



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년09월21일
(11) 등록번호 10-2581782
(24) 등록일자 2023년09월19일

- (51) 국제특허분류(Int. Cl.)
G06F 16/2452 (2019.01) G06F 16/901 (2019.01)
G06N 3/08 (2023.01)
- (52) CPC특허분류
G06F 16/24522 (2019.01)
G06F 16/9024 (2019.01)
- (21) 출원번호 10-2021-0013643
- (22) 출원일자 2021년01월29일
심사청구일자 2021년01월29일
- (65) 공개번호 10-2022-0109979
- (43) 공개일자 2022년08월05일
- (56) 선행기술조사문헌
A. M. Weiner et al., Visualizing Cost-Based XQuery Optimization, IEEE 26th International Conference on Data Engineering, pp1165-1168(2010)*
A. Prasad et al., Enhancement of Natural Language to SQL Query Conversion using Machine Learning Techniques, International Journal of Advanced Computer Science and Applications Vol.11(2020)*
L. Jiang et al., Grammer-aware Parallelization for Scalable XPath Querying, ACM SIGPLAN Notices Vol.52, pp371-383(2017)*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
- (72) 발명자
한옥신
경상북도 포항시 남구 청암로 77 창의IT융합공학과 (지곡동, 포항공과대학교)
강혁규
경상북도 포항시 남구 효자동길5번길 23, 201호(효자동)
나인혁
서울특별시 서대문구 연희로32길 48, 104동 105호(연희동, 연희동성원아파트)
- (74) 대리인
특허법인이룸리온

전체 청구항 수 : 총 7 항

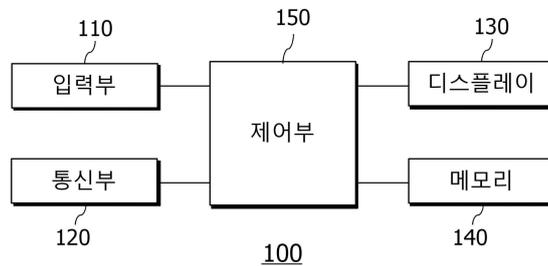
심사관 : 권계민

(54) 발명의 명칭 자연어 질의를 QGM 질의로 변환하는 방법 및 장치

(57) 요약

본 발명은 자연어 질의를 QGM 질의로 변환하는 방법 및 장치에 관한 것이다. 본 발명의 일 실시예에 따른 방법은 전자 장치에 의해 수행되며, 데이터베이스에 대한 자연어 인터페이스를 구현하는 방법으로서, 데이터베이스 및 자연어 질의의 입력에 대해 QGM 질의를 생성하는 단계를 포함한다.

대표도 - 도2



(52) CPC특허분류
G06N 3/08 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103171
과제번호	2018-0-01398-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	SW컴퓨팅산업원천기술개발사업(SW스타랩)
연구과제명	대화 가능하고 자동으로 튜닝하는 DBMS의 개발
기여율	1/1
과제수행기관명	포항공과대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

전자 장치에 의해 수행되는 방법으로서,

데이터베이스 및 자연어 질의의 입력에 대해 인공 신경망과 푸시다운 트랜스듀서를 이용하여 QGM 질의를 생성함으로써, 데이터베이스에 대한 자연어 인터페이스를 구현하는 단계를 포함하며,

상기 인공 신경망은 데이터베이스 및 자연어 질의를 입력 받아 문자열을 출력하고,

상기 푸시다운 트랜스듀서는 상기 인공 신경망에서 출력된 문자열을 입력 받아 상기 QGM 질의를 출력하며,

상기 인공 신경망은 RA-트랜스포머를 이용하여 상기 데이터베이스 및 자연어 질의를 인코딩하며 상기 푸시다운 트랜스듀서의 상태와 큐를 이용하여 문자열을 자기회귀적으로 생성하고,

상기 RA-트랜스포머는 스키마 연결(Schema linking) 정보 관계가 자연어 질의 및 스키마의 단어가 함께 입력되고 자연어 질의를 의존 분석(dependency parsing)한 정보 관계를 인코딩하는 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 인공 신경망은 트랜스포머 및 다중 퍼셉트론 신경망을 통해 상기 푸시다운 트랜스듀서의 상태와 큐를 이용하는 방법.

청구항 5

제1항에 있어서,

상기 푸시다운 트랜스듀서는 키폴럼들을 먼저 생성한 후에 다른 정보들을 생성하는 순서로 QGM 질의를 생성하는 방법.

청구항 6

제5항에 있어서,

상기 푸시다운 트랜스듀서는 키폴럼들을 먼저 생성한 후 큐에 삽입하고, 큐에서 추출된 키폴럼들에 대해 관련 정보를 생성하며, 부분 질의 박스 정보를 스택으로 관리하는 방법.

청구항 7

데이터베이스에 대한 자연어 인터페이스를 구현하는 장치로서,

인공 신경망과 푸시다운 트랜스듀서를 저장한 메모리; 및

데이터베이스 및 자연어 질의의 입력에 대해 상기 인공 신경망과 푸시다운 트랜스듀서를 이용하여 QGM 질의를 생성함으로써, 데이터베이스에 대한 자연어 인터페이스를 구현하도록 제어하는 제어부;를 포함하며,

상기 인공 신경망은 데이터베이스 및 자연어 질의를 입력 받아 문자열을 출력하고,

상기 푸시다운 트랜스듀서는 상기 인공 신경망에서 출력된 문자열을 입력 받아 상기 QGM 질의를 출력하며,

상기 인공 신경망은 RA-트랜스포머를 이용하여 상기 데이터베이스 및 자연어 질의를 인코딩하며 상기 푸시다운 트랜스듀서의 상태와 큐를 이용하여 문자열을 자기회귀적으로 생성하고,

상기 RA-트랜스포머는 스키마 연결(Schema linking) 정보 관계가 자연어 질의 및 스키마의 단어가 함께 입력되고 자연어 질의를 의존 분석(dependency parsing)한 정보 관계를 인코딩하는 장치.

청구항 8

삭제

청구항 9

삭제

청구항 10

제7항에 있어서,

상기 푸시다운 트랜스듀서는 컬럼들을 먼저 생성한 후에 다른 정보들을 생성하는 상향식으로 QGM 질의를 생성하는 장치.

청구항 11

제7항에 있어서,

상기 푸시다운 트랜스듀서는 컬럼들을 먼저 생성한 후 큐에 삽입하고, 큐에서 추출된 컬럼들에 대해 관련 정보를 생성하며, 부분 질의 박스 정보를 스택으로 관리하는 장치.

발명의 설명

기술 분야

[0001] 본 발명은 데이터베이스에 대한 자연어 인터페이스를 위한 기술에 관한 것으로서, 더욱 상세하게는 데이터베이스에 대한 자연어 질의를 QGM 질의로 변환하는 방법 및 장치에 관한 것이다.

배경 기술

[0003] 도 1은 SQL 질의(도 1(a) 참조)와 그와 동일 의미를 갖는 QGM 질의(도 1(b) 참조)를 각각 나타낸다.

[0004] 즉, 도 1을 참조하면, QGM(Query Graph Model)은 IBM의 데이터베이스인 DB2에서 SQL 질의를 표현하는 그래프 형태의 데이터 구조이다. 이러한 QGM은 형태 변환이 쉬우며, 직관적으로 표현될 수 있는 특징을 갖는다.

[0005] 종래에 데이터베이스에 대한 자연어 질의를 SQL 질의로 변환하기 위한 기술이 존재하나, 해당 자연어 질의를 QGM 질의로 변환하는 기술은 전무한 실정이다. 이는 QGM 질의가 그래프 구조로 되어 있어, 이를 수학적으로 적절히 모델링하는 것은 매우 어려운 문제이기 때문이다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) KR 10-2020-0080822 A

발명의 내용

해결하려는 과제

[0008] 본 발명은 데이터베이스에 대한 자연어 질의를 데이터베이스 언어의 일종인 QGM 질의로 변환하기 위한 방법 및 장치를 제공하는데 그 목적이 있다.

[0009] 다만, 본 발명이 해결하고자 하는 과제는 이상에서 언급한 과제에 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0011] 상기와 같은 과제를 해결하기 위한 본 발명의 일 실시예에 따른 방법은 전자 장치에 의해 수행되며, 데이터베이스에 대한 자연어 인터페이스를 구현하는 방법으로서, 데이터베이스 및 자연어 질의의 입력에 대해 인공 신경망과 푸시다운 트랜스듀서를 이용하여 QGM 질의를 생성하는 단계를 포함한다.

[0012] 또한, 본 발명의 일 실시예에 따른 장치는 데이터베이스에 대한 자연어 인터페이스를 구현하는 장치로서, 인공 신경망과 푸시다운 트랜스듀서를 저장한 메모리; 및 데이터베이스 및 자연어 질의의 입력에 대해 상기 인공 신경망과 푸시다운 트랜스듀서를 이용하여 QGM 질의를 생성하도록 제어하는 제어부;를 포함한다.

[0013] 상기 인공 신경망은 데이터베이스 및 자연어 질의를 입력 받아 문자열을 출력할 수 있으며, 상기 푸시다운 트랜스듀서는 상기 인공 신경망에서 출력된 문자열을 입력 받아 상기 QGM 질의를 출력할 수 있다.

[0014] 상기 인공 신경망은 RA-트랜스포머를 이용하여 상기 데이터베이스 및 자연어 질의를 인코딩할 수 있으며, 상기 푸시다운 트랜스듀서의 상태와 큐를 이용하여 문자열을 자기회귀적으로 생성할 수 있다.

[0015] 상기 인공 신경망은 트랜스포머 및 다중 퍼셉트론 신경망을 통해 상기 푸시다운 트랜스듀서의 상태와 큐를 이용할 수 있다.

[0016] 상기 푸시다운 트랜스듀서는 컬럼들을 먼저 생성한 후에 다른 정보들을 생성하는 상향식으로 QGM 질의를 생성할 수 있다.

[0017] 상기 푸시다운 트랜스듀서는 컬럼들을 먼저 생성한 후 큐에 삽입하고, 큐에서 추출된 컬럼들에 대해 관련 정보를 생성하며, 부분 질의 박스 정보를 스택으로 관리할 수 있다.

발명의 효과

[0019] 상기와 같이 구성되는 본 발명은 푸시다운 트랜스듀서(pushdown transducer)와 인공 신경망(artificial neural network)를 이용함에 따라 데이터베이스에 대한 자연어 질의를 QGM 질의로 보다 쉽고 정확하게 변환할 수 있는 이점이 있다.

[0020] 즉, 본 발명은 유효한 QGM 질의의 생성을 보장할 수 있으며, 인공 신경망이 그래프 구조를 이해하여 그 출력을 생성할 수 있는 이점이 있다.

[0021] 본 발명에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0023] 도 1은 SQL 질의와 그와 동일 의미를 갖는 QGM 질의를 각각 나타낸다.

도 2는 본 발명을 개괄적으로 표현한 모식도이다.

도 3은 본 발명의 일 실시예에 따른 학습데이터 변환 장치(100)의 블록 구성도를 나타낸다.

도 4는 푸시다운 트랜스듀서의 전이 규칙을 전체적으로 나타낸다.

도 5 내지 도 9는 각각 도 3을 부분 별로 나누어 나타낸다.

도 10은 인공 신경망 모델의 구조를 나타낸다.

도 11은 q_{nl} , q_{sq1} , d , 및 q_{qgm} 에 대한 예시를 나타낸다.

도 12 내지 도 15는 도 11에 따른 각 단계의 결과를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0024] 본 발명의 상기 목적과 수단 및 그에 따른 효과는 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해질 것이며, 그에 따라 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다.
- [0025] 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며, 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 경우에 따라 복수형도 포함한다. 본 명세서에서, "포함하다", "구비하다", "마련하다" 또는 "가지다" 등의 용어는 언급된 구성요소 외의 하나 이상의 다른 구성요소의 존재 또는 추가를 배제하지 않는다.
- [0026] 본 명세서에서, "또는", "적어도 하나" 등의 용어는 함께 나열된 단어들 중 하나를 나타내거나, 또는 둘 이상의 조합을 나타낼 수 있다. 예를 들어, "A 또는 B", "A 및 B 중 적어도 하나"는 A 또는 B 중 하나만을 포함할 수 있고, A와 B를 모두 포함할 수도 있다.
- [0027] 본 명세서에서, "예를 들어" 등에 따르는 설명은 인용된 특성, 변수, 또는 값과 같이 제시한 정보들이 정확하게 일치하지 않을 수 있고, 허용 오차, 측정 오차, 측정 정확도의 한계와 통상적으로 알려진 기타 요인을 비롯한 변형과 같은 효과로 본 발명의 다양한 실시 예에 따른 발명의 실시 형태를 한정하지 않아야 할 것이다.
- [0028] 본 명세서에서, 어떤 구성요소가 다른 구성요소에 '연결되어' 있다거나 '접속되어' 있다고 기재된 경우, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성 요소에 '직접 연결되어' 있다거나 '직접 접속되어' 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해될 수 있어야 할 것이다.
- [0029] 본 명세서에서, 어떤 구성요소가 다른 구성요소의 '상에' 있다거나 '접하여' 있다고 기재된 경우, 다른 구성요소에 상에 직접 맞닿아 있거나 또는 연결되어 있을 수 있지만, 중간에 또 다른 구성요소가 존재할 수 있다고 이해되어야 할 것이다. 반면, 어떤 구성요소가 다른 구성요소의 '바로 위에' 있다거나 '직접 접하여' 있다고 기재된 경우에는, 중간에 또 다른 구성요소가 존재하지 않는 것으로 이해될 수 있다. 구성요소 간의 관계를 설명하는 다른 표현들, 예를 들면, '~사이에'와 '직접 ~사이에' 등도 마찬가지로 해석될 수 있다.
- [0030] 본 명세서에서, '제1', '제2' 등의 용어는 다양한 구성요소를 설명하는데 사용될 수 있지만, 해당 구성요소는 위 용어에 의해 한정되어서는 안 된다. 또한, 위 용어는 각 구성요소의 순서를 한정하기 위한 것으로 해석되어서는 안되며, 하나의 구성요소와 다른 구성요소를 구별하는 목적으로 사용될 수 있다. 예를 들어, '제1구성요소'는 '제2구성요소'로 명명될 수 있고, 유사하게 '제2구성요소'도 '제1구성요소'로 명명될 수 있다.
- [0031] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어는 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또한, 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0033] 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 일 실시예를 상세히 설명하도록 한다.
- [0034] 도 2는 본 발명을 개괄적으로 표현한 모식도이다.
- [0035] 본 발명의 일 실시예에 따른 변환 장치(100)는 데이터베이스에 대한 자연어 질의를 직관적이고 쉽게 표현하기 위해 QGM 질의로 변환(번역)하는 장치이다. 즉, 도 2에 도시된 바와 같이, 변환 장치(100)는 푸시다운 트랜스듀서(pushdown transducer)와 인공 신경망(artificial neural network)를 이용하여, 데이터베이스에 대한 자연어 질의를 데이터베이스 언어의 일종인 QGM 질의로 변환하는 인터페이스 장치이다.

- [0036] 이러한 인터페이스는 자연어를 SQL로 번역하는 문제에 해당할 수 있다. 즉, 변환 장치(100)는 데이터베이스 d에 대한 자연어 질의 q_{nl} 이 주어졌을 때, 이를 대답할 수 있는 SQL 질의 q_{sql} 을 구한 후, q_{sql} 과 의미적으로 동치인 QGM 질의인 q_{qgm} 을 구함으로써, 해당 인터페이스를 구현할 수 있다.
- [0037] 푸시다운 트랜스듀서(Pushdown Transducer)는 추상의 자동 머신(auto machine)으로서, 유한개의 상태(State)와 스택(Stack)을 가져 입력 문자열을 출력 문자열로 변환한다. 본 발명에서는 편의상 스택 및 큐(Queue)를 각각 한 개씩 사용하는 것으로 설명하나, 본 발명이 이에 한정되는 것은 아니다.
- [0038] 푸시다운 트랜스듀서 수학적으로 8개의 쌍, 즉 $\langle Q, \Sigma, \Delta, \Gamma, \delta, q_0, Z_0, F \rangle$ 로 정의될 수 있다. 이때, Q는 가능한 상태의 집합을 나타내며, Σ, Δ 및 Γ 는 각각 입력, 출력 및 큐(스택) 심볼의 집합을 나타낸다. 또한, $\delta: Q \times \Sigma \times \Gamma \rightarrow Q \times \Gamma \times \Delta$ 는 전이 규칙을 나타내며, q_0, Z_0 및 F는 각각 시작 상태, 시작 큐(스택) 심볼 및 최종 상태를 나타낸다.
- [0039] QGM 질의는 그래프 구조로 되어 있어, 그래프 구조의 생성을 이해하는 모델을 수학적으로 모델링하는 것은 쉽지 않은 문제이다. 이에 따라, 본 발명은 이러한 문제를 푸시다운 트랜스듀서를 이용하여 해결한다. 즉, 변환 장치(100)는 입력 문자열을 QGM 질의로 변환하는 푸시다운 트랜스듀서를 이용한다.
- [0040] 구체적으로, 변환 장치(100)는 d와 q_{nl} 을 입력으로 받아 문자열을 출력하는 인공 신경망과, 인공 신경망에서 출력된 문자열을 입력으로 받아 q_{qgm} 을 출력하는 푸시다운 트랜스듀서를 포함한다. 즉, 인공 신경망은 d와 q_{nl} 을 이해하여, 푸시다운 트랜스듀서가 q_{qgm} 을 출력하게끔 하는 문자열을 생성하도록 학습된다. 푸시다운 트랜스듀서는 그래프 구조를 만들기 위해 정점 정보를 스택 및 큐에 저장하며, 해당 정점과 연결된 정보들을 생성할 때 그 저장된 정보들을 사용하도록 구현된다. 스택은 현재 얼마나 중첩된(Nesting) 박스를 생성하고 있는지를 저장하고, 큐는 컬럼들을 저장한다. 푸시다운 트랜스듀서는 이러한 스택 및 큐 정보를 인공 신경망에 제공함으로써, 인공 신경망이 수학적으로 그래프 생성을 모델링하는 자기회귀적(autoregressive) 확률 모델이 되도록 한다.
- [0041] 특히, 본 발명은 푸시다운 트랜스듀서를 이용하여 QGM 질의를 변환함에 따라 다음의 두 가지 이점을 갖는다.
- [0042] 첫째로, 유효한 QGM 질의의 생성을 보장할 수 있다. 즉, 본 발명에서 인공 신경망은 주어진 입력에 대해 출력 문자열의 확률 분포를 예측할 수 있으며, 이 중에서 유효한 QGM 질의의 생성을 보장하는 가장 확률 높은 문자열을 선택할 수 있다.
- [0043] 둘째로, 인공 신경망이 그래프 구조를 이해하여 그 출력을 생성할 수 있다. 즉, 인공 신경망은 문자열을 생성하는 동시에, 생성한 문자열이 입력으로 들어간 푸시다운 트랜스듀서의 큐 및 상태 정보를 이용한다. 푸시다운 트랜스듀서는 질의에 사용될 컬럼을 먼저 생성하여 큐에 저장하고, 해당 컬럼과 관련된 정보를 생성할 때 큐에서 컬럼을 추출하여 생성한다. 이러한 큐를 이용함으로써 어떤 정점과 연결된 부분을 생성하고 있는지를 인공신경망이 모델링하게끔 할 수 있다. 이때, 상태는 현재 QGM 질의의 어느 부분을 생성하고 있는지를 인공신경망이 모델링하게끔 할 수 있다.
- [0044] 또한, 푸시다운 트랜스듀서는 상향식으로 QGM 질의를 생성하도록 구현되었다. 이때, 상향식이란, 질의에서 어떤 컬럼들이 사용되는지를 우선 정하고 이후 다른 정보들을 생성하는 방식이다. 즉, 자연어 질의를 이해할 때에 자연어 질의에 어떤 개체(즉, 컬럼)들이 등장하는지를 먼저 파악한 후에 세세한 조건들을 파악하는 방식이다.
- [0045] 인공 신경망은 입력층(input layer)과 출력층(output layer) 사이에 하나 이상의 은닉층(hidden layer)을 포함하며, 이러한 은닉층을 통해 입력층과 출력층 간의 다양한 비선형적 관계(non-linear relationship)를 학습하여 모델링할 수 있다. 즉, 학습된 인공 신경망은 입력층에 입력되는 입력 데이터와 출력층에 출력되는 출력 데이터 간의 관계를 하나 이상의 은닉층으로 표현한다.
- [0046] 이때, 입력층, 다수의 은닉층 및 출력층은 각각 다수의 노드(N)를 포함하는데, 전단 층의 노드(전단 노드)들은 후단 층의 노드(후단 노드)들과 다양한 방식으로 연결된다. 이때, 각 노드는 가중치(W)를 가지며, 각 후단 노드의 노드 값은 그와 연결된 전단 노드의 노드 값 및 가중치에 의해 결정될 수 있다.
- [0047] 가령, 인공 신경망은 다양한 딥러닝(deep learning) 기법에 따라 학습되어 복수개의 은닉층을 가지는 딥러닝 모델(deep learning model)일 수 있다. 이때, 각 은닉층은 적어도 하나 이상의 필터로 이루어지며, 각 필터는 가중치(weight)의 매트릭스(matrix)를 가진다. 즉, 해당 필터의 매트릭스에서 각 원소(픽셀)는 가중치의 값에 해당할 수 있다.

- [0048] 예를 들어, 딥 러닝 기법은 Deep Neural Network(DNN), Convolutional Neural Network(CNN), Recurrent Neural Network(RNN), Restricted Boltzmann Machine(RBM), Deep Belief Network(DBN), Deep Q-Networks, 오토인코더(Auto-encoder), 변이형 오토인코더(Variational Auto-encoder) 등을 포함할 수 있으나, 이에 한정되는 것은 아니다.
- [0049] 도 3은 본 발명의 일 실시예에 따른 학습데이터 변환 장치(100)의 블록 구성도를 나타낸다.
- [0050] 이러한 변환 장치(100)는 푸시다운 트랜스듀서와 인공 신경망을 기반으로 데이터베이스에 대한 자연어 질의를 QGM 질의로 변환하는 컴퓨팅(computing)이 가능한 전자 장치로서, 도 3에 도시된 바와 같이, 입력부(110), 통신부(120), 디스플레이(130), 메모리(140) 및 제어부(150)를 포함할 수 있다. 즉, 변환 장치(100)는 푸시다운 트랜스듀서와 인공 신경망을 이용하여 데이터베이스에 대한 자연어 질의를 QGM 질의로 변환하는 장치일 수 있다.
- [0051] 예를 들어, 전자 장치는 데스크탑 PC(desktop personal computer), 랩탑 PC(laptop personal computer), 태블릿 PC(tablet personal computer), 넷북 컴퓨터(netbook computer), 워크스테이션(workstation), PDA(personal digital assistant), 스마트폰(smartphone), 스마트패드(smarpad), 또는 휴대폰(mobile phone) 등일 수 있으나, 이에 한정되는 것은 아니다.
- [0052] 입력부(110)는 다양한 사용자의 입력에 대응하여, 입력데이터를 발생시키며, 다양한 입력수단을 포함할 수 있다. 예를 들어, 입력부(110)는 키보드(key board), 키패드(key pad), 돔 스위치(dome switch), 터치 패널(touch panel), 터치 키(touch key), 터치 패드(touch pad), 마우스(mouse), 메뉴 버튼(menu button) 등을 포함할 수 있으나, 이에 한정되는 것은 아니다.
- [0053] 통신부(120)는 다른 장치와의 통신을 수행하는 구성이다. 예를 들어, 통신부(120)는 5G(5th generation communication), LTE-A(long term evolution-advanced), LTE(long term evolution), 블루투스, BLE(bluetooth low energy), NFC(near field communication), 와이파이(WiFi) 통신 등의 무선 통신을 수행하거나, 케이블 통신 등의 유선 통신을 수행할 수 있으나, 이에 한정되는 것은 아니다. 가령, 통신부(120)는 자연어 질의, 데이터베이스에 대한 정보 등을 타 장치로부터 수신할 수 있으며, 생성한 QGM 질의 등을 타 장치로 송신할 수 있다. 이 경우,
- [0054] 물론, 자연어 질의, 데이터베이스에 대한 정보 등은 입력부(110)를 통해 입력되거나, 메모리(140)에 기 저장될 수도 있다.
- [0055] 디스플레이(130)는 다양한 영상 데이터를 화면으로 표시하는 것으로서, 비발광형 패널이나 발광형 패널로 구성될 수 있다. 예를 들어, 디스플레이(130)는 액정 디스플레이(LCD; liquid crystal display), 발광 다이오드(LED; light emitting diode) 디스플레이, 유기 발광 다이오드(OLED; organic LED) 디스플레이, 마이크로 전자 기계 시스템(MEMS; micro electro mechanical systems) 디스플레이, 또는 전자 종이(electronic paper) 디스플레이 등을 포함할 수 있으나, 이에 한정되는 것은 아니다. 또한, 디스플레이(130)는 입력부(110)와 결합되어 터치 스크린(touch screen) 등으로 구현될 수 있다.
- [0056] 메모리(140)는 변환 장치(100)의 동작에 필요한 각종 정보를 저장한다. 가령, 저장 정보로는 자연어 질의, 데이터베이스, QGM 질의, 푸시다운 트랜스듀서, 인공 신경망 등에 대한 정보 외에도, 후술할 QGM 질의 변환 동작에 관련된 프로그램 정보 등이 포함될 수 있으나, 이에 한정되는 것은 아니다. 예를 들어, 메모리(140)는 그 유형에 따라 하드디스크 타입(hard disk type), 마그네틱 매체 타입(Sagnetic media type), CD-ROM(compact disc read only memory), 광기록 매체 타입(Optical Media type), 자기-광 매체 타입(Sagnetto-optical media type), 멀티미디어 카드 마이크로 타입(Sultimedia card micro type), 플래시 저장부 타입(flash memory type), 롬 타입(read only memory type), 또는 램 타입(random access memory type) 등을 포함할 수 있으나, 이에 한정되는 것은 아니다. 또한, 메모리(140)는 그 용도/위치에 따라 캐시(cache), 버퍼, 주기억장치, 또는 보조기억장치이거나 별도로 마련된 저장 시스템일 수 있으나, 이에 한정되는 것은 아니다.
- [0057] 제어부(150)는 변환 장치(100)의 다양한 제어 동작을 수행할 수 있다. 즉, 제어부(150)는 후술할 QGM 질의 변환 동작의 수행을 제어할 수 있으며, 변환 장치(100)의 나머지 구성, 즉 입력부(110), 통신부(120), 디스플레이(130), 메모리(140) 등의 동작을 제어할 수 있다. 예를 들어, 제어부(150)는 하드웨어인 프로세서(processor) 또는 해당 프로세서에서 수행되는 소프트웨어인 프로세스(process) 등을 포함할 수 있으나, 이에 한정되는 것은 아니다.
- [0058] 제어부(150)는 메모리(140) 저장된 인공 신경망에 대한 정보를 기반으로 인공 신경망의 동작을 제어할 수 있다.

또한, 제어부(150)는 메모리(140) 저장된 푸시다운 트랜스듀서에 대한 정보를 기반으로 푸시다운 트랜스듀서의 동작을 제어할 수 있다. 즉, QGM 질의를 생성하기 위해, 제어부(150)는 인공 신경망을 학습시키거나, 학습된 인공 신경망 및 푸시다운 트랜스듀서를 이용하도록 제어할 수 있다.

- [0059] 이하, 본 발명에 따른 QGM 질의 변환 동작에 대해 보다 상세하게 설명하도록 한다.
- [0060] QGM 질의 생성 동작은 푸시다운 트랜스듀서와 인공 신경망을 이용하여, 데이터베이스에 대한 자연어 질의를 QGM 질의로 변환하는 동작이다. 이를 위한 푸시다운 트랜스듀서 및 인공 신경망의 보다 상세한 동작은 다음과 같다.
- [0061] <QGM 질의를 생성하는 푸시다운 트랜스듀서의 동작>
- [0062] 도 4는 푸시다운 트랜스듀서의 전이 규칙을 전체적으로 나타내며, 도 5 내지 도 9는 각각 도 3을 부분 별로 나누어 나타낸다. 즉, 도 5는 메인 박스의 사영 컬럼들을 만들어내는 전이 규칙, 도 6은 메인 박스의 사영 컬럼 이외의 컬럼들을 만들어내는 전이 규칙, 도 7은 박스 생성 및 박스의 사영 정보를 만들어내는 전이 규칙, 도 8은 박스의 술부(predicate)을 만들어내는 전이 규칙, 도 9는 정렬 정보를 만들어내는 전이 규칙을 각각 나타낸다.
- [0063] 도 4 내지 도 9에서, 원은 상태를 나타내며, q_0 부터 순서대로 이름을 붙이되, 가독성을 위해 일부 명시적으로 이름붙이지 않은 상태들이 존재한다. 붉은 원은 컬럼을 큐에 삽입하는 상태이며, 파란 원은 큐에서 컬럼을 추출하는 상태이다. 파란 테두리 원은 인공 신경망이 큐에서 다음에 추출될 컬럼을 활용하는 상태로, 해당 컬럼과 관련된 부분들을 생성하는 상태이다. 여러 액션이 동일한 시작 상태 및 도착 상태를 가질 경우, 가독성을 위해 하나의 화살표로 표기하였다.
- [0064] 먼저, 푸시다운 트랜스듀서는, 도 5 및 도 6에 도시된 전이 규칙에 따라, QGM 질의에서 쓰일 컬럼 및 테이블들을 선택하여 큐에 삽입한다. 즉, 도 5에 도시된 바와 같이, 메인 박스의 사영(projection) 컬럼을 선택하며, 도 6에 도시된 바와 같이, 그 이외의 컬럼들을 선택한다. 이 두 단계를 명시적으로 구분하는 것은 푸시다운 트랜스듀서의 성능 향상에 도움이 되기 때문이다.
- [0065] 사영 컬럼을 선택하는 방식과 그 이외의 컬럼들을 선택하는 방식은 차이가 있는데, 전자는 컬럼을 총 몇 개 선택할 것인지를 우선 정하고(q_1), 정해진 개수에 따라 컬럼을 순서대로 선택한다. 컬럼은 컬럼 이름과 테이블 이름의 순서쌍으로 나타내며, 붉은 원들에서 출발하는 화살표는 컬럼 개수만큼의 액션을 나타낸다.
- [0066] 사영 컬럼 이외의 컬럼을 선택하는 방식은, 먼저 컬럼의 선택 여부를 결정한다(q_2). 선택하기로 결정했으면 컬럼을 선택하고, q_2 상태로 다시 돌아가서 컬럼 선택 여부를 결정을 반복한다. 만일, 선택하지 않기로 결정한다면 다음 단계로 넘어간다.
- [0067] 이후, 푸시다운 트랜스듀서는, 도 7에 도시된 전이 규칙에 따라, 박스를 생성하여 박스의 사영 정보를 만든다. 먼저 박스가 SELECT 박스인지, GROUP 박스인지를 선택한다(q_4). 본 발명에서, 푸시다운 트랜스듀서는 정확도를 높이기 위해 의도적으로 간소화되었다. 이에 따라, INTERSECT, EXCEPT, UNION 박스를 지원하지 않고, 이러한 질의들은 IN 혹은 NOT IN 등을 이용해 다른 방식으로 표현한다. 일부 UNION 질의에 대해서는 변환이 불가능한 질의가 있지만, 극소수이다. q_5 상태에서는 몇 개의 사영 정보를 만들어낼지를 정한다. 정해진 개수만큼, 큐에서 컬럼을 추출하여 어떤 집계 함수(aggregator)를 적용할지를 결정한다. 즉, q_5 하위에 표기된 파란 원들은(집계 함수를 사용하지 않는다는 선택지 포함) 집계 함수의 개수 + 1 만큼의 액션 개수를 갖는다.
- [0068] 이후, 푸시다운 트랜스듀서는, 도 8에 도시된 전이 규칙에 따라, 박스의 술부(predicate)을 만들어 낸다. 본 발명에서, 푸시다운 트랜스듀서는 의도적으로 테이블이 어떻게 조인(join)되는지에 대한 술부는 만들어내지 않는다. 마찬가지로, 전이 규칙을 간소화하여 정확도를 높이기 위해, 조인 관련 술부는 테이블의 외래키 관계를 이용해 QGM 질의 생성 이후 추론 가능하다. q_6 상태에서는 술부 있는지 없는지를 선택하여, 있다면 q_7 으로, 없다면 q_{12} 으로 이동한다. q_7 내지 q_{10} 에서는 좌변, 연산자 및 우변으로 이루어진 술부 절을 하나 생성한다. q_7 은 좌변의 집계 함수를 선택하고, q_8 은 연산자를, q_9 은 우변이 값인지 부분질의 박스(subquery box)인지를 선택한다. 값인 경우, q_{10} 으로 이동한다. 부분질의 박스인 경우, q_5 로 이동하면서 새로운 박스를 생성한다. 이 경우, 스택에 임의의 기호 하나를 추가하여 부분질의 박스라는 표시를 한다. 질의의 값이 정확히 무엇인지 찾아내는 것은 본 발명의 범위가 아니며, 푸시다운 트랜스듀서에서는 찾지 않는다. 이후, q_{10} 에서는 술부 절이 더 있

는지 없는지를 선택하여, 있다면 q_{11} 에서 두 술부를 잇는 접속사(AND, OR)를 선택하고 상술한 단계를 반복한다.

[0069] 이후, 푸시다운 트랜스듀서는, 도 9에 도시된 전이 규칙 및 종료 규칙에 따라, 정렬 정보(ORDER BY)를 만들고 종료한다. 즉, q_{12} 에서 정렬하는지 안하는지를 선택하며, 정렬하는 경우에 q_{13} 로 이동하여 컬럼에 적용할 집계 함수를 선택하고, q_{14} 에서는 오름차순 정렬인지 내림차순 정렬인지를 선택한다.

[0070] q_{15} 는 QGM 질의 생성의 종료 여부를 선택하는데, 생성 중인 박스가 메인 박스인 경우에 종료하고, 부분질의 박스인 경우에 부모 박스의 생성으로 돌아간다. 스택에서 값을 추출하여 값이 없다면 메인 박스이고, 값이 있다면 부분질의 박스라는 것을 알 수 있다.

[0071] <푸시다운 트랜스듀서에 입력될 문자열을 생성하는 인공 신경망의 동작>

[0072] 한편, 인공 신경망은 입력되는 데이터베이스(d) 및 자연어 질의(q_{nl})에 따라 출력 문자열의 확률 분포를 예측하는 모델이다. 즉, 인공 신경망은 $P[(\sigma_1, \dots, \sigma_n) | d, q_{nl}]$ 을 모델링하는 확률 모델이다. 이때, $(\sigma_1, \dots, \sigma_n)$ 는 푸시다운 트랜스듀서에 입력될 문자열이다. 인공 신경망은 다음과 같이 자기회귀적으로 모델링한다.

[0073]
$$P[(\sigma_1, \dots, \sigma_n) | d, q_{nl}] = \prod_{i=1}^n P[\sigma_i | (\sigma_1, \dots, \sigma_{i-1}), d, q_{nl}]$$

[0074] 푸시다운 트랜스듀서의 상태 정보와 큐 정보를 이용하여 모델링 하기 위해, 각각을 다음과 같이 분해한다.

[0075]
$$P[\sigma_i | (\sigma_1, \dots, \sigma_{i-1}), d, q_{nl}] = P[\sigma_i | q_i, \gamma_i, (\sigma_1, \dots, \sigma_{i-1}), d, q_{nl}] \cdot P[q_i, \gamma_i | (\sigma_1, \dots, \sigma_{i-1}), d, q_{nl}]$$

[0076] q_i 는 $(\sigma_1, \dots, \sigma_{i-1})$ 를 입력으로 받고 나서의 상태이며, γ_i 는 이때의 큐에서 다음으로 추출될 심볼이다.

[0077] $P[q_i, \gamma_i | (\sigma_1, \dots, \sigma_{i-1}), d, q_{nl}]$ 는 1로 정해져 있기 때문에(입력 문자열이 정해지면 상태와 큐 심볼은 정해지기 때문), 인공 신경망은 앞선 항을 모델링한다.

[0078] 도 10은 인공 신경망 모델의 구조를 나타낸다.

[0079] 도 10을 참조하면, 우선 RA-트랜스포머(Peter Shaw, Jakob Uszkoreit, and Ashish Vaswani. Self-Attention with Relative Position Representations. NAACL 2018)를 사용하여 q_{nl} 및 d를 인코딩한다. 인코딩 방식은 RATSQ(Bailin Wang, Richard Shin, Xiaodong Liu, Oleksandr Polozov, Matthew Richardson. RAT-SQL: Relation-Aware Schema Encoding and Linking for Text-to-SQL Parsers. ACL 2020)과 유사하게 스키마 연결(Schema linking) 정보 관계를 자연어 질의와 스키마의 단어들과 함께 입력한다. 자연어 질의를 의존 분석(dependency parsing)한 정보 관계 또한 인코딩한다. 인코딩된 벡터들을 트랜스포머 디코더에 넘겨 트랜스포머 디코더는 문자열을 자기회귀적으로 생성한다. 생성 시, 트랜스포머 디코더의 마지막 레이어에서 다중 퍼셉트론을 이용해 상태 및 큐 심볼 또한 입력으로 사용한다. 소프트맥스를 적용하여 각 문자들의 확률 분포를 계산하는데, 푸시다운 트랜스듀서가 받아들이지 못하는 문자에 대해서는 마스킹을 하여 확률 분포가 0이 되도록 한다.

[0080] <실시예>

[0081] 도 11은 q_{nl} , q_{sql} , d, 및 q_{qgm} 에 대한 예시를 나타낸다.

[0082] 도 11을 참조하면, q_{qgm} 은 q_{sql} 과 동일한 의미를 가지는데, 예시를 위해 간소화되어 쉽게 생성될 수 있게 만들었다. 본 예시는 푸시다운 트랜스듀서를 중심으로 설명한다.

[0083] 도 12 내지 도 15는 도 11에 따른 각 단계의 결과를 나타낸다. 즉, 도 12는 q_1 내지 q_3 상태를 거친 중간 결과, 도 13은 q_4 및 q_5 를 거치며 사영 정보를 생성한 중간 결과, 도 14는 q_6 내지 q_9 을 거쳐 술부 절 하나 및 부분 질의 박스를 생성한 결과, 도 15는 최종 결과를 각각 나타낸다.

[0084] 먼저, 도 12에 도시된 바와 같이, 푸시다운 트랜스듀서가 q_1 내지 q_3 상태를 통해 컬럼들을 생성해 내어 큐에 컬럼들을 삽입한다. 즉, student 테이블의 age 컬럼과 allergy_type 테이블의 allergytype 컬럼 2개가 생성되어, 큐에 3개 컬럼이 삽입된다.

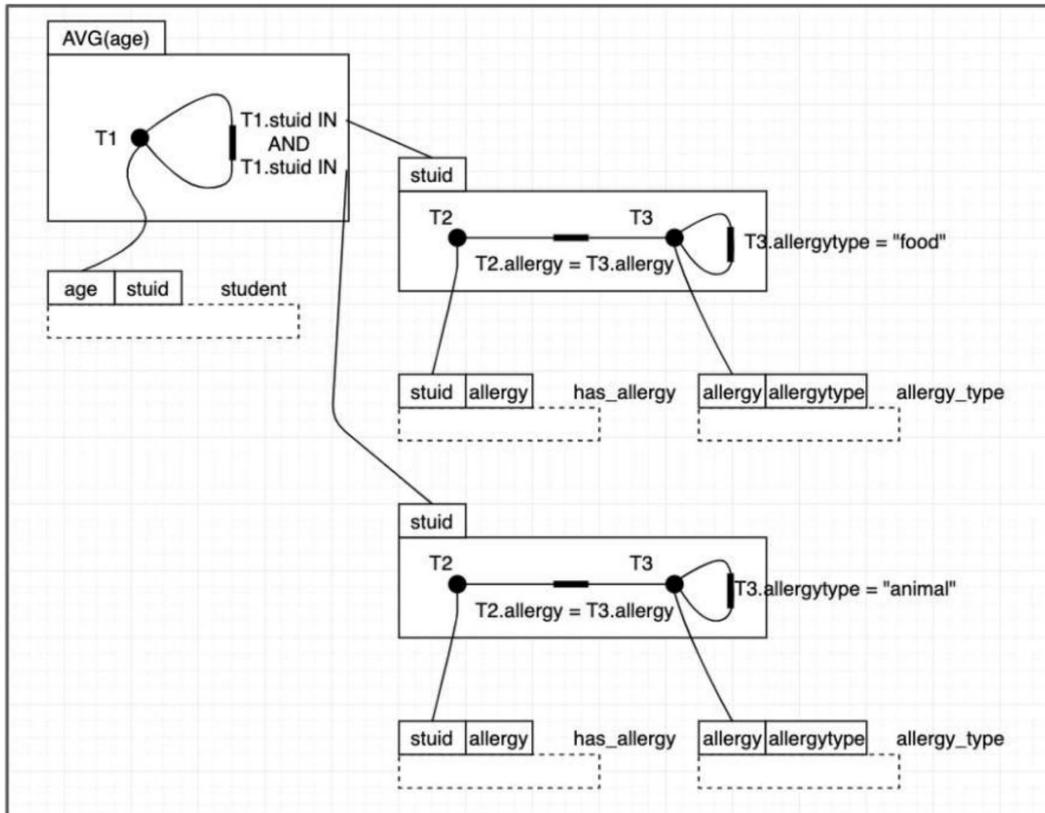
도면

도면1

```

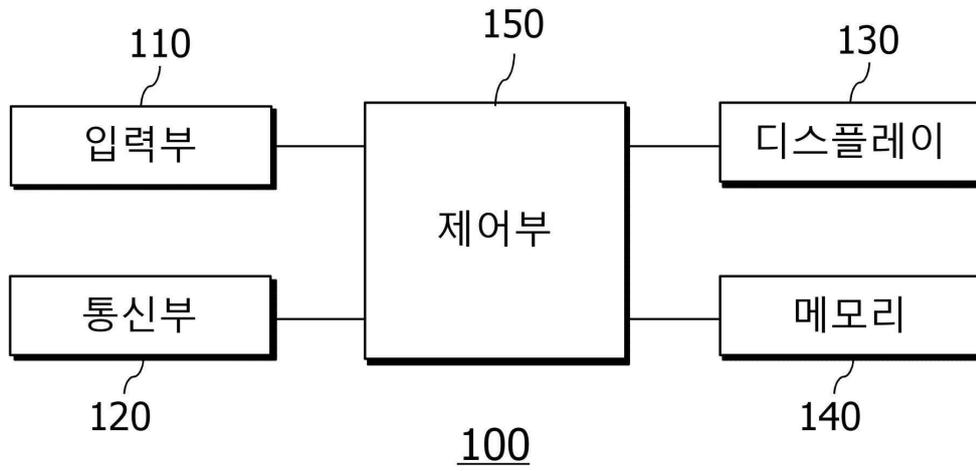
SELECT avg(age)
FROM student
WHERE stuid IN (SELECT T1.stuid
                FROM has_allergy AS T1
                JOIN allergy_type AS T2
                ON T1.allergy = T2.allergy
                WHERE T2.allergytype = "food")
AND IN (
SELECT T1.stuid
FROM has_allergy AS T1
JOIN allergy_type AS T2
ON T1.allergy = T2.allergy
WHERE T2.allergytype = "animal")
    
```

(a)

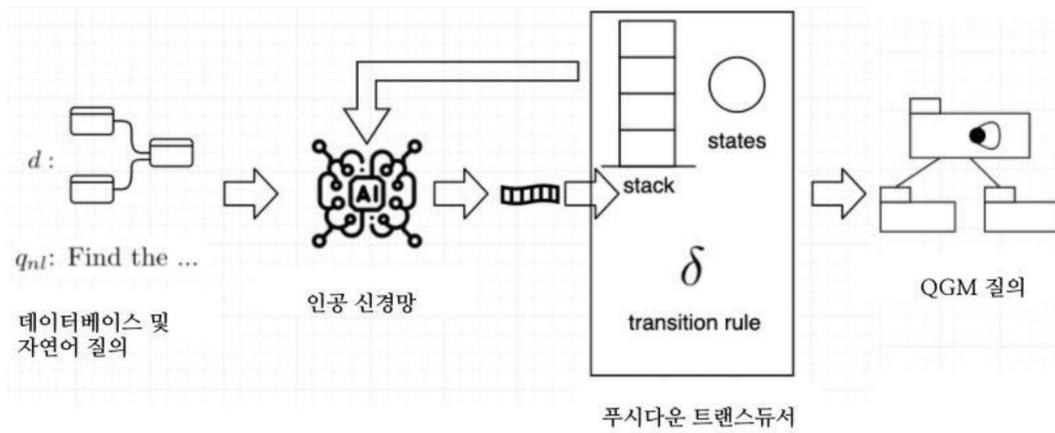


(b)

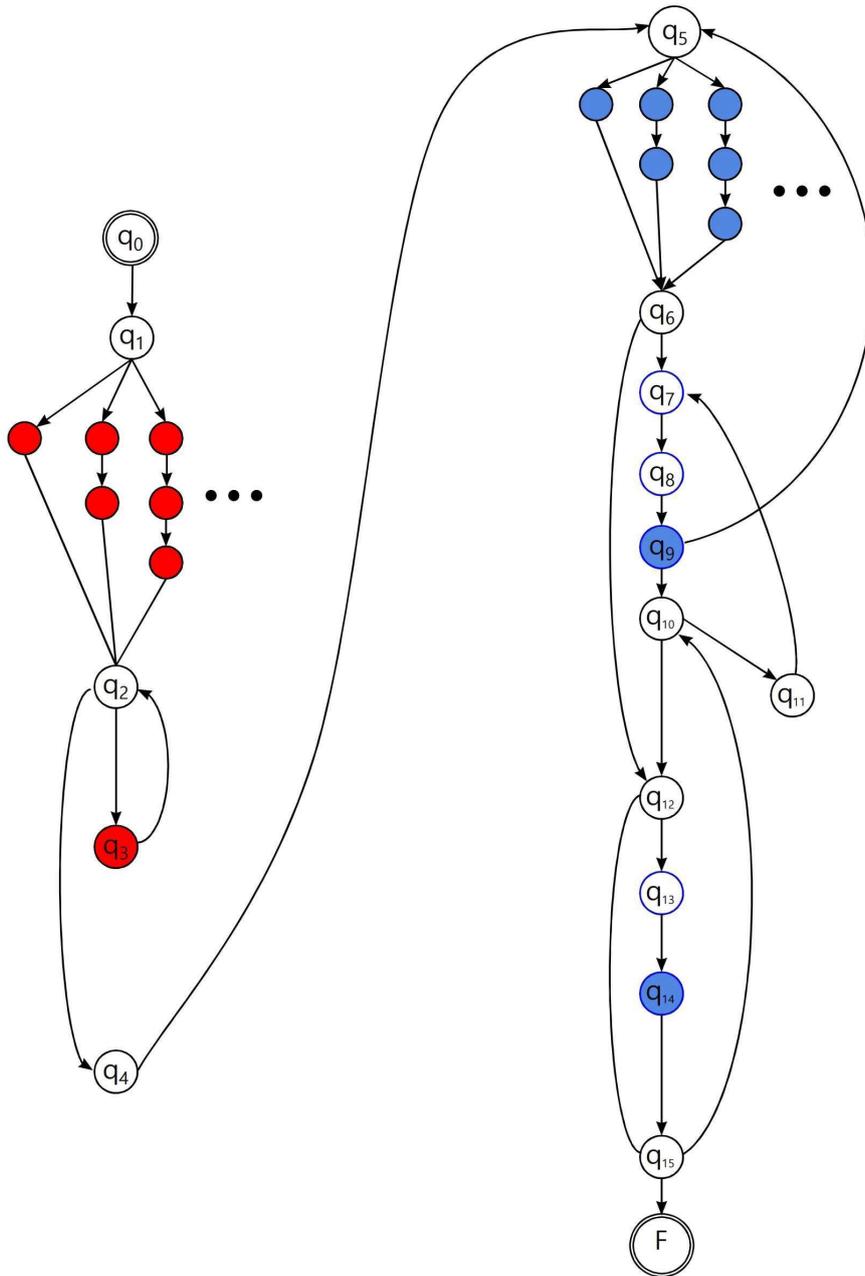
도면2



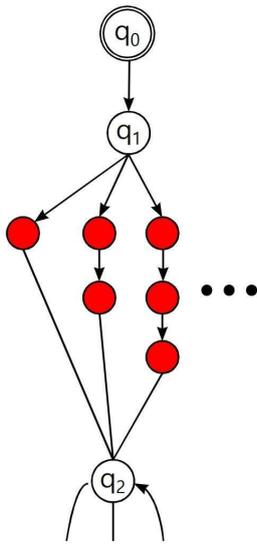
도면3



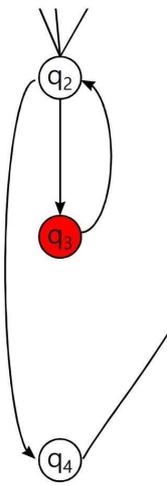
도면4



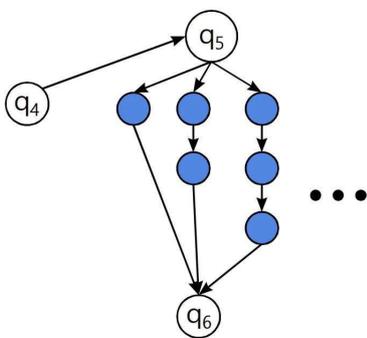
도면5



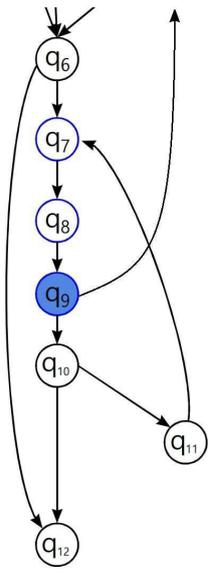
도면6



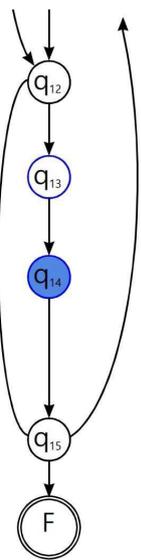
도면7



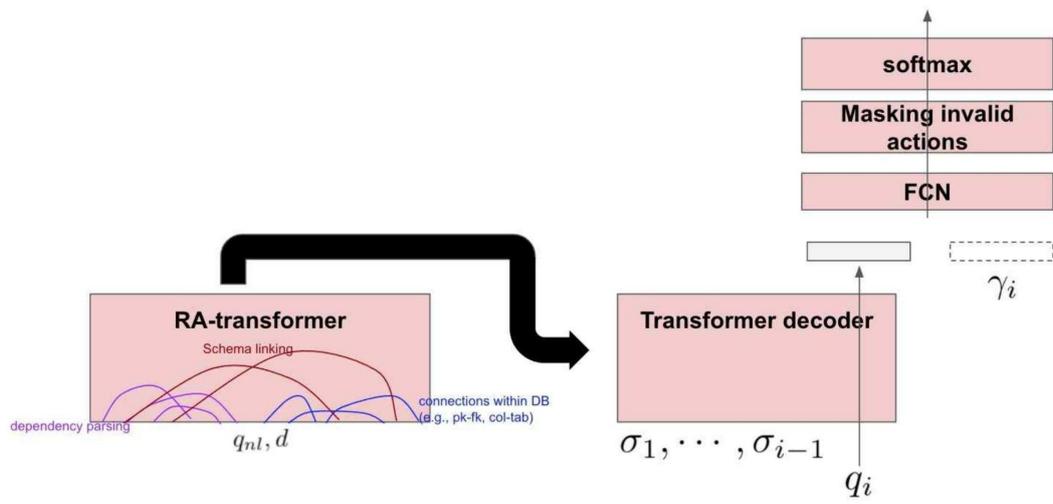
도면8



도면9



도면10

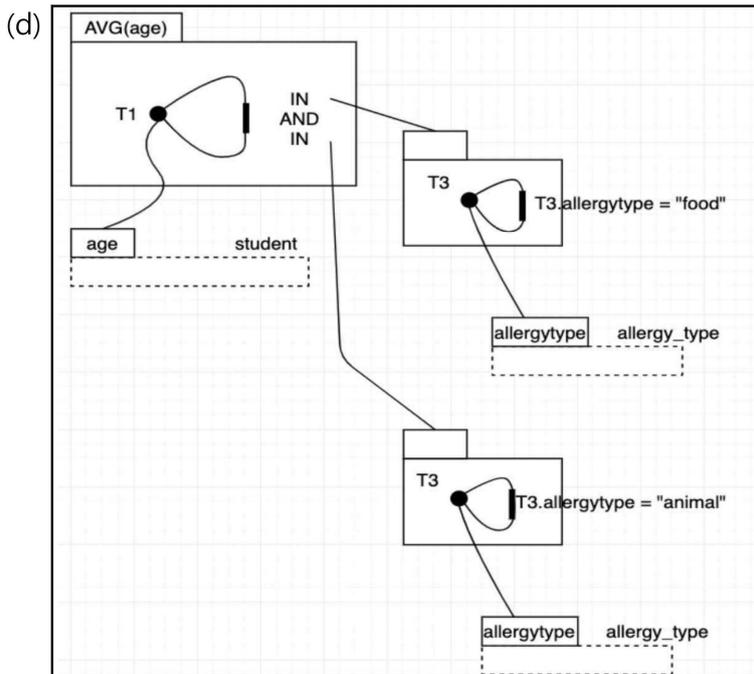
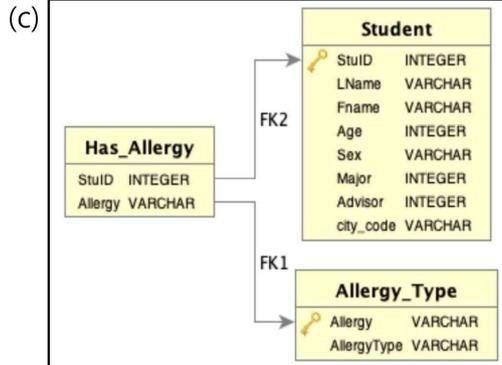


도면11

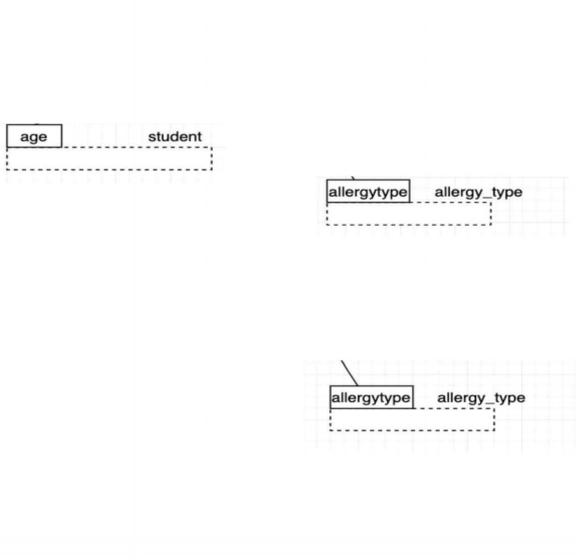
(a) "Find the average age of the students who have allergies with food and animal types"

```

(b) SELECT avg(age) FROM student
WHERE stuid IN (SELECT T1.stuid
                FROM has_allergy AS T1 JOIN allergy_type AS T2
                ON T1.allergy = T2.allergy
                WHERE T2.allergytype = "food")
AND IN (SELECT T1.stuid
        FROM has_allergy AS T1 JOIN allergy_type AS T2
        ON T1.allergy = T2.allergy
        WHERE T2.allergytype = "animal")
    
```



도면12

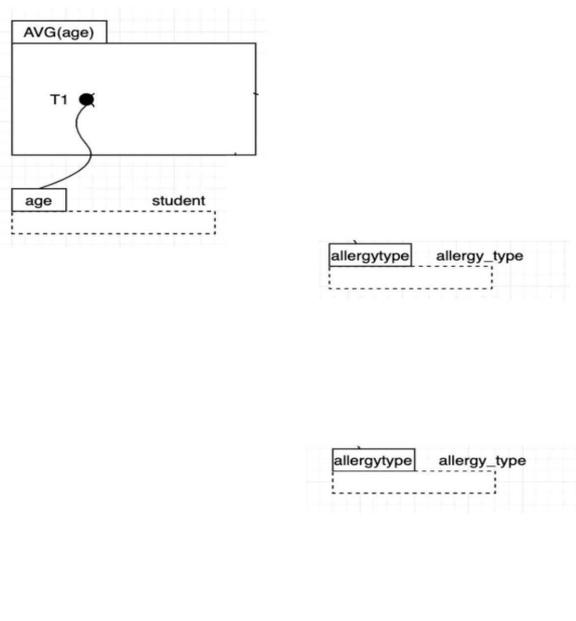


allergytype, allergy_type
allergytype, allergy_type
age, student

큐

스택

도면13

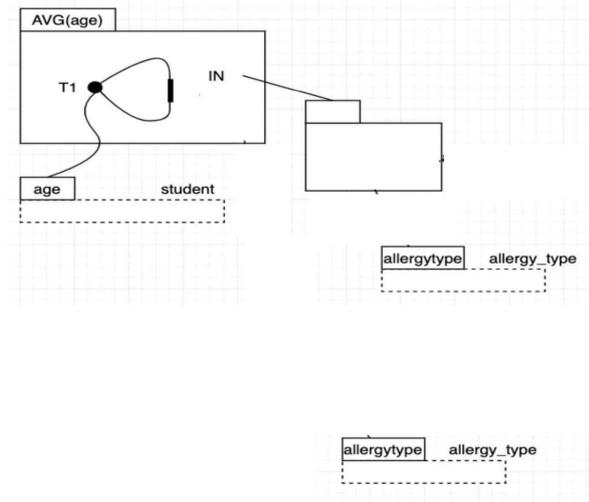


allergytype, allergy_type
allergytype, allergy_type

큐

스택

도면14



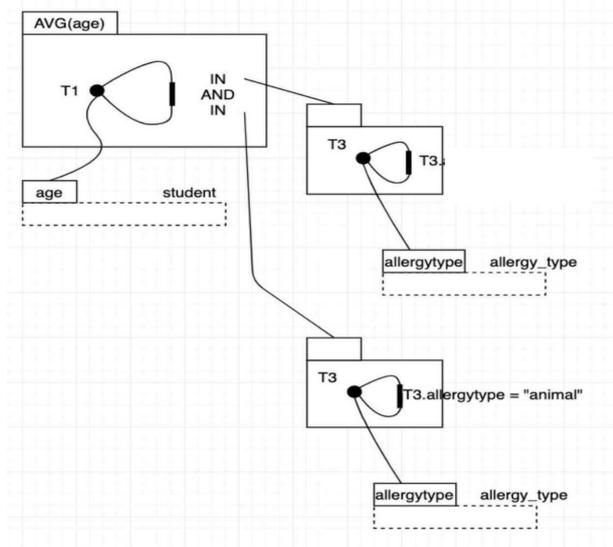
allergytype, allergy_type
allergytype, allergy_type

큐

[Subquery]

스택

도면15



큐

스택