

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2020년 6월 4일 (04.06.2020)

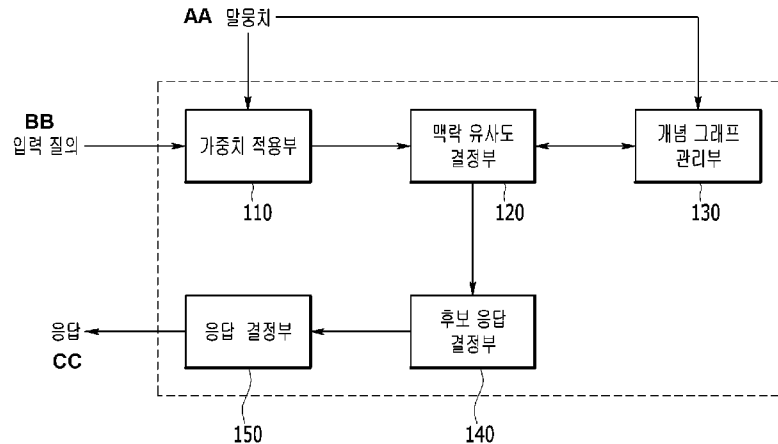


(10) 국제공개번호
WO 2020/111314 A1

- (51) 국제특허분류: *G06F 16/00* (2019.01) *G06F 17/27* (2006.01)
- (21) 국제출원번호: PCT/KR2018/014860
- (22) 국제출원일: 2018년 11월 28일 (28.11.2018)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2018-0148013 2018년 11월 27일 (27.11.2018) KR
- (71) 출원인: 한국과학기술원 (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) [KR/KR]; 34141 대전시 유성구 대학로 291, Daejeon (KR).
- (72) 발명자: 맹성현 (MYAENG, Sung Hyon); 34141 대전시 유성구 대학로 291, Daejeon (KR). 김경민 (KIM, Kyung-Min); 34141 대전시 유성구 대학로 291, Daejeon (KR).
- (74) 대리인: 유미특허법인 (YOU ME PATENT AND LAW FIRM); 06134 서울시 강남구 테헤란로 115, 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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, 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.

(54) Title: CONCEPTUAL GRAPH-BASED QUERY-RESPONSE APPARATUS AND METHOD

(54) 발명의 명칭: 개념 그래프 기반 질의응답 장치 및 방법



110 ... Weight application unit
120 ... Context similarity determination unit
130 ... Conceptual graph management unit
140 ... Candidate response determination unit
150 ... Response determination unit
AA ... Corpus
BB ... Input query
CC ... Response

(57) Abstract: A query-response apparatus comprises: a weight application unit for extracting a corpus context word from a corpus, applying a weight to the corpus context word according to a weight rule, extracting a query context word from an input query, and applying a weight to the query context word according to the weight rule; a context similarity determination unit for generating a plurality of corpus sentence embedding vectors for the weight-applied corpus and a query sentence embedding vector for the weight-applied input query, and determining a target sentence embedding vector from among the plurality of corpus sentence embedding vectors on the basis of vector similarities between the plurality of corpus sentence embedding vectors and the query sentence embedding vector; a candidate response determination unit for determining a target conceptual graph corresponding to the target sentence embedding vector,



WO 2020/111314 A1

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

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

calculating a graph matching score for each of one or more concepts included in the target conceptual graph, and determining one or more candidate concepts corresponding to the input query; and a response determination unit for determining a preset number of concepts from among the candidate concepts as a response to the input query.

(57) 요약서: 질의응답 장치로서, 말뭉치에서 말뭉치 맥락 단어를 추출하고, 가중치 규칙에 따라 상기 말뭉치 맥락 단어에 가중치를 적용하고, 입력 질의에서 질의 맥락 단어를 추출하고, 상기 가중치 규칙에 따라 상기 질의 맥락 단어에 가중치를 적용하는 가중치 적용부, 가중치가 적용된 말뭉치에 대한 복수의 말뭉치 문장 임베딩 벡터들 및 가중치가 적용된 입력 질의에 대한 질의 문장 임베딩 벡터를 생성하고, 상기 복수의 말뭉치 문장 임베딩 벡터들과 상기 질의 문장 임베딩 벡터 사이의 벡터 유사도를 기준으로 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 대상 문장 임베딩 벡터를 결정하는 맥락 유사도 결정부, 상기 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를 결정하고, 상기 대상 개념 그래프에 포함된 적어도 하나 이상의 개념들 각각에 대한 그래프 매칭 점수를 계산하여, 상기 입력 질의에 대응하는 적어도 하나 이상의 후보 개념들을 결정하는 후보 응답 결정부, 그리고 상기 후보 개념들 중에서 미리 설정된 수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정하는 응답 결정부를 포함한다.

명세서

발명의 명칭: 개념 그래프 기반 질의응답 장치 및 방법

기술분야

- [1] 본 발명은 개념 그래프 기반 질의응답 기술에 관한 것이다.

배경기술

- [2] 질의응답을 위해 다양한 방법들이 연구되고 있다.
- [3] 우선, 두 개의 개념 그래프를 이용하여 확장 그래프를 생성하고, 외부로부터 입력되는 질문을 토대로 생성한 질문 그래프와 확장 그래프를 토대로 정답 노드를 찾아 질문에 대한 정답을 도출하는, 개념 그래프 매칭을 이용한 질의응답 방법이 있다. 이 질의응답 방법을 이용하여 질의에 응답하는 경우 질문 그래프와 모든 문서 그래프간 매칭이 필요하기 때문에 오랜 시간이 소요되어, 질의 응답 속도가 느려지는 문제점이 있다.
- [4] 최근에는 질의 및 질의에 대한 응답 문장들로 구성된 말뭉치의 임베딩 벡터를 생성하고, 임베딩 벡터들 사이의 벡터 유사도를 기준으로 질의에 대한 응답을 결정하는 방법이 활용되었다. 그러나, 이러한 방법은 임베딩 벡터를 생성하는 과정에서 미리 추출된 맥락 단어만을 활용하므로, 질의 및 말뭉치에서 맥락 단어와 일반 단어간의 관계성을 활용하지 못해 일반 단어로만 구성된 질의에 대해 말뭉치 내에서 검색 성능이 떨어지는 문제가 있었다.

발명의 상세한 설명

기술적 과제

- [5] 본 발명은 개념 그래프 기반 질의응답 장치에서 임베딩 벡터를 생성하는데 있어, 질의 및 말뭉치 내 문장들을 구성하는 단어들 중 맥락 단어에 높은 가중치를 적용하여 질의에 대한 응답을 제공하는 기술을 제공한다.

과제 해결 수단

- [6] 본 발명의 일 실시예에 따른 질의응답 장치는 말뭉치에서 말뭉치 맥락 단어를 추출하고, 가중치 규칙에 따라 상기 말뭉치 맥락 단어에 가중치를 적용하고, 입력 질의에서 질의 맥락 단어를 추출하고, 상기 가중치 규칙에 따라 상기 질의 맥락 단어에 가중치를 적용하는 가중치 적용부, 가중치가 적용된 말뭉치에 대한 복수의 말뭉치 문장 임베딩 벡터들 및 가중치가 적용된 입력 질의에 대한 질의 문장 임베딩 벡터를 생성하고, 상기 복수의 말뭉치 문장 임베딩 벡터들과 상기 질의 문장 임베딩 벡터 사이의 벡터 유사도를 기준으로 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 대상 문장 임베딩 벡터를 결정하는 맥락 유사도 결정부, 상기 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를 결정하고, 상기 대상 개념 그래프에 포함된 적어도 하나 이상의 개념들 각각에 대한 그래프 매칭 점수를 계산하여, 상기 입력 질의에 대응하는 적어도 하나 이상의 후보 개념들을 결정하는 후보 응답 결정부, 그리고 상기 후보 개념들 중에서 미리 설정된

수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정하는 응답 결정부를 포함한다.

- [7] 상기 가중치 적용부는 텍스트 분석 알고리즘을 이용하여 상기 말뭉치 및 상기 입력 질의로부터 상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어를 추출한다.
- [8] 상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어는 특정 카테고리에 해당하는 단어이고, 상기 가중치 적용부는 상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어에 미리 설정된 가중치를 적용한다.
- [9] 상기 특정 카테고리는 시간 카테고리 또는 장소 카테고리 중 적어도 하나를 포함한다.
- [10] 상기 미리 설정된 가중치는 상기 말뭉치 및 상기 입력 질의에 포함된 단어들 중에서, 상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어에 해당하지 않는 단어에 적용되는 가중치보다 높다.
- [11] 상기 맥락 유사도 결정부는 상기 가중치가 적용된 말뭉치 및 상기 가중치가 적용된 입력 질의에 대해 벡터화 알고리즘을 이용하여 상기 복수의 말뭉치 문장 임베딩 벡터들 및 상기 질의 문장 임베딩 벡터를 생성한다.
- [12] 상기 맥락 유사도 결정부는 벡터 유사도 판단 알고리즘을 이용하여 상기 복수의 말뭉치 문장 임베딩 벡터들과 상기 질의 문장 임베딩 벡터 사이의 벡터 유사도를 각각 계산하고, 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 상기 벡터 유사도가 높은 문장 임베딩 벡터를 상기 대상 문장 임베딩 벡터로 결정한다.
- [13] 상기 질의응답 장치는 상기 말뭉치로부터 개념, 관계 및 속성을 추출하고, 상기 추출한 개념, 관계 및 속성을 이용하여 상기 말뭉치에 대한 복수의 개념 그래프들을 생성하고, 상기 말뭉치 내 동일한 문장에서 비롯된 개념 그래프와 말뭉치 문장 임베딩 벡터를 대응시켜 저장하는 개념 그래프 관리부를 더 포함하고, 상기 후보 응답 결정부는 상기 복수의 개념 그래프들 중에서, 상기 대상 문장 임베딩 벡터와 동일한 문장에서 비롯된 개념 그래프를 상기 대상 개념 그래프로 결정한다.
- [14] 상기 응답 결정부는 상기 복수의 후보 개념들 중에서 상기 그래프 매칭 점수가 높은 상기 미리 설정된 수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정한다.
- [15] 본 발명의 일 실시예에 따른 질의응답 장치가 입력 질의의 응답을 결정하는 방법은 입력 질의에서 추출한 맥락 단어에 가중치 규칙에 따라 가중치를 적용하고, 말뭉치에서 추출한 맥락 단어에 상기 가중치 규칙에 따라 가중치를 적용하는 단계, 가중치가 적용된 입력 질의에 대한 질의 문장 임베딩 벡터를 생성하고, 가중치가 적용된 말뭉치에 대한 복수의 말뭉치 문장 임베딩 벡터들을 생성하는 단계, 상기 질의 문장 임베딩 벡터와 상기 복수의 말뭉치 문장 임베딩 벡터들 사이의 벡터 유사도를 각각 계산하고, 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 상기 벡터 유사도가 높은 문장 임베딩 벡터를 대상 문장 임베딩

벡터로 결정하는 단계, 상기 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를 결정하고, 상기 대상 개념 그래프에 포함된 적어도 하나 이상의 개념들 각각에 대한 그래프 매칭 점수를 계산하여, 상기 입력 질의에 대응하는 적어도 하나 이상의 후보 개념들을 결정하는 단계, 그리고 상기 후보 개념들 중에서 미리 설정된 수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정하는 단계를 포함한다.

- [16] 상기 가중치를 적용하는 단계는 텍스트 분석 알고리즘을 이용하여 상기 입력 질의 및 상기 말뭉치로부터 맥락 단어를 추출하고, 추출한 맥락 단어에 미리 설정된 가중치를 적용한다.
- [17] 상기 추출한 맥락 단어는 시간 정보 카테고리 또는 장소 정보 카테고리 중 적어도 하나에 해당한다.
- [18] 상기 미리 설정된 가중치는 상기 입력 질의 및 상기 말뭉치에 포함된 단어들 중에서, 상기 추출한 맥락 단어에 해당하지 않는 단어에 적용되는 가중치보다 높다.
- [19] 상기 질의 문장 임베딩 벡터 및 상기 복수의 말뭉치 문장 임베딩 벡터들을 생성하는 단계는 상기 가중치가 적용된 입력 질의 및 상기 가중치가 적용된 말뭉치에 대해 벡터화 알고리즘을 이용하여 상기 질의 문장 임베딩 벡터 및 상기 복수의 말뭉치 문장 임베딩 벡터들을 생성한다.
- [20] 상기 벡터 유사도가 높은 임베딩 벡터를 대상 문장 임베딩 벡터로 결정하는 단계는 벡터 유사도 판단 알고리즘을 이용하여 상기 복수의 말뭉치 문장 임베딩 벡터들과 상기 질의 문장 임베딩 벡터 사이의 벡터 유사도를 각각 계산하고, 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 상기 벡터 유사도가 가장 높은 임베딩 벡터를 상기 대상 문장 임베딩 벡터로 결정한다.
- [21] 상기 응답 결정 방법은 상기 대상 개념 그래프를 결정하는 단계 이전에, 상기 말뭉치로부터 개념, 관계 및 속성을 추출하는 단계, 상기 추출한 개념, 관계 및 속성을 이용하여 상기 말뭉치에 대한 복수의 개념 그래프들을 생성하는 단계, 그리고 상기 말뭉치 내 동일한 문장에서 비롯된 개념 그래프와 문장 임베딩 벡터를 대응시켜 저장하는 단계를 더 포함한다.
- [22] 상기 대상 개념 그래프를 결정하는 단계는 상기 복수의 개념 그래프들 중에서, 상기 대상 문장 임베딩 벡터와 동일한 문장에서 비롯된 개념 그래프를 상기 대상 개념 그래프로 결정한다.
- [23] 상기 입력 질의에 대한 응답으로 결정하는 단계는 상기 후보 개념들 중에서 상기 그래프 매칭 점수가 높은 상기 미리 설정된 수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정한다.

발명의 효과

- [24] 본 발명에 따르면, 질의 및 말뭉치 내 문장들을 구성하는 단어들 중 특정 카테고리에 해당하는 맥락 단어에 일반 단어보다 높은 가중치를 적용하는바,

맥락 단어를 일반 단어와 효율적으로 구분할 수 있고, 이를 통해 질의에 대한 응답 결정의 성능을 향상시킬 수 있다.

도면의 간단한 설명

- [25] 도 1은 본 발명의 실시예에 따른 질의응답 장치를 도시한 도면이다.
- [26] 도 2는 가중치 적용부가 문장의 맥락 단어들에 가중치를 적용하고, 맥락 유사도 결정부가 가중치가 적용된 문장을 이용하여 문장 임베딩 벡터를 생성하는 예시를 설명하는 도면이다.
- [27] 도 3은 맥락 유사도 결정부가 문장 임베딩 벡터 사이의 유사도를 계산하는 방법을 설명하는 도면이다.
- [28] 도 4는 개념 그래프 관리부가 개념 그래프들을 생성하는 방법을 설명하는 도면이다.
- [29] 도 5는 질의응답 장치가 입력 질의의 응답을 결정하는 방법을 설명하는 도면이다.

발명의 실시를 위한 형태

- [30] 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.
- [31] 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.
- [32] 본 명세서에서, "맥락 단어"는 질의 또는 말뭉치 내 문장들에서 장소 또는 시간 중 적어도 하나에 대한 정보를 포함하는 단어를 지칭한다. 예를 들면, 문장 "In 1705, who was the first scientist to be knighted?"에서, "맥락 단어"는 시간에 대한 정보를 나타내는 단어 "1705"이고, "맥락 단어"가 아닌 이외의 단어들 "In", ",", "who", "was", "the", "first", "scientist", "to", "be", knighted" 및 "?"는 일반 단어에 해당한다.
- [33] 또한, 본 명세서에서, "문장 임베딩 벡터"는 문장에 대한 함축적인 의미를 나타내는 고정 길이의 k차원의 벡터를 지칭한다. 따라서, 본 명세서에서, "말뭉치 문장 임베딩 벡터"는 말뭉치를 구성하는 문장들에 대한 함축적인 의미를 나타내는 고정 길이의 k차원의 벡터를 지칭하며, "질의 문장 임베딩 벡터"는 문장 형태의 질의에 대한 함축적인 의미를 나타내는 고정 길이의 k차원의 벡터를 지칭한다.
- [34] 이하 도면을 참조로 하여 본 발명의 실시예에 따른 개념 그래프 기반 질의응답 장치 및 이를 이용한 질의응답 방법에 대해 설명한다.

- [35] 도 1은 본 발명의 실시예에 따른 질의응답 장치를 도시한 도면이고, 도 2는 가중치 적용부가 문장의 맥락 단어들에 가중치를 적용하고, 맥락 유사도 결정부가 가중치가 적용된 문장을 이용하여 문장 임베딩 벡터를 생성하는 예시를 설명하는 도면이고, 도 3은 맥락 유사도 결정부가 문장 임베딩 벡터 사이의 유사도를 계산하는 방법을 설명하는 도면이고, 도 4는 개념 그래프 관리부가 개념 그래프들을 생성하는 방법을 설명하는 도면이다.
- [36] 도 1을 참고하면, 질의응답 장치(100)는 적어도 하나의 프로세서에 의해 구동되는 컴퓨팅 디바이스일 수 있으며, 가중치 적용부(110), 맥락 유사도 결정부(120), 개념 그래프 관리부(130), 후보 응답 결정부(140) 및 응답 결정부(150)를 포함한다. 본 발명의 실시예에서는 설명의 편의를 위하여 상기의 구성 요소만을 언급하고 있으나, 질의응답에 필요한 추가 구성 요소들을 포함할 수 있다.
- [37] 가중치 적용부(110)는 말뭉치에서 말뭉치 맥락 단어를 추출하고, 입력 질의에서 질의 맥락 단어를 추출한다.
- [38] 구체적으로, 가중치 적용부(110)는 텍스트 분석 알고리즘을 이용하여 말뭉치 및 입력 질의로부터 말뭉치 맥락 단어 및 질의 맥락 단어를 각각 추출한다.
- [39] 이 경우, 말뭉치 맥락 단어 및 질의 맥락 단어는 특정 카테고리에 해당하는 단어이고, 특정 카테고리는 시간 카테고리 또는 장소 카테고리 중 적어도 하나를 포함할 수 있다.
- [40] 예를 들면, 가중치 적용부(110)는 GeoNames와 같은 지명 사전 데이터베이스(미도시)와 연동하여 개체명 인식 알고리즘(Named-Entity Recognition, NER)을 이용해 말뭉치 또는 입력 질의로부터 장소에 대한 정보를 포함하는 맥락 단어를 추출할 수 있다.
- [41] 또한, 가중치 적용부(110)는 규칙 기반 방법으로 말뭉치 또는 입력 질의로부터 시간에 대한 정보를 포함하는 맥락 단어를 추출할 수 있다. 특정 문장에서 단어를 추출하는 방법은 이미 공지된 기술이므로 본 명세서에서는 자세한 설명을 생략한다.
- [42] 본 명세서에서 "맥락 단어"에 대해 앞서 정의한 바에 따라, 말뭉치 맥락 단어들은 말뭉치를 구성하는 문장들에서 장소 또는 시간 중 적어도 하나에 대한 정보를 포함하는 단어이고, 질의 맥락 단어는 문장 형태의 질의에서 장소 또는 시간 중 적어도 하나에 대한 정보를 포함하는 단어를 의미한다.
- [43] 가중치 적용부(110)는 가중치 규칙에 따라 말뭉치 맥락 단어 및 질의 맥락 단어에 가중치를 적용한다.
- [44] 구체적으로, 가중치 적용부(110)는 말뭉치 맥락 단어 및 질의 맥락 단어에 미리 설정된 가중치를 적용하며, 미리 설정된 가중치는 말뭉치 및 입력 질의에 포함된 단어들 중에서, 말뭉치 맥락 단어 및 질의 맥락 단어에 해당하지 않는 단어에 적용되는 가중치보다 높을 수 있다.
- [45] 맥락 유사도 결정부(120)는 가중치가 적용된 말뭉치 및 가중치가 적용된 입력

질의에 대해 벡터화 알고리즘을 이용하여 복수의 말뭉치 문장 임베딩 벡터들 및 질의 문장 임베딩 벡터를 생성한다.

- [46] 도 2는 가중치 적용부가 문장의 맥락 단어들에 가중치를 적용하고, 맥락 유사도 결정부가 가중치가 적용된 문장을 이용하여 문장 임베딩 벡터를 생성하는 예시를 설명하는 도면이다.
- [47] 도 2를 참고하면, 예를 들어 문장 "H.Gilbert is an American 20c singer"에 대해 맥락 유사도 결정부가 문장 임베딩 벡터를 생성하는 경우, 가중치 적용부(110)는 문장을 구성하는 단어들 중 장소에 대한 정보를 포함하는 맥락 단어인 "American" 및 시간에 대한 정보를 포함하는 맥락 단어인 "20c"에 일반 단어들보다 높은 가중치를 적용할 수 있다. 예를 들면, 가중치 적용부(110)는 일반 단어들에 0.5의 가중치를 적용하고, 맥락 단어들에 0.5 내지 1의 가중치를 적용할 수 있다.
- [48] 맥락 유사도 결정부(120)는 문장을 구성하는 각 단어들에 대한 단어 임베딩 벡터를 대상으로 맥락 어텐션을 이용하여 맥락 단어들에 일반 단어들 보다 높은 가중치를 적용한다. 도 2에서, L_Att는 장소 정보 카테고리에 해당하는 맥락 단어의 단어 임베딩 벡터에 높은 가중치를 처리하는 맥락 어텐션이고, T_Att는 시간 정보 카테고리에 해당하는 맥락 단어의 단어 임베딩 벡터에 높은 가중치를 처리하는 맥락 어텐션이다.
- [49] 이후, 맥락 유사도 결정부(120)는 단어들 사이의 유사성을 결정하고, 문장 전체에 대한 문장 임베딩 벡터를 생성한다. 문장 임베딩 벡터 생성 과정에서 일반 단어보다 맥락 단어들에 높은 가중치가 적용되는바, 맥락 단어들과 관련된 단어 임베딩 벡터는 일반 단어와 관련된 임베딩 벡터보다 높은 가중치로 학습되어 생성된 문장 임베딩 벡터에서 맥락 단어들은 높은 영향을 미치게 된다.
- [50] 맥락 단어들과 일반 단어들에 동일한 가중치가 적용되는 경우, 학습 데이터에서 맥락 단어들의 사용 빈도가 일반 단어들에 비해 적은 편이기 때문에 신경망 분류기(예를 들면, Bi-LSTM)가 맥락 단어들을 잘못 분류할 확률이 높으나, 위와 같이 생성된 문장 임베딩 벡터는 맥락 단어들이 일반 단어들보다 높은 영향을 가지게 되어 신경망 분류기에서 문장을 구성하는 맥락 단어들을 보다 용이하게 분류할 수 있고, 이는 결과적으로 질의에 대한 응답 결정의 성능을 향상시킬 수 있다.
- [51] 맥락 유사도 결정부(120)는 복수의 말뭉치 문장 임베딩 벡터들과 질의 문장 임베딩 벡터 사이의 벡터 유사도를 기준으로 복수의 말뭉치 문장 임베딩 벡터들 중에서 대상 문장 임베딩 벡터를 결정한다.
- [52] 도 3은 맥락 유사도 결정부가 문장 임베딩 벡터 사이의 유사도를 계산하는 방법을 설명하는 도면이다.
- [53] 구체적으로, 맥락 유사도 결정부(120)는 벡터 유사도 판단 알고리즘을 이용하여 복수의 말뭉치 문장 임베딩 벡터들과 질의 문장 임베딩 벡터 사이의 벡터 유사도를 각각 계산한다.

- [54] 예를 들면, 도 3을 참고하면, 맥락 유사도 결정부(120)는 복수의 말뭉치 문장 임베딩 벡터 중 어느 하나와 질의 문장 임베딩 벡터 사이의 수학적 유사도를 벡터 유사도 판단 알고리즘을 이용하여 계산할 수 있고, 계산된 수학적 유사도를 두 벡터 사이의 벡터 유사도로 결정할 수 있다.
- [55] 이후, 맥락 유사도 결정부(120)는 복수의 말뭉치 문장 임베딩 벡터들 중에서 벡터 유사도가 높은 문장 임베딩 벡터를 질의 문장 임베딩 벡터와 유사한 맥락을 가진 대상 문장 임베딩 벡터로 결정한다. 예를 들면, 맥락 유사도 결정부(120)는 복수의 말뭉치 문장 임베딩 벡터들 중에서 벡터 유사도가 가장 높은 문장 임베딩 벡터를 대상 문장 임베딩 벡터로 결정할 수 있다.
- [56] 개념 그래프 관리부(130)는 말뭉치로부터 개념, 관계 및 속성을 추출하고, 추출한 개념, 관계 및 속성을 이용하여 말뭉치에 대한 복수의 개념 그래프들을 생성하고, 말뭉치 내 동일한 문장에서 비롯된 개념 그래프와 말뭉치 문장 임베딩 벡터를 대응시켜 저장한다.
- [57] 구체적으로, 개념 그래프 관리부(130)는 말뭉치 내 문장들로부터 개념, 관계 및 속성을 각각 추출한다. 특정 문장으로부터 개념, 관계 및 속성을 추출하는 방법은 다양한 방법으로 실행할 수 있으므로, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.
- [58] 도 4는 개념 그래프 관리부가 개념 그래프들을 생성하는 방법을 설명하는 도면이다.
- [59] 도 4를 참고하면, 예를 들면, 말뭉치 내 문장 "The word 'robot' firstly written in a play" (from wikipedia document titled 'robot')에 대해, 개념 그래프 관리부(130)는 개념 및 관계로 {<robot, is_a, word> : Wikipedia:robot), (<robot, appear, play> : Wikipedia:robot)}를 추출하고, 이에 대한 개념 그래프를 생성할 수 있다.
- [60] 이후, 개념 그래프 관리부(130)는 문장 "The word 'robot' firstly written in a play" (from wikipedia document titled 'robot')의 말뭉치 문장 임베딩 벡터와 개념 그래프를 대응시켜 저장할 수 있다. 개념 그래프 관리부(130)는 말뭉치 내 문장들 각각에 대해 위와 같은 방법으로 말뭉치 문장 임베딩 벡터와 개념 그래프를 대응시켜 저장할 수 있다.
- [61] 후보 응답 결정부(140)는 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를 결정한다.
- [62] 구체적으로, 후보 응답 결정부(140)는 복수의 개념 그래프들 중에서, 대상 문장 임베딩 벡터와 동일한 문장에서 비롯된 개념 그래프를 대상 개념 그래프로 결정한다. 이를 위해 후보 응답 결정부(140)는 대상 문장 임베딩 벡터를 이용하여 개념 그래프 관리부(130)에 저장된 복수의 개념 그래프들을 검색할 수 있다.
- [63] 후보 응답 결정부(140)는 대상 개념 그래프에 포함된 적어도 하나 이상의 개념들 각각에 대한 그래프 매칭 점수를 계산하여, 입력 질의에 대응하는 적어도 하나 이상의 후보 개념들을 결정한다. 이때, 본 명세서에서는 그래프 매칭 점수는 센터피스(center-piece) 알고리즘 등을 사용하여 구하는 것을 예로 하여

설명하나, 반드시 이와 같이 한정되는 것은 아니다. 또한, 센터피스 알고리즘은 이미 알려진 알고리즘으로, 본 명세서에서는 상세한 설명은 생략한다.

- [64] 응답 결정부(150)는 후보 개념들 중에서 미리 설정된 수만큼의 개념을 입력 질의에 대한 응답으로 결정한다.
- [65] 구체적으로, 응답 결정부(150)는 복수의 후보 개념들 중에서 그래프 매칭 점수가 높은 미리 설정된 수만큼의 개념을 입력 질의에 대한 응답으로 결정한다. 예를 들면, 응답 결정부(150)는 계산한 그래프 매칭 점수를 토대로 대상 개념 그래프와 의미적으로 가까운 상위 미리 설정된 수만큼의 개념들을 추출하고, 추출한 개념들을 입력 질의에 대한 응답으로 결정하여 사용자에게 제공할 수 있다.
- [66] 도 5는 질의응답 장치가 입력 질의의 응답을 결정하는 방법을 설명하는 도면이다.
- [67] 도 5에서, 도 1 내지 도 4와 동일한 내용은 그 자세한 설명을 생략한다.
- [68] 도 5를 참고하면, 질의응답 장치(100)는 입력 질의 및 말뭉치에서 맥락 단어를 추출한다(S100).
- [69] 구체적으로, 질의응답 장치(100)는 텍스트 분석 알고리즘을 이용하여 입력 질의 및 말뭉치로부터 맥락 단어를 각각 추출한다. 이 경우, 추출한 맥락 단어는 시간 정보 카테고리 또는 장소 정보 카테고리 중 적어도 하나에 해당한다.
- [70] 질의응답 장치(100)는 추출한 맥락 단어에 가중치 규칙에 따라 가중치를 적용한다(S101).
- [71] 구체적으로, 질의응답 장치(100)는 추출한 맥락 단어에 미리 설정된 가중치를 적용하며, 미리 설정된 가중치는 입력 질의 및 말뭉치에 포함된 단어들 중에서, 추출한 맥락 단어에 해당하지 않는 단어에 적용되는 가중치보다 높을 수 있다.
- [72] 질의응답 장치(100)는 가중치가 적용된 입력 질의에 대한 질의 문장 임베딩 벡터를 생성하고, 가중치가 적용된 말뭉치에 대한 복수의 말뭉치 문장 임베딩 벡터들을 생성한다(S103).
- [73] 구체적으로, 질의응답 장치(100)는 가중치가 적용된 입력 질의 및 가중치가 적용된 말뭉치에 대해 벡터화 알고리즘을 이용하여 질의 문장 임베딩 벡터 및 말뭉치 문장 임베딩 벡터들을 각각 생성한다.
- [74] 질의응답 장치(100)는 질의 문장 임베딩 벡터와 말뭉치 문장 임베딩 벡터들 사이의 벡터 유사도를 각각 계산한다(S105).
- [75] 구체적으로, 질의응답 장치(100)는 벡터 유사도 판단 알고리즘을 이용하여 복수의 말뭉치 문장 임베딩 벡터들과 질의 문장 임베딩 벡터 사이의 벡터 유사도를 각각 계산한다.
- [76] 질의응답 장치(100)는 말뭉치 문장 임베딩 벡터들 중에서 벡터 유사도가 높은 문장 임베딩 벡터를 대상 문장 임베딩 벡터로 결정한다(S107).
- [77] 구체적으로, 질의응답 장치(100)는 복수의 말뭉치 문장 임베딩 벡터들 중에서 벡터 유사도가 가장 높은 임베딩 벡터를 대상 문장 임베딩 벡터로 결정한다.

- [78] 질의응답 장치(100)는 말뭉치 내 문장들에 대한 개념 그래프를 생성하고, 생성한 개념 그래프를 말뭉치 문장 임베딩 벡터와 대응시켜 저장한다(S109).
- [79] 구체적으로, 질의응답 장치(100)는 말뭉치로부터 개념, 관계 및 속성을 추출하고, 추출한 개념, 관계 및 속성을 이용하여 말뭉치에 대한 복수의 개념 그래프들을 생성한다. 이후, 질의응답 장치(100)는 말뭉치 내 동일한 문장에서 비롯된 개념 그래프와 문장 임베딩 벡터를 대응시켜 저장한다.
- [80] 한편, 비록 도 5에서는 대상 문장 임베딩 벡터를 결정한 이후 개념 그래프와 문장 임베딩 벡터를 대응시켜 저장하였으나, 이후 단계에서 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를 결정하기 이전에 개념 그래프와 문장 임베딩 벡터를 대응시켜 저장하면 족하다.
- [81] 질의응답 장치(100)는 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를 결정한다(S111).
- [82] 구체적으로, 질의응답 장치(100)는 복수의 개념 그래프들 중에서, 대상 문장 임베딩 벡터와 동일한 문장에서 비롯된 개념 그래프를 대상 개념 그래프로 결정한다.
- [83] 질의응답 장치(100)는 대상 개념 그래프에 포함된 적어도 하나 이상의 개념들 각각에 대한 그래프 매칭 점수를 계산하여, 입력 질의에 대응하는 적어도 하나 이상의 후보 개념들을 결정한다(S113).
- [84] 질의응답 장치(100)는 후보 개념들 중에서 미리 설정된 수만큼의 개념을 입력 질의에 대한 응답으로 결정한다(S115).
- [85] 구체적으로, 질의응답 장치(100)는 복수의 후보 개념들 중에서 그래프 매칭 점수가 높은 미리 설정된 수만큼의 개념을 입력 질의에 대한 응답으로 결정한다.
- [86] 본 발명에 따르면, 질의 및 말뭉치 내 텍스트를 구성하는 단어들 중 특정 카테고리에 해당하는 단어들인 맥락 표현 단어들에 일반 단어보다 높은 가중치를 적용하는바, 맥락 표현 단어들을 일반 단어와 효율적으로 구분할 수 있고, 이를 통해 질의에 대한 응답 결정의 성능을 향상시킬 수 있다.
- [87] 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.
- [88]
- [89]

청구범위

[청구항 1]

질의응답 장치로서,
말뭉치에서 말뭉치 맥락 단어를 추출하고, 가중치 규칙에 따라
상기 말뭉치 맥락 단어에 가중치를 적용하고, 입력 질의에서 질의
맥락 단어를 추출하고, 상기 가중치 규칙에 따라 상기 질의 맥락
단어에 가중치를 적용하는 가중치 적용부,
가중치가 적용된 말뭉치에 대한 복수의 말뭉치 문장 임베딩
벡터들 및 가중치가 적용된 입력 질의에 대한 질의 문장 임베딩
벡터를 생성하고, 상기 복수의 말뭉치 문장 임베딩 벡터들과 상기
질의 문장 임베딩 벡터 사이의 벡터 유사도를 기준으로 상기
복수의 말뭉치 문장 임베딩 벡터들 중에서 대상 문장 임베딩
벡터를 결정하는 맥락 유사도 결정부,
상기 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를
결정하고, 상기 대상 개념 그래프에 포함된 적어도 하나 이상의
개념들 각각에 대한 그래프 매칭 점수를 계산하여, 상기 입력
질의에 대응하는 적어도 하나 이상의 후보 개념들을 결정하는
후보 응답 결정부, 그리고
상기 후보 개념들 중에서 미리 설정된 수만큼의 개념을 상기 입력
질의에 대한 응답으로 결정하는 응답 결정부
를 포함하는 질의응답 장치.

[청구항 2]

제1항에 있어서,
상기 가중치 적용부는
텍스트 분석 알고리즘을 이용하여 상기 말뭉치 및 상기 입력
질의로부터 상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어를
추출하는 질의응답 장치.

[청구항 3]

제1항에 있어서,
상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어는
특정 카테고리에 해당하는 단어이고,
상기 가중치 적용부는
상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어에 미리 설정된
가중치를 적용하는 질의응답 장치.

[청구항 4]

제3항에 있어서,
상기 특정 카테고리는
시간 카테고리 또는 장소 카테고리 중 적어도 하나를 포함하는
질의응답 장치.

[청구항 5]

제3항에 있어서,
상기 미리 설정된 가중치는

상기 말뭉치 및 상기 입력 질의에 포함된 단어들 중에서, 상기 말뭉치 맥락 단어 및 상기 질의 맥락 단어에 해당하지 않는 단어에 적용되는 가중치보다 높은 질의응답 장치.

[청구항 6]

제1항에 있어서,

상기 맥락 유사도 결정부는

상기 가중치가 적용된 말뭉치 및 상기 가중치가 적용된 입력 질의에 대해 벡터화 알고리즘을 이용하여 상기 복수의 말뭉치 문장 임베딩 벡터들 및 상기 질의 문장 임베딩 벡터를 생성하는 질의응답 장치.

[청구항 7]

제1항에 있어서,

상기 맥락 유사도 결정부는

벡터 유사도 판단 알고리즘을 이용하여 상기 복수의 말뭉치 문장 임베딩 벡터들과 상기 질의 문장 임베딩 벡터 사이의 벡터 유사도를 각각 계산하고, 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 상기 벡터 유사도가 높은 문장 임베딩 벡터를 상기 대상 문장 임베딩 벡터로 결정하는 질의응답 장치.

[청구항 8]

제1항에 있어서,

상기 말뭉치로부터 개념, 관계 및 속성을 추출하고, 상기 추출한 개념, 관계 및 속성을 이용하여 상기 말뭉치에 대한 복수의 개념 그래프들을 생성하고, 상기 말뭉치 내 동일한 문장에서 비롯된 개념 그래프와 말뭉치 문장 임베딩 벡터를 대응시켜 저장하는 개념 그래프 관리부

를 더 포함하고,

상기 후보 응답 결정부는

상기 복수의 개념 그래프들 중에서, 상기 대상 문장 임베딩 벡터와 동일한 문장에서 비롯된 개념 그래프를 상기 대상 개념 그래프로 결정하는 질의응답 장치.

[청구항 9]

제1항에 있어서,

상기 응답 결정부는

상기 복수의 후보 개념들 중에서 상기 그래프 매칭 점수가 높은 상기 미리 설정된 수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정하는 질의응답 장치.

[청구항 10]

질의응답 장치가 입력 질의의 응답을 결정하는 방법으로서, 입력 질의에서 추출한 맥락 단어에 가중치 규칙에 따라 가중치를 적용하고, 말뭉치에서 추출한 맥락 단어에 상기 가중치 규칙에 따라 가중치를 적용하는 단계,

가중치가 적용된 입력 질의에 대한 질의 문장 임베딩 벡터를 생성하고, 가중치가 적용된 말뭉치에 대한 복수의 말뭉치 문장

임베딩 벡터들을 생성하는 단계,
 상기 질의 문장 임베딩 벡터와 상기 복수의 말뭉치 문장 임베딩 벡터들 사이의 벡터 유사도를 각각 계산하고, 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 상기 벡터 유사도가 높은 문장 임베딩 벡터를 대상 문장 임베딩 벡터로 결정하는 단계,
 상기 대상 문장 임베딩 벡터에 대응하는 대상 개념 그래프를 결정하고, 상기 대상 개념 그래프에 포함된 적어도 하나 이상의 개념들 각각에 대한 그래프 매칭 점수를 계산하여, 상기 입력 질의에 대응하는 적어도 하나 이상의 후보 개념들을 결정하는 단계, 그리고
 상기 후보 개념들 중에서 미리 설정된 수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정하는 단계를 포함하는 응답 결정 방법.

[청구항 11]

제10항에 있어서,
 상기 가중치를 적용하는 단계는
 텍스트 분석 알고리즘을 이용하여 상기 입력 질의 및 상기 말뭉치로부터 맥락 단어를 추출하고, 추출한 맥락 단어에 미리 설정된 가중치를 적용하는 응답 결정 방법.

[청구항 12]

제11항에 있어서,
 상기 추출한 맥락 단어는
 시간 정보 카테고리 또는 장소 정보 카테고리 중 적어도 하나에 해당하는 응답 결정 방법.

[청구항 13]

제11항에 있어서,
 상기 미리 설정된 가중치는
 상기 입력 질의 및 상기 말뭉치에 포함된 단어들 중에서, 상기 추출한 맥락 단어에 해당하지 않는 단어에 적용되는 가중치보다 높은 응답 결정 방법.

[청구항 14]

제10항에 있어서,
 상기 질의 문장 임베딩 벡터 및 상기 복수의 말뭉치 문장 임베딩 벡터들을 생성하는 단계는
 상기 가중치가 적용된 입력 질의 및 상기 가중치가 적용된 말뭉치에 대해 벡터화 알고리즘을 이용하여 상기 질의 문장 임베딩 벡터 및 상기 복수의 말뭉치 문장 임베딩 벡터들을 생성하는 응답 결정 방법.

[청구항 15]

제10항에 있어서,
 상기 벡터 유사도가 높은 임베딩 벡터를 대상 문장 임베딩 벡터로 결정하는 단계는
 벡터 유사도 판단 알고리즘을 이용하여 상기 복수의 말뭉치 문장

임베딩 벡터들과 상기 질의 문장 임베딩 벡터 사이의 벡터 유사도를 각각 계산하고, 상기 복수의 말뭉치 문장 임베딩 벡터들 중에서 상기 벡터 유사도가 가장 높은 임베딩 벡터를 상기 대상 문장 임베딩 벡터로 결정하는 응답 결정 방법.

[청구항 16]

제10항에 있어서,
상기 대상 개념 그래프를 결정하는 단계 이전에,
상기 말뭉치로부터 개념, 관계 및 속성을 추출하는 단계,
상기 추출한 개념, 관계 및 속성을 이용하여 상기 말뭉치에 대한 복수의 개념 그래프들을 생성하는 단계, 그리고
상기 말뭉치 내 동일한 문장에서 비롯된 개념 그래프와 문장 임베딩 벡터를 대응시켜 저장하는 단계
를 더 포함하는 응답 결정 방법.

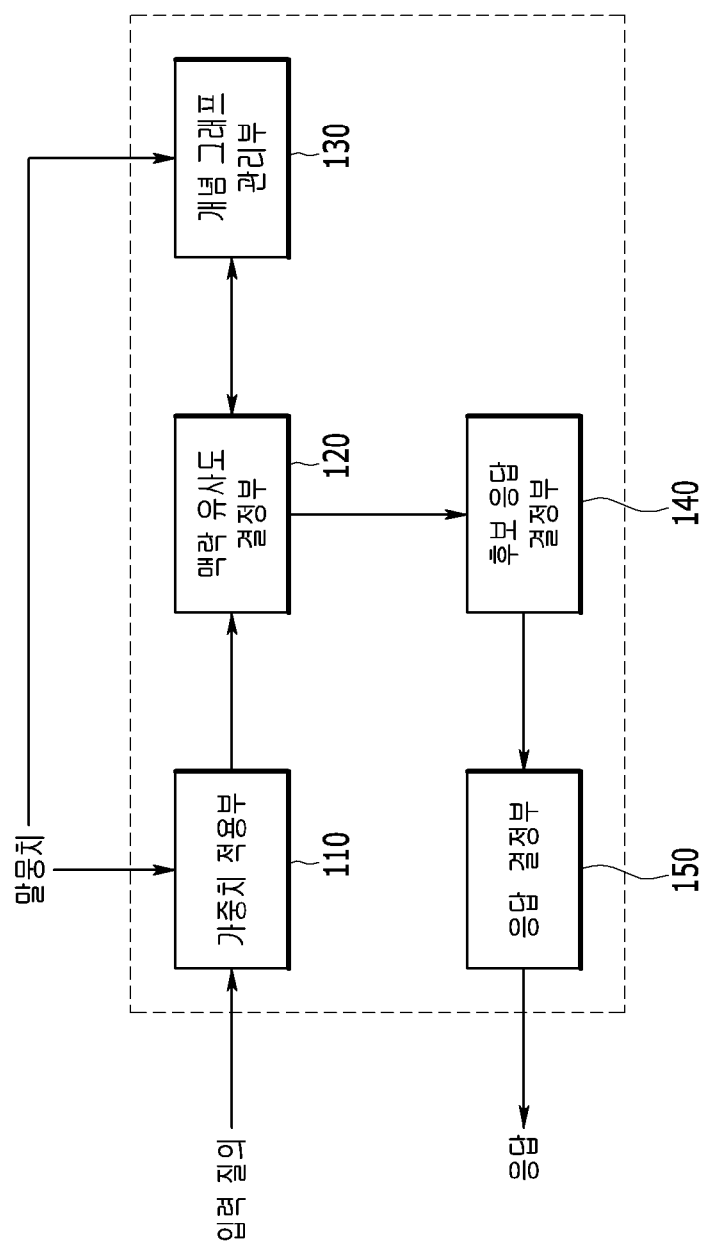
[청구항 17]

제16항에 있어서,
상기 대상 개념 그래프를 결정하는 단계는
상기 복수의 개념 그래프들 중에서, 상기 대상 문장 임베딩 벡터와 동일한 문장에서 비롯된 개념 그래프를 상기 대상 개념 그래프로 결정하는 응답 결정 방법.

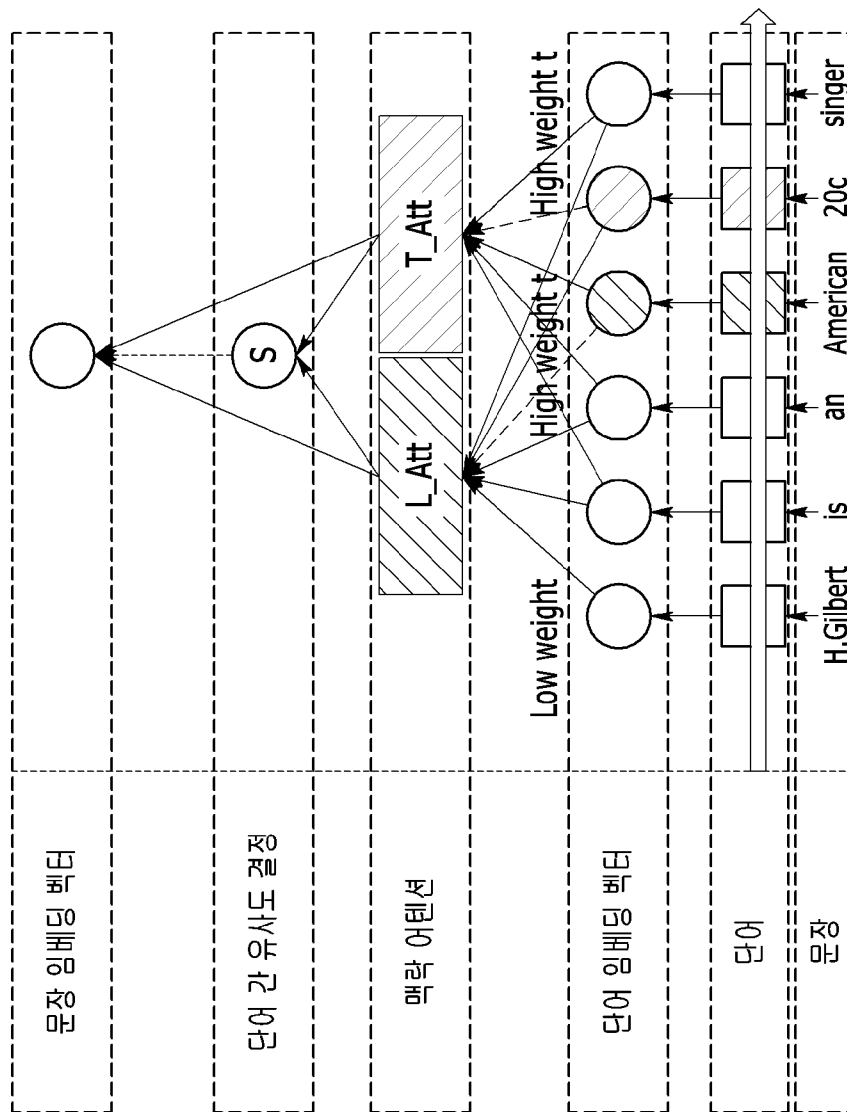
[청구항 18]

제10항에 있어서,
상기 입력 질의에 대한 응답으로 결정하는 단계는
상기 후보 개념들 중에서 상기 그래프 매칭 점수가 높은 상기 미리 설정된 수만큼의 개념을 상기 입력 질의에 대한 응답으로 결정하는 응답 결정 방법.

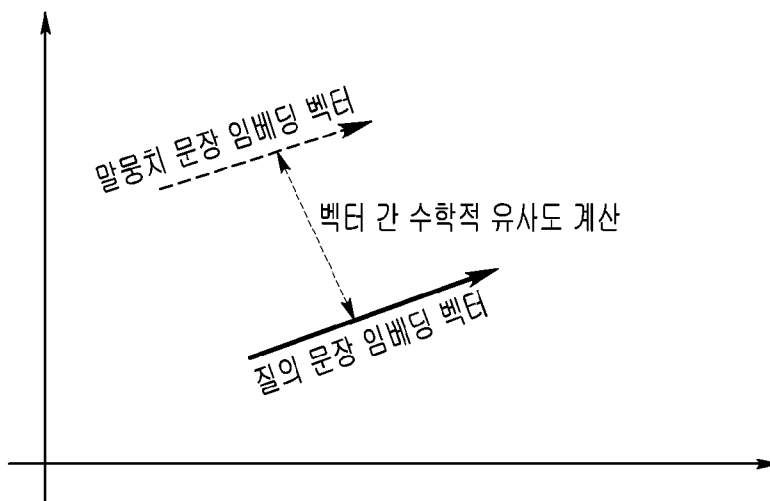
[Fig. 1]



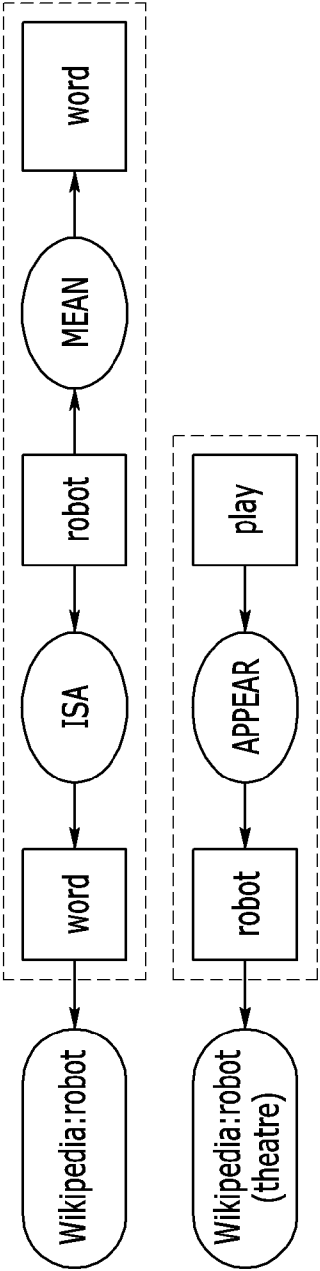
[Fig. 2]



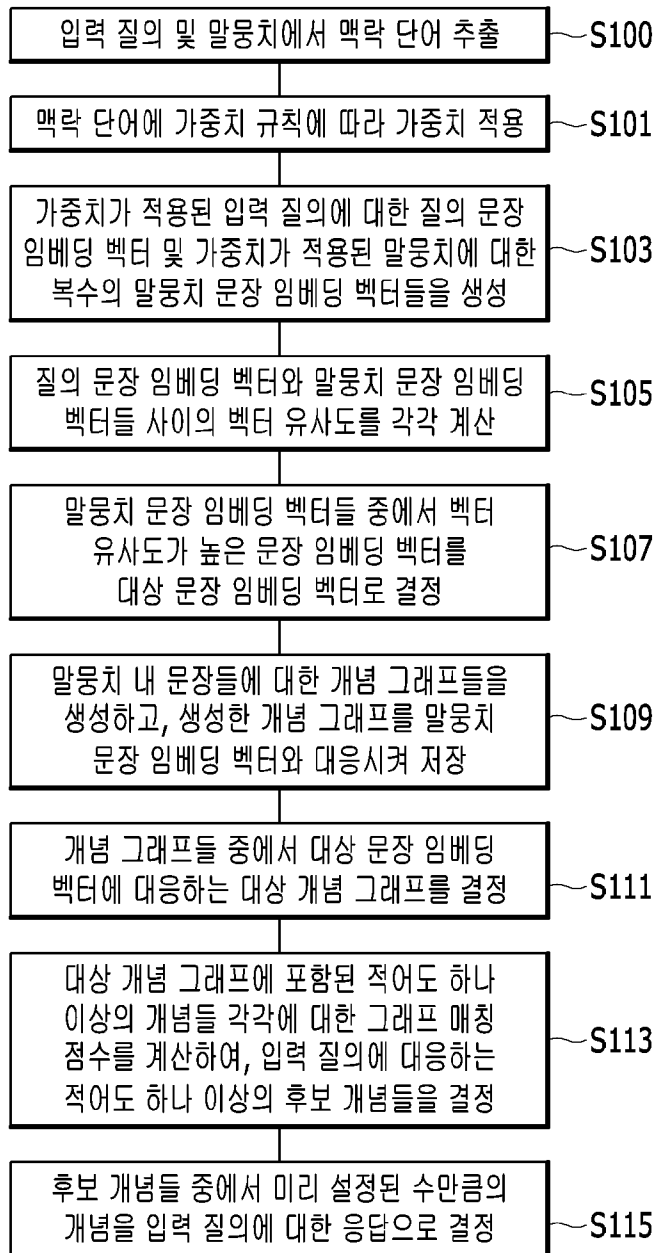
[Fig. 3]



[Fig. 4]



[Fig. 5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2018/014860

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/00(2019.01)i, G06F 17/27(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 16/00; G06F 17/27; G06F 17/30; G06N 3/08

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: question, response, embedding vector, context, weight, concept graph, similarity

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2018-0092808 A (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) 20 August 2018 See claim 1 and paragraphs [0028], [0031]-[0034], [0037], [0064].	1-18
Y	US 2011-0302156 A1 (VADLAMANI, Viswanath et al.) 08 December 2011 See claim 11.	1-18
A	KR 10-2018-0125746 A (DONG-A UNIVERSITY RESEARCH FOUNDATION FOR INDUSTRY-ACADEMY COOPERATION) 26 November 2018 See claims 1-7 and figure 8.	1-18
A	KR 10-2017-0096282 A (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) 24 August 2017 See claims 1-8 and figure 3.	1-18
A	US 2016-0357855 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 08 December 2016 See claims 1-9 and figure 2.	1-18



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

14 AUGUST 2019 (14.08.2019)

Date of mailing of the international search report

16 AUGUST 2019 (16.08.2019)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsu-ro, Seo-gu,
Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2018/014860

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2018-0092808 A	20/08/2018	None	
US 2011-0302156 A1	08/12/2011	CN 102279871 A US 9177057 B2	14/12/2011 03/11/2015
KR 10-2018-0125746 A	26/11/2018	None	
KR 10-2017-0096282 A	24/08/2017	None	
US 2016-0357855 A1	08/12/2016	US 2016-0358094 A1	08/12/2016

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06F 16/00(2019.01)i, G06F 17/27(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06F 16/00; G06F 17/27; G06F 17/30; G06N 3/08

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: 질의(question), 응답(response), 임베딩 벡터(embedding vector), 맥락(context), 가중치(weight), 개념그래프(concept graph), 유사도(similarity)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2018-0092808 A (한국과학기술원) 2018.08.20 청구항 1 및 [0028], [0031]-[0034], [0037], [0064] 참조.	1-18
Y	US 2011-0302156 A1 (VISWANATH VADLAMANI 등) 2011.12.08 청구항 11 참조.	1-18
A	KR 10-2018-0125746 A (동아대학교 산학협력단) 2018.11.26 청구항 1-7 및 도 8 참조.	1-18
A	KR 10-2017-0096282 A (한국과학기술원) 2017.08.24 청구항 1-8 및 도 3 참조.	1-18
A	US 2016-0357855 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2016.12.08 청구항 1-9 및 도 2 참조.	1-18

☐ 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

2019년 08월 14일 (14.08.2019)

국제조사보고서 발송일

2019년 08월 16일 (16.08.2019)

ISA/KR의 명칭 및 우편주소



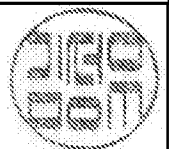
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

김성우

전화번호 +82-42-481-3348



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2018-0092808 A	2018/08/20	없음	
US 2011-0302156 A1	2011/12/08	CN 102279871 A US 9177057 B2	2011/12/14 2015/11/03
KR 10-2018-0125746 A	2018/11/26	없음	
KR 10-2017-0096282 A	2017/08/24	없음	
US 2016-0357855 A1	2016/12/08	US 2016-0358094 A1	2016/12/08