



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0046124  
(43) 공개일자 2019년05월07일

(51) 국제특허분류(Int. Cl.)  
G06F 17/28 (2006.01)

(52) CPC특허분류  
G06F 17/289 (2013.01)

(21) 출원번호 10-2017-0139323

(22) 출원일자 2017년10월25일

심사청구일자 없음

(71) 출원인

한국전자통신연구원

대전광역시 유성구 가정로 218 (가정동)

(72) 발명자

김운

대전광역시 유성구 용산2로 30, 107동 605호(용산동, 경남아너스빌 1단지 아파트)

김영길

대전광역시 서구 둔산북로 160, 2동 1301호(둔산동, 한마루아파트)

(74) 대리인

성병기

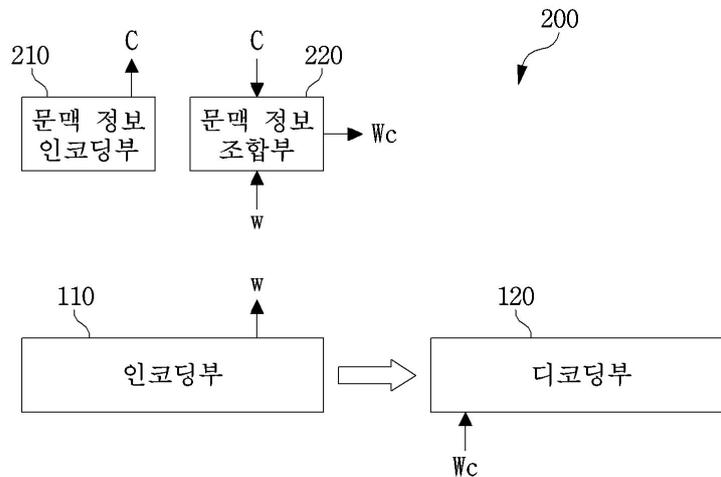
전체 청구항 수 : 총 1 항

(54) 발명의 명칭 문맥 정보에 기반한 실시간 자동통역 방법 및 장치

(57) 요약

본 개시는 문맥 정보에 기반하는 자동 통역을 실시간으로 제공하는 방법 및 장치에 대한 것이다. 본 개시의 일 실시 예에 따른 문맥 정보를 이용한 실시간 자동 통역 방법은, 현재 발화 내용을 인코딩하는 단계; 상기 현재 발화 내용에 관련된 상기 문맥 정보를 인코딩하는 단계; 상기 문맥 정보의 인코딩된 결과에 기초하여, 상기 현재 발화 내용의 인코딩된 결과를 보정하는 단계; 및 보정된 상기 현재 발화 내용의 인코딩된 결과에 기초하여, 상기 현재 발화 내용을 디코딩하는 단계를 포함할 수 있다.

대표도



이 발명을 지원한 국가연구개발사업  
과제고유번호 2016-0-00014  
부처명 미래창조과학부  
연구관리전문기관 정보통신기술진흥센터(IITP)  
연구사업명 산업원천기술개발사업  
연구과제명 지식증강형 실시간 동시통역 원천기술 개발  
기 여 율 1/1  
주관기관 ETRI  
연구기간 2017.01.01 ~ 2017.12.31

---

## 명세서

### 청구범위

#### 청구항 1

문맥 정보를 이용한 실시간 자동 통역 방법에 있어서,

현재 발화 내용을 인코딩하는 단계;

상기 현재 발화 내용에 관련된 상기 문맥 정보를 인코딩하는 단계;

상기 문맥 정보의 인코딩된 결과에 기초하여, 상기 현재 발화 내용의 인코딩된 결과를 보정하는 단계; 및

보정된 상기 현재 발화 내용의 인코딩된 결과에 기초하여, 상기 현재 발화 내용을 디코딩하는 단계를 포함하는, 문맥 정보를 이용한 실시간 자동 통역 방법.

### 발명의 설명

#### 기술 분야

[0001] 본 개시는 자동 통역 기술에 대한 것이며, 구체적으로는 문맥 정보에 기반하는 자동 통역을 실시간으로 제공하는 방법 및 장치에 대한 것이다.

#### 배경 기술

[0002] 실시간 자동 통역 기술은 발화자의 원시 언어(original language)로 된 음성 데이터를 입력 받아 실시간으로 청자의 목적 언어(target language)로 자동 번역하여 출력하는 기술을 의미한다.

[0003] 자동 통역 기술의 종래기술에 있어서, 통계적 학습 번역(Statistical Machine Translation) 기반의 자동 통역 방식은 발화가 종료된 후에야 번역 처리가 가능하므로, 실시간으로 통역을 제공하는 것이 어려웠다. 또한, 심화 학습(deep learning) 기반의 자동 통역 방식은 순차(sequence) 심화 학습 모델을 또는 이중 순차(sequence to sequence) 심화 학습 모델을 이용하여 실시간 통역을 제공할 수 있지만, 학습 데이터가 충분하지 않거나 학습 데이터와 입력 데이터의 차이가 큰 경우에는 성능 저하가 심한 문제가 있었다.

### 발명의 내용

#### 해결하려는 과제

[0004] 본 개시의 기술적 과제는 발화자의 입력 데이터와 문맥 정보를 함께 사용함으로써, 통역 정확도를 개선한 실시간 자동 통역 방법 및 장치를 제공하는 것이다.

[0005] 본 개시의 기술적 과제는 발화자의 주제, 키워드, 이전 발화 내용 등을 문맥 정보로서 이용하고, 컨볼루션 네트워크 방식을 적용함으로써, 통역 정확도를 개선한 실시간 자동 통역 방법 및 장치를 제공하는 것이다.

[0006] 본 개시에서 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급하지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

#### 과제의 해결 수단

[0007] 본 개시의 일 양상에 따른 문맥 정보를 이용한 실시간 자동 통역 방법은, 현재 발화 내용을 인코딩하는 단계; 상기 현재 발화 내용에 관련된 상기 문맥 정보를 인코딩하는 단계; 상기 문맥 정보의 인코딩된 결과에 기초하여, 상기 현재 발화 내용의 인코딩된 결과를 보정하는 단계; 및 보정된 상기 현재 발화 내용의 인코딩된 결과에 기초하여, 상기 현재 발화 내용을 디코딩하는 단계를 포함할 수 있다.

[0008] 본 개시에 대하여 위에서 간략하게 요약된 특징들은 후술하는 본 개시의 상세한 설명의 예시적인 양상일 뿐이며, 본 개시의 범위를 제한하는 것은 아니다.

**발명의 효과**

- [0009] 본 개시에 따르면, 발화자의 입력 데이터와 문맥 정보를 함께 사용함으로써, 통역 정확도를 개선한 실시간 자동 통역 방법 및 장치가 제공될 수 있다.
- [0010] 본 개시에 따르면, 발화자의 주제, 키워드, 이전 발화 내용 등을 문맥 정보로서 이용하고, 컨볼루션 네트워크 방식을 적용함으로써, 통역 정확도를 개선한 실시간 자동 통역 방법 및 장치가 제공될 수 있다.
- [0011] 본 개시에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

**도면의 간단한 설명**

- [0012] 도 1은 본 개시가 적용될 수 있는 심화 학습 기반의 자동 번역 방식을 설명하기 위한 도면이다.  
 도 2는 본 개시에 따른 문맥 정보를 이용하는 심화 학습 기반의 자동 번역 방식을 설명하기 위한 도면이다.  
 도 3은 본 개시에 따른 문맥 정보를 이용하는 심화 학습 기반의 자동 번역 방식의 일례를 설명하기 위한 도면이다.  
 도 4는 본 개시에 따른 다중 타입의 문맥 정보를 이용하는 예시에 있어서의 문맥 정보 인코딩부 및 문맥 정보 조합부를 설명하기 위한 도면이다.  
 도 5는 본 개시에 따른 문맥 정보 조합부의 예시를 설명하기 위한 도면이다.  
 도 6은 본 개시에 따른 문맥 정보를 이용하는 실시간 자동 번역 방법을 설명하기 위한 흐름도이다.

**발명을 실시하기 위한 구체적인 내용**

- [0013] 이하에서는 첨부한 도면을 참고로 하여 본 개시의 실시 예에 대하여 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나, 본 개시는 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시 예에 한정되지 않는다.
- [0014] 본 개시의 실시 예를 설명함에 있어서 공지 구성 또는 기능에 대한 구체적인 설명이 본 개시의 요지를 흐릴 수 있다고 판단되는 경우에는 그에 대한 상세한 설명은 생략한다. 그리고, 도면에서 본 개시에 대한 설명과 관계없는 부분은 생략하였으며, 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0015] 본 개시에 있어서, 어떤 구성요소가 다른 구성요소와 "연결", "결합" 또는 "접속"되어 있다고 할 때, 이는 직접적인 연결관계뿐만 아니라, 그 중간에 또 다른 구성요소가 존재하는 간접적인 연결관계도 포함할 수 있다. 또한 어떤 구성요소가 다른 구성요소를 "포함한다" 또는 "가진다"고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 배제하는 것이 아니라 또 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0016] 본 개시에 있어서, 제1, 제2 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용되며, 특별히 언급되지 않는 한 구성요소들간의 순서 또는 중요도 등을 한정하지 않는다. 따라서, 본 개시의 범위 내에서 일 실시 예에서의 제1 구성요소는 다른 실시 예에서 제2 구성요소라고 칭할 수도 있고, 마찬가지로 일 실시 예에서의 제2 구성요소를 다른 실시 예에서 제1 구성요소라고 칭할 수도 있다.
- [0017] 본 개시에 있어서, 서로 구별되는 구성요소들은 각각의 특징을 명확하게 설명하기 위함이며, 구성요소들이 반드시 분리되는 것을 의미하지는 않는다. 즉, 복수의 구성요소가 통합되어 하나의 하드웨어 또는 소프트웨어 단위로 이루어질 수도 있고, 하나의 구성요소가 분산되어 복수의 하드웨어 또는 소프트웨어 단위로 이루어질 수도 있다. 따라서, 별도로 언급하지 않더라도 이와 같이 통합된 또는 분산된 실시 예도 본 개시의 범위에 포함된다.
- [0018] 본 개시에 있어서, 다양한 실시 예에서 설명하는 구성요소들이 반드시 필수적인 구성요소들은 의미하는 것은 아니며, 일부는 선택적인 구성요소일 수 있다. 따라서, 일 실시 예에서 설명하는 구성요소들의 부분집합으로 구성되는 실시 예도 본 개시의 범위에 포함된다. 또한, 다양한 실시 예에서 설명하는 구성요소들에 추가적으로 다른 구성요소를 포함하는 실시 예도 본 개시의 범위에 포함된다.
- [0019] 이하에서는, 본 개시에 따른 다양한 예시들에 대해서 설명한다.
- [0020] 본 개시는 실시간 자동 통역에 있어서 문맥 정보를 이용하는 방법 및 장치에 대한 다양한 예시들을 포함한다.

보다 구체적으로는 실시간으로 입력되는 음성 정보를 인식하고, 인식된 음성 정보에 대해서 인공지능 기반으로 실시간 자동 통역을 수행함에 있어서, 인식된 음성 정보에 연관된 또는 주변의 문맥 정보를 이용하는 방안에 대해서 설명한다.

- [0021] 본 개시에 있어서 문맥 정보를 이용한 실시간 자동 통역은, 이미 발화한 문장이나 기존에 발화한 유사 문장을 이용한 심화 학습 기반의 실시간 자동 통역 방식을 포함한다.
- [0022] 보다 구체적으로, 문맥 정보는 발화자의 주제, 발화자의 이전 발화, 발화자의 이전 발화 번역문, 발화자의 핵심 키워드 등을 포함할 수 있다. 이러한 다양한 문맥 정보는 인코딩되어 현재 발화 내용의 번역에 이용될 수 있다. 또한, 이러한 다양한 문맥 정보를 이용한 실시간 자동 통역에 컨볼루션(convolution) 네트워크 학습 기법을 적용함으로써, 실시간 자동 통역의 성능을 극대화할 수 있다.
- [0023] 실시간 자동 통역 기술은 발화자의 원시 언어(original language)로 된 음성 데이터를 입력 받아 실시간으로 청자의 목적 언어(target language)로 자동 번역하여 출력하는 기술을 의미한다. 즉, 실시간 자동 통역 시스템은 실시간으로 원시 언어의 음성 데이터가 입력되는 상황(예를 들어, 서로 다른 언어를 사용하는 사람들의 대화하는 경우, 다른 언어를 사용하는 사람의 강의를 듣는 경우)에서, 청자의 언어 이해를 돕기 위해서 목적 언어로 된 음성 데이터를 출력하는 시스템 또는 장치이다.
- [0024] 실시간으로 입력되는 음성 데이터를 처리하는 종래의 자동 통번역 시스템은, 한 언어로 된 음성 발화를 입력으로 받아 음성 인식을 거친 후 음성 인식된 결과를 텍스트로 변환하여 자동 통번역 시스템의 입력으로 사용한다. 이러한 자동 통번역 시스템은, 입력된 텍스트 정보를 규칙 기반이나 통계적 학습 기반(Statistical Machine Translation, SMT) 방식을 사용하여 한 언어로부터 다른 언어로(예를 들어, 한국어로부터 영어로) 자동 번역하고, 번역된 결과를 다시 음성 합성하여 사용자에게 전달할 수 있다. 이러한 자동 통번역 시스템은 그 과정이 복잡하고, 한 문장이나 사람의 발화가 끝날 때까지 번역 입력이 들어오지 않아 기다려야 하는 단점이 있어 실시간성이 부족하다. 또한, 번역 방법에 있어서 규칙 기반이나 통계적 방법인 SMT는 그 성능이 높지 않아 실시간 자동 통역에 적합하지 않다.
- [0025] 또한, 종래의 실시간 자동 통역 시스템으로서 심화 학습(Deep Learning) 기반의 통역 방식은, 음성 인식이나 자동 번역에 모두 사용 가능하며 그 성능이 우수하여 산업 전반에 큰 영향을 주고 있다. 심화 학습은 이론적으로 충분한 학습 데이터가 제공되는 경우 원하는 성능을 얻을 수 있다. 그러나, 학습 데이터가 적절하지 않거나 불충분한 경우에 원하는 성능을 얻기 어렵다. 예를 들어, 여성의 음성 데이터만으로 학습된 경우 남성 음성을 인식의 성능이 떨어지고, 학습 데이터 포함되지 않은 음성 데이터는 통번역이 불가능하거나 그 성능이 현저하게 떨어지게 된다.
- [0026] 이러한 심화 학습 기반의 자동 통역 방식의 문제점을 해결하기 위해서, 본 개시의 다양한 예시들에 따르면 문맥 정보를 이용하여 실시간 자동 통역의 성능을 높일 수 있다.
- [0027] 먼저 심화 학습 기반의 동작 원리에 대해서 설명하고, 본 개시의 특징에 대해서 후술한다.
- [0028] 도 1은 본 개시가 적용될 수 있는 심화 학습 기반의 자동 번역 방식을 설명하기 위한 도면이다.
- [0029] 심화 학습 기반의 자동 번역 방식은, 학습 데이터를 이용하여 학습 모델을 생성한 후, 생성된 학습 모델을 이용하여 한 언어로 된 입력 문장 원시 언어 문장을 타 언어인 목적언어로 변환하는 것으로 요약할 수 있다. 이때, 학습 단계에서는 크게 인코딩(encoding) 및 디코딩(decoding)을 거칠 수 있다.
- [0030] 심화 학습 기반의 자동 번역 시스템의 일례로서, 도 1에서는 이중 순차(sequence to sequence) 심화 학습 기반 자동 번역 구조(100)를 나타낸다. 예를 들어, 심화 학습 기반 자동 번역 구조(100)는 인코딩부(110)에 해당하는 인코딩 노드들(111, 112, 113, 114), 및 디코딩부(120)에 해당하는 디코딩 노드들(121, 122, 123, 124, 125)을 포함할 수 있다. 이러한 복수의 노드(111, 112, 113, 114, 121, 122, 123, 124, 125)은, 예를 들어, RNN(Recurrent Neural Network) 또는 RNN의 특수 형태인 LSTM(Long Short Term Memory) 네트워크를 구성하는 노드들에 해당할 수 있다.
- [0031] 도 1의 예시에서는 원시 언어가 한국어이고 목적 언어가 영어인 경우를 나타낸다. 예를 들어, 한국어로 된 "나는 학교에 간다." 라는 음성 데이터를 영어로 된 "I go to school."로 변환하는 예시를 가정한다. 도 1의 예시에서  $w$ 는 입력 문장을 인코딩하여 얻은 결과 벡터값을 나타낸다. 구체적으로, 도 1과 같은 심화 학습 기반의 자동 번역 시스템을 이용하면, "나는 학교에 간다." 라는 음성 데이터를 인코딩하여  $w$  라는 벡터 형태의 출력 값을 생성할 수 있다. 그리고, 디코딩 단계에서는 인코딩 결과 값인  $w$  벡터와 이전 노드의 값을 이용하여 다음 노

드의 값을 생성할 수 있다.

- [0032] 보다 구체적으로, 인코딩부(110)에서는 학습 데이터의 원시 언어 문장을 인코딩할 수 있다. 원시 언어 문장은 단어 또는 토큰이라는 요소들로 이루어져 있는데, 이런 요소들은 순차적으로 심화 학습 구조에 입력될 수 있다. 도 1의 예시에서 입력 문장은 4 개의 단어 또는 토큰, 즉, "나는" "학교에" "간다" "."로 구성될 수 있다. 이러한 입력 문장의 요소들이 심화 학습 과정의 인코딩부(110)에서 벡터 형태로 변환되어 표현되며, 각 벡터 형태로 표현된 단어는 순차적인 심화 학습 모델인 RNN(Recurrent Neural Network) 또는 RNN 방법의 특수 형태인 LSTM(Long Short Term Memory) 네트워크에 의해 인코딩될 수 있다. 즉, 입력된 원시 언어 문장의 "나는"이라는 요소는 노드(111)에서, "학교에"라는 요소는 노드(112)에서, "간다"라는 요소는 노드(113)에서, "."라는 요소는 노드(114)에서 인코딩되어, 최종적으로  $w$  라는 하나의 벡터 형태의 출력값을 생성할 수 있다. 이러한 출력 값을 디코딩부(120)에서 목적 언어 문장의 각각의 요소들을 생성하는데 관여할 수 있다.
- [0033] 디코딩부(120)도 인코딩부(110)와 같이 순차적 학습 모델인 RNN 또는 LSTM 기반의 네트워크에 의해서 표현될 수 있다. 즉, 디코딩부(120)는 인코딩 출력 값( $w$ )에 근거하여 목적 언어의 단어 또는 토큰을 하나씩 생성할 수 있다. 디코딩부(120)의 첫 번째 노드(121)에서, 인코딩 결과인 출력값( $w$ )에 기초하여 목적 언어의 첫 번째 요소(단어 또는 토큰)가 생성될 수 있다. 도 1의 예시에서 디코딩부(120)의 첫 번째 노드(121)에서 생성된 단어는 "I"인 경우를 나타낸다. 그 다음으로, 이미 생성된 단어 또는 토큰이 인코딩된 출력값( $w$ )과 함께 그 다음 단어 또는 토큰 생성에 순차적으로 및 직접적으로 관여할 수 있다. 예를 들어, 디코딩부(120)의 두 번째 노드(122)에서는 이전 노드(121)의 결과값인 "I" 및 인코딩 출력값( $w$ )에 기초하여 "go"라는 요소를 생성할 수 있다. 디코딩부(120)의 세 번째 노드(123)에서는 이전 노드(122)의 결과값인 "go" 및 인코딩 출력값( $w$ )에 기초하여 "to"라는 요소를 생성할 수 있다. 디코딩부(120)의 네 번째 노드(124)에서는 이전 노드(123)의 결과값인 "to" 및 인코딩 출력값( $w$ )에 기초하여 "school"이라는 요소를 생성할 수 있다. 디코딩부(120)의 다섯 번째 노드(125)에서는 이전 노드(124)의 결과값인 "school" 및 인코딩 출력값( $w$ )에 기초하여 "<EOF>" (즉, End Of Frame)이라는 요소를 생성할 수 있다.
- [0034] 최종적으로, "나는 학교에 간다."라는 입력 문장에 대한 인코딩 벡터 값  $w$ 에 기초하여, 디코딩부(120)에서는 "I go to school."이라는 번역 결과를 출력할 수 있다.
- [0035] 이와 같은 심화 학습 기반의 실시간 자동 번역에 있어서, 학습 모델을 이용한 학습 과정에 있어서 원시 언어 입력 문장 및 목적 언어 정답 문장(즉, 원시 언어 문장의 정확한 번역 결과)이 모두 주어지며, 번역 결과가 정답 문장과 가장 유사하게 생성될 수 있도록 심화 학습 네트워크의 각 노드를 조절하면서 반복적인 학습 과정을 거치게 된다. 심화 학습 네트워크의 노드를 조절한다는 것은, 인코딩과 디코딩의 각각의 RNN 또는 LSTM 노드의 파라미터와 바이어스(bias) 값을 조절하는 것을 포함할 수 있다.
- [0036] 도 1을 참조하여 설명한 바와 같이, 원시 언어 입력 문장은 인코딩 과정을 거치면서 디코딩 과정에 많은 영향을 주며, 번역 성능과 직결된다.
- [0037] 본 개시에서는 발화에 의해 실시간으로 생성되는 입력 문장에 추가적으로 문맥 정보를 인코딩 단계에서 이용함으로써, 디코딩 단계에서 보다 적절한 목적 언어 단어 또는 토큰이 생성되도록 하는 방안에 대해서 설명한다.
- [0038] 도 2는 본 개시에 따른 문맥 정보를 이용하는 심화 학습 기반의 자동 번역 방식을 설명하기 위한 도면이다.
- [0039] 도 2의 문맥 정보를 이용하는 심화 학습 기반 자동 번역 구조(200)는, 도 1의 인코딩부(110) 및 디코딩부(120)에 대응하는 구성요소들을 포함하고, 이에 추가적으로 문맥 정보 인코딩부(210) 및 문맥 정보 조합부(220)를 포함할 수 있다.
- [0040] 문맥 정보 인코딩부(210)는 다양한 타입의 문맥 정보를 인코딩하여 인코딩 결과 값을 출력할 수 있다. 문맥 정보의 타입은, 발화자의 현재 발화 내용의 주제, 현재 발화 내용에 관련된 이전 문장, 이전 발화 내용의 번역 결과, 현재 발화 내용의 키워드 등을 포함할 수 있다. 문맥 정보 인코딩부(210)의 출력에 해당하는 인코딩된 문맥 정보는 "C"로 표현할 수 있다. 또는 문맥 정보 인코딩부(210)의 출력은 문맥 정보의 타입에 따라서 구분될 수도 있다. 예를 들어, 발화자의 현재 발화 내용의 주제에 해당하는 제 1 타입의 문맥 정보의 인코딩 결과는 C1으로 표현되고, 현재 발화 내용에 관련된 이전 문장에 해당하는 제 2 타입의 문맥 정보의 인코딩 결과는 C2로 표현되고, 이전 발화 내용의 번역 결과에 해당하는 제 3 타입의 문맥 정보의 인코딩 결과는 C3로 표현되고, 현재 발화 내용의 키워드에 해당하는 제 4 타입의 문맥 정보의 인코딩 결과는 C4로 표현될 수 있다. 본 개시의 범위는 문맥 정보 타입의 위와 같은 예시들로 제한되는 것은 아니며, 통역 또는 번역에 고려되는 다른 문맥 정보도 문맥 정보 인코딩부(210)에 의해서 인코딩될 수 있다.

- [0041] 문맥 정보 조합부(220)는 문맥 정보 인코딩부(210)의 출력(즉, 인코딩된 문맥 정보 C)와, 현재 발화 내용(또는 현재 입력 문장)의 인코딩 결과 값(즉, w 벡터)를 입력으로 받아서, 인코딩된 문맥 정보에 기초하여 보정된 인코딩 결과 값인 "Wc"를 출력할 수 있다. 예를 들어, 문맥 정보 조합부(220)는 문맥 정보 인코딩 값(C)에 포함된 다양한 타입의 문맥 정보 인코딩 결과(예를 들어, C1, C2, C3, C4)를 합산하거나, 제곱을 취하여 합산하거나, 또는 소정의 임계값 이상인 값만 조합하는 등의 처리를 할 수 있다.
- [0042] 도 3은 본 개시에 따른 문맥 정보를 이용하는 심화 학습 기반의 자동 번역 방식의 일례를 설명하기 위한 도면이다.
- [0043] 도 3의 문맥 정보를 이용하는 심화 학습 기반 자동 번역 구조(300)는, 도 1의 인코딩부(110) 및 디코딩부(120)에 대응하는 구성요소들을 포함하고, 이에 추가적으로 문맥 정보 인코딩부(310) 및 문맥 정보 조합부(320)를 포함할 수 있다.
- [0044] 도 3의 예시에서 문맥 정보 인코딩부(310) 및 문맥 정보 조합부(320)는, 도 2의 예시의 문맥 정보 인코딩부(210) 및 문맥 정보 조합부(220)의 일 실시예에 해당할 수 있다. 즉, 도 3의 예시에서는 설명의 명료성을 위해서, 문맥 정보 인코딩부(210)가 현재 발화 내용에 관련된 이전 문장에 해당하는 문맥 정보를 인코딩하는 것을 대표적으로 나타낸다. 여기서, 이전 문장은 현재 발화 내용의 직전 발화 내용으로 제한되는 것은 아니며, 현재 발화 내용 이전의 다수의 문장 중에서 현재 발화 내용과 관련된 하나 이상의 문장에 해당할 수 있다.
- [0045] 인코딩부(110)는 발화자의 원시 언어의 현재 발화 내용에 해당하는 "나는 학교에 간다."라는 입력 문장을 인코딩하여 출력값 w를 생성할 수 있다.
- [0046] 문맥 정보 인코딩부(310)는 현재 발화 내용에 관련된 이전 문장에 해당하는 "비록 학교가 멀지만"이라는 문맥 정보를 인코딩하여 출력값 C를 생성할 수 있다. 예를 들어, 이전 문장의 3 개의 요소에 해당하는 "비록" "학교가" "멀지만"이라는 단어가 문맥 정보 인코딩부(310)에서 벡터 형태로 변환되어 표현되며, 각 벡터 형태로 표현된 단어는 순차적인 심화 학습 모델인 RNN 또는 LSTM 네트워크의 각각의 노드들에 의해 인코딩될 수 있다. 즉, 이전 문장의 "비록"이라는 요소는 문맥 정보 인코딩 노드(311)에서, "학교가"라는 요소는 문맥 정보 인코딩 노드(312)에서, "멀지만"이라는 요소는 문맥 정보 인코딩 노드(313)에서 인코딩되어, 최종적으로 C 라는 하나의 벡터 형태의 출력값을 생성할 수 있다. 예를 들어, C 라는 문맥 정보 인코딩 결과는, "비록 학교가 멀지만"이라는 이전 문장이 "나는 학교에 간다."라는 입력 문장과 관계에 있어서 양보의 의미(즉, 그런데도, 또는 그럼에도 불구하고라는 의미)를 부여하는 것을 나타내는 벡터 값으로 결정될 수 있다.
- [0047] 문맥 정보 조합부(320)는, 인코딩부(110)의 출력값 w와 문맥 정보 인코딩부(310)의 출력값 C를 조합하여, 문맥 정보를 반영한 보정된 인코딩 결과 값인 Wc를 출력할 수 있다.
- [0048] 예를 들어, 도 1의 예시와 같이 "나는 학교에 간다."라는 입력 문장에 대한 문맥 정보 없이(즉, 인코딩 결과 w 만에 기초하여) 번역하는 경우 "I go to school."의 결과가 출력되지만, "비록 학교가 멀지만"이라는 문맥 정보를 고려하면 "나는 학교에 간다."라는 입력 문장에서 화자의 의지를 표현하는 것이 보다 적절할 수 있다. 즉, 화자의 의도는 "(그런데도) 나는 학교에 간다."는 것을 문맥 정보를 통하여 파악할 수 있다. 이와 같이 문맥 정보가 양보의 의미를 포함하는 경우에, 입력 문장의 번역에 있어서 "그런데도"와 같은 화자의 의지를 표현하는 요소(예를 들어, 영어에서 "still"이라는 단어)가 추가된다면 보다 적절한 번역이 될 수 있다. 이를 위해서, 문맥 정보 조합부(320)에서는 문맥 정보 인코딩부(310)의 결과 값인 C에 기초하여, "나는 학교에 간다"라는 입력 문장에 대한 인코딩된 벡터 값 w에, 양보의 의미를 부가하여 보정된 인코딩 결과 값인 Wc를 생성할 수 있다.
- [0049] 이에 따라, 디코딩부(120)에서는 보정된 인코딩 결과 값(Wc)에 근거하여 목적 언어의 단어 또는 토큰을 하나씩 생성할 수 있다. 예를 들어, 디코딩부(120)의 첫 번째 노드(121)에서, 보정된 인코딩 결과 값인 Wc에 기초하여 목적 언어의 첫 번째 요소(단어 또는 토큰)가 생성될 수 있다. 도 3의 예시에서 디코딩부(120)의 첫 번째 노드(121)에서 생성된 단어는 "I"인 경우를 나타낸다. 그 다음으로, 이미 생성된 단어 또는 토큰이 보정된 인코딩 결과 값(Wc)과 함께 그 다음 단어 또는 토큰 생성에 순차적으로 및 직접적으로 관여할 수 있다. 예를 들어, 디코딩부(120)의 두 번째 노드(122)에서는 이전 노드(121)의 결과값인 "I" 및 보정된 인코딩 결과 값(Wc)에 기초하여 "still"이라는 요소를 생성할 수 있다. 디코딩부(120)의 세 번째 노드(123)에서는 이전 노드(122)의 결과값인 "still" 및 보정된 인코딩 결과 값(Wc)에 기초하여 "go"라는 요소를 생성할 수 있다. 디코딩부(120)의 네 번째 노드(124)에서는 이전 노드(123)의 결과값인 "to" 및 보정된 인코딩 결과 값(Wc)에 기초하여 "to"라는 요소를 생성할 수 있다. 디코딩부(120)의 다섯 번째 노드(125)에서는 이전 노드(124)의 결과값인 "to" 및 보정된 인코딩 결과 값(Wc)에 기초하여 "school"이라는 요소를 생성할 수 있다. 디코딩부(120)의 여섯 번째 노드(126)

에서는 이전 노드(125)의 결과값인 "school" 및 보정된 인코딩 결과 값(Wc)에 기초하여 "<EOF>" (즉, End Of Frame)이라는 요소를 생성할 수 있다.

- [0050] 최종적으로, "나는 학교에 간다."라는 입력 문장에 대한 인코딩 벡터 값 w에 대해서 "비록 학교가 멀지만"이라는 문맥 정보에 대한 인코딩 벡터 값 C에 기초하여 보정된 인코딩 결과 값 Wc에 기초하여, 디코딩부(120)에서는 "I still go to school."이라는 번역 결과를 출력할 수 있다.
- [0051] 도 4는 본 개시에 따른 다중 타입의 문맥 정보를 이용하는 예시에 있어서의 문맥 정보 인코딩부 및 문맥 정보 조합부를 설명하기 위한 도면이다.
- [0052] 도 4의 예시에서 문맥 정보 인코딩부(410) 및 문맥 정보 조합부(420)는, 도 2의 예시의 문맥 정보 인코딩부(210) 및 문맥 정보 조합부(220)의 일 실시예에 해당할 수 있다.
- [0053] 문맥 정보 인코딩부(410)는 복수의 타입의 문맥 정보에 대해서 타입 별로 문맥 정보 인코딩 결과 값을 생성할 수 있다. 예를 들어, 문맥 정보 인코딩부(410)는, 제 1 타입 문맥 정보를 인코딩하는 제 1 타입 문맥 정보 인코딩부(411), 제 2 타입 문맥 정보를 인코딩하는 제 2 타입 문맥 정보 인코딩부(412), 제 3 타입 문맥 정보를 인코딩하는 제 3 타입 문맥 정보 인코딩부(413), 제 4 타입 문맥 정보를 인코딩하는 제 4 타입 문맥 정보 인코딩부(414)를 포함할 수 있다.
- [0054] 본 개시의 범위는 문맥 정보 타입의 예시 및 개수에 의해서 제한되지 않으며, 도 4의 예시 중의 일부 타입의 문맥 정보 인코딩부를 포함할 수도 있고, 다른 타입의 문맥 정보 인코딩부를 더 포함할 수도 있다. 또한, 도 4의 예시에서는 문맥 정보의 타입 별로 하나의 타입 문맥 정보 인코딩부를 포함하는 것을 나타내지만, 본 개시의 범위는 이에 제한되지 않으며 동일한 타입의 문맥 정보에 대해서, 복수의 타입 문맥 정보 인코딩부가 포함될 수도 있다.
- [0055] 예를 들어, 제 1 타입 문맥 정보 인코딩부(411)는 발화자의 현재 발화 내용의 주제에 해당하는 제 1 타입의 문맥 정보를 인코딩한 결과 값인 C1을 출력할 수 있다. 보다 구체적으로, 발화자의 현재 발화 내용의 주제는, 발화의 내용을 기 분류한 주제 범주 중에서 어느 하나의 범주로 미리 발화자에 의해 선택될 수도 있다. 또한, 발화자의 현재 발화 내용의 주제는, 발화의 내용을 기 분류한 주제와의 유사도를 계산하여 분류할 수도 있다. 예를 들어, 발화의 내용을 기 분류한 주제 또는 주제 범주에 대해서 하나 이상의 토픽 식별자(topic\_id) 또는 주제 코드 값이 부여될 수 있으며, 선택된 또는 계산에 의해 분류된 주제 또는 주제 범주는 특정 topic\_id 값으로 표현될 수 있다. 현재 발화 내용의 주제인 topic\_id 값이 노드(411\_1)에 입력 및 처리되어 최종적으로 인코딩된 문맥 정보 C1가 생성될 수 있다.
- [0056] 제 2 타입 문맥 정보 인코딩부(412)는 현재 발화 내용에 관련된 이전 문장에 해당하는 제 2 타입의 문맥 정보를 인코딩한 결과 값인 C2를 출력할 수 있다. 보다 구체적으로, 발화자의 현재 발화 내용이 "나는 학교에 간다."인 경우, 이와 관련된 이전 발화 내용으로 "비록 학교가 멀지만"을 문맥 정보로서 인코딩하여 이용할 수 있다. 일반적으로, 현재 발화 내용은 발화자의 이전 발화의 연장선 상에 있으며, 현재 발화의 모호한 내용이나 단어가 이전 발화에는 정확하게 표현될 수 있으며, 문장 구성 상으로 보더라도 현재 발화에 없는 전치사구 또는 주격과 같은 중요한 정보가 들어 있다. 예를 들면, "오늘 제가 여러분을 여기에 모이자고 한 것은"와 같은 내용은 길지만 문장의 주격 요소로서 가능하며, 이 내용이 없으면 이후의 내용을 이해하기 어렵다. 또한, 서로 다른 언어의 어순이 달라질 수 있으므로(예를 들면, 한국어에서 영어로 번역하는 경우 일부 문장의 경우 어순이 도치되므로) 이전 발화의 내용이 현재 발화 내용을 번역하는데 많은 도움이 된다. 이전 발화 내용의 단어들이 "비록" "학교가" "멀지만"은 요소별로 노드(412\_1, 412\_2, 412\_3)에 입력 및 처리되어 최종적으로 인코딩된 문맥 정보 C2가 생성될 수 있다.
- [0057] 제 3 타입 문맥 정보 인코딩부(413)는 이전 발화 내용의 번역 결과에 해당하는 제 3 타입의 문맥 정보를 인코딩한 결과 값인 C3을 출력할 수 있다. 보다 구체적으로, 현재 발화 내용이 "나는 학교에 간다."인 경우, 이와 관련된 이전 발화 내용으로 "비록 학교가 멀지만"의 번역 결과인 "even school far"를 문맥 정보로서 인코딩하여 이용할 수 있다. 이전 발화 내용의 번역 결과에는 현재 발화 내용의 번역에 유용한 정보가 포함될 수 있으며, 동일한 요소에 대해서 동일한 번역 결과를 참조할 수 있는 등 번역의 정확성을 높이기 위한 유용한 문맥정보로서 사용될 수 있다. 이전 발화 내용의 번역문의 단어들이 "even" "school" "far"는 요소별로 노드(413\_1, 413\_2, 413\_3)에 입력 및 처리되어 최종적으로 인코딩된 문맥 정보 C3가 생성될 수 있다.
- [0058] 제 4 타입 문맥 정보 인코딩부(414)는 현재 발화 내용의 키워드에 해당하는 제 4 타입의 문맥 정보를 인코딩한 결과 값인 C4를 출력할 수 있다. 보다 구체적으로, 발화자의 핵심 키워드를 인코딩하여 활용할 수 있다. 대부분

의 발화는 특정 주제와 관련되기 때문에, 해당 주제에 적합한 키워드들이 문맥 정보로서 사용될 수 있다. 이러한 키워드들은 발화의 시작부터 끝까지 발화자의 내용에 자주 등장하게 된다. 따라서, 발화자의 내용 중 단어들의 빈도 및 중요도를 실시간 계산하여 키워드를 추가하여 문맥정보로 활용하면 현재 발화자의 내용을 번역하는데 많은 도움을 줄 수 있다. 또한, 이런 키워드는 대응되는 대역어를 함께 문맥 정보로 활용함으로써 번역 과정에 대역어 선택 오류를 줄임으로써 번역 성능에 기여하는 효과를 가져올 수 있다. 현재 발화 내용의 키워드에 해당하는 "kw1" "kw2" "kw2"는 각각 노드(414\_1, 414\_2, 414\_3)에 입력 및 처리되어 최종적으로 인코딩된 문맥 정보 C4가 생성될 수 있다.

- [0059] 이와 같이 다양한 타입의 문맥 정보의 하나 이상을 이용하여 대화의 문맥을 제대로 이해하고 번역하는데 모호성을 줄임으로써 번역 성능 향상 효과를 가져올 수 있다.
- [0060] 도 5는 본 개시에 따른 문맥 정보 조합부의 예시를 설명하기 위한 도면이다.
- [0061] 도 5의 예시에서 문맥 정보 조합부(520)는, 도 2의 예시의 문맥 정보 조합부(220)의 일 실시예에 해당할 수 있다.
- [0062] 예를 들어, 문맥 정보 조합부(520)는 문맥 정보 인코딩 결과 값 입력부(521), 컨볼루션 학습부(522), 현재 발화 내용 인코딩 결과 값 입력부(523), 및 보정부(524)를 포함할 수 있다.
- [0063] 문맥 정보 인코딩 결과 값 입력부(521)는, 문맥 정보 인코딩부(210, 310, 410)의 출력에 해당하는 인코딩된 문맥 정보를 입력 받을 수 있다. 인코딩된 문맥 정보는 하나 이상의 타입에 대한 문맥 정보를 포함할 수 있다. 도 5의 예시에서는 4 가지 타입에 대한 문맥 정보 인코딩 결과 값인 C1, C2, C3, C4가 입력되는 경우를 가정한다. 예를 들어, C1, C2, C3, C4는 각각 제 1 타입 문맥 정보 인코딩 결과 값 입력 노드(521\_1), 제 2 타입 문맥 정보 인코딩 결과 값 입력 노드(521\_2), 제 3 타입 문맥 정보 인코딩 결과 값 입력 노드(521\_3), 제 4 타입 문맥 정보 인코딩 결과 값 입력 노드(521\_4)를 통하여 문맥 정보 조합부(520)에 제공될 수 있다.
- [0064] 이와 같이 하나 이상의 문맥 정보가 제공되는 경우, 어떤 문맥 정보가 현재 발화 내용의 번역에 어떤 영향을 미치는지를 직접적으로 결정하지 않고, 인코딩된 문맥 정보의 유용성을 선별하여 번역 성능 향상의 효과를 극대화할 수 있다. 이를 위해서, 본 개시에 따른 문맥 정보 조합부(520)는 컨볼루션(convolution) 네트워크를 이용한 학습을 통해서 최종적인 문맥 정보 값을 얻을 수 있다. 이러한 최종적인 결과 값은 보정부(524)로 전달될 수 있다.
- [0065] 구체적으로 컨볼루션 학습부(522)는 인코딩된 문맥 정보 입력 값(C1, C2, C3, C4)을 다양한 루트를 통하여 확장 및 조합할 수 있다. 예를 들어, 5개의 서로 다른 조합 노드(522\_1, 522\_2, 522\_3, 522\_4, 522\_5)는, C1, C2, C3, C4 중의 하나 이상을 서로 다른 기준에 따라서 조합할 수 있다. 예를 들어, 어떤 조합 노드에 입력되는 문맥 정보 타입들의 조합은 다른 조합 노드에 입력되는 문맥 정보 타입들의 조합과 동일할 수도 있고 다를 수도 있다. 또한, 어떤 조합 노드에서 적용되는 조합 기준은 다른 조합 노드의 조합 기준과 동일할 수도 있고 다를 수도 있다.
- [0066] 예를 들어, 제 1 조합 노드(522\_1)에서는 C1, C2, C3를 조합할 수 있고, 제 2 조합 노드(522\_2)에서는 C1, C2, C3, C4를 조합할 수 있고, 제 3 조합 노드(522\_3)에서는 C1, C2, C3, C4를 조합할 수 있고, 제 4 조합 노드(522\_4)에서는 C2, C3, C4를 조합할 수 있고, 제 5 조합 노드(522\_5)에서는 C1, C3, C4를 조합할 수 있다. 이 경우, 제 2 조합 노드(522\_2)와 제 3 조합 노드(522\_3)에 입력되는 인코딩된 문맥 정보의 타입의 조합은 동일하지만, 각각의 조합 노드(522\_2, 522\_3)에서 적용되는 조합 방식(예를 들어, 단순 합산, 계급의 합산, 소정의 임계값 이상인 값만 합산)이 상이할 수도 있고, 조합에 적용되는 파라미터나 바이어스 값이 상이할 수도 있다.
- [0067] 본 개시의 범위는 조합 노드의 개수, 조합 노드에 입력되는 문맥 정보의 타입들의 조합 등에 의해서 제한되는 것은 아니며, 다양한 방식으로 다양한 문맥 정보를 조합하는 예시들을 포함할 수 있다. 또한, 어떤 문맥 정보는 다른 문맥 정보와 조합되지 않고 보정부(524)로 전달될 수도 있다.
- [0068] 또한, 컨볼루션 학습부(522)에 포함되는 조합 노드들(522\_1, 522\_2, 522\_3, 522\_4, 522\_5)은, 문맥 정보 인코딩 결과 값 입력부(521)에 입력된 벡터 값들로부터 유용한 결과 값을 추출하는 은닉층(hidden layer)으로서 구성되어, 은닉층에 포함되는 여러 형태의 노드로 확장될 수 있다. 이와 같이 확장된 은닉층은 최종 노드 값을 보정부(524)로 출력할 수도 있다.
- [0069] 보정부(524)에서는 컨볼루션 학습부(522)로부터 출력되는 결과 값을 기초로, 현재 발화 내용 인코딩 결과 값 입력부(523)로부터 전달되는 현재 발화 내용 인코딩 결과 값(w)에 대한 보정을 수행할 수 있다. 이에 따라, 현재

발화 내용 인코딩 결과 값(w)이 최종적인 최적의 문맥 정보(C)에 의해서 보정되어, 보정된 인코딩 결과 값인 Wc가 출력될 수 있다.

- [0070] 도 6은 본 개시에 따른 문맥 정보를 이용하는 실시간 자동 번역 방법을 설명하기 위한 흐름도이다.
- [0071] 도 6의 예시와 같은 방법은 본 개시에 따른 문맥 정보 기반 실시간 자동 통역 장치(이하, 장치)에 의해서 수행될 수 있다.
- [0072] 단계 S610에서 장치는 현재 발화 내용에 대한 인코딩을 수행할 수 있다. 예를 들어, 원시 언어의 현재 발화 내용을 음성 인식하여 입력 문장을 구성하고, 입력 문장의 요소들에 대한 심화 학습 모델에 기초한 하나 이상의 노드를 통한 인코딩을 수행하여, 현재 발화 내용 인코딩 결과(w)를 생성할 수 있다.
- [0073] 단계 S620에서 장치는 현재 발화 내용에 관련된 문맥 정보를 인코딩할 수 있다. 예를 들어, 현재 발화 내용의 주제, 이전 발화 내용, 이전 발화 내용의 번역 결과, 현재 발화 내용의 키워드 등의 하나 이상의 타입의 문맥 정보의 인코딩 결과 값을 생성할 수 있다. 또한, 복수의 문맥 정보의 인코딩 결과 값을 컨볼루션 방식으로 조합하여 최종적으로 인코딩된 문맥 정보(C)를 생성할 수도 있다.
- [0074] 단계 S630에서 장치는 단계 S620의 결과물인 인코딩된 문맥 정보에 기초하여, 단계 S610의 결과물인 인코딩된 현재 발화 내용을 보정할 수 있다. 즉, 현재 발화 내용에 관련된 문맥 정보를 이용하여, 상황에 맞는 최적의 번역 결과를 기대할 수 있는 형태로 보정된 현재 발화 내용 인코딩 결과 값(Wc)을 생성할 수 있다.
- [0075] 단계 S640에서 장치는 보정된 현재 발화 내용 인코딩 결과값(Wc)에 기초하여, 현재 발화 내용을 디코딩할 수 있다. 예를 들어, 문맥 정보에 기초하여 보정된 원시 언어의 현재 발화 내용 인코딩 결과 값(Wc)을 이용하여 목적 언어로 된 최초 노드의 디코딩 결과 값을 생성하고, 보정된 현재 발화 내용 인코딩 결과 값(Wc) 및 이전 노드의 디코딩 결과 값을 이용하여 다음 노드의 결과 값을 순차적으로 생성할 수 있다.
- [0076] 본 개시의 예시적인 방법들은 설명의 명확성을 위해서 동작의 시리즈로 표현되어 있지만, 이는 단계가 수행되는 순서를 제한하기 위한 것은 아니며, 필요한 경우에는 각각의 단계가 동시에 또는 상이한 순서로 수행될 수도 있다. 본 개시에 따른 방법을 구현하기 위해서, 예시하는 단계에 추가적으로 다른 단계를 포함하거나, 일부의 단계를 제외하고 나머지 단계를 포함하거나, 또는 일부의 단계를 제외하고 추가적인 다른 단계를 포함할 수도 있다.
- [0077] 본 개시의 다양한 실시 예는 모든 가능한 조합을 나열한 것이 아니고 본 개시의 대표적인 양상을 설명하기 위한 것이며, 다양한 실시 예에서 설명하는 사항들은 독립적으로 적용되거나 또는 둘 이상의 조합으로 적용될 수도 있다.
- [0078] 또한, 본 개시의 다양한 실시 예는 하드웨어, 펌웨어(firmware), 소프트웨어, 또는 그들의 결합 등에 의해 구현될 수 있다. 하드웨어에 의한 구현의 경우, 하나 또는 그 이상의 ASICs(Application Specific Integrated Circuits), DSPs(Digital Signal Processors), DSPDs(Digital Signal Processing Devices), PLDs(Programmable Logic Devices), FPGAs(Field Programmable Gate Arrays), 범용 프로세서(general processor), 컨트롤러, 마이크로 컨트롤러, 마이크로 프로세서 등에 의해 구현될 수 있다.
- [0079] 본 개시의 범위는 다양한 실시 예의 방법에 따른 동작이 장치 또는 컴퓨터 상에서 실행되도록 하는 소프트웨어 또는 머신-실행가능한 명령들(예를 들어, 운영체제, 애플리케이션, 펌웨어(firmware), 프로그램 등), 및 이러한 소프트웨어 또는 명령 등이 저장되어 장치 또는 컴퓨터 상에서 실행 가능한 비-일시적 컴퓨터-판독가능 매체(non-transitory computer-readable medium)를 포함한다.

**부호의 설명**

- [0080] 100 심화 학습 기반 자동 번역 구조
- 110 인코딩부
- 111, 112, 113, 114 인코딩 노드
- 120 디코딩부
- 121, 122, 123, 124, 125, 126 디코딩 노드
- 210, 310, 410 문맥 정보 인코딩부

220, 320, 420, 520 문맥 정보 조합부

311, 312, 313 문맥 정보 인코딩 노드

411, 412, 413, 414 타입 문맥 정보 인코딩부

411\_1, 412\_1, 412\_2, 412\_3, 413\_1, 413\_2, 413\_3, 414\_1, 414\_2, 414\_3 타입 문맥 정보 인코딩 노드

521 문맥 정보 인코딩 결과 값 입력부

521\_1, 521\_2, 521\_3, 521\_4 타입 문맥 정보 인코딩 결과 값 입력부

522 컨볼루션 학습부

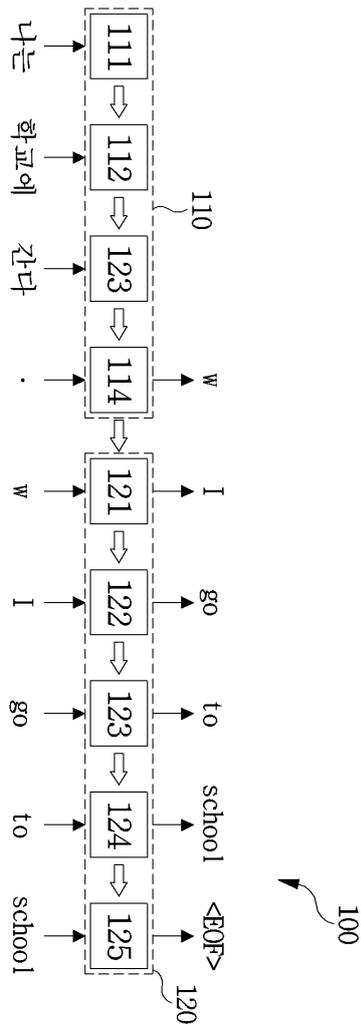
522\_1, 522\_2, 522\_3, 522\_4, 522\_5 조합 노드

523 현재 발화 내용 인코딩 결과 값 입력부

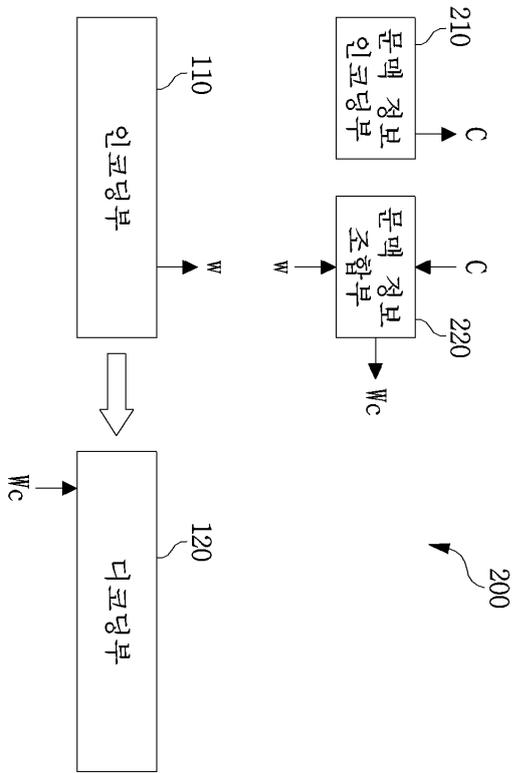
524 보정부

**도면**

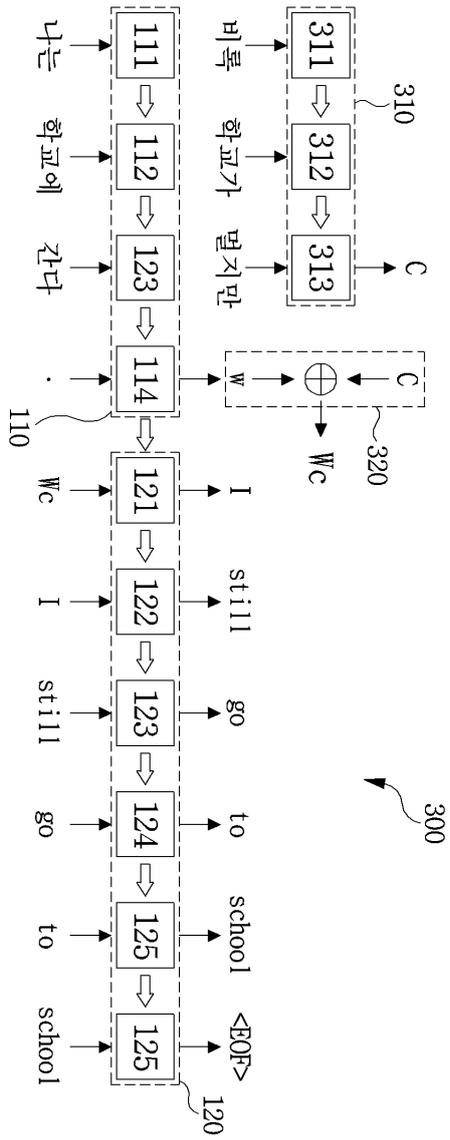
**도면1**



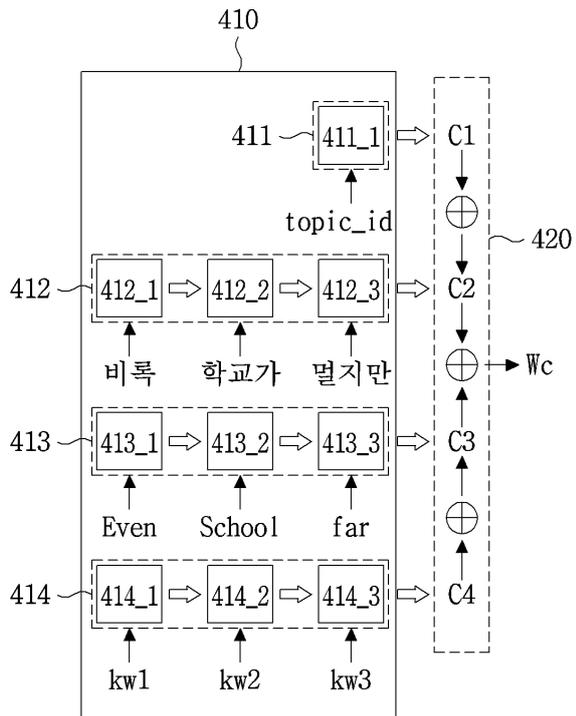
도면2



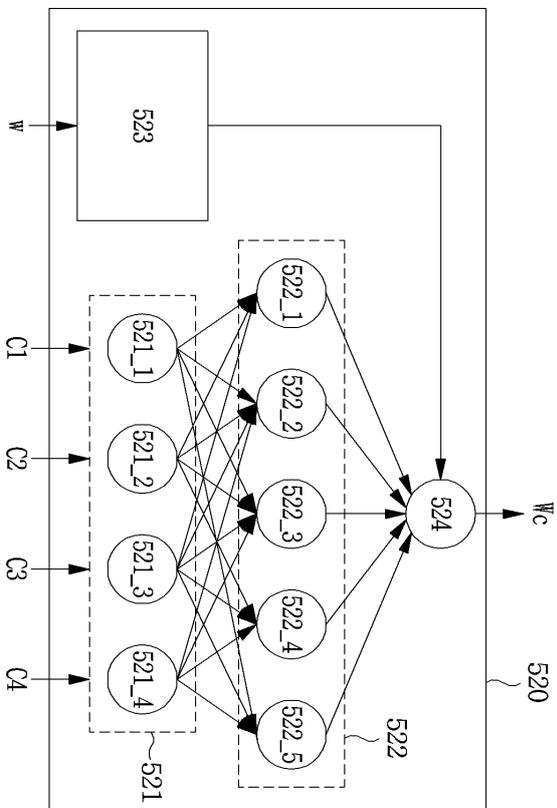
도면3



도면4



도면5



도면6

