

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2016년 7월 28일 (28.07.2016)



(10) 국제공개번호
WO 2016/117920 A1

- (51) 국제특허분류:
G06F 17/30 (2006.01) G06F 17/20 (2006.01)
- (21) 국제출원번호: PCT/KR2016/000579
- (22) 국제출원일: 2016년 1월 20일 (20.01.2016)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2015-0009518 2015년 1월 20일 (20.01.2015) KR
10-2015-0139189 2015년 10월 2일 (02.10.2015) KR
- (71) 출원인: 한국과학기술원 (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) [KR/KR]; 34141 대전시 유성구 대학로 291, Daejeon (KR).
- (72) 발명자: 최기선 (CHOI, Key-Sun); 34141 대전시 유성구 대학로 291, Daejeon (KR). 함영균 (HAHM, Young-gyun); 34141 대전시 유성구 대학로 291, Daejeon (KR). 서지우 (SEO, Ji Woo); 34141 대전시 유성구 대학로 291, Daejeon (KR).

(74) 대리인: 양성보 (YANG, Sungbo); 06099 서울시 강남구 선릉로 125길 14 삼성빌딩 2층 피앤티특허법률사무소, Seoul (KR).

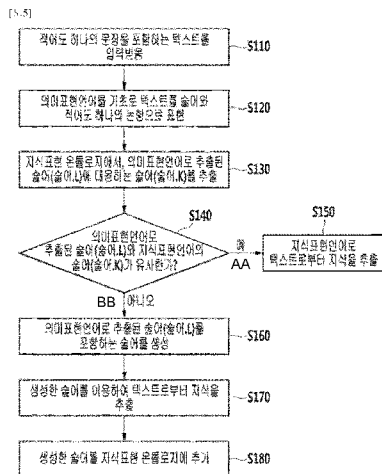
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR),

[다음 쪽 계속]

(54) Title: KNOWLEDGE REPRESENTATION EXPANSION METHOD AND APPARATUS

(54) 발명의 명칭: 지식표현 확장 방법 및 장치



- S110 ... Receive an input of text including at least one sentence
S120 ... Represent the text with a predicate and at least one argument on the basis of a meaning representation language
S130 ... Extract a predicate (predicate.K) corresponding to a predicate (predicate.L) extracted in a meaning representation language, in a knowledge representation ontology
S140 ... Is the predicate (predicate.L) extracted in the meaning representation language similar to the predicate (predicate.K) of the knowledge representation language?
S150 ... Extract knowledge from the text in the knowledge representation language
S160 ... Generate a predicate including the predicate (predicate.L) extracted in the meaning representation language
S170 ... Extract knowledge from the text using the generated predicate
S180 ... Add the generated predicate to the knowledge representation ontology
AA ... Yes
BB ... No

(57) Abstract: A knowledge representation expansion apparatus includes: a predicate-argument structure analyzing unit for extracting a predicate and at least one argument from a text using a meaning representation language; an ontology unit for representing knowledge using a knowledge representation language, which is a structured format understandable by a computer, and for extracting a second predicate corresponding to a first predicate, which is extracted from the predicate-argument structure analyzing unit; and a knowledge representation unit for representing knowledge extracted from the text using the first predicate, when the similarity of the first predicate and the second predicate is equal to or less than a threshold value.

(57) 요약서: 지식표현 확장 장치로서, 의미표현언어를 이용하여 텍스트에서 술어와 적어도 하나의 논항을 추출하는 술어-논항 구조 해석부, 컴퓨터가 이해할 수 있는 구조화된 포맷인 지식표현언어를 이용하여 지식을 표현하는 온톨로지부, 그리고 상기 온톨로지부에서 상기 술어-논항 구조 해석부에서 추출된 제 1 술어에 대응하는 제 2 술어를 추출하고, 상기 제 1 술어와 상기 제 2 술어의 유사도가 기준값 이하인 경우, 상기 제 1 술어를 이용하여 상기 텍스트로부터 추출된 지식을 표현하는 지식표현부를 포함한다.



OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, **공개:**
ML, MR, NE, SN, TD, TG).

— 국제조사보고서와 함께 (조약 제 21 조(3))

명세서

발명의 명칭: 지식표현 확장 방법 및 장치

기술분야

- [1] 본 발명은 지식표현 확장 방법 및 장치에 관한 것이다.

배경기술

- [2] 최근 시맨틱 웹(semantic web)과 빅데이터 기반으로 질의응답 시스템에 대한 연구가 활발하다. 시맨틱 웹은 인터넷과 같은 분산환경에서 정보들 사이의 관계와 의미 정보(Semanteme)를 컴퓨터가 처리할 수 있는 온톨로지로 표현하는 의미론적인 웹이다. 또한 온톨로지 기반 지식데이터베이스를 구축하는 많은 연구들이 진행되고 있다. 그러나 전통적으로 지식은 자연 언어로 작성되어 있으며, 특히나 몇몇 연구에 의하면 구조화된 데이터베이스보다 비구조 데이터에서 많은 지식이 포함되어 있다고 알려져 있다. 따라서 자연 언어 텍스트를 포함하는 비구조 데이터로부터 온톨로지 스키마의 인스턴스들을 자동으로 생성하는 연구들이 지식데이터베이스를 확장을 위해 진행되고 있다.
- [3] 특히, 시맨틱 웹은 웹의 지식을 컴퓨터가 이해할 수 있는 구조화된 포맷, 즉 RDF(Resource Description Framework) 트리플로 표현해야 하고, 이를 위해 지식 요소들의 다양한 속성들을 충분히 설명할 수 있는 프로퍼티(property)를 갖고 있는 온톨로지가 요구된다. RDF 트리플은 월드 와이드 웹 컨소시엄(World Wide Web Consortium, W3C)이 관장하는 국제 표준으로서, 지식과 정보를 서브젝트[Subject(resource)], 프레디키트[Predicate(property)], 오브젝트[(Object(literal))]의 세 쌍으로 나타내는 형식이다. 여기서, 프로퍼티는 RDF 트리플의 프레디키트에 해당하며, 서브젝트와 오브젝트 사이의 관계에 해당한다.
- [4] 시맨틱 웹의 최신 기술인 디비피디아(DBpedia)는 백과사전적 텍스트인 위키피디아로부터 자동 구축된 지식데이터베이스이다. 디비피디아는 위키피디아의 지식을 표현하기 위하여 위키피디아의 인포박스에서 기원한 디비피디아 온톨로지를 사용한다. 그러나, 디비피디아 온톨로지는 위키피디아의 요약된 지식을 표현하기에는 충분하다고 볼 수 있지만, 위키피디아 텍스트상의 모든 지식을 표현할 수 있다고 보장하기는 어렵다. 따라서 자연어 텍스트에서 나타난 지식 요소들의 다양한 속성들을 표현할 수 있는 온톨로지가 필요하고, 이를 기초로 자동으로 지식데이터베이스를 구축하여 지식을 확장하는 기술이 필요하다.

발명의 상세한 설명

기술적 과제

- [5] 본 발명이 해결하려는 과제는 지식표현 확장 방법 및 장치로서, 지식표현 온톨로지에서 사용 중인 지식표현언어로 어느 텍스트로부터 추출한 지식을

표현할 수 없는 경우, 의미표현언어를 이용하여 지식표현을 확장하는 방법에 관한 것이다.

과제 해결 수단

- [6] 본 발명의 한 실시예에 따른 지식표현 확장 장치로서, 의미표현언어를 이용하여 텍스트에서 술어와 적어도 하나의 논항을 추출하는 술어-논항 구조 해석부, 컴퓨터가 이해할 수 있는 구조화된 포맷인 지식표현언어를 이용하여 지식을 표현하는 온톨로지부, 그리고 상기 온톨로지부에서 상기 술어-논항 구조 해석부에서 추출된 제1술어에 대응하는 제2술어를 추출하고, 상기 제1술어와 상기 제2술어의 유사도가 기준값 이하인 경우, 상기 제1술어를 이용하여 상기 텍스트로부터 추출된 지식을 표현하는 지식 표현부를 포함한다.
- [7] 상기 지식 표현부는 상기 온톨로지부에서 상기 적어도 하나의 논항에 관계된 상기 제2술어를 추출할 수 있다.
- [8] 상기 지식 표현부는 상기 지식표현언어의 도메인들 중에서 상기 적어도 하나의 논항에 부여된 어휘 타입과 기준값이상으로 유사한 제1도메인을 추출하고, 상기 지식표현언어의 범위들 중에서 상기 적어도 하나의 논항에 부여된 어휘 타입과 기준값이상으로 유사한 제1범위를 추출하며, 상기 제1도메인과 상기 제1범위에 관련된 술어를 상기 제2술어로 추출할 수 있다.
- [9] 상기 지식 표현부는 상기 제1술어와 상기 적어도 하나의 논항 중 임의 논항에 관련된 정보가 결합된 문자열을 생성하고, 상기 문자열을 상기 온톨로지부의 지식표현언어에 추가할 수 있다.
- [10] 상기 지식표현언어는 RDF(Resource Description Framework) 삼항 관계로 표현되는 언어일 수 있다.
- [11] 본 발명의 다른 실시예에 따른 장치가 지식표현을 확장하는 방법으로서, 적어도 하나의 문장을 포함하는 텍스트를 입력받는 단계, 의미표현언어를 기초로 상기 텍스트를 제1술어와 적어도 하나의 논항으로 표현하는 단계, 지식표현 온톨로지에서, 상기 제1술어에 대응하는 제2술어를 추출하는 단계, 상기 제1술어와 상기 제2술어의 유사도를 비교하는 단계, 그리고 상기 유사도가 기준값 이하인 경우, 상기 텍스트로부터 추출된 지식을 상기 제1술어를 이용하여 표현하는 단계를 포함한다.
- [12] 상기 제1술어에 대응하는 제2술어를 추출하는 단계는 상기 적어도 하나의 논항에 부여된 어휘 타입을 이용하여 상기 지식표현 온톨로지에서 상기 제1술어에 대응하는 상기 제2술어를 추출할 수 있다.
- [13] 상기 지식표현 온톨로지는 지식을 서브젝트(subject), 프레디케이트(Predicate), 오브젝트(object)의 삼항 관계로 표현하는 지식표현언어를 사용하고, 상기 제1술어에 대응하는 제2술어를 추출하는 단계는 상기 지식표현언어의 서브젝트들 중에서 상기 적어도 하나의 논항에 부여된 어휘 타입과 기준값이상으로 유사하고, 상기 지식표현언어의 오브젝트들 중에서 상기

적어도 하나의 논항에 부여된 어휘 타입과 기준값이상으로 유사한 프레디키트를 상기 제2술어로 추출할 수 있다.

- [14] 상기 제1술어를 이용하여 표현하는 단계는 상기 제1술어와 상기 적어도 하나의 논항 중 임의의 논항에 관련된 정보가 결합된 문자열을 생성하고, 상기 문자열을 이용하여 상기 텍스트로부터 추출한 지식을 표현할 수 있다.
- [15] 상기 방법은 상기 문자열을 상기 지식표현 온톨로지의 지식표현언어에 추가하는 단계를 더 포함할 수 있다.
- [16] 본 발명의 또 다른 실시예에 따른 장치가 지식표현을 확장하는 방법으로서, 텍스트의 술어-논항 구조를 해석하는 단계, 상기 텍스트의 술어-논항 구조를 지식표현언어의 삼항 관계로 매칭하는 단계, 그리고 매칭 유사도를 기초로 상기 텍스트의 술어-논항 구조에서 추출된 제1술어를 상기 지식표현언어의 술어로 추가하는 단계를 포함한다.
- [17] 상기 지식표현언어의 술어로 추가하는 단계는 상기 지식표현언어의 삼항 관계에서, 상기 텍스트의 술어-논항 구조의 제1술어에 매칭된 제2술어를 추출하는 단계, 상기 제1술어와 상기 제2술어의 유사도를 비교하는 단계, 그리고 상기 유사도가 기준값 이하인 경우, 상기 제1술어를 상기 지식표현언어에 추가하는 단계를 포함할 수 있다.
- [18] 상기 방법은 상기 제1술어를 이용하여 상기 텍스트를 삼항 관계로 표현하는 단계를 더 포함할 수 있다.
- [19] 상기 지식표현언어의 삼항 관계로 매칭하는 단계는 상기 텍스트의 술어-논항 구조에서 추출된 논항들과 상기 삼항 관계의 도메인 및 범위의 유사도를 기초로 상기 텍스트의 술어-논항 구조를 상기 삼항 관계로 매칭할 수 있다.

발명의 효과

- [20] 본 발명의 실시예에 따르면 지식표현 온톨로지에서도 사용 중인 지식표현언어로 어느 텍스트로부터 추출한 지식을 표현할 수 없는 경우, 의미표현언어를 이용하여 지식표현을 확장할 수 있다. 즉, 본 발명에 실시예에 따르면 지식표현 온톨로지가 웹 텍스트로부터 지식데이터베이스를 구축할 때 충분한 커버리지를 갖지 못하는 문제를 해결할 수 있다.
- [21] 본 발명의 실시예에 따르면 문장 의미 술어-논항 구조 기반으로 자연 언어와 같은 비구조 데이터에 포함된 지식을 컴퓨터가 이해할 수 있는 포맷의 지식표현언어로 표현하여 지식데이터베이스를 빠르고 쉽게 확장할 수 있다.
- [22] 본 발명의 실시예에 따르면 지식데이터베이스의 "관계" 온톨로지가 확충되어, 지식표현력을 높일 수 있고, CGC(Collaboratively Generated Content) 지향 지식 형태 및 해석에 적용될 수 있다.

도면의 간단한 설명

- [23] 도 1은 본 발명의 한 실시예에 따른 의미표현언어의 예시이다.
- [24] 도 2는 본 발명의 한 실시예에 따른 지식표현 확장 장치의 블록도이다.

- [25] 도 3은 본 발명의 한 실시예에 따른 술어-논항 구조를 해석한 결과를 설명하는 예시 도면이다.
- [26] 도 4는 본 발명의 한 실시예에 따른 삼항 관계 지식표현 구조를 설명하는 예시 도면이다.
- [27] 도 5는 본 발명의 한 실시예에 따른 지식표현 확장 방법의 흐름도이다.
- [28] 도 6은 본 발명의 한 실시예에 따른 지식표현 확장 방법을 예시하는 흐름도이다.
- [29] 도 7은 본 발명의 한 실시예에 따른 예시문의 술어-논항 구조를 해석한 결과를 설명하는 도면이다.
- [30] 도 8은 본 발명의 한 실시예에 따른 예시문의 삼항 관계 지식 표현 구조를 설명하는 도면이다.

발명의 실시를 위한 최선의 형태

- [31] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [32] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [33] 지식데이터베이스는 지식표현언어로 구조화된 정보를 저장한다. 온톨로지(ontology)는 컴퓨터가 이해할 수 있는 구조화된 포맷으로 지식을 표현한다. 지식표현언어는 다양할 수 있으나, 예를 들면 RDF 트리플일 수 있다. RDF 트리플은 지식과 정보를 서브젝트[Subject(resource)], 프레디케이트[Predicate(property)], 오브젝트[(Object(literal))]의 삼항 관계로 나타내는 형식이다. RDF 트리플 중 프레디케이트 또는 프로퍼티는 술어로서, 주어(서브젝트) 자리에 있는 개체(entity)와 목적어(오브젝트) 자리에 있는 개체 또는 값(value) 사이의 관계(relationship)/속성(property)을 나타낸다.
- [34] 이렇게 온톨로지는 구조화된 정보에 국한하므로, 구조화되지 않은 지식원(knowledge source)에서 추출된 지식을 표현하기 어려운 한계가 있다. 특히, 링크드 데이터(linked data)의 중심인 디비피디아에 대한 온톨로지의 커버리지 계산을 통해 텍스트로부터 충분한 지식을 추출할 수 있는지 검토한 결과, 구조화되지 않은 텍스트를 지식원으로 하여 새로운 지식을 추출할 때 표현력이 제한되는 것을 알 수 있다.
- [35] 다음에서 의미표현언어를 기초로 지식표현을 확장하는 방법에 대해 설명한다. 즉, 텍스트로부터 추출된 지식을 현재의 지식표현언어로 표현할 수 없는 경우,

새로운 온톨로지 인스턴스를 생성하여 지식표현을 확장하는 방법에 대해 설명한다.

[36] 도 1은 본 발명의 한 실시예에 따른 의미표현언어의 예시이다.

[37] 도 1을 참고하면, 다음과 같은 질의문을 예로 들어 설명한다. 질의문의 "이것"은 "인터페론(interferon)"이다.

[38]

[39] 질의문: 이것은 바이러스에 감염된 동물 세포가 생성하는 당단백질이다. 바이러스의 감염과 증식을 저지하는작용을 한다. 유전공학의 발달로 대량 생산되며, B형 감염이나 헤르페스(포진) 따위의 바이러스 질병 치료에 쓰인다.

[40] 정답: 인터페론

[41]

[42] 지식데이터베이스의 온톨로지는 이것(인터페론)이 "당단백질(glycoprotein)"이라는 타입(type)을 구조화된 정보(RDF)로 표현할 수 있다. 하지만, 구조화되지 않은 질의문에서 "감염된", "생성하는", "저지하는", "작용을 한다", "대량 생산되며", "치료에 쓰인다" 등의 술어가 중요한 정보이지만 지식표현언어로 이들을 표현하기 어렵다.

[43] 본 발명은 의미표현언어를 이용하여 지식의 표현력을 높인다. 여기서, 의미표현언어는 술어(Property/Predicate)와 논항(Argument)의 관계를 기초로 문장의 의미를 표현하는 언어이다. 술어-논항 구조란 술어가 문장을 구성하면서 요구하는 논항들의 관계를 나타낸다. 논항의 수는 술어에 따라 결정된다. 어느 술어는 절이나 문장을 만들기 위해 한 개의 필수적인 논항을 요구할 수 있고, 어느 술어는 두 개 또는 세 개의 논항을 요구할 수 있다.

[44] 의미표현언어는 디비피디아 온톨로지로는 표현하기 어려운 특정 개체에 대한 원인, 결과, 의견, 행동, 상태 등에 대해 기술할 수 있다. 예를 들면, 술어-논항 구조는 프레임넷(FrameNet)을 이용해 추출될 수 있으나, 이에 한정되지 않는다. 프레임넷은 문장에서 어휘들이 어떻게 사용되는가를 시멘틱 프레임(Semantic-Frame)의 형태로 어노테이션하여 구축된 언어자원이다.

[45] 도 1을 참고하면, 질의문은 RDF 구조(structure)의 프레임넷 구조 그래프로 표현될 수 있다. 이와 같이, 질의문은 술어-논항 구조로 표현될 수 있다. 예를 들어, "감염된"은 프레임넷의 "Influence_of_event_on_cognizer"로 표현될 수 있고, "생성하는(생성하다)"은 프레임넷의 "Creating"으로 표현될 수 있고, "저지하는(저지하다)"은 프레임넷의 "Intercepting"으로 표현될 수 있으며, "치료하다"는 프레임넷의 "Cure"으로 표현될 수 있다.

[46] 도 2는 본 발명의 한 실시예에 따른 지식표현 확장 장치의 블록도이고, 도 3은 본 발명의 한 실시예에 따른 술어-논항 구조를 해석한 결과를 설명하는 예시 도면이며, 도 4는 본 발명의 한 실시예에 따른 삼항 관계 지식표현 구조를 설명하는 예시 도면이다.

[47] 도 2를 참고하면, 지식표현 확장 장치(앞으로 "장치"라고 한다)(100)는 텍스트

입력부(110), 술어-논항 구조 해석부(130), 지식표현 온톨로지부(150) 그리고 지식 표현부(170)를 포함한다.

[48] 텍스트 입력부(110)는 적어도 하나의 문장을 포함하는 텍스트를 입력받는다.

[49] 술어-논항 구조 해석부(130)는 의미표현언어를 기초로 텍스트를 술어와 적어도 하나의 논항으로 구분한다. 의미표현언어는 문장의 임의 단어(예를 들면, 술어에 해당하는 단어)에 반드시 있어야 하는 적어도 하나의 논항을 지정하고, 술어-논항 구조를 이용하여 문장의 의미를 표현한다. 도 3을 참고하면, 술어-논항 구조 해석부(130)는 텍스트에서 술어(술어.L)를 찾고, 술어에 해당하는 적어도 하나의 논항(논항1 ~ 논항n)을 찾는다. 이때, 술어-논항 구조 해석부(130)는 논항 각각의 어휘 타입(T.1 ~ T.n)을 출력할 수 있다. 예를 들어, 의미표현언어는 프레임넷(FrameNet)일 수 있다. 프레임넷을 이용하여 술어-논항 구조를 해석하는 경우, 술어-논항 구조 해석부(130)는 문장에서 프레임 대상(Frame target)을 파악하고 프레임 요소(Frame element)를 찾는다. 여기서, 프레임 대상이 문장의 술어에 해당하고, 프레임 요소가 술어에 관계된 논항에 해당한다. 술어-논항 구조 해석부(130)는 프레임넷 해석 결과에 대한 주석(annotation) 텍스트를 출력할 수 있다.

[50] 지식표현 온톨로지부(150)는 컴퓨터가 이해할 수 있는 구조화된 포맷으로 지식을 표현한다. 이를 위해, 지식표현 온톨로지부(150)는 지식표현언어를 이용하여 지식 요소들의 속성들을 설명한다. 예를 들면, 지식표현언어는 RDF(Resource Description Framework)일 수 있고, 지식은 RDF 트리플, 즉 삼항 관계<S,P,O>로 표현된다. 지식표현 온톨로지부(150)는 미리 정의된 삼항 관계로 텍스트를 표현한다. 도 4를 참고하면, 지식표현언어는 RDF일 수 있고, <도메인(Domain, D), 술어(프레디케이트), 범위(Range, R)>로 표현될 수 있다. 여기서, 도메인(D)은 술어에 관계된 도메인의 클래스로서, 삼항 관계에서 서브젝트(Subject)의 클래스에 해당한다. 범위(R)는 술어에 관계된 범위의 클래스로서, 삼항 관계에서 오브젝트(object)의 클래스에 해당한다. 예를 들어, 디비피디아 온톨로지는 문장("철수는 한국에서 1944년에 태어났다.")으로부터 <사람:"철수", dbo:birthPlace, 장소:"한국">와 <사람:"철수", dbo:birthDay, 시간:"1944년">를 지식표현 삼항 관계로 추출할 수 있다.

[51] 지식 표현부(170)는 텍스트의 술어-논항 구조를 지식표현 온톨로지부(150)의 포맷으로 변환한다. 지식 표현부(170)는 지식표현의 유사성을 비교하여 술어-논항 구조 해석부(130)에서 해석된 지식이 지식표현 온톨로지부(150)의 포맷으로 표현될 수 있는지 판단한다. 술어-논항 구조 해석부(130)에서 해석된 지식이 지식표현 온톨로지부(150)의 포맷으로 충분히 표현 가능한 경우, 지식 표현부(170)는 지식표현 온톨로지부(150)의 포맷으로, 텍스트로부터 지식을 추출한다. 만약, 술어-논항 구조 해석부(130)에서 해석된 지식이 지식표현 온톨로지부(150)의 포맷으로 충분히 표현 불가능한 경우, 지식 표현부(170)는 술어-논항 구조 해석부(130)에서 해석된 지식을 이용하여 텍스트를 표현한다.

따라서, 지식 표현부(170)는 미리 정의된 삼항 관계로 텍스트의 의미를 제대로 표현하기 어려운 경우, 의미표현언어를 기초로 텍스트로부터 지식을 추출한다. 그리고, 지식 표현부(170)는 의미표현언어를 이용하여 생성한 속성(온톨로지 인스턴스, 술어에 해당함)을 지식표현 온톨로지부(150)에 전달할 수 있다. 지식표현 온톨로지부(150)는 의미표현언어를 이용하여 생성한 정보(온톨로지 인스턴스)를 지식표현언어에 추가할 수 있다.

- [52] 이와 같이, 지식표현 확장 장치(100)는 의미표현언어를 이용하여 지식표현 온톨로지의 지식표현을 확장할 수 있다.
- [53] 도 5는 본 발명의 한 실시예에 따른 지식표현 확장 방법의 흐름도이다.
- [54] 도 5를 참고하면, 장치(100)는 적어도 하나의 문장을 포함하는 텍스트를 입력받는다(S110).
- [55] 장치(100)는 의미표현언어를 기초로 텍스트를 술어와 적어도 하나의 논항으로 표현한다(S120). 장치(100)는 도 3과 같이 텍스트에서 술어(술어.L)와 술어의 논항(논항1 ~ 논항n)을 찾는다. 이때, 장치(100)는 논항 각각의 어휘 타입(T.1 ~ T.n)을 출력할 수 있다.
- [56] 장치(100)는 지식표현 온톨로지에서, 의미표현언어로 추출된 술어(술어.L)에 대응하는 술어(술어.K)를 추출한다(S130). 장치(100)는 텍스트의 술어-논항 구조를 지식표현언어의 삼항 관계로 매칭한다. 장치(100)는 술어-논항 구조 해석 결과에 따라, 삼항 관계 지식표현의 도메인(D)과 범위(R)에 해당하는 논항들이 확보되면, 도 4와 같이 도메인(D)과 범위(R)에 해당하는 술어(술어.K)를 추출할 수 있다. 장치(100)는 논항의 어휘 타입과 같거나 유사한 도메인(D)과 범위(R)를 찾을 수 있다.
- [57] 장치(100)는 의미표현언어로 추출된 술어(술어.L)와 지식표현언어의 술어(술어.K)의 유사도를 판단한다(S140). 이때, 장치(100)는 의미표현언어로 추출된 술어(술어.L)와 논항의 어휘 타입이 결합된 문자열과, 지식표현언어의 술어(술어.K)의 유사도를 판단할 수 있다.
- [58] 유사도를 판단하는 방법으로는, 1) 문자열 수준에서의 유사도 (Edit distance), 2) 단어 의미상의 유사도 (언어 자원을 사용한 개념 계층구조를 활용한 유사도 측정), 3) 코퍼스 기반의 단어 유사도 측정 등의 방법이 있다. 1) 문자열 수준에서의 유사도를 측정하기 위해서, 하나의 문자열이 대상 문자열로 변환하기 위해서 거치는 편집 작업의 수를 계산하는 방식이 있으며, 전통적으로는 레벤슈타인 편집거리 (Levenshtein Distance) 와 같은 방법이 있다. 2) 단어 의미상의 유사도는 워드넷 (WordNet)과 같은 의미 어휘 데이터베이스를 사용하여, 어휘 사이의 유사도를 계층 구조 속에서 거리를 측정하여 계산한다. 전통적으로는 경로 유사도 (Path similarity) 와 같이 워드넷 계층 구조의 노드 사이의 최소 거리를 측정하는 방법, Leacock & Chodorow 유사도와 같이 노드 사이의 최소 거리 및 최대 깊이를 측정하는 방법, Wu & Palmer 유사도와 같이 노드의 깊이 및 노드간의 최소 상위 노드와의 거리를 활용하는 방법 등이 있다.

3)의 코퍼스 기반의 단어 유사도 측정의 경우, 코퍼스 내에서 각각의 단어가 차원 공간에서 특정 벡터값을 갖도록 계산하여, 유사한 벡터공간의 단어들간의 유사도를 측정하는 방법이다. 최근에는 워드 임베딩(Word embedding)을 사용한 접근법이 사용되고 있다.

- [59] 유사한 경우, 장치(100)는 이미 저장된 지식표현언어를 이용하여 텍스트로부터 지식을 추출한다(S150). 의미표현언어로 해석된 지식이 지식표현 온톨로지의 포맷으로 충분히 표현 가능하기 때문에, 장치(100)는 지식표현언어의 포맷으로 텍스트의 지식을 표현한다. 즉, 장치(100)는 의미표현언어로 추출된 술어(술어.L)가 지식표현언어의 술어(술어.K)와 기준값 이상으로 유사하므로, 지식표현을 확장할 필요 없이, 지식표현언어의 포맷으로 입력 텍스트를 충분히 표현할 수 있다고 판단한다. 지식은 <도메인(D)에 해당하는 어휘, 술어.K, 범위(R)에 해당하는 어휘>로 표현될 수 있다.
- [60] 유사하지 않은 경우, 장치(100)는 의미표현언어로 추출된 술어(술어.L)를 포함하는 술어를 생성한다(S160).
- [61] 장치(100)는 생성한 술어를 이용하여 텍스트로부터 지식을 추출한다(S170). 즉, 장치(100)는 텍스트를 지식표현 온톨로지에 존재하는 삼항 관계로 표현할 수 있으면, 저장된 지식표현 온톨로지 기반으로 입력 텍스트를 표현하고, 지식표현 온톨로지로 표현할 수 없는 경우, 입력 텍스트를 술어-논항 구조의 술어를 이용하여 확장된 삼항 관계로 표현한다. 지식은 <도메인(D)에 해당하는 어휘, 술어.L, 범위(R)에 해당하는 어휘> 또는 <도메인(D)에 해당하는 어휘, 술어.L + 범위(R)에 해당하는 어휘 타입, 범위(R)에 해당하는 어휘>로 표현될 수 있다.
- [62] 장치(100)는 생성한 술어를 지식표현 온톨로지에 추가한다(S180). 생성한 술어는 새로운 지식표현 인스턴스로 추가된다.
- [63] 다음에서, 예시문("철수는 한국에서 1944년에 태어났다.")에서 지식을 추출하는 방법을 예로 들어 설명한다.
- [64] 도 6은 본 발명의 한 실시예에 따른 지식표현 확장 방법을 예시하는 흐름도이고, 도 7은 본 발명의 한 실시예에 따른 예시문의 술어-논항 구조를 해석한 결과를 설명하는 도면이며, 도 8은 본 발명의 한 실시예에 따른 예시문의 삼항 관계 지식 표현 구조를 설명하는 도면이다.
- [65] 도 6을 참고하면, 장치(100)는 텍스트("철수는 한국에서 1944년에 태어났다.")를 입력받는다(S210).
- [66] 장치(100)는 도 7과 같이, 의미표현언어를 기초로 텍스트를 술어와 논항으로 분류한다(S220). 술어("태어났다")에 대한 논항이 "누가", "언제", "어디에서"인 경우, 논항에 해당하는 문자열은 "철수", "한국", "1944년"이다. 프레임넷을 이용하는 경우, 프레임 대상은 "태어났다"이고, 프레임 술어 클래스(Class)는 "being_born"이다. 프레임 술어 클래스("being_born")에 대한 프레임 논항은 "Child", "Place", "Time"으로 정해져 있으므로, 프레임 논항-문자열 쌍은 Child-철수, Place-한국, Time-1944년이다. 그리고 논항에 대한 어휘 타입도

정해져 있으며, "Child"의 어휘 타입은 "사람(people)"이고, "Place"의 어휘 타입은 "장소(place)"이며, "Time"의 어휘 타입은 "시간(time)"일 수 있다.

- [67] 장치(100)는 논항과 삼항 관계의 도메인을 비교하여, 논항들 중에서 삼항 관계의 도메인에 매칭되는 논항을 추출한다(S230). 장치(100)는 논항들의 어휘 타입과 유사한 삼항 관계의 도메인을 찾을 수 있다. 장치(100)는 술어-논항 구조를 삼항 관계로 변환하기 위해 논항에 관계된 도메인/범위를 찾는데, 논항-도메인 유사도 측정을 먼저 할 수 있다. 장치(100)는 논항의 어휘 타입 중 "사람"이 삼항 관계의 도메인인 "people"과 유사하다고 판단할 수 있다.
- [68] 장치(100)는 논항과 삼항 관계의 범위를 비교하여, 논항들 중에서 삼항 관계의 범위에 매칭되는 논항을 추출한다(S240). 장치(100)는 논항의 어휘 타입 중 "시간"이 삼항 관계의 범위인 "Time"과 유사하다고 판단할 수 있다.
- [69] 장치(100)는 삼항 관계 지식표현에서 요구되는 서브젝트(도메인)와 오브젝트(범위)를 추출했으므로, 서브젝트(도메인)와 오브젝트(범위)에 관계된 술어(프레디케이트)를 추출한다(S250). 도 8을 참고하면, 도메인 "people"과 범위 "Time"에 관계된 술어(프레디케이트)는 "birthday"이다.
- [70] 장치(100)는 의미표현언어의 술어("being_born")와 삼항 관계의 술어("birthday")의 유사도를 측정한다(S260). 이때, 장치(100)는 술어("being_born")에 관련 논항/관련 논항의 어휘타입/관련 범위인 "time"을 결합하여, 결합한 문자열("being_bornTime")을 생성하고, "being_bornTime"과 "birthday"를 비교할 수 있다.
- [71] 술어가 유사한 경우, 장치(100)는 삼항 관계의 술어("birthday")를 이용하여 텍스트로부터 추출된 지식을 표현한다(S270). 텍스트로부터 추출된 지식은 <철수, birthday, 1994년>일 수 있고, "철수"와 "1994년"은 URI가 링크될 수 있다.
- [72] 술어가 유사하지 않은 경우, 장치(100)는 의미표현언어의 술어("being_born")를 이용하여 텍스트로부터 추출된 지식을 표현한다(S280). 즉, 장치(100)는 지식표현언어에서 현재 정의된 술어("birthday")가 문장의 의미를 충분히 표현하지 못하므로, 삼항 관계의 술어 대신, 의미표현언어의 술어를 이용한다. 여기서 새롭게 생성된 술어는 "being_born"을 포함하는 문자열일 수 있고, 예를 들면 "being_bornTime"일 수 있다. 텍스트로부터 추출된 지식은 확장된 삼항 관계로 표현되고, 예를 들면 <철수, being_born, 1994년> 또는 <철수, being_bornTime, 1994년>일 수 있다. "철수"와 "1994년"은 URI가 링크될 수 있다.
- [73] 장치(100)는 새로운 술어를 도메인 "people"과 범위 "Time"에 관계된 술어로 저장한다. 여기서 새로운 술어는 "being_born"을 포함하는 문자열이고, 예를 들면 "being_bornTime"일 수 있다.
- [74] 지식표현언어에서 현재 정의된 술어("birthday")는 "1944년"과 유사한 시간 정보를 포함하나, "1944년"은 태어난 해일 뿐, "birthday"는 아니므로 불충분한 지식을 표현할 수 있다. 따라서, 장치(100)는 "birthday" 대신, "being_born" 또는 더 상세히 "being_bornTime"을 술어로 교체할 수 있다.

- [75] 이와 같이, 장치(100)는 지식표현언어의 한정된 표현력을 의미표현언어를 이용하여 자동을 확장할 수 있고, 이를 통해, 좀더 정확한 지식을 추출할 수 있는 지식표현언어를 구축할 수 있다.
- [76] 한편, 장치(100)는 논항의 어휘 타입 중 "장소"가 삼항 관계의 범위인 "Place"와 유사하다고 판단할 수 있다. 도메인 "people"과 범위 "Place"에 관계된 술어(프레디케이트)는 "birthplace"이다. 장치(100)는 위에서 설명한 방법과 같은 방법으로, "birthplace"를 그대로 이용하거나, "being_bornPlace" 등으로 확장된 술어를 이용하여 지식을 추출할 수 있다.
- [77] 장치(100)는 디비피디아 뿐만 아니라, 온톨로지 기반의 지식데이터베이스의 지식표현력을 확장할 수 있다. 장치(100)는 프레임넷과 같이 문장의 어느 단어에 대한 분류가 지정된 포맷으로 온톨로지화 되어 있고, 단어에 관계된 논항들이 지정되어 있는 의미표현언어에 확장될 수 있다.
- [78] 이와 같이, 본 발명의 실시예에 따르면 지식표현 온톨로지에서 사용 중인 지식표현언어로 어느 텍스트로부터 추출한 지식을 표현할 수 없는 경우, 의미표현언어를 이용하여 지식표현을 확장할 수 있다. 즉, 본 발명에 실시예에 따르면 지식표현 온톨로지가 웹 텍스트로부터 지식데이터베이스를 구축할 때 충분한 커버리지를 갖지 못하는 문제를 해결할 수 있다.
- [79] 본 발명의 실시예에 따르면 문장 의미 술어-논항 구조 기반으로 자연 언어와 같은 비구조 데이터에 포함된 지식을 컴퓨터가 이해할 수 있는 포맷의 지식표현언어로 표현하여 지식데이터베이스를 빠르고 쉽게 확장할 수 있다.
- [80] 본 발명의 실시예에 따르면 지식데이터베이스의 "관계" 온톨로지가 확충되어, 지식표현력을 높일 수 있고, CGC(Collaboratively Generated Content) 지향 지식 형태 및 해석에 적용될 수 있다.
- [81] 지식표현 확장 장치(100)는 도 1부터 도 8을 참고로 설명한 지식표현 확장 방법을 수행하기 위한 명령어(instructions)를 저장하고 있거나, 저장 장치로부터 명령어를 로드하여 일시 저장하는 메모리, 메모리에 저장되어 있거나 로드된 명령어를 실행하여 본 발명의 지식표현 확장 방법을 처리하는 프로세서, 그리고 통신장치를 포함한다. 도 1부터 도 8을 참고로 설명한 지식표현 확장 방법을 수행하기 위한 명령어(instructions)는 프로세서가 처리할 수 있는 프로그램으로 구현된다.

발명의 실시를 위한 형태

- [82] 이상에서 설명한 본 발명의 실시예는 장치 및 방법을 통해서만 구현이 되는 것은 아니며, 본 발명의 실시예의 구성에 대응하는 기능을 실현하는 프로그램 또는 그 프로그램이 기록된 기록 매체를 통해 구현될 수도 있다.
- [83] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의

권리범위에 속하는 것이다.

청구범위

- [청구항 1] 지식표현 확장 장치로서,
 의미표현언어를 이용하여 텍스트에서 술어와 적어도 하나의 논항을
 추출하는 술어-논항 구조 해석부,
 컴퓨터가 이해할 수 있는 구조화된 포맷인 지식표현언어를 이용하여
 지식을 표현하는 온톨로지부, 그리고
 상기 온톨로지부에서 상기 술어-논항 구조 해석부에서 추출된 제1술어에
 대응하는 제2술어를 추출하고, 상기 제1술어와 상기 제2술어의 유사도가
 기준값 이하인 경우, 상기 제1술어를 이용하여 상기 텍스트로부터 추출된
 지식을 표현하는 지식 표현부
 를 포함하는 지식표현 확장 장치.
- [청구항 2] 제1항에서,
 상기 지식 표현부는
 상기 온톨로지부에서 상기 적어도 하나의 논항에 관계된 상기 제2술어를
 추출하는 지식표현 확장 장치.
- [청구항 3] 제2항에서,
 상기 지식 표현부는
 상기 지식표현언어의 도메인들 중에서 상기 적어도 하나의 논항에
 부여된 어휘 타입과 기준값이상으로 유사한 제1도메인을 추출하고, 상기
 지식표현언어의 범위들 중에서 상기 적어도 하나의 논항에 부여된 어휘
 타입과 기준값이상으로 유사한 제1범위를 추출하며, 상기 제1도메인과
 상기 제1범위에 관련된 술어를 상기 제2술어로 추출하는 지식표현 확장
 장치.
- [청구항 4] 제3항에서,
 상기 지식 표현부는
 상기 제1술어와 상기 적어도 하나의 논항 중 임의 논항에 관련된 정보가
 결합된 문자열을 생성하고, 상기 문자열을 상기 온톨로지부의
 지식표현언어에 추가하는 지식표현 확장 장치.
- [청구항 5] 제1항에서,
 상기 지식표현언어는 RDF(Resource Description Framework) 삼항 관계로
 표현되는 언어인 지식표현 확장 장치.
- [청구항 6] 장치가 지식표현을 확장하는 방법으로서,
 적어도 하나의 문장을 포함하는 텍스트를 입력받는 단계,
 의미표현언어를 기초로 상기 텍스트를 제1술어와 적어도 하나의
 논항으로 표현하는 단계,
 지식표현 온톨로지부에서, 상기 제1술어에 대응하는 제2술어를 추출하는
 단계,

상기 제1술어와 상기 제2술어의 유사도를 비교하는 단계, 그리고
상기 유사도가 기준값 이하인 경우, 상기 텍스트로부터 추출된 지식을
상기 제1술어를 이용하여 표현하는 단계
를 포함하는 지식표현 확장 방법.

[청구항 7] 제6항에서,
상기 제1술어에 대응하는 제2술어를 추출하는 단계는
상기 적어도 하나의 논항에 부여된 어휘 타입을 이용하여 상기 지식표현
온톨로지에서 상기 제1술어에 대응하는 상기 제2술어를 추출하는
지식표현 확장 방법.

[청구항 8] 제6항에서,
상기 지식표현 온톨로지는 지식을 서브젝트(subject),
프레디케이트(Predicate), 오브젝트(object)의 삼항 관계로 표현하는
지식표현언어를 사용하고,
상기 제1술어에 대응하는 제2술어를 추출하는 단계는
상기 지식표현언어의 서브젝트들 중에서 상기 적어도 하나의 논항에
부여된 어휘 타입과 기준값이상으로 유사하고, 상기 지식표현언어의
오브젝트들 중에서 상기 적어도 하나의 논항에 부여된 어휘 타입과
기준값이상으로 유사한 프레디케이트를 상기 제2술어로 추출하는
지식표현 확장 방법

[청구항 9] 제6항에서,
상기 제1술어를 이용하여 표현하는 단계는
상기 제1술어와 상기 적어도 하나의 논항 중 임의 논항에 관련된 정보가
결합된 문자열을 생성하고, 상기 문자열을 이용하여 상기 텍스트로부터
추출한 지식을 표현하는 지식표현 확장 방법

[청구항 10] 제9항에서,
상기 문자열을 상기 지식표현 온톨로지의 지식표현언어에 추가하는 단계
를 더 포함하는 지식표현 확장 방법

[청구항 11] 장치가 지식표현을 확장하는 방법으로서,
텍스트의 술어-논항 구조를 해석하는 단계,
상기 텍스트의 술어-논항 구조를 지식표현언어의 삼항 관계로 매칭하는
단계, 그리고
매칭 유사도를 기초로 상기 텍스트의 술어-논항 구조에서 추출된
제1술어를 상기 지식표현언어의 술어로 추가하는 단계
를 포함하는 지식표현 확장 방법.

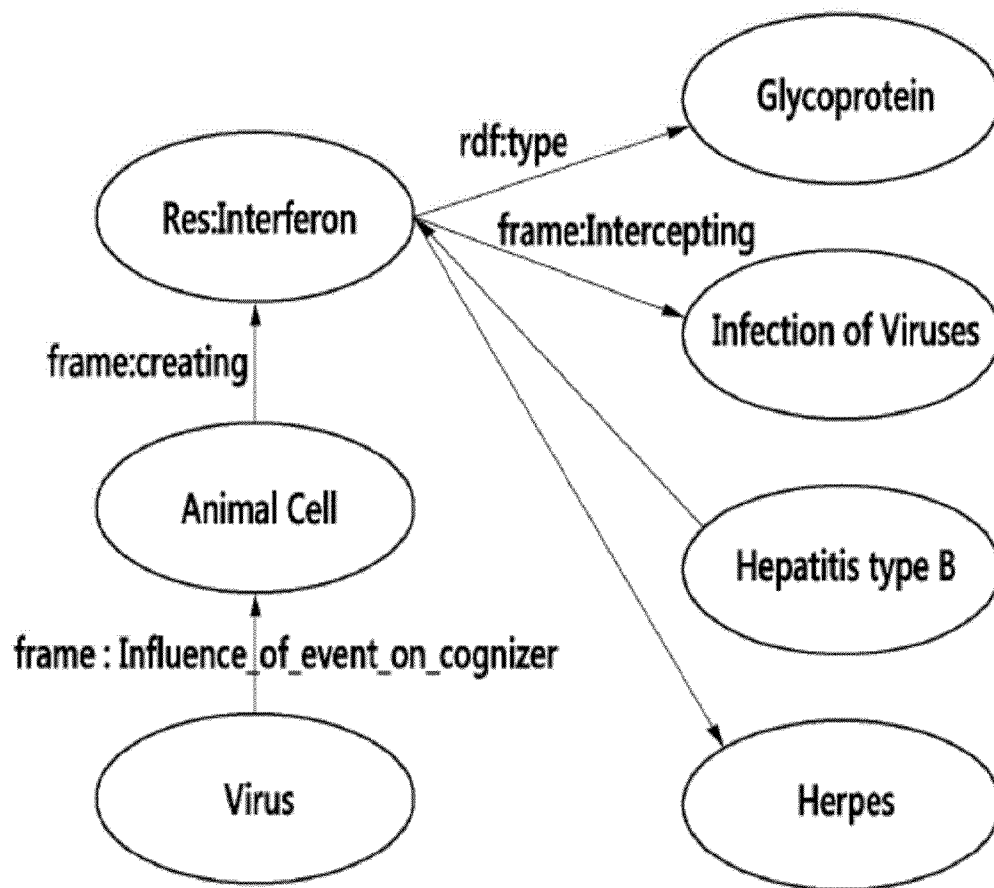
[청구항 12] 제11항에서,
상기 지식표현언어의 술어로 추가하는 단계는
상기 지식표현언어의 삼항 관계에서, 상기 텍스트의 술어-논항 구조의
제1술어에 매칭된 제2술어를 추출하는 단계,

상기 제1술어와 상기 제2술어의 유사도를 비교하는 단계, 그리고
상기 유사도가 기준값 이하인 경우, 상기 제1술어를 상기 지식표현언어에
추가하는 단계
를 포함하는 지식표현 확장 방법.

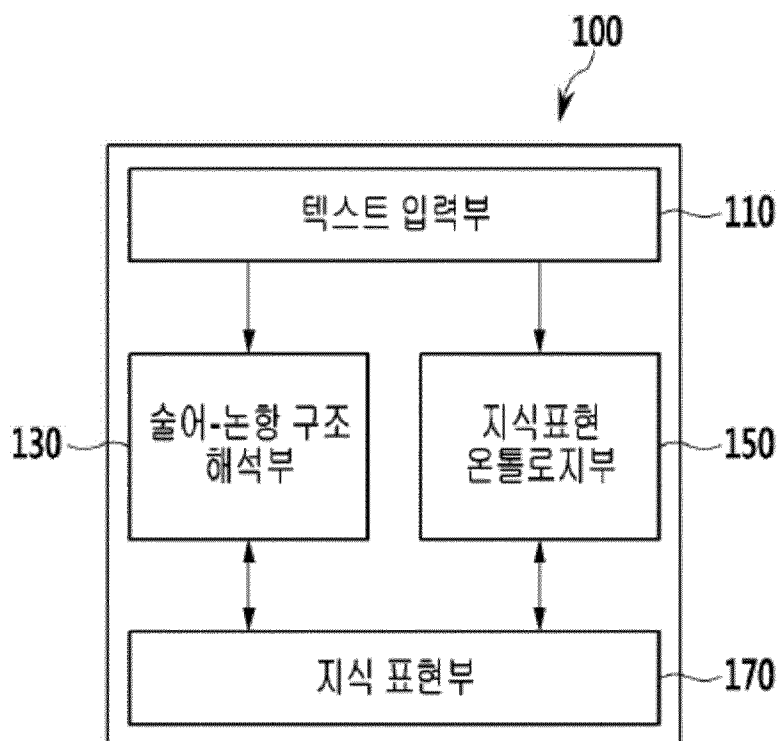
[청구항 13] 제11항에서,
상기 제1술어를 이용하여 상기 텍스트를 삼항 관계로 표현하는 단계
를 더 포함하는 지식표현 확장 방법.

[청구항 14] 제11항에서,
상기 지식표현언어의 삼항 관계로 매칭하는 단계는
상기 텍스트의 술어-논항 구조에서 추출된 논항들과 상기 삼항 관계의
도메인 및 범위의 유사도를 기초로 상기 텍스트의 술어-논항 구조를 상기
삼항 관계로 매칭하는 지식표현 확장 방법.

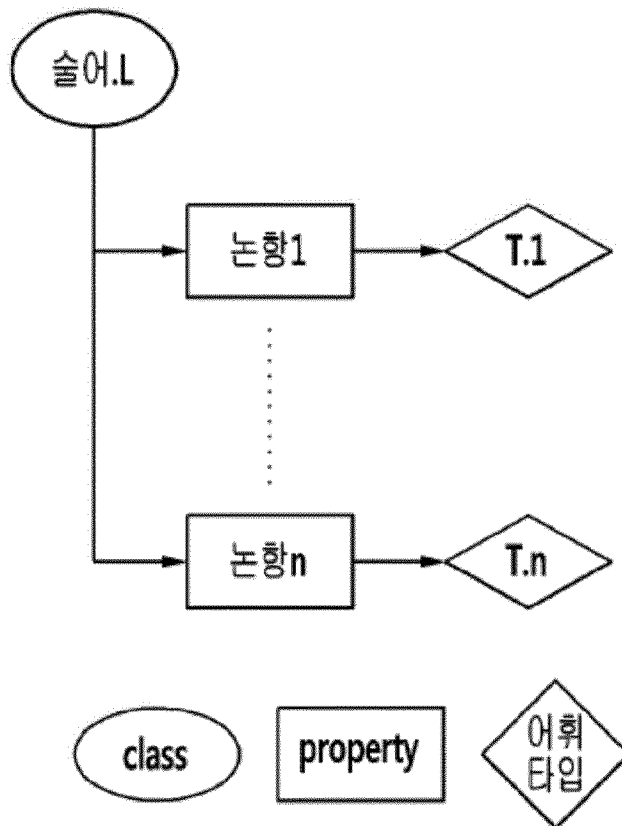
[도1]



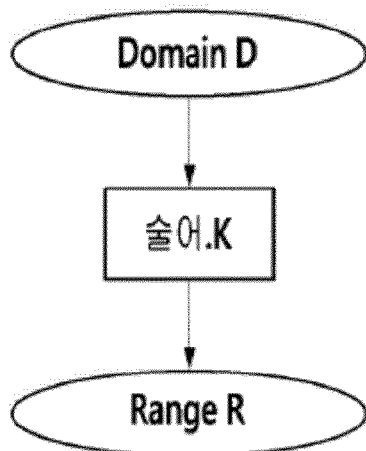
[도2]



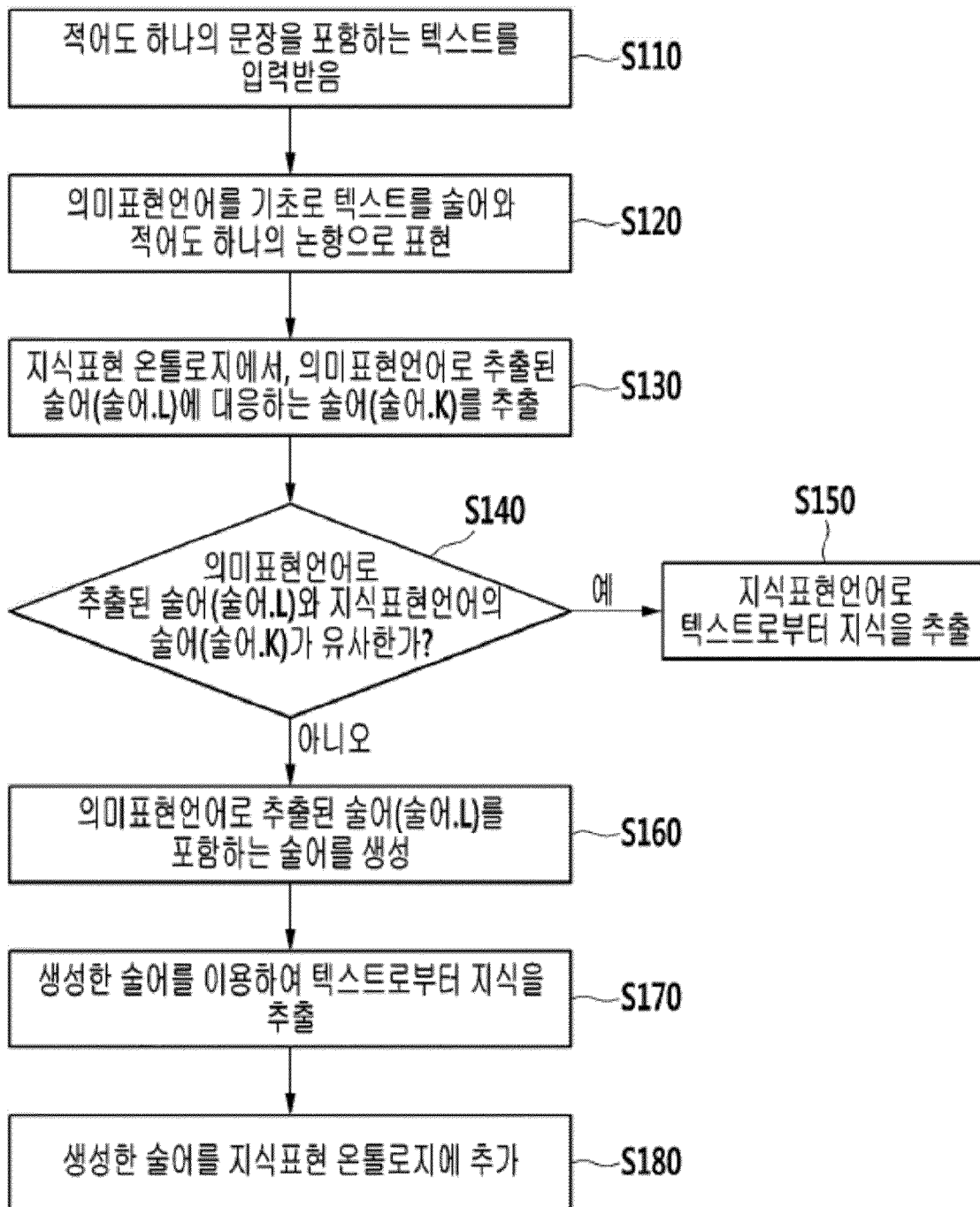
[도3]



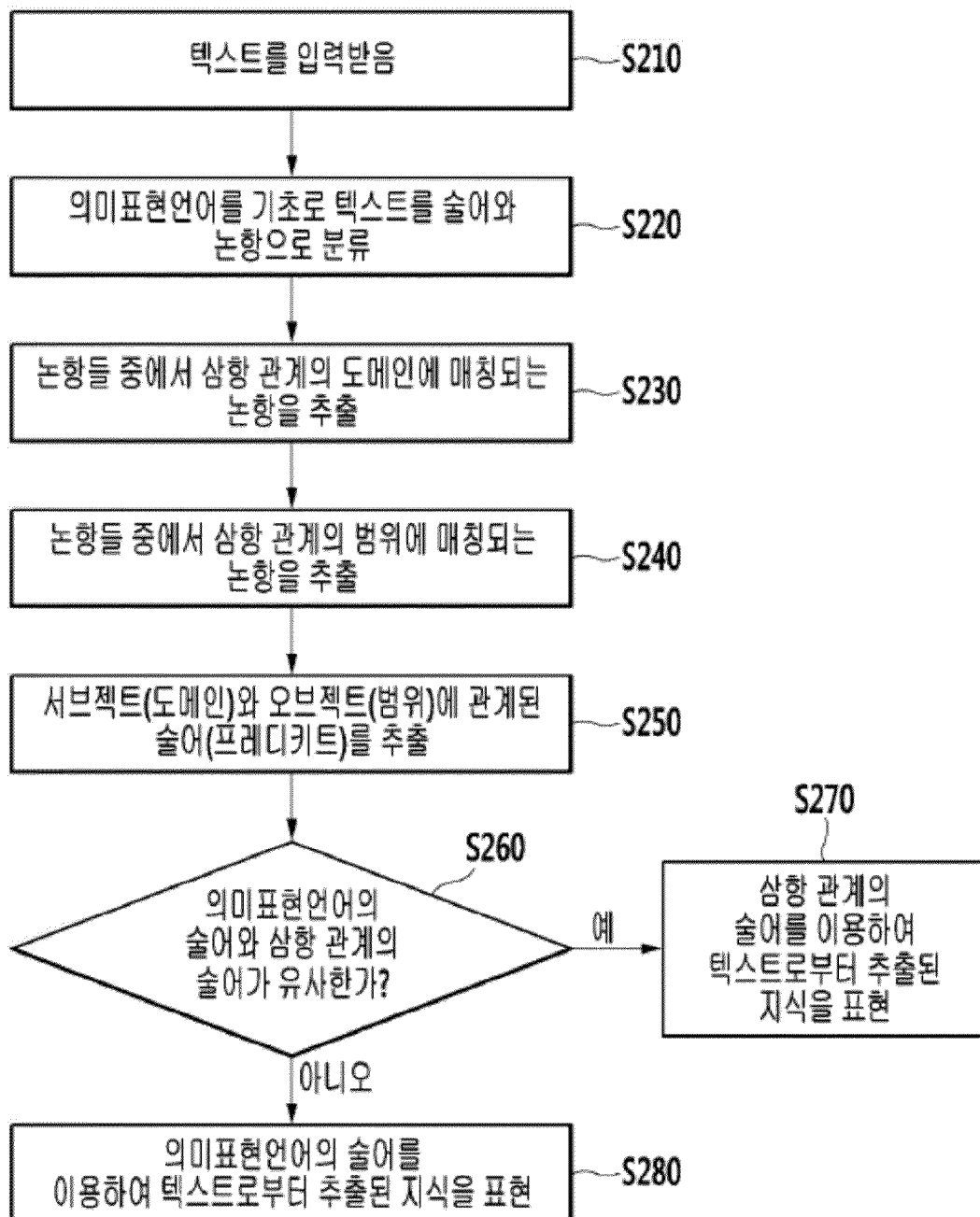
[도4]



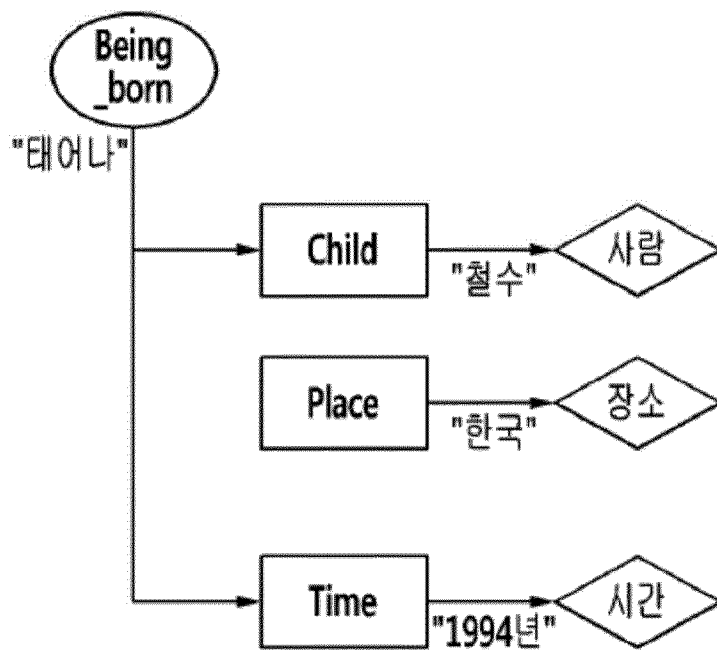
[도5]



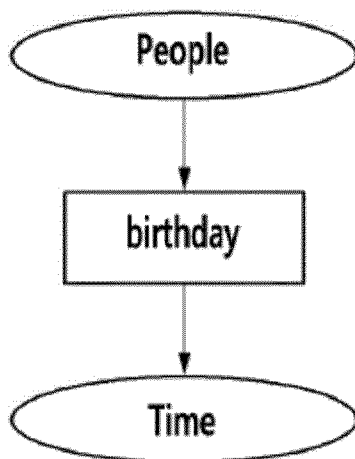
[도6]



[도7]



[도8]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2016/000579**A. CLASSIFICATION OF SUBJECT MATTER****G06F 17/30(2006.01)i, G06F 17/20(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 17/30; G06F 17/20; G06F 17/00; G06F 7/02; G06F 7/00; G06F 12/00; G06F 17/27

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Utility models and applications for Utility models: IPC as above

Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: predicate-argument calculus, knowledge representation, ontology, comparison, RDF.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2014-0052328 A (SK TELECOM CO., LTD.) 07 May 2014 See paragraphs [0010] and [0013]-[0014]; claims 1, 4 and 6; and figure 1.	1-14
A	US 2011-0153673 A1 (BOSCHEE, Elizabeth Megan et al.) 23 June 2011 See paragraphs [0036]-[0037] and [0039]; claims 1 and 19; and figure 2.	1-14
A	US 2004-0111408 A1 (CAUDILL, Maureen et al.) 10 June 2004 See paragraphs [0014], [0103], [0111] and [0180]; and figures 4 and 7.	1-14
A	US 2013-0232143 A1 (AIT-MOKHTAR, Salah) 05 September 2013 See paragraphs [0039]-[0045] and [0055]-[0066]; and figures 1-2.	1-14
A	US 2011-0225167 A1 (BHATTACHARJEE, Bishwaranjan et al.) 15 September 2011 See paragraphs [0077]-[0083] and figure 11.	1-14
A	WO 2011-129481 A1 (KOREA INSTITUTE OF SCIENCE & TECHNOLOGY INFORMATION) 20 October 2011 See paragraphs [0042]-[0055] and [0057]-[0066]; and figures 1-2.	1-14



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

18 APRIL 2016 (18.04.2016)

Date of mailing of the international search report

19 APRIL 2016 (19.04.2016)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2016/000579

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2014-0052328 A	07/05/2014	NONE	
US 2011-0153673 A1	23/06/2011	US 2009-0100053 A1 US 7890539 B2 US 8260817 B2	16/04/2009 15/02/2011 04/09/2012
US 2004-0111408 A1	10/06/2004	AU 2002-235314 A1 US 2002-0129015 A1 US 6766316 B2 US 7496561 B2 WO 02-057961 A2 WO 02-057961 A3	30/07/2002 12/09/2002 20/07/2004 24/02/2009 25/07/2002 09/10/2003
US 2013-0232143 A1	05/09/2013	US 8527517 B1	03/09/2013
US 2011-0225167 A1	15/09/2011	US 8458191 B2	04/06/2013
WO 2011-129481 A1	20/10/2011	EP 2378467 A1 KR 10-0978740 B1 US 2011-0258230 A1	19/10/2011 30/08/2010 20/10/2011

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06F 17/30(2006.01)i, G06F 17/20(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06F 17/30; G06F 17/20; G06F 17/00; G06F 7/02; G06F 7/00; G06F 12/00; G06F 17/27

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 술어-논항 해석, 지식표현, 온톨로지, 비교, RDF.

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-2014-0052328 A (에스케이텔레콤 주식회사) 2014.05.07 단락 [0010]과 [0013]-[0014]; 청구항 1, 4와 6; 및 도면 1 참조.	1-14
A	US 2011-0153673 A1 (ELIZABETH MEGAN BOSCHKE 등) 2011.06.23 단락 [0036]-[0037]과 [0039]; 청구항 1과 19; 및 도면 2 참조.	1-14
A	US 2004-0111408 A1 (MAUREEN CAUDILL 등) 2004.06.10 단락 [0014], [0103], [0111]과 [0180]; 및 도면 4와 7 참조.	1-14
A	US 2013-0232143 A1 (SALAH AIT-MOKHTAR) 2013.09.05 단락 [0039]-[0045]와 [0055]-[0066]; 및 도면 1-2 참조.	1-14
A	US 2011-0225167 A1 (BISHWARANJAN BHATTACHARJEE 등) 2011.09.15 단락 [0077]-[0083] 및 도면 11 참조.	1-14
A	WO 2011-129481 A1 (한국과학기술정보연구원) 2011.10.20 단락 [0042]-[0055]과 [0057]-[0066]; 및 도면 1-2 참조.	1-14

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.

☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“&” 동일한 대응특허문헌에 속하는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일

2016년 04월 18일 (18.04.2016)

국제조사보고서 발송일

2016년 04월 19일 (19.04.2016)

ISA/KR의 명칭 및 우편주소



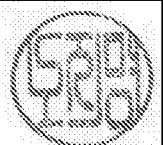
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

노지명

전화번호 +82-42-481-8528



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2014-0052328 A	2014/05/07	없음	
US 2011-0153673 A1	2011/06/23	US 2009-0100053 A1 US 7890539 B2 US 8260817 B2	2009/04/16 2011/02/15 2012/09/04
US 2004-0111408 A1	2004/06/10	AU 2002-235314 A1 US 2002-0129015 A1 US 6766316 B2 US 7496561 B2 WO 02-057961 A2 WO 02-057961 A3	2002/07/30 2002/09/12 2004/07/20 2009/02/24 2002/07/25 2003/10/09
US 2013-0232143 A1	2013/09/05	US 8527517 B1	2013/09/03
US 2011-0225167 A1	2011/09/15	US 8458191 B2	2013/06/04
WO 2011-129481 A1	2011/10/20	EP 2378467 A1 KR 10-0978740 B1 US 2011-0258230 A1	2011/10/19 2010/08/30 2011/10/20