(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0031327
(43) 공개일자 2025년03월07일

(51) 국제특허분류(Int. Cl.)
G06F 16/332 (2025.01) *G06F 16/33* (2025.01)
G06F 16/36 (2019.01) *G06F 40/268* (2020.01)
G06F 40/289 (2020.01) *G06N 5/04* (2023.01)
(52) CPC특허분류
G06F 16/3329 (2025.01)
G06F 16/3325 (2019.01)
(21) 출원번호 10-2023-0112630
(22) 출원일자 2023년08월28일
심사청구일자 2023년08월28일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
이상근
서울특별시 강남구 선릉로 221(도곡동, 도곡렉슬아파트)
목원람
서울특별시 성북구 고려대로22길 25 (안암동5가) 3층 308호 02855
(74) 대리인
윤귀상

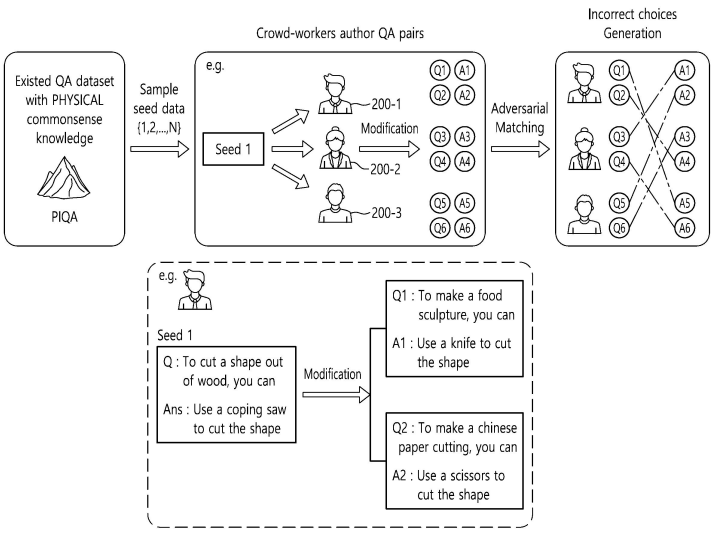
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **크라우드소싱을 이용한 물리적 상식 추론 벤치마크를 생성하는 시스템 및 방법**

(57) 요 약

본 발명은 크라우드소싱을 이용한 물리적 상식 추론 벤치마크(PIQA: Physical Interaction Question Answer)를 생성하는 시스템에 대한 것으로, 미리 마련된 PIQA로부터 추출된 "질문-응답" 쌍(pair)을 수신하고, 수신된 "질문-응답" 쌍을 기반으로 복수개의 수정된 "질문-응답" 쌍을 생성하는 크라우드 단말; 및 상기 수정된 "질문-응답" 쌍에서 서로 다른 쌍에 포함되는 질문과 응답을 결합하여 새로운 "질문-응답" 쌍을 생성하는 서버를 포함한다.

대 표 도 - 도3



(52) CPC특허분류
G06F 16/3334 (2019.01)
G06F 16/36 (2019.01)
G06F 40/268 (2020.01)
G06F 40/289 (2020.01)
G06N 5/04 (2023.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711193973
과제번호 2019-0-00079-005
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신방송혁신인재양성
연구과제명 인공지능대학원지원(고려대학교)
기 여 율 1/2
과제수행기관명 고려대학교산학협력단
연구기간 2023.01.01 ~ 2023.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711188552
과제번호 2021R1A2C3010430
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)
연구과제명 미플랫폼: 자기지도학습과 연합학습으로 강화되는 작은 인공지능 플랫폼
기 여 율 1/2
과제수행기관명 고려대학교
연구기간 2023.03.01 ~ 2024.02.29

명 세 서

청구범위

청구항 1

크라우드소싱을 이용한 물리적 상식 추론 벤치마크(PIQA: Physical Interaction Question Answer)를 생성하는 시스템에 있어서,

미리 마련된 PIQA로부터 추출된 "질문-응답" 쌍(pair)을 수신하고, 수신된 "질문-응답" 쌍을 기반으로 복수개의 수정된 "질문-응답" 쌍을 생성하는 크라우드 단말; 및

상기 수정된 "질문-응답" 쌍에서 서로 다른 쌍에 포함되는 질문과 응답을 결합하여 새로운 "질문-응답" 쌍을 생성하는 서버를 포함하는, 크라우드소싱을 이용한 PIQA 생성 시스템.

청구항 2

제1항에 있어서,

상기 크라우드 단말은,

상기 수신된 "질문-응답" 쌍에 포함되는 응답 구문(statement)을 기반으로 대상(object) 및 상기 대상의 활용(utilization)을 식별하고,

상기 응답 구문에서 상기 식별된 대상 또는 상기 활용에 대응하는 표현을 수정하여 수정된 응답 구문을 생성하고,

상기 수정된 응답 구문에 매칭되는 새로운 질문을 생성함으로써 상기 수정된 "질문-응답" 쌍을 생성함을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템.

청구항 3

제2항에 있어서,

상기 크라우드 단말은,

상기 수신된 "질문-응답" 쌍에 포함되는 질문 구문 및 응답 구문에 공통적으로 포함되는 명사(noun)/동사(verb)를 확인하고, 상기 식별된 대상 및 활용 중 적어도 하나의 요소가 상기 명사/동사에 포함될 경우 해당 요소를 다른 표현으로 변경함을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템.

청구항 4

제3항에 있어서,

상기 크라우드 단말은,

상기 응답 구문에서 상기 식별된 대상 또는 상기 활용에 대응하는 표현을 상기 응답 구문의 상황(situation) 내에서 그럴듯한 대안(plausible alternatives)이 되는 표현으로 수정하고, 상기 추출된 명사/동사를 포함하도록 상기 수정된 응답 구문을 생성함을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템.

청구항 5

제2항에 있어서,

상기 크라우드 단말은,

상기 대상 및 활용 중 적어도 하나의 요소가 식별되지 않을 경우, 해당 데이터를 스킵(skip)하고,

상기 식별된 대상 및 활용의 쌍이 하나 이상일 경우, 상기 활용이 하나인 대상만을 식별함을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템.

청구항 6

크라우드소싱을 이용한 물리적 상식 추론 벤치마크(PIQA: Physical Interaction Question Answer)를 생성하는 시스템의 PIQA 생성 방법에 있어서,

크라우드 단말이 미리 마련된 PIQA로부터 추출된 "질문-응답" 쌍(pair)을 수신하고, 수신된 "질문-응답" 쌍을 기반으로 복수개의 수정된 "질문-응답" 쌍을 생성하는 단계; 및

서버가 상기 수정된 "질문-응답" 쌍에서 서로 다른 쌍에 포함되는 질문과 응답을 결합하여 새로운 "질문-응답" 쌍을 생성하는 단계;를 포함하는, 크라우드소싱을 이용한 PIQA 생성 시스템의 PIQA 생성 방법.

청구항 7

제6항에 있어서.

상기 "질문-응답" 쌍을 생성하는 단계는,

상기 수신된 "질문-응답" 쌍에 포함되는 응답 구문(statement)을 기반으로 대상(object) 및 상기 대상의 활용(utilization)을 식별하는 단계;

상기 응답 구문에서 상기 식별된 대상 또는 상기 활용에 대응하는 표현을 수정하여 수정된 응답 구문을 생성하는 단계; 및

상기 수정된 응답 구문에 매칭되는 새로운 질문을 생성함으로써 상기 수정된 "질문-응답" 쌍을 생성하는 단계;를 포함함을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템의 PIQA 생성 방법.

청구항 8

제7항에 있어서,

상기 수신된 "질문-응답" 쌍에 포함되는 질문 구문 및 응답 구문에 공통적으로 포함되는 명사(noun)/동사(verb)를 확인하고, 상기 식별된 대상 및 활용 중 적어도 하나의 요소가 상기 명사/동사에 포함될 경우 해당 요소를 다른 표현으로 변경하는 단계를 더 포함함을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템의 PIQA 생성 방법.

청구항 9

제8항에 있어서,

상기 수정된 응답 구문을 생성하는 단계는,

상기 응답 구문에서 상기 식별된 대상 또는 상기 활용에 대응하는 표현을 상기 응답 구문의 상황(situation) 내에서 그럴듯한 대안(plausible alternatives)이 되는 표현으로 수정하고, 상기 추출된 명사/동사를 포함하도록 상기 수정된 응답 구문을 생성하는 것임을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템의 PIQA 생성 방법.

청구항 10

제7항에 있어서,

상기 대상 및 활용 중 적어도 하나의 요소가 식별되지 않을 경우, 해당 데이터를 스킵(skip)하는 단계; 및

상기 식별된 대상 및 활용의 쌍이 하나 이상일 경우, 상기 활용이 하나인 대상만을 식별하는 단계를 더 포함함을 특징으로 하는, 크라우드소싱을 이용한 PIQA 생성 시스템의 PIQA 생성 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 크라우드소싱(crowdsourcing)을 이용한 물리적 상식 추론 벤치마크를 생성하는 시스템 및 방법에 관한 것이다.

배 경 기 술

[0003] 인간이 일상생활에서 결정을 내리고 의사소통을 하는데 상식추론 능력이 필요한 것처럼, 기계가 현실 세계에 대한 깊은 이해를 하기 위해서는 상식 지식을 습득하고 추론하는 능력이 필요하다. 하지만 기계가 인간 수준의 상식추론 능력을 갖고 있는지는 의문이므로, 새로운 상황에 일반화 시킬 수 있는 상식추론 능력을 필요로 하는 벤치마크들이 다수 개발되었다.

[0004] 최근 자연어처리 분야의 빠른 발전 속도에 따라 사전학습 언어모델들이 높은 수준의 자연어 이해 능력을 갖게 되면서 자연어이해와 자연어추론 벤치마크에서 뛰어난 성능을 보여주었다. 그러나 언어모델들의 높은 상식추론 능력이 과대평가되어 있음을 보여주는 연구들이 대거 등장하고 있으며, 이는 언어모델들이 실제로 상식추론 객관식 질의응답에 강건한 이해 능력을 갖고 있는 것이 아니라 데이터셋에 있는 주석 아티팩트(Annotation artifacts)에 의존하여 문제를 쉽게 해결하고 있는 것은 아닌지 의문을 제기한다.

[0005] 따라서 이러한 주석 아티팩트를 포함하지 않도록 하는 크라우드소싱 기법의 개발이 요구된다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국공개특허공보 제10-2020-0135892호

발명의 내용

해결하려는 과제

[0008] 본 발명은 상기와 같은 문제를 해결하기 위해 안출된 것으로, 본 발명의 목적은 크라우드소싱을 이용한 물리적 상식 추론 벤치마크를 생성하는 시스템 및 방법을 제공하는 것이다.

과제의 해결 수단

[0010] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 크라우드소싱을 이용한 물리적 상식 추론 벤치마크(PIQA)를 생성하는 시스템으로서, 미리 마련된 PIQA로부터 추출된 "질문-응답" 쌍(pair)을 수신하고, 수신된 "질문-응답" 쌍을 기반으로 복수개의 수정된 "질문-응답" 쌍을 생성하는 크라우드 단말; 및 상기 수정된 "질문-응답" 쌍에서 서로 다른 쌍에 포함되는 질문과 응답을 결합하여 새로운 "질문-응답" 쌍을 생성하는 서버를 포함한다.

[0011] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 크라우드소싱을 이용한 물리적 상식 추론 벤치마크

(PIQA)를 생성하는 시스템의 PIQA 생성 방법으로서, 크라우드 단말이 미리 마련된 PIQA로부터 추출된 "질문-응답" 쌍(pair)을 수신하고, 수신된 "질문-응답" 쌍을 기반으로 복수개의 수정된 "질문-응답" 쌍을 생성하는 단계; 및 서버가 상기 수정된 "질문-응답" 쌍에서 서로 다른 쌍에 포함되는 질문과 응답을 결합하여 새로운 "질문-응답" 쌍을 생성하는 단계;를 포함한다.

발명의 효과

[0013] 상술한 본 발명의 일측면에 따르면, 크라우드소싱을 이용한 물리적 상식 추론 벤치마크를 생성하는 시스템 및 방법을 제공함으로써, 단어 간의 세밀하고 내재적인 관계까지 고려하여 주석 아티팩트를 포함하지 않고 공정하게 평가할 수 있는 벤치마크를 얻을 수 있다. 뿐만 아니라 과대평가 없이 언어모델의 상식추론 능력을 정확하게 측정할 수 있고, 앞으로 진행될 연구에 대한 상식추론 능력의 결핍을 찾아내는 도구가 될 수 있다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 실시예에 따른 물리적 상식 추론 벤치마크를 생성하는 시스템을 나타낸 도면,

도 2는 본 발명의 실시예에 따른 크라우드 단말의 내부 블록을 나타낸 장치도,

도 3은 본 발명의 실시예에 따른 물리적 상식 추론 벤치마크를 생성하는 전체 프로세스를 나타낸 도면,

그리고, 도 4는 본 발명의 실시예에 따른 크라우드 단말의 물리적 상식 추론 벤치마크를 생성하는 동작을 나타낸 순서도.

발명을 실시하기 위한 구체적인 내용

[0016] 후술하는 본 발명에 대한 상세한 설명은, 본 발명이 실시될 수 있는 특정 실시예를 예시로서 도시하는 첨부 도면을 참조한다. 이들 실시예는 당업자가 본 발명을 실시할 수 있기에 충분하도록 상세히 설명된다. 본 발명의 다양한 실시예는 서로 다르지만 상호 배타적일 필요는 없음이 이해되어야 한다. 예를 들어, 여기에 기재되어 있는 특정 형상, 구조 및 특성은 일 실시예와 관련하여 본 발명의 정신 및 범위를 벗어나지 않으면서 다른 실시예로 구현될 수 있다. 또한, 각각의 개시된 실시예 내의 개별 구성요소의 위치 또는 배치는 본 발명의 정신 및 범위를 벗어나지 않으면서 변경될 수 있음이 이해되어야 한다. 따라서, 후술하는 상세한 설명은 한정적인 의미로서 취하려는 것이 아니며, 본 발명의 범위는, 적절하게 설명된다면, 그 청구항들이 주장하는 것과 균등한 모든 범위와 더불어 첨부된 청구항에 의해서만 한정된다. 도면에서 유사한 참조부호는 여러 측면에 걸쳐서 동일하거나 유사한 기능을 지칭한다.

[0017] 본 발명에 따른 구성요소들은 물리적인 구분이 아니라 기능적인 구분에 의해서 정의되는 구성요소들로써 각각이 수행하는 기능들에 의해서 정의될 수 있다. 각각의 구성요소들은 하드웨어 또는 각각의 기능을 수행하는 프로그램 코드 및 프로세싱 유닛으로 구현될 수 있을 것이며, 두 개 이상의 구성요소의 기능이 하나의 구성요소에 포함되어 구현될 수도 있을 것이다. 따라서 이하의 실시예에서 구성요소에 부여되는 명칭은 각각의 구성요소를 물리적으로 구분하기 위한 것이 아니라 각각의 구성요소가 수행되는 대표적인 기능을 암시하기 위해서 부여된 것이며, 구성요소의 명칭에 의해서 본 발명의 기술적 사상이 한정되지 않는 것임에 유의하여야 한다.

[0018] 이하에서는 도면들을 참조하여 본 발명의 바람직한 실시예들을 보다 상세하게 설명하기로 한다.

[0019] 도 1은 본 발명의 실시예에 따른 물리적 상식 추론 벤치마크를 생성하는 시스템을 나타낸 도면이고, 도 2는 본 발명의 실시예에 따른 크라우드 단말의 내부 블록을 나타낸 장치도이고, 도 3은 본 발명의 실시예에 따른 물리적 상식 추론 벤치마크를 생성하는 전체 프로세스를 나타낸 도면이다.

[0020] 본 발명의 실시예에 따른 물리적 상식 추론 벤치마크인 PIQA(Physical Interaction Question Answer)를 생성하는 시스템은 서버(100)와 복수개의 크라우드 단말(200-1, 200-2, 200-3, ...)을 포함한다. 여기서 크라우드 단말이라 함은 크라우드소싱 플랫폼을 통해 작업을 수행하는 크라우드 작업자들(workers)이 작업을 수행하기 위해 사용하는 장치를 의미한다.

[0021] 도 3에 나타낸 바와 같이 서버(100)는 물리적 상식 추론을 평가하기 위해 미리 마련된 데이터 셋인 PIQA로부터 "질문(Question)-응답(Answer)" 쌍(pair)을 추출하고, 추출된 "질문-응답" 쌍을 초기 데이터(seed data)로써

복수의 크라우드 작업자들, 즉 크라우드 단말(200-1, 200-2, 200-3)로 전송한다.

[0022] 크라우드 단말(200-1, 200-2, 200-3) 각각은 수신된 "질문-응답" 쌍을 기반으로 복수개의 수정(modification)된 "질문-응답" 쌍을 생성하고, 상기 수정된 "질문-응답" 쌍을 서버(100)로 전송한다.

[0023] 서버(100)는 상기 수정된 "질문-응답" 쌍에서 서로 다른 쌍에 포함되는 질문과 응답을 결합하여 새로운 "질문-응답" 쌍을 생성한다.

[0024] 예를 들어 설명하면, 서버(100)로부터 수신되는 초기 데이터를 통해 크라우드 단말(200-1)이 "Q1-A1", "Q2-A2" 쌍을 생성하고, 크라우드 단말(200-2)이 "Q3-A3", "Q4-A4" 쌍을 생성하고, 크라우드 단말(200-3)이 "Q5-A5", "Q6-A6" 쌍을 생성한다고 가정하면, 서버(100)는 경쟁적 매칭(Adversarial matching)을 통해 "Q1-A5", "Q2-A4", "Q3-A1", "Q4-A6", "Q5-A2", "Q6-A3" 쌍을 생성한다.

[0025] 한편, 도 2에 도시된 크라우드 단말(200)은 수신부(210), 식별부(220) 및 생성부(230)를 포함하며, 이하에서는 상기 크라우드 단말(200)이 수정된 "질문-응답" 쌍을 생성하는 과정을 보다 상세히 설명하도록 한다.

[0026] 수신부(210)는 서버(100)로부터 전송되는 초기 데이터, 즉 미리 마련된 PIQA로부터 추출된 "질문-응답" 쌍을 수신하고, 식별부(220)는 상기 수신된 "질문-응답" 쌍에 포함되는 응답 구문(statement)을 기반으로 대상(object) 및 상기 대상의 활용(utilization)을 식별한다.

[0027] 생성부(230)는 상기 수신부(210)를 통해 수신된 "질문-응답" 쌍의 응답 구문에서 상기 식별된 대상 또는 상기 활용에 대응하는 표현을 수정하여 수정된 응답 구문을 생성하고, 상기 수정된 응답 구문을 위한 새로운 질문을 생성하여 수정된 "질문-응답" 쌍을 생성한다.

[0028] 이하에서는 크라우드 단말이 수정된 "질문-응답" 쌍을 생성하는 과정에 대해 적절한 예시를 통해 단계별로 설명하도록 한다.

[0029] - 제1 단계

[0030] 질문: To **cut** a **shape** out of wood, you can.

[0031] 응답: Use a copy saw to **cut** the **shape**.

[0032] 크라우드 단말은 위와 같은 "질문-응답" 쌍이 수신되면, 상기 질문 구문과 상기 응답 구문에 공통적으로 포함되는 명사(noun)/동사(verb)를 확인한다. 이때 크라우드 단말은 명사의 단수/복수와 동사의 시제 변화를 고려하지 않고 상기 명사/동사에 해당하는 단어를 확인한다. 위 예시에서 질문 구문과 응답 구문에 공통적으로 포함되는 명사/동사는 'cut', 'shape'이다.

[0033] - 제2 단계

[0034] 크라우드 단말은 "질문-응답" 쌍의 응답 구문에서 대상과 상기 대상의 활용을 식별한다. 여기서 대상의 활용이라 함은 응답 구문의 상황(situation)에서 상기 대상이 어떻게 사용되는지를 의미한다. 이때 크라우드 단말은 식별된 대상과 활용 중 적어도 하나의 요소가 제1 단계에서 확인된 명사/동사에 포함될 경우 해당 요소를 다른 표현으로 변경한다. 위 예시에서 식별된 대상은 'copy saw'이고, 활용은 'cut'이나 상기 활용에 해당하는 단어가 제1 단계에서 확인된 단어, 즉 'cut'에 포함되므로 해당 요소를 다른 표현인 'use'로 변경한다.

[0035] 또한 크라우드 단말은 상기 대상 및 활용 중 적어도 하나의 요소가 식별되지 않을 경우, 해당 데이터를 스킵(skip)하고, 상기 식별된 대상 및 활용의 쌍이 하나 이상일 경우, 상기 활용이 하나인 대상만을 식별한다.

[0036] - 제3 단계

[0037] 크라우드 단말은 제2 단계에서 식별된 대상과 활용에 대하여 응답 구문의 상황 내에서 그럴듯한 대안(plausible alternatives)이 되는 표현들을 브레인스토밍한다. 위 예시에서 식별된 대상인 'copy saw'에 대한 그럴듯한 대안이 되는 표현은 일례로 'knife', 'scissors', 'cutter' 등이 될 수 있고, 식별된 활용인 'use'에 대한 그럴듯한 대안이 되는 표현은 일례로 'utilize', 'work with', 'make use of' 등이 될 수 있다.

[0038] - 제4 단계

[0039] 크라우드 단말은 "질문-응답" 쌍의 응답 구문에서 제2 단계에서 식별된 대상 또는 활용에 대한 표현을 제3 단계에서 브레인스토밍한 표현들 중 하나의 표현으로 대체하여 수정된 응답 구문인 응답 1을 생성한다. 이때 크라우드 단말은 상기 대상 또는 활용에 대한 표현을 브레인스토밍한 표현들 중 가능한한 가장 적은 수정에 해당하는

표현으로 대체하며, 상기 응답 1은 제1 단계에서 확인된 명사/동사 모두를 포함하도록 생성된다. 이를 통해 생성된 응답 1의 예시는 다음과 같다.

[0040] 응답 1: use a **knife** to **cut** the **shape**.

[0041] 또한 크라우드 단말은 응답 1의 생성시 직접적인 부정 표현, 상식에 어긋나는 표현, 과도한 수정에 유의하여 응답 1을 생성하도록 한다.

[0042] - 제5 단계

[0043] 크라우드 단말은 제4 단계에서 수정된 응답 구문인 응답 1에 매칭되는 질문 구문, 즉 질문 1을 생성함으로써 "질문 1-응답 1" 쌍을 생성한다. 응답 1에 매칭되는 질문 1의 예시는 다음과 같다.

[0044] 질문 1: To make a **food sculpture**, you can.

[0045] - 제6 단계

[0046] 크라우드 단말은 앞서 설명한 제4 단계 및 제5 단계를 반복하여 "질문 2-응답 2" 쌍을 생성한다. "질문 2-응답 2" 쌍의 예시는 다음과 같다.

[0047] 질문 2: To make a **Chinese paper cutting**, you can.

[0048] 응답 2: use a **scissors** to **cut** the **shape**.

[0049] 도 4는 본 발명의 실시예에 따른 크라우드 단말의 물리적 상식 추론 벤치마크를 생성하는 동작을 나타낸 순서도이다.

[0050] 크라우드 단말은 "질문-응답" 쌍이 수신되면,(S402) 질문 구문과 응답 구문에 공통적으로 포함되는 명사/동사를 확인한다.(S404)

[0051] 또한 크라우드 단말은 응답 구문에서 대상과 상기 대상의 활용을 식별하고,(S406) 상기 식별된 대상과 활용에 대하여 응답 구문의 상황 내에서 그럴듯한 대안이 되는 표현들을 브레인스토밍한다.(S408)

[0052] 이후 크라우드 단말은 S402에서 수신된 "질문-응답" 쌍의 응답 구문에서 대상 또는 활용에 대한 표현을 상기 브레인스토밍한 표현들 중 하나의 표현으로 대체하여 응답 1을 생성하고, 상기 응답 1에 매핑되는 질문 1을 생성하여 "질문 1-응답 1" 쌍을 생성한다.(S410)

[0053] 크라우드 단말은 S410의 "질문 1-응답 1" 쌍 생성과 동일한 과정을 통해 "질문 2-응답 2" 쌍을 생성한다.(S412)

[0054] 위와 같은 본 발명의 크라우드소싱을 이용한 PIQA 생성 방법은 다양한 컴퓨터 구성요소를 통하여 수행될 수 있는 프로그램 명령어의 형태로 구현되어 컴퓨터 판독 가능한 기록 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능한 기록 매체는 프로그램 명령어, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다.

[0055] 상기 컴퓨터 판독 가능한 기록 매체에 기록되는 프로그램 명령어는 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수도 있다.

[0056] 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD 와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 ROM, RAM, 플래시 메모리 등과 같은 프로그램 명령어를 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다.

[0057] 프로그램 명령어의 예에는, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 상기 하드웨어 장치는 본 발명에 따른 처리를 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0058] 이상에서는 본 발명의 다양한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

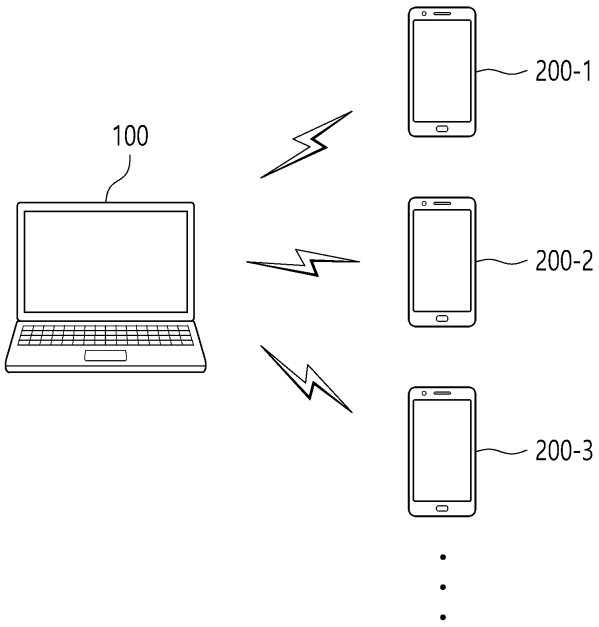
부호의 설명

[0060] 100: 서버

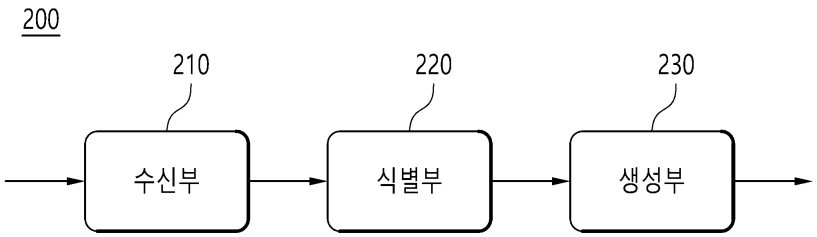
200, 200-1, 200-2, 200-3: 크라우드 단말

도면

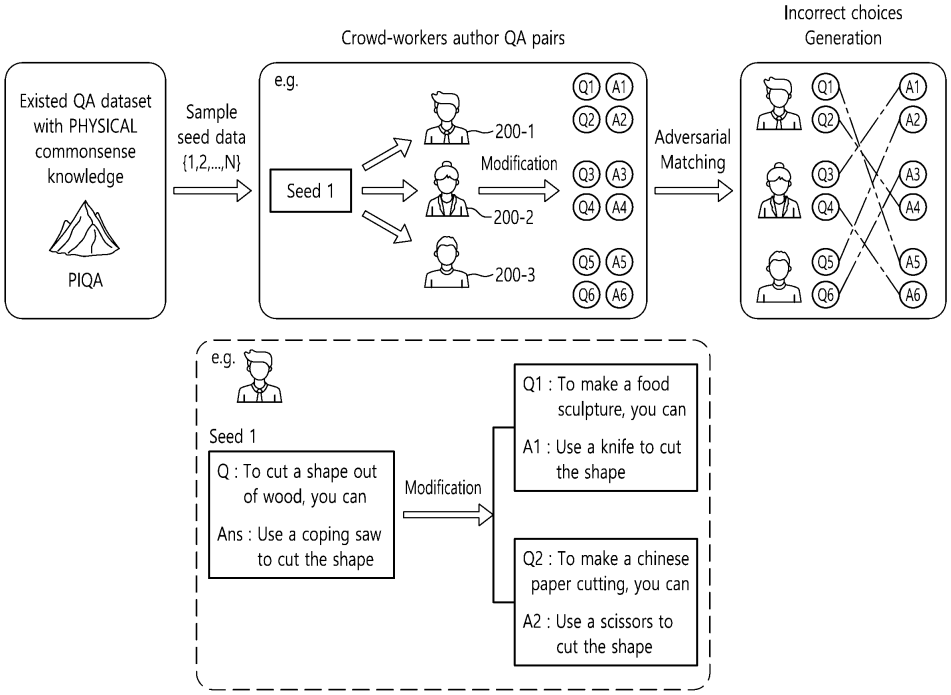
도면1



도면2

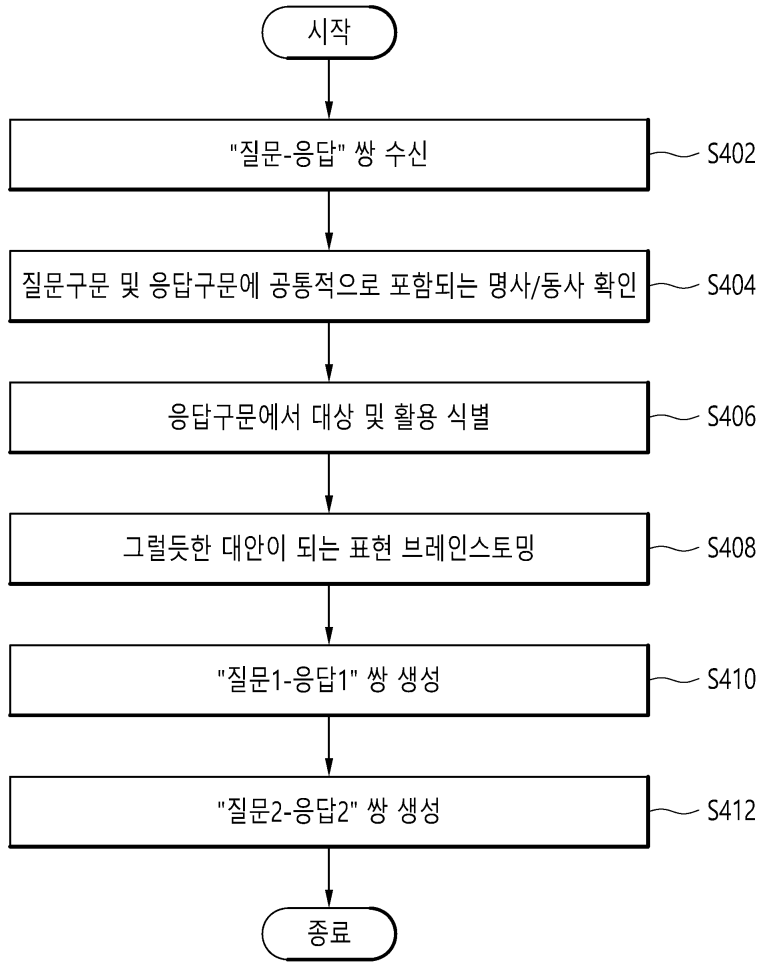


도면3


Crowd-workers author QA pairs
Incorrect choices Generation
Existed QA dataset with PHYSICAL commonsense knowledge
PIQA
Sample seed data {1,2,...,N}
e.g.
Seed 1
200-1
Modification
200-2
200-3
Adversarial Matching
Q1
A1
Q2
A2
Q3
A3
Q4
A4
Q5
A5
Q6
A6

e.g.
Seed 1
Q : To cut a shape out of wood, you can
Ans : Use a coping saw to cut the shape
Modification
Q1 : To make a food sculpture, you can
A1 : Use a knife to cut the shape
Q2 : To make a chinese paper cutting, you can
A2 : Use a scissors to cut the shape

도면4


시작
"질문-응답" 쌍 수신
S402
질문구문 및 응답구문에 공통적으로 포함되는 명사/동사 확인
S404
응답구문에서 대상 및 활용 식별
S406
그럴듯한 대안이 되는 표현 브레인스토밍
S408
"질문1-응답1" 쌍 생성
S410
"질문2-응답2" 쌍 생성
S412
종료