인한 후, 확인 결과를 상기 개체명 사전에 저장할 수 있다.
- [0016] 개체명 및 마이닝 규칙 검색모듈은, 용어의 개체명을 개체명 사전으로부터 검색하는 개체명 검색부, 및 마이닝 패턴에 대응되는 마이닝 규칙을 마이닝 규칙 데이터베이스로부터 검색하는 마이닝 규칙 검색부를 포함할 수 있다.
- [0017] 개체명 사전 갱신모듈은, 검색 결과, 용어가 개체명 사전으로부터 검색되었는지 판단하는 검색결과 판단부, 판단 결과, 용어가 개체명 사전으로부터 검색되지 않고 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색되는 경우 마이닝 규칙을 이용하여 용어의 개체명을 추정하고 사용자의 선택에 따라 개체명 사전에 저장하는 개체명 사전 갱신부, 검색결과 판단부의 판단 결과에 따라 개체명 사전 갱신부에 개체명 특성을 정의하여 저장되도록 제어하는 제어부를 포함할 수 있다.
- [0018] 개체명 사전 갱신부는, 개체명의 특성을 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 개체명으로 분류되는 용어, 용어의 식별자, 대표 용어와 대표 용어의 식별자로 구성된 전거 데이터를 포함하여 정의하고 하나의 포맷으로 연결하여 저장할 수 있다.
- [0019] 개체명 사전은, 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 개체명으로 분류되는 용어, 용어의 식별자, 대표 용어와 대표 용어의 식별자로 구성된 전거 데이터를 하나의 포맷으로 연결하여 저장할 수 있다.
- [0020] 마이닝 패턴 생성모듈은, 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 인식된 용어를 상기 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 개체명 인식부, 및 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 패턴 변환부를 포함할 수 있다.
- [0021] 본 발명의 또 다른 측면에 따르면, 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈, 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 마이닝 패턴 생성모듈, 인식된 용어와 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 개체명 사전과 마이닝 규칙 데이터베이스에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈, 및 마이닝 패턴에 대응되는 마이닝 규칙이 마이닝 규칙 데이터베이스로부터 검색되지 않고 인식된 용어에 대응되는 개체명이 개체명 사전으로부터 검색되는 경우, 인식된 용어의 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 대응되는 마이닝 규칙을 생성한 후, 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장하는 마이닝 규칙 데이터베이스 갱신 모듈을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치를 제공한다.
- [0022] 본 발명의 또 다른 측면에 따르면, 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스가 저장되는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈, 입력된 마이닝 패턴과 마이닝 패턴에 포함되는 용어를 이용하여 대응되는 마이닝 규칙과 개체명을 마이닝 규칙 데이터베이스와 개체명 사전에서 각각 검색하는 개체명 및 마이닝 규칙 검색모듈, 및 마이닝 패턴에 대응되는 마이닝 규칙이 마이닝 규칙 데이터베이스로부터 검색되지 않고 용어에 대응되는 개체명이 개체명 사전으로부터 검색되는 경우, 용어의 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 대응되는 마이닝 규칙을 생성한 후, 사용자의 선택에 따라 마이닝 규칙 데이터베이스에

저장하는 마이닝 규칙 데이터베이스 갱신모듈을 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치를 제공한다.

- [0023] 마이닝 규칙 데이터베이스 갱신모듈은, 마이닝 패턴에 대응되는 마이닝 규칙을 표시하는 검증 인터페이스와 연결되며, 검증 인터페이스를 통해 사용자로부터 마이닝 패턴과 대응되는 마이닝 규칙을 확인한 후, 확인 결과를 마이닝 규칙 데이터베이스에 저장할 수 있다.
- [0024] 개체명 및 마이닝 규칙 검색모듈은, 용어의 개체명을 상기 개체명 사전으로부터 검색하는 개체명 검색부, 및 마이닝 패턴에 대응되는 마이닝 규칙을 마이닝 규칙 데이터베이스로부터 검색하는 마이닝 규칙 검색부를 포함할 수 있다.
- [0025] 마이닝 규칙 데이터베이스 갱신모듈은, 검색 결과, 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색되었는지 판단하는 검색결과 판단부, 판단 결과, 마이닝 패턴이 검색되지 않고 용어가 개체명 사전으로부터 검색되는 경우 검색된 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 관계명에 대응되는 마이닝 규칙을 생성한 후 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장하는 마이닝 규칙 데이터베이스 갱신부, 및 검색결과 판단부의 판단 결과에 따라 마이닝 규칙 데이터베이스 갱신부에 마이닝 규칙을 정의하여 저장되도록 제어하는 제어부를 포함할 수 있다.
- [0026] 마이닝 규칙 데이터베이스 갱신부는, 마이닝 규칙을 상기 온톨로지 스키마의 관계명(Property), 관계명에 대응되는 하나의 마이닝 패턴, 마이닝 패턴과 대응되는 하나 이상의 RDF 트리플을 포함하여 정의하고 하나의 포맷으로 연결하여 저장할 수 있다.
- [0027] 마이닝 패턴 생성모듈은, 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 인식된 용어를 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 개체명 인식부, 및 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 마이닝 패턴 변환부를 포함할 수 있다.
- [0028] 본 발명의 또 다른 측면에 따르면, 미리 저장된 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스를 이용한 개체명 사전 갱신 방법에 있어서, 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계, 인식된 용어와 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 미리 저장된 개체명 사전과 상기 마이닝 규칙 데이터베이스로부터 각각 검색하는 단계, 개체명 사전과 마이닝 규칙 데이터베이스 검색 결과, 인식된 용어 또는 마이닝 패턴과 일치하는 용어 또는 마이닝 패턴이 검색되었는지 판단하는 단계, 판단 결과, 일치하는 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색되었으나 일치하는 용어가 개체명 사전으로부터 검색되지 않은 경우 상기 마이닝 패턴에 대응되는 마이닝 규칙과 온톨로지 스키마를 이용하여 용어의 개체명을 추정하는 단계, 및 추정된 개체명을 사용자의 선택에 따라 개체명 사전에 저장하는 단계를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법을 제공한다.
- [0029] 추정된 개체명을 사용자의 선택에 따라 개체명 사전에 저장하는 단계는, 인식된 용어와 추정된 개체명을 검증 인터페이스에 표시하여 사용자로부터 검증을 받을 수 있다.
- [0030] 개체명 사전은, 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 개체명으로 분류되는 용어, 용어의 식별자, 대표 용어와 대표 용어의 식별자로 구성된 전거 데이터를 하나의 포맷으로 연결하여 저장할 수 있다.
- [0031] 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계는, 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 인식된 용어를 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 단계, 및 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 단계를 포함할 수 있다.
- [0032] 본 발명의 또 다른 측면에 따르면, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법을 수행하는 프로그램이 기록되고 전자 장치에서 판독 가능한 기록 매체를 제공한다.
- [0033] 본 발명의 또 다른 측면에 따르면, 저장된 온톨로지 스키마와 결합되고 개념(Class)과 연결된 용어의 개체명이 정의된 개체명 사전, 온톨로지 스키마와 결합되고 관계명과 연결된 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙이 정의된 마이닝 규칙 데이터베이스를 이용한 개체명 사전 갱신 방법에 있어서, 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계, 인식된 용어와 마이닝 패턴을 이용하여 대응되는 개체명과 마이닝 규칙을 미리 저장된 개체명 사전과 마이닝 규칙 데이터베이스

이으로부터 각각 검색하는 단계, 개체명 사전과 마이닝 규칙 데이터베이스 검색 결과, 인식된 용어 또는 마이닝 패턴과 일치하는 용어 또는 마이닝 패턴이 검색되었는지 판단하는 단계, 판단 결과, 일치하는 용어가 개체명 사전에서 검색되었으나 일치하는 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색되지 않은 경우 용어의 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 대응되는 마이닝 규칙을 생성하는 단계, 및 추정된 관계명과 마이닝 규칙을 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장하는 단계를 포함하는 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법을 제공한다.

[0034] 추정된 관계명과 마이닝 규칙을 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장하는 단계는, 추정된 관계명, 마이닝 패턴 및 마이닝 규칙을 검증 인터페이스에 표시하여 사용자로부터 검증을 받을 수 있다.

[0035] 마이닝 규칙 데이터베이스는, 마이닝 규칙을 온톨로지 스키마의 관계명(Property), 관계명에 대응되는 하나의 마이닝 패턴, 마이닝 패턴과 대응되는 하나 이상의 RDF 트리플을 포함하여 정의하고 하나의 포맷으로 연결하여 저장할 수 있다.

[0036] 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하는 단계는, 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 인식된 용어를 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타내는 단계, 및 개체명을 포함하는 텍스트를 마이닝 패턴으로 변환하는 단계를 포함할 수 있다.

[0037] 본 발명의 또 다른 측면에 따르면, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법을 수행하는 프로그램이 기록되고 전자 장치에서 판독 가능한 기록 매체를 제공한다.

발명의 효과

[0038] 본 발명에 따르면, 개체명 사전과 마이닝 규칙을 온톨로지 스키마와 연계하여 언어 자원을 통일적 방식으로 관리할 수 있는 효과가 있다.

[0039] 또한, 온톨로지 스키마와 연계된 개체명 사전을 이용하여 마이닝 규칙 데이터베이스에 저장되지 않은 마이닝 규칙을 추정하고, 온톨로지 스키마와 연계된 마이닝 규칙 데이터베이스를 이용하여 개체명을 추정하여 개체명 사전 또는 마이닝 규칙 데이터베이스를 확장할 수 있는 효과가 있다.

도면의 간단한 설명

[0040] 도 1은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치의 일 실시예를 나타내는 구성도.

도 2는 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치의 일 실시예를 나타내는 구성도.

도 3은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치 또는 마이닝 규칙 데이터베이스 갱신 장치의 마이닝 패턴 생성모듈의 일 실시예를 나타내는 구성도.

도 4는 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치 또는 마이닝 규칙 데이터베이스 갱신 장치의 개체명 및 마이닝 규칙 검색모듈의 일 실시예를 나타내는 구성도.

도 5는 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치의 개체명 사전 갱신모듈의 일 실시예를 나타내는 구성도.

도 6은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치의 마이닝 규칙 데이터베이스 갱신모듈의 일 실시예를 나타내는 구성도.

도 7은 본 발명의 다른 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법 또는 마이닝 규칙 데이터베이스 갱신 방법의 일 실시예를 나타내는 흐름도.

도 8은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙 데이터베이스의 연결 관계를 설명하기 위한 개념도.

발명을 실시하기 위한 구체적인 내용

- [0041] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나 이는 본 발명의 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0042] 이제 본 발명의 실시예에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치 및 방법, 마이닝 규칙 데이터베이스 갱신 장치 및 방법, 이를 기록한 전자 장치에 의해 판독 가능한 기록매체에 대하여 도면을 참조하여 상세하게 설명하고, 도면 부호에 관계없이 동일하거나 대응하는 구성요소는 동일한 참조 번호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.
- [0043] 도 1은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치(100)의 일 실시예를 나타내는 구성도를 개략적으로 도시한 도면이고, 도 8은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙 데이터베이스의 연결관계를 설명하기 위한 개념도이다.
- [0044] 본 발명의 실시예에 따르면, 도 1에 도시한 바와 같이, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치(100)는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈(110), 마이닝 패턴 생성모듈(120), 개체명 및 마이닝 규칙 검색모듈(130) 및 개체명 사전 갱신모듈(140)을 포함한다.
- [0045] 또한, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치(100)는 추정된 개체명을 사용자에게 검증받기 위하여, 추정된 개체명과 용어를 표시하는 검증 인터페이스(150)를 포함하여 구성될 수 있다.
- [0046] 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈(110)은 개체명 사전과 마이닝 규칙 데이터베이스를 저장할 수 있다.
- [0047] 도 8에 도시한 바와 같이, 개체명 사전에 저장된 개체명과 마이닝 규칙 데이터베이스의 마이닝 규칙은 온톨로지 스키마의 개념(Class)과 연결된 용어 및 관계명과 연결된 마이닝 패턴과 RDF 트리플 형태로 형성될 수 있다.
- [0048] 예를 들어, 온톨로지 스키마의 개념(Class)인 종교, 사람, 종목 등은 개체명 사전의 키 값인 개체명으로 정의되고, 구체적인 인스턴트는 개체명 사전의 용어로 정의될 수 있다.
- [0049] 또한, 온톨로지 스키마의 관계명(Property)인 직업, 신앙, 분야 등은 마이닝 규칙 데이터베이스의 키 값인 관계명으로 정의되고, 각각의 관계명에 대응되는 마이닝 패턴, RDF 트리플과 연결될 수 있다.
- [0050] 따라서, 개체명 사전 및 마이닝 규칙 데이터베이스는 온톨로지 스키마의 개념 또는 관계명과 연계되어 형성될 수 있다.
- [0051] 개체명 사전은 온톨로지 스키마와 결합되고 용어의 개체명이 정의되며, 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 개체명으로 분류되는 용어, 용어의 식별자, 대표 용어와 대표 용어의 식별자로 구성된 전거 데이터를 하나의 포맷으로 연결하여 저장할 수 있다.
- [0052] 마이닝 규칙 데이터베이스는 온톨로지 스키마와 결합되고 관계명에 대응되는 마이닝 패턴과 리소스 기술 프레임워크(RDF: Resource Description Framework) 트리플로 구성된 마이닝 규칙을 저장할 수 있다.
- [0053] 마이닝 규칙은 온톨로지 스키마의 관계명(Property), 관계명에 대응되는 하나의 마이닝 패턴, 마이닝 패턴과 대응되는 하나 이상의 RDF 트리플을 포함하여 정의되며, 마이닝 규칙 데이터베이스는 온톨로지 스키마의 관계명(Property), 마이닝 패턴, RDF 트리플을 하나의 포맷으로 연결하여 저장할 수 있다.
- [0054] 또한, 본 발명의 다른 실시예에 따르면, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치는 마이닝 패턴 생성모듈이 생략되고, 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈, 개체명 및 마이닝 규칙 검색모듈 및 개체명 사전 갱신모듈을 포함하여 구성될 수 있다.
- [0055] 즉, 마이닝 패턴이 직접 개체명 및 마이닝 규칙 검색모듈로 입력되는 경우 마이닝 패턴 생성모듈의 기능이 필요하지 않으므로 마이닝 패턴 생성모듈은 생략될 수 있다.
- [0056] 도 3은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치 또는 마이닝 규칙 데이터베이스 갱신 장치의 마이닝 패턴 생성모듈의 일 실시예를 나타내는 구성도이다.

고, 도 4는 본 발명의 개체명 및 마이닝 규칙 검색모듈의 일 실시예를 나타내는 구성도이고, 도 5는 본 발명의 개체명 사전 갱신모듈의 일 실시예를 나타내는 구성도를 개략적으로 도시한 도면이다.

- [0057] 본 발명의 실시예에 따르면, 도 3에 도시한 바와 같이, 마이닝 패턴 생성모듈(120)은 텍스트로부터 용어를 인식하여 마이닝 패턴으로 변환하고, 개체명 인식부(122) 및 마이닝 패턴 변환부(124)를 포함하여 구성될 수 있다.
- [0058] 개체명 인식부(122)는 개체명 사전과 연동하여 텍스트로부터 형태소 분석, 구문 분석, 의미 분석을 사용하여 용어를 인식하고, 인식된 용어를 개체명 사전에 저장된 용어에 대응되는 개체명으로 나타낼 수 있다.
- [0059] 마이닝 패턴 변환부(124)는 개체명을 포함하여 표시된 텍스트를 마이닝 패턴으로 변환하는 기능을 한다.
- [0060] 본 발명의 실시예에 따르면, 도 4에 도시한 바와 같이, 개체명 및 마이닝 규칙 검색모듈(130)은 개체명 검색부(132) 및 마이닝 규칙 검색부(134)를 포함하여 구성될 수 있다.
- [0061] 즉, 개체명 검색부(132)는 인식된 용어의 개체명을 개체명 사전으로부터 검색하고, 마이닝 규칙 검색부(134)는 마이닝 패턴에 대응되는 마이닝 규칙을 마이닝 규칙 데이터베이스로부터 검색할 수 있다.
- [0062] 개체명 사전 갱신모듈(140)은 인식된 용어에 대응되는 개체명이 개체명 사전으로부터 검색되지 않고 마이닝 패턴에 대응되는 마이닝 규칙이 마이닝 규칙 데이터베이스로부터 검색되는 경우, 마이닝 규칙을 이용하여 용어의 개체명을 추정하고, 추정된 용어의 개체명을 사용자의 선택에 따라 개체명 사전에 저장할 수 있다.
- [0063] 사용자는 검증 인터페이스(150)를 통해 표시된 용어와 추정된 개체명 또는 개체명 특성이 적합한지 여부를 판단하고 개체명 사전에 저장할 것인지 여부를 결정할 수 있다.
- [0064] 개체명 사전 갱신모듈(140)은 검색결과 판단부(142), 개체명 사전 갱신부(144) 및 제어부(146)를 포함하여 구성될 수 있다.
- [0065] 검색결과 판단부(142)는 개체명 검색부의 검색 결과를 확인하여 인식된 용어가 개체명 사전으로부터 검색되었는지 판단하고, 개체명 사전 갱신부(144)는 검색결과 판단부(142)의 판단 결과 용어가 개체명 사전으로부터 검색되지 않고 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색되는 경우, 마이닝 규칙을 이용하여 용어의 개체명을 추정하고 사용자의 선택에 따라 개체명 사전에 추정된 개체명과 용어를 저장할 수 있다.
- [0066] 제어부(146)는 검색결과 판단부(142) 및 개체명 사전 갱신부(144)를 제어하여, 검색결과 판단부(142)의 판단 결과에 따라 개체명 사전 갱신부(144)에 개체명 특성을 정의하여 저장되도록 제어할 수 있다.
- [0067] 특히, 개체명 사전 갱신부(144)는 개체명의 특성을 온톨로지 스키마의 개념(Class)에 해당되는 개체명, 개체명으로 분류되는 용어, 용어의 식별자, 대표 용어와 대표 용어의 식별자로 구성된 전거 데이터를 포함하여 정의하고 하나의 포맷으로 연결하여 저장할 수 있다.
- [0068] 예를 들어, "김연아가 조니위어를 자기의 이상형이라고 했다는 얘기가 있다."는 문장이 입력되면, 마이닝 패턴 생성모듈은 '김연아', '조니위어' 등의 용어를 인식하여 마이닝 패턴으로 변환한다.
- [0069] 인식된 용어인 '김연아'와 '조니위어'를 개체명 사전으로부터 검색하면 '김연아'의 개체명은 '[사람]'으로 검색되지만 '조니위어'의 개체명은 검색되지 않는 경우, 문장을 마이닝 패턴으로 변환하면 "[사람1] jcs [?] jco %self의 %ideal_type이라고"로 변환될 수 있다.
- [0070] 변환된 마이닝 패턴과 일치하는 마이닝 패턴에 대응되는 마이닝 규칙을 마이닝 규칙 데이터베이스로부터 검색하면 '[사람1] jcs [사람2] jco %self의 %ideal_type이라고' -> '[사람1]-이상형-[사람2]'가 검색되고, 이로부터 '조니위어'의 개체명을 '[사람]'으로 추정할 수 있다.
- [0071] 개체명 사전 갱신모듈(140)은 아래 [표 1]과 같이, 용어, 추정된 개체명과 식별자 등을 검증 인터페이스(150)에 표시하여 사용자의 확인을 받을 수 있다.

표 1

추정 개체명	용어	식별자	확인
사람	조니위어	http://ontoframe.kr/Person/0010	<input checked="" type="checkbox"/>
사람	조니위커	http://ontoframe.kr/Person/0011	<input type="checkbox"/>

- [0073] 도 2는 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙

데이터베이스 갱신 장치(200)의 일 실시예를 나타내는 구성도를 개략적으로 도시한 도면이다.

- [0074] 본 발명의 실시예에 따르면, 도 2에 도시한 바와 같이, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치(200)는 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈(210), 마이닝 패턴 생성모듈(220), 개체명 및 마이닝 규칙 검색모듈(230) 및 마이닝 규칙 데이터베이스 갱신모듈(240)을 포함하여 구성될 수 있다.
- [0075] 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈(210), 마이닝 패턴 생성모듈(220) 및 개체명 및 마이닝 규칙 검색모듈(230)은 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치(100)의 구성과 동일한 것으로 동일한 기능을 수행하며, 차이가 있는 마이닝 규칙 데이터베이스 갱신모듈(240)의 구성 및 기능을 중심으로 설명하도록 한다.
- [0076] 도 6은 본 발명의 일 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치(200)의 마이닝 규칙 데이터베이스 갱신모듈(240)의 일 실시예를 나타내는 구성도를 개략적으로 도시한 도면이다.
- [0077] 본 발명의 실시예에 따르면, 도 6에 도시한 바와 같이, 마이닝 규칙 데이터베이스 갱신모듈(240)은 검색결과 판단부(242), 마이닝 규칙 데이터베이스 갱신부(244) 및 제어부(246)를 포함하여 구성될 수 있다.
- [0078] 마이닝 규칙 데이터베이스 갱신모듈(240)은 인식된 용어의 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 대응되는 마이닝 규칙을 생성한 후, 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장할 수 있다.
- [0079] 또한, 마이닝 규칙 데이터베이스 갱신모듈(240)은 마이닝 패턴에 대응되는 마이닝 규칙을 표시하는 검증 인터페이스(250)와 연결될 수 있다.
- [0080] 검증 인터페이스(250)는 추정된 관계명, 마이닝 패턴, 마이닝 규칙 등이 표시될 수 있고, 사용자의 확인에 따라 마이닝 규칙 데이터베이스에 저장될 수 있다.
- [0081] 마이닝 규칙 데이터베이스 갱신모듈(240)의 검색결과 판단부(242)는 개체명 및 마이닝 규칙 검색모듈(220)의 검색결과 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색되었는지 판단할 수 있다.
- [0082] 마이닝 규칙 데이터베이스 갱신부(244)는 마이닝 규칙 데이터베이스로부터 검색되었는 판단 결과를 이용하여 마이닝 규칙 데이터베이스를 갱신할 수 있다.
- [0083] 즉, 마이닝 규칙 데이터베이스 갱신부(244)는 마이닝 패턴이 검색되지 않고 용어가 개체명 사전으로부터 검색되는 경우 검색된 개체명과 온톨로지 스키마를 이용하여 관계명을 추정하고 관계명에 대응되는 마이닝 규칙을 생성한 후 사용자의 선택에 따라 마이닝 규칙 데이터베이스에 저장할 수 있다.
- [0084] 마이닝 규칙 데이터베이스 갱신부(244)는 마이닝 규칙을 온톨로지 스키마의 관계명(Property), 관계명에 대응되는 하나의 마이닝 패턴, 마이닝 패턴과 대응되는 하나 이상의 RDF 트리플을 포함하여 정의하고 하나의 포맷으로 연결하여 저장할 수 있다.
- [0085] 제어부(246)는 검색결과 판단부(242) 및 마이닝 규칙 데이터베이스 갱신부(244)의 기능을 제어하며, 검색결과 판단부(242)의 판단 결과에 따라 마이닝 규칙을 정의하여 마이닝 규칙 데이터베이스 갱신부(244)에 저장되도록 제어할 수 있다.
- [0086] 또한, 본 발명의 다른 실시예에 따르면, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치는 마이닝 패턴 생성모듈이 생략되고, 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈, 개체명 및 마이닝 규칙 검색모듈 및 마이닝 규칙 데이터베이스 갱신모듈을 포함하여 구성될 수 있다.
- [0087] 즉, 마이닝 패턴이 직접 개체명 및 마이닝 규칙 검색모듈로 입력되는 경우 마이닝 패턴 생성모듈의 기능이 필요하지 않으므로 마이닝 패턴 생성모듈은 생략될 수 있다.
- [0088] 도 7은 본 발명의 다른 측면에 따른 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법 또는 마이닝 규칙 데이터베이스 갱신 방법의 일 실시예를 나타내는 흐름도이다.
- [0089] 본 발명의 실시예에 따르면, 도 7에 도시한 바와 같이, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 방법 또는 마이닝 규칙 데이터베이스 갱신 방법은 텍스트의 용어를 인식하고

(S710), 텍스트를 마이닝 패턴으로 변환하며(S720), 용어와 마이닝 패턴을 개체명 사전 또는 마이닝 규칙 데이터베이스(DB)로부터 검색할 수 있다(S730).

- [0090] 상기의 검색 결과, 일치하는 용어와 마이닝 패턴이 검색되었는지 여부를 판단하고(S740), 판단 결과에 따라 개체명 사전을 갱신하거나 마이닝 규칙 데이터베이스를 갱신할 수 있다.
- [0091] 판단 결과, 용어가 개체명 사전으로부터 검색되고 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색되지 않은 경우, 개체명을 이용하여 용어의 관계명을 추정하고(S750a), 관계명에 대응되는 마이닝 규칙을 생성하며(S760a), 사용자 검증을 거친 후(770a), 마이닝 규칙 데이터베이스에 저장하여 마이닝 규칙 데이터베이스를 갱신할 수 있다(S780a).
- [0092] 판단 결과, 용어가 개체명 사전으로부터 검색되지 않고 마이닝 패턴이 마이닝 규칙 데이터베이스로부터 검색된 경우, 마이닝 패턴에 대응되는 마이닝 규칙을 이용하여 개체명을 추정하고(S750b), 사용자 검증을 거친 후(S760b), 개체명 사전에 저장하여 개체명 사전을 갱신할 수 있다(S770b).
- [0093] 따라서, 사회, 경제적 변화에 따라 새로운 용어가 출현하거나 종래에 사용하던 용어가 사용되지 않는 경우 또는 개념 사이의 새로운 관계 체계를 규명해야할 필요가 있는 경우 개체명 사전 또는 마이닝 규칙 데이터베이스를 갱신할 수 있다.
- [0094] 본 발명의 또 다른 측면에 따르면, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법을 수행하는 프로그램이 기록되고 전자 장치에서 관독 가능한 기록 매체에 기록될 수 있다.
- [0095] 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법은 컴퓨터 프로그램으로 작성 가능하며, 프로그램을 구성하는 코드들 및 코드 세그먼트들은 당해 분야의 컴퓨터 프로그래머에 의하여 용이하게 추론될 수 있다.
- [0096] 또한, 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 방법은 컴퓨터가 읽을 수 있는 정보저장매체(Computer Readable Medium)에 저장되고, 컴퓨터에 의하여 읽혀지고 실행됨으로써 개체명 사전 또는 마이닝 규칙 데이터베이스를 갱신할 수 있다.

산업상 이용가능성

- [0097] 본 발명은 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 또는 마이닝 규칙 데이터베이스 갱신 장치 및 이를 이용하여 데이터를 모델링하는 장치 등에 적용할 수 있다.

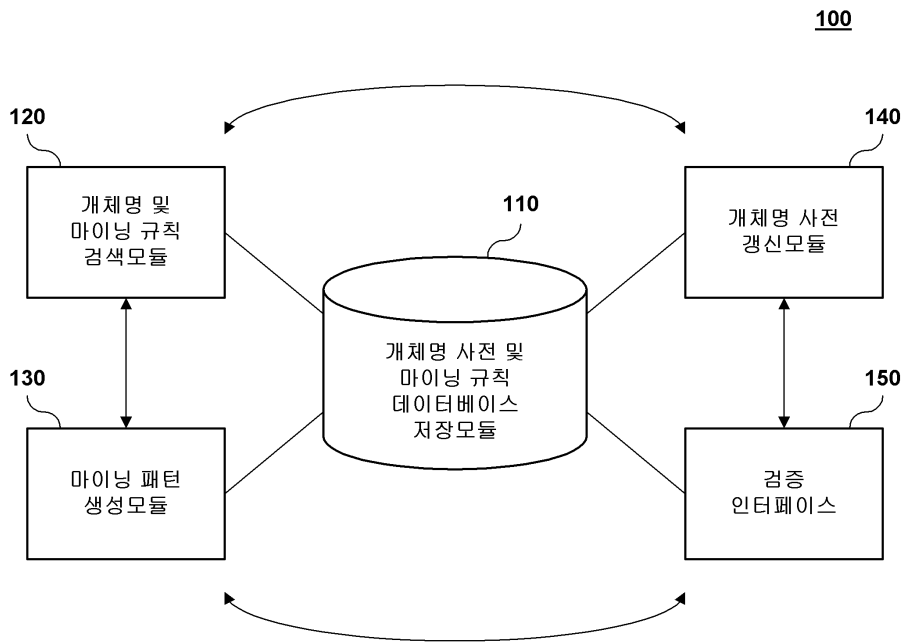
부호의 설명

- [0098] 100: 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 개체명 사전 갱신 장치
- 200: 온톨로지 스키마와 결합된 개체명 사전 및 마이닝 규칙을 이용한 마이닝 규칙 데이터베이스 갱신 장치
- 110, 210: 개체명 사전 및 마이닝 규칙 데이터베이스 저장모듈
- 120, 220: 마이닝 패턴 생성모듈
- 122: 개체명 인식부
- 124: 마이닝 패턴 변환부
- 130, 230: 개체명 및 마이닝 규칙 검색모듈
- 132: 개체명 검색부
- 134: 마이닝 규칙 검색부
- 140: 개체명 사전 갱신모듈
- 142: 검색결과 판단부
- 144: 개체명 사전 갱신부

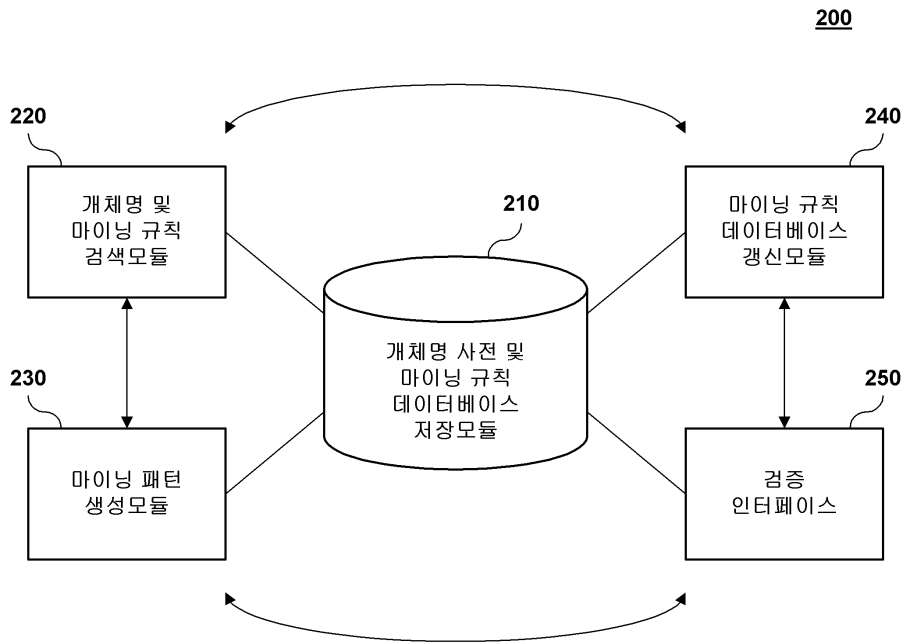
- 146: 제어부
- 150: 검증 인터페이스
- 240: 마이닝 규칙 데이터베이스 갱신모듈
- 242: 검색결과 판단부
- 244: 마이닝 규칙 데이터베이스 갱신부
- 246: 제어부

도면

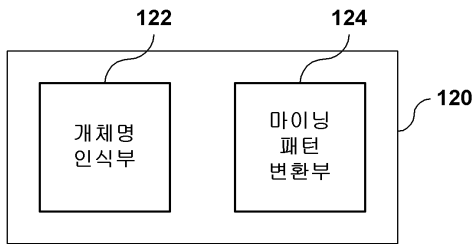
도면1



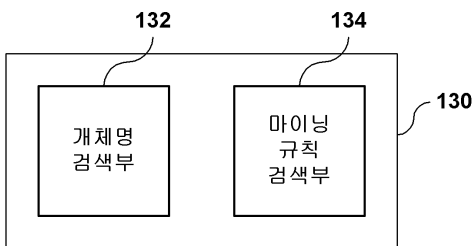
도면2



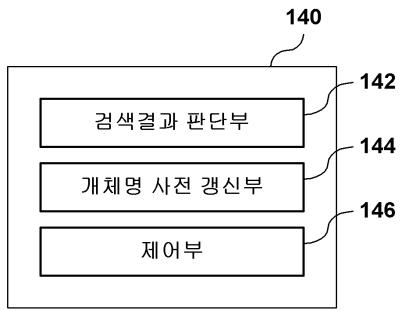
도면3



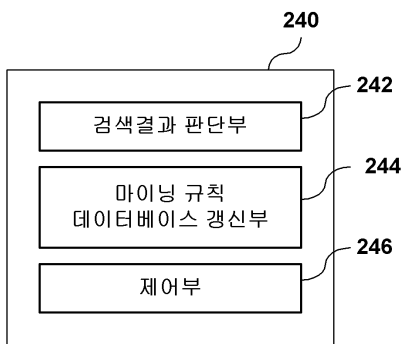
도면4



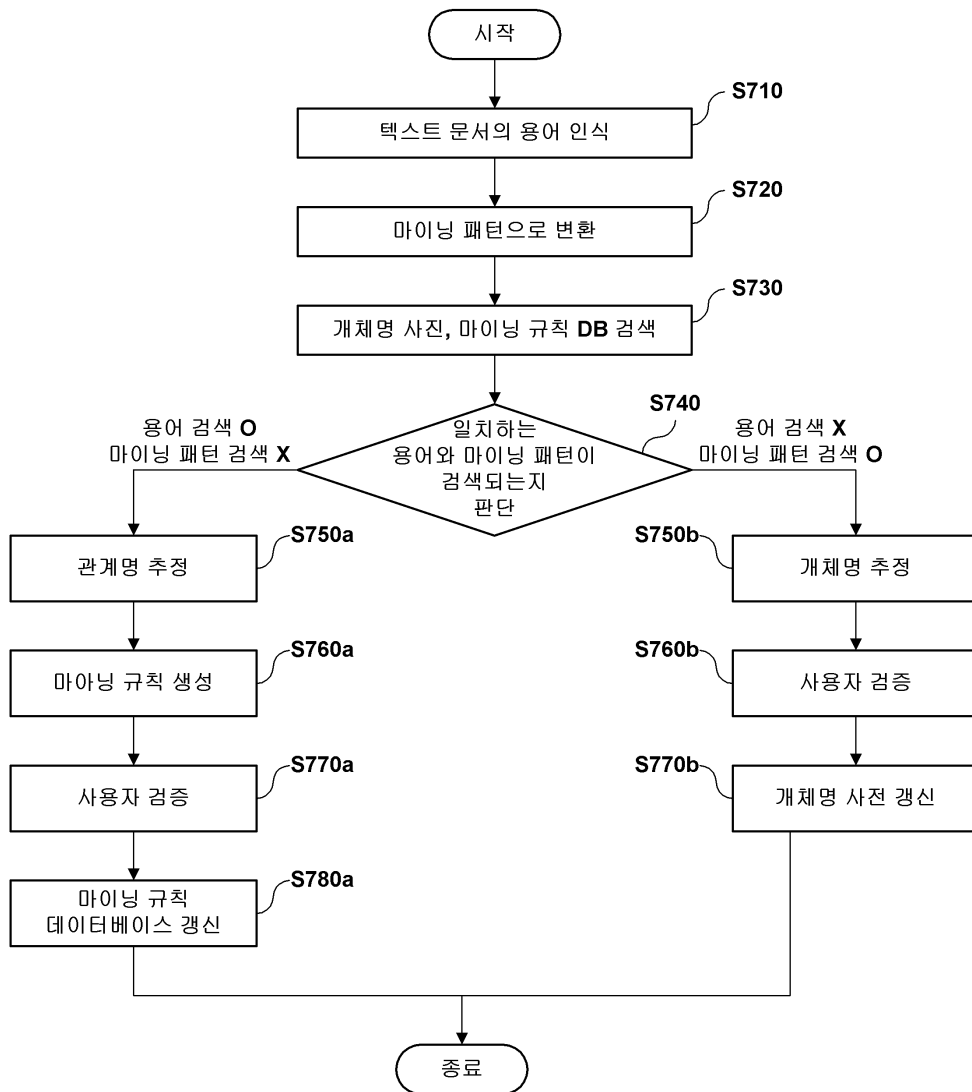
도면5



도면6



도면7



도면8

