



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0058263
(43) 공개일자 2020년05월27일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2006.01) G06F 16/22 (2019.01)
(52) CPC특허분류
G06N 3/08 (2013.01)
G06F 16/22 (2019.01)
(21) 출원번호 10-2019-0021846
(22) 출원일자 2019년02월25일
심사청구일자 2019년02월25일
(30) 우선권주장
1020180139072 2018년11월13일 대한민국(KR)

(71) 출원인
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
(72) 발명자
최재식
울산광역시 울주군 언양읍 유니스트길 50
권순재
울산광역시 울주군 언양읍 유니스트길 50
(뒷면에 계속)
(74) 대리인
특허법인 무한

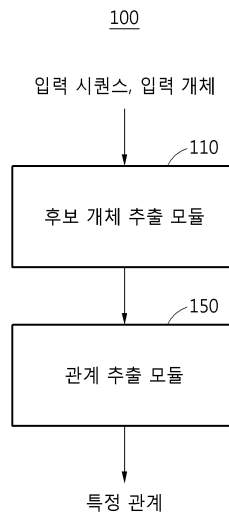
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 신경망 기반 자동 슬롯 채우기 기술 및 장치

(57) 요약

신경망 기반 자동 슬롯 채우기 기술 및 장치가 제공된다. 일 실시예에 따른 슬롯 채우기 장치는 입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출하는 후보 개체 추출 모듈 및 후보 개체를 미리 학습된 제2 신경망에 입력하여 입력 개체와 후보 개체 사이의 특정 관계를 추출하는 관계 예측 모듈을 포함하고, 제1 신경망과 제2 신경망은 입력 레이어를 공유하여 학습된다.

대표도 - 도1



(72) 발명자

임성우

울산광역시 울주군 언양읍 유니스트길 50

이세현

울산광역시 울주군 언양읍 유니스트길 50

이 발명을 지원한 국가연구개발사업

과제고유번호 1711065111

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 ICT융합산업원천기술개발사업

연구과제명 (지능정보-총괄/1세부) 자율지능 디지털 동반자 프레임워크 및 응용 연구개발

기 여 율 1/1

주관기관 전자부품연구원

연구기간 2017.04.01 ~ 2020.12.31

명세서

청구범위

청구항 1

입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 상기 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출하는 단계; 및

상기 후보 개체를 미리 학습된 제2 신경망에 입력하여 상기 입력 개체와 상기 후보 개체 사이의 상기 특정 관계를 추출하는 단계

를 포함하고,

상기 제1 신경망과 상기 제2 신경망은 입력 레이어를 공유하여 학습되는

슬롯 채우기 방법.

청구항 2

제1항에 있어서,

상기 제1 신경망과 상기 제2 신경망은

상기 제1 신경망의 손실함수와 상기 제2 신경망의 손실함수를 결합하여 생성된 새로운 손실함수를 사용하여 학습되는,

슬롯 채우기 방법.

청구항 3

제1항에 있어서,

상기 제1 신경망과 상기 제2 신경망은 동일한 단어 임베딩을 사용하는 방식으로 상기 입력 레이어의 자유 변수를 공유하여 학습되는,

슬롯 채우기 방법.

청구항 4

제3항에 있어서,

상기 단어 임베딩은 상기 제1 신경망과 상기 제2 신경망의 데이터베이스에 록업 테이블 형태로 기록되고, 학습하는 과정에서 상기 록업 테이블을 갱신(update)하는,

슬롯 채우기 방법.

청구항 5

제1항에 있어서,

상기 입력 시퀀스는 토큰 단위로 구성되고,

상기 후보 개체를 추출하는 단계는

상기 입력 시퀀스를 상기 토큰 별로 상기 제1 신경망에 입력하여 상기 토큰에 대응하는 태그로 구성되는 태그

시퀀스를 출력하는 단계; 및
상기 태그 시퀀스에 기초하여 상기 후보 개체를 추출하는 단계
를 포함하는,
슬롯 채우기 방법.

청구항 6

제5항에 있어서,
상기 특정 관계를 추출하는 단계는
상기 후보 개체를 상기 제2 신경망에 입력하여 상기 후보 개체와 상기 입력 개체 사이의 상기 특정 관계를 갖는
정도에 기초하여, 상기 후보 개체를 재순위화하는 단계;
상기 재순위화된 상기 후보 개체와 상기 입력 개체 사이의 상기 특정 관계를 추출하는 단계; 및
상기 특정 관계에 대응하는 수치를 추출하는 단계
를 포함하는,
슬롯 채우기 방법.

청구항 7

제1항에 있어서,
상기 입력 시퀀스는
텍스트 시퀀스를 포함하는,
슬롯 채우기 방법.

청구항 8

제1항에 있어서,
상기 제1 신경망은
장단기 메모리(Long short-term memory: LSTM) 방식의 순환 신경망을 포함하는,
슬롯 채우기 방법.

청구항 9

제1항에 있어서,
상기 제2 신경망은
구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 포함하는,
슬롯 채우기 방법.

청구항 10

하드웨어와 결합되어 제1항 내지 제9항 중 어느 하나의 항의 방법을 실행시키기 위하여 매체에 저장된 컴퓨터

프로그램.

청구항 11

입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 상기 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출하는 후보 개체 추출 모듈; 및

상기 후보 개체를 미리 학습된 제2 신경망에 입력하여 상기 입력 개체와 상기 후보 개체 사이의 상기 특정 관계를 추출하는 관계 예측 모듈

을 포함하고,

상기 제1 신경망과 상기 제2 신경망은 입력 레이어를 공유하여 학습되는

슬롯 채우기 장치.

청구항 12

제11항에 있어서,

상기 제1 신경망과 상기 제2 신경망은

상기 제1 신경망의 손실함수와 상기 제2 신경망의 손실함수를 결합하여 생성된 새로운 손실함수를 사용하여 학습되는,

슬롯 채우기 장치.

청구항 13

제11항에 있어서,

상기 제1 신경망과 상기 제2 신경망은 동일한 단어 임베딩을 사용하는 방식으로 상기 입력 레이어의 자유 변수를 공유하여 학습되는,

슬롯 채우기 장치.

청구항 14

제13항에 있어서,

상기 단어 임베딩은 상기 제1 신경망과 상기 제2 신경망의 데이터베이스에 록업 테이블 형태로 기록되고, 학습하는 과정에서 상기 록업 테이블을 갱신(update)하는,

슬롯 채우기 장치.

청구항 15

제11항에 있어서,

상기 입력 시퀀스는 토큰 단위로 구성되고,

상기 후보 개체를 추출 모듈은

상기 입력 시퀀스를 상기 토큰 별로 상기 제1 신경망에 입력 받아 상기 토큰에 대응하는 태그로 구성되는 태그 시퀀스를 출력하고,

상기 태그 시퀀스에 기초하여 상기 후보 개체를 추출하는,
슬롯 채우기 장치.

청구항 16

제15항에 있어서,
상기 관계 예측 모듈은
상기 후보 개체를 상기 제2 신경망에 입력하여 상기 후보 개체와 상기 입력 개체 사이의 상기 특정 관계를 갖는
정도에 기초하여, 상기 후보 개체를 재순위화하고,
상기 재순위화된 상기 후보 개체와 상기 입력 개체의 상기 특정 관계와 상기 특정 관계에 대응하는 수치를 추출
하는,
슬롯 채우기 장치.

청구항 17

제11항에 있어서,
상기 입력 시퀀스는
텍스트 시퀀스를 포함하는,
슬롯 채우기 장치.

청구항 18

제11항에 있어서,
상기 제1 신경망은
장단기 메모리(Long short-term memory: LSTM) 방식의 순환 신경망을 포함하는,
슬롯 채우기 장치.

청구항 19

제11항에 있어서,
상기 제2 신경망은
구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 포함하는,
슬롯 채우기 장치.

발명의 설명

기술 분야

아래의 실시예들은 신경망 기반 자동 슬롯 채우기 기술 및 장치에 관한 것이다.

배경 기술

[0001]

- [0003] 지식 베이스(Knowledge Base)는 데이터의 비정형적인 관계를 유연하게 표현할 수 있는 효과적인 수단이다. 지식 베이스는 자동 질의 시스템 등 다양한 인공지능 기술을 현실화 하는데 매우 중요하게 사용되고 있다. 그럼에도 불구하고 이런 방대한 지식 베이스를 수동으로 구축하고 유지하는 것은 현실적으로 매우 어렵다. 따라서, 컴퓨터 소프트웨어가 책과 인터넷 등 자료를 스스로 읽어, 중요한 관계를 자동으로 추출할 수 있는 기술(Relation Extraction)은 지식 베이스를 자동으로 확장(Knowledge Base Population)하는데 매우 중요한 기술이다.
- [0004] 관계를 자동으로 추출하는 다양한 기술이 존재하지만, 그 중에서 슬롯 채우기(Slot-Filling) 문제는 관계(예, 본사의 위치)와 키워드(예, 구글)가 주어졌을 때, 주어진 문서에서 키워드와 관계를 갖는 답(예, 캘리포니아)을 찾는 문제로 정의된다. 예를 들어, "구글은 지난 5월부터 캘리포니아 본사 주변에서 차량 공유 서비스를 시범 운영 ..."이라는 문장을 읽고, 구글의 본사가 캘리포니아에 있다는 관계를 찾아, 지식 베이스에 추가하는 작업이다. 슬롯 채우기 문제를 해결하면 기존의 지식 베이스에 저장된 관계가 없는 경우에도, 자연어 문서만으로 새로운 관계를 추출하여 지식 베이스를 확장할 수 있으며, 추론 알고리즘을 통하여 새로운 지식(관계)을 도출할 수도 있다. 이렇게 확장된 지식 베이스는 개인화된 자동 질의 서비스의 질을 향상 시키고, 의사나 변호사가 환자나 판례에 관련한 방대한 자료를 찾는 데 걸리는 시간을 현저하게 줄일 수 있다.
- [0005] 기존 슬롯 채우기 방법들은 개체명 인식(Named-entity recognition; NER) 등의 모듈을 이용하여 입력 개체에 대한 후보 구문을 추출하고, 후보 구문에 대하여 관계 추출기 등의 모듈을 사용하여 정답을 유추하는 접근 방법을 주로 채택하였다. 하지만 해당 방법들은 각 모듈의 학습과 예측 과정이 완전히 분리되어, 후보 구문 추출 과정과 관계 예측 과정에서 공통된 자식을 표현하기 힘들고, 슬롯 채우기 문제 전체를 최적화할 수 없었다.

발명의 내용

해결하려는 과제

- [0007] 실시예들은 후보 개체 추출 모듈과 관계 예측 모듈을 학습하는 과정에서 자유 변수를 공유하고자 한다.
- [0008] 실시예들은 후보 개체 추출 모듈과 관계 예측 모듈을 학습하는 과정에서 각 모듈의 손실함수를 결합하고자 한다.
- [0009] 실시예들은 후보 개체 추출 모듈과 관계 예측 모듈을 학습하는 과정에서 단어에 대응하는 임베딩을 데이터베이스의 룩업 테이블에서 기술하고, 이를 학습 과정에서 업데이트하고자 한다.
- [0010] 실시예들은 후보 개체 추출 모듈에 장단기 메모리(Long short-term memory: LSTM) 방식의 순환 신경망을 사용하고 있다.
- [0011] 실시예들은 관계 예측 모듈에 구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 사용하고 있다.

과제의 해결 수단

- [0013] 일 실시예에 따른 슬롯 채우기 방법은 입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 상기 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출하는 단계; 및 상기 후보 개체를 미리 학습된 제2 신경망에 입력하여 상기 입력 개체와 상기 후보 개체 사이의 상기 특정 관계를 추출하는 단계를 포함하고, 상기 제1 신경망과 상기 제2 신경망은 입력 레이어를 공유하여 학습된다.
- [0014] 상기 제1 신경망과 상기 제2 신경망은 상기 제1 신경망의 손실함수와 상기 제2 신경망의 손실함수를 결합하여 생성된 새로운 손실함수를 사용하여 학습될 수 있다.
- [0015] 상기 제1 신경망과 상기 제2 신경망은 동일한 단어 임베딩을 사용하는 방식으로 상기 입력 레이어의 자유 변수를 공유하여 학습될 수 있다.
- [0016] 상기 단어 임베딩은 상기 제1 신경망과 상기 제2 신경망의 데이터베이스에 룩업 테이블 형태로 기록되고, 학습하는 과정에서 상기 룩업 테이블을 갱신(update)할 수 있다.

- [0017] 상기 입력 시퀀스는 토큰 단위로 구성되고, 상기 후보 개체를 추출하는 단계는 상기 입력 시퀀스를 상기 토큰 별로 상기 제1 신경망에 입력하여 상기 토큰에 대응하는 태그로 구성되는 태그 시퀀스를 출력하는 단계; 및 상기 태그 시퀀스에 기초하여 상기 후보 개체를 추출할 수 있다.
- [0018] 상기 특정 관계를 추출하는 단계는 상기 후보 개체의 상기 입력 개체와의 상기 특정 관계를 갖는 정도에 기초하여, 상기 후보 개체를 재순위화하는 단계; 상기 재순위화된 상기 후보 개체와 상기 입력 개체 사이의 상기 특정 관계를 추출하는 단계; 및 상기 특정 관계에 대응하는 수치를 추출하는 단계를 포함할 수 있다.
- [0019] 상기 입력 시퀀스는 텍스트 시퀀스를 포함할 수 있다.
- [0020] 상기 제1 신경망은 장단기 메모리(Long short-term memory: LSTM) 방식의 순환 신경망을 포함할 수 있다.
- [0021] 제1항에 있어서, 상기 제2 신경망은 구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 포함할 수 있다.
- [0022] 일 실시예에 따른 슬롯 채우기 장치는 입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 상기 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출하는 후보 개체 추출 모듈; 및 상기 후보 개체를 미리 학습된 제2 신경망에 입력하여 상기 입력 개체와 상기 후보 개체 사이의 상기 특정 관계를 추출하는 관계 예측 모듈을 포함하고, 상기 제1 신경망과 상기 제2 신경망은 입력 레이어를 공유하여 학습된다.
- [0023] 상기 제1 신경망과 상기 제2 신경망은 상기 제1 신경망의 손실함수와 상기 제2 신경망의 손실함수를 결합하여 생성된 새로운 손실함수를 사용하여 학습될 수 있다.
- [0024] 상기 제1 신경망과 상기 제2 신경망은 동일한 단어 임베딩을 사용하는 방식으로 상기 입력 레이어의 자유 변수를 공유하여 학습될 수 있다.
- [0025] 상기 단어 임베딩은 상기 제1 신경망과 상기 제2 신경망의 데이터베이스에 록업 테이블 형태로 기록되고, 학습하는 과정에서 상기 록업 테이블을 갱신(update)할 수 있다.
- [0026] 상기 입력 시퀀스는 토큰 단위로 구성되고, 상기 후보 개체를 추출 모듈은 상기 입력 시퀀스를 상기 토큰 별로 상기 제1 신경망에 입력하여 상기 토큰에 대응하는 태그로 구성되는 태그 시퀀스를 출력하고, 상기 태그 시퀀스에 기초하여 상기 후보 개체를 추출할 수 있다.
- [0027] 상기 관계 예측 모듈은 상기 후보 개체의 상기 입력 개체와의 상기 특정 관계를 갖는 정도에 기초하여, 상기 후보 개체를 재순위화하고, 상기 재순위화된 상기 후보 개체와 상기 입력 개체의 상기 특정 관계와 상기 특정 관계에 대응하는 수치를 추출할 수 있다.
- [0028] 상기 입력 시퀀스는 텍스트 시퀀스를 포함할 수 있다.
- [0029] 상기 제1 신경망은 장단기 메모리(Long short-term memory: LSTM) 방식의 순환 신경망을 포함할 수 있다.
- [0030] 상기 제2 신경망은 구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 포함할 수 있다.

발명의 효과

- [0032] 실시예들은 후보 개체 추출 모듈과 관계 예측 모듈을 학습하는 과정에서 자유 변수를 공유할 수 있다.
- [0033] 실시예들은 후보 개체 추출 모듈과 관계 예측 모듈을 학습하는 과정에서 각 모듈의 손실함수를 결합할 수 있다.
- [0034] 실시예들은 후보 개체 추출 모듈과 관계 예측 모듈을 학습하는 과정에서 단어에 대응하는 임베딩을 데이터베이스의 록업 테이블에서 기술하고, 이를 학습 과정에서 업데이트할 수 있다.
- [0035] 실시예들은 후보 개체 추출 모듈에 장단기 메모리(Long short-term memory: LSTM) 방식의 순환 신경망을 사용할 수 있다.
- [0036] 실시예들은 관계 예측 모듈에 구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 사용할 수 있다.

도면의 간단한 설명

- [0038] 도 1은 일 실시예에 따른 슬롯 채우기 장치의 동작 방법을 설명하기 위한 도면이다.
- 도 2는 일 실시예에 따른 후보 개체 추출 모듈의 구체적인 동작 방법을 설명하기 위한 도면이다.
- 도 3은 일 실시예에 따른 관계 예측 모듈의 구체적인 동작 방법을 설명하기 위한 도면이다.
- 도 4는 일 실시예에 따른 슬롯 채우기 방법의 순서도이다.
- 도 5는 일 실시예에 따른 슬롯 채우기 방법을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0039] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0040] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 변경, 균등물, 또는 대체물을 포함한다.
- [0041] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만, 예를 들어 본 발명의 개념에 따른 권리 범위로부터 이탈되지 않은 채, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0042] 어떤 구성요소가 다른 구성요소에 “연결되어” 있다거나 “접속되어” 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 “직접 연결되어” 있다거나 “직접 접속되어” 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성요소들 간의 관계를 설명하는 표현들, 예를 들어 “~사이에”와 “바로~사이에” 또는 “~에 직접 이웃하는” 등도 마찬가지로 해석되어야 한다.
- [0043] 본 명세서에서 사용한 용어는 단지 특정한 실시예들을 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, “포함하다” 또는 “가지다” 등의 용어는 실시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0044] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0045] 이하, 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0047] 도 1은 일 실시예에 따른 슬롯 채우기 장치의 동작 방법을 설명하기 위한 도면이다.
- [0048] 슬롯 채우기 기술은 입력 시퀀스(input sequence)를 기반으로, 입력 개체(input entity)(예를 들어, 인명, 기관명, 지명, 시간, 날짜, 종교 등)에 대해 특정 관계(예를 들어, 탄생일, 사망일, 설립일, 종교, 배우자, 자녀, 자매 등)를 갖는 개체들을 찾아서 반환하는 기술일 수 있다. 예를 들어, 입력 개체가 ‘버락 오바마’ 이고, 질의-관계가 ‘배우자’ 라면, “버락 오바마 정권의 영부인인 미셸 오바마가 한국을 방문했다” 라는 입력 시퀀스

를 기반으로, ‘미셸 오바마’가 ‘버락 오바마’의 배우자라는 것을 출력할 수 있다. 슬롯 채우기 기술은 관계 추출(relation extraction) 기술로 지칭될 수 있다. 개체는 인명, 기관명, 지명, 시간, 날짜 등 고유한 의미를 가지는 것들로, 형태소, 품사가 아닌 의미가 있는 단어들이 수 있다.

- [0049] 기존 슬롯 채우기 방법들은 개체명 인식(Named-entity recognition; NER) 등의 모듈을 이용하여 입력 개체에 대한 후보 구문을 추출하고, 후보 구문에 대하여 관계 추출기 등의 모듈을 사용하여 정답을 유추하는 접근 방법을 주로 채택하였다. 하지만 해당 방법들은 각 모듈의 학습과 예측 과정이 완전히 분리되어, 후보 구문 추출 과정과 관계 예측 과정에서 공통된 자식을 표현하기 힘들고, 슬롯 채우기 문제 전체를 최적화할 수 없었다.
- [0050] 도 1을 참조하면, 일 실시예에 따른 슬롯 채우기 장치(100)는 후보 개체 추출 모듈(110) 및 관계 예측 모듈(150)을 포함한다.
- [0051] 후보 개체 추출 모듈(110)은 미리 학습된 제1 신경망(미도시)을 포함할 수 있다. 후보 개체 추출 모듈(110)은 입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출할 수 있다. 입력 시퀀스는 토큰 단위로 구성되고, 자연어 데이터인 텍스트 시퀀스를 포함할 수 있다. 특정 관계는 입력 개체와 후보 개체 사이의 탄생일, 사망일, 설립일, 종교, 배우자, 자녀, 자매 등을 포함할 수 있다. 또한, 특정 관계는 본 명세서에 기재된 예시 뿐만 아니라 학습 단계에서 미리 결정될 수 있다. 후보 개체는 입력 개체와 특정 관계를 갖을 수 있는 입력 시퀀스에 포함된 개체일 수 있다. 후보 개체 추출 모듈(110)의 구체적인 동작 방법은 아래에서 도 2를 참조하여 상세히 설명된다.
- [0052] 관계 예측 모듈(150)은 미리 학습된 제2 신경망(미도시)을 포함할 수 있다. 관계 예측 모듈(150)은 후보 개체를 미리 학습된 제2 신경망에 입력하여 입력 개체와 후보 개체 사이의 특정 관계를 추출할 수 있다. 관계 예측 모듈(150)의 구체적인 동작 방법은 아래에서 도 3을 참조하여 상세히 설명된다.
- [0053] 슬롯 채우기 장치(100)의 후보 개체 추출 모듈(110)의 제1 신경망과 관계 예측 모듈(150)의 제2 신경망은 입력 레이어를 공유하여 학습된다. 후보 개체 추출 모듈(110)의 제1 신경망과 관계 예측 모듈(150)의 제2 신경망을 학습하는 과정에서 자유 변수를 공유하기 때문에 공통된 자식을 표현할 수 있고, 슬롯 채우기 문제 전체를 최적화할 수 있다.
- [0054] 제1 신경망과 제2 신경망은 동일한 단어 임베딩을 사용하는 방식으로 입력 레이어의 자유 변수를 공유할 수 있다. 후보 개체 추출 모듈(110)의 제1 신경망과 관계 예측 모듈(150)의 제2 신경망은 슬롯 채우기 방법을 구현하기 위해서 단어를 벡터 형식으로 변환하는 단어 임베딩 모델을 사용할 수 있다. 단어 임베딩(word embedding)은 단어의 의미를 벡터로 표현한 것일 수 있다. 예를 들어, 대용량 말뭉치를 입력으로 말뭉치 내 각 단어들을 n 차원의 실수 벡터 공간상에 사상(mapping)하여 단어의 의미를 파악하는 비지도 학습(unsupervised learning)의 일종일 수 있다. 단어 임베딩은 일반적으로 입력 말뭉치를 토큰(token) 단위로 분할한 다음 의미적 연관성이 높은 토큰들을 유사한 실수 벡터 값으로 생성할 수 있고, 단어(word) 단위 즉, 띄어쓰기 단위로 토큰을 생성할 수 있다. 또는, 한국어에서는 형태소 단위로 토큰을 구성하여 단어 임베딩을 학습하는 방식이 사용될 수 있다. 단어 임베딩 방법에 관한 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0055] 또한, 제1 신경망과 제2 신경망은 단어에 대응하는 임베딩을 데이터베이스의 룩업 테이블에서 기술하고, 이를 학습 과정에서 업데이트할 수 있다. 후보 개체 추출 모듈(110)과 관계 예측 모듈(150)은 동일한 단어 임베딩을 사용하는 방식으로 입력 레이어의 자유 변수를 공유하고, 단어에 대응하는 임베딩을 데이터베이스의 룩업 테이블에서 기술하고, 이를 학습 과정에서 업데이트하여 두 모듈이 슬롯 채우기 문제에 필요한 공통된 지식을 공유하게 만들 수 있다.
- [0056] 제1 신경망과 상기 제2 신경망은 제1 신경망의 손실함수와 제2 신경망의 손실함수를 결합하여 생성된 새로운 손실함수를 사용하여 학습될 수 있다. 새로운 손실함수를 생성하는 구체적인 동작 방법은 아래에서 도 4를 참조하여 상세히 설명된다.
- [0058] 도 2는 일 실시예에 따른 후보 개체 추출 모듈의 구체적인 동작 방법을 설명하기 위한 도면이다.
- [0059] 도 2를 참조하면, 일 실시예에 따른 후보 개체 추출 모듈(110)은 입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출할 수 있다. 제1 신경망은 순환 신경망(recurrent neural network: RNN)을 포함할 수 있다. 예를 들어, 제1 신경망은 장단기 메모

리(Long short-term memory: LSTM) 방식의 순환 신경망을 포함할 수 있다. 또는, 후보 개체 추출 모듈(110)은 양방향 LSTM-CRF 모델을 포함할 수 있다.

[0060] 제1 신경망은 품사 태깅(Part-of-speech tagging), 개체명 인식(Named Entity Recognition), 레머타이제이션(Lemmatization), 의존 구문 분석(dependency parsing)등 자연어 처리를 위한 알고리즘들을 수행할 수 있도록 학습될 수 있다.

[0061] 후보 개체 추출 모듈(110)은 입력 시퀀스를 토큰 별로 제1 신경망에 입력 받아 토큰에 대응하는 태그로 구성되는 태그 시퀀스를 출력할 수 있다. 예를 들어, 입력 시퀀스는 "President Barack Obama and First Lady Michelle Obama welcome Donald Trump at White House"일 수 있고, 입력 개체는 "Barack Obama"일 수 있다. 이하에서, 설명의 편의를 위하여 "President Barack Obama and First Lady Michelle Obama welcome Donald Trump at White House"에 기초하여 설명한다. 토큰은 시퀀스에 포함되는 단어 단위, 즉 띄어쓰기 단위로 결정될 수 있다. 예를 들어, 입력 시퀀스는 "President Barack Obama and First Lady Michelle Obama welcome Donald Trump at White House"는 14개의 토큰으로 이뤄질 수 있다.

[0062] 후보 개체 추출 모듈(110)은 학습된 제1 신경망을 이용하여, 입력 개체 "Barack Obama"에 기초하여 입력 시퀀스의 각 토큰에 대응하는 태그로 구성되는 태그 시퀀스 "B-cand E1 E1 0 0 0 B-cand I-cand 0 0 0 0 0"를 출력할 수 있다. "B-cand"는 후보 개체의 시작을 의미하는 태그이고, "I-cand"는 연속되는 후보 개체를 의미하는 태그이고, "E1"은 입력 개체를 의미하는 태그이며, "0"는 후보 개체도 입력 개체도 아닌 의미를 갖지 않는 태그일 수 있다.

[0063] 후보 개체 추출 모듈(110)은 태그 시퀀스에 기초하여 후보 개체를 추출할 수 있다. 태그 시퀀스 "B-cand E1 E1 0 0 0 B-cand I-cand 0 0 0 0 0"에 기초하여 "President", "Michelle Obama"를 후보 개체로 추출할 수 있다.

[0065] 도 3은 일 실시예에 따른 관계 예측 모듈의 구체적인 동작 방법을 설명하기 위한 도면이다.

[0066] 도 3을 참조하면, 일 실시예에 따른 관계 예측 모듈(150)은 후보 개체를 제2 신경망에 입력하여 후보 개체와 입력 개체 사이의 특정 관계를 갖는 정도에 기초하여, 후보 개체를 재순위화할 수 있다. 제1 신경망은 구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 포함할 수 있다. PCNN은 CNN 모델을 확장한 것으로, CNN에서 사용하는 최대 풀링 레이어(max pooling layer)를 구분적 최대 풀링 레이어(piecewise max pooling layer)로 확장하였다는 것이 큰 차이점일 수 있다.

[0067] 관계 예측 모듈(150)은 재순위화된 후보 개체와 입력 개체 사이의 특정 관계를 추출할 수 있고, 나아가 특정 관계에 대응하는 수치도 추출할 수 있다. 예를 들어, 관계 예측 모듈(150)은 후보 개체 "president"와 "Michelle Obama"를 제2 신경망에 입력하여, 재순위화할 수 있고, 그 결과에 기초하여 후보 개체 "president"는 입력 개체 "Barack Obama"와 "title"의 관계에 있고, 후보 개체 "Michelle Obama"는 입력 개체 "Barack Obama"와 "spouse"의 관계에 있음을 추출할 수 있다. 또한, 후보 개체 "president"는 입력 개체 "Barack Obama"와 "title"의 관계에 있음과 함께 0.87의 점수를 가짐을, 마찬가지로 "Michelle Obama"는 입력 개체 "Barack Obama"와 "spouse"의 관계에 있음과 함께 0.46의 점수를 가짐을 추출할 수 있다.

[0068] 일 실시예에 따른 관계 예측 모듈(150)은 후보 개체와 입력 개체 사이의 특정 관계에 대응하는 점수가 임계값 이상인 경우에 후보 개체의 특정 관계를 결정할 수 있다. 예를 들어, 후보 개체와 입력 개체 사이의 특정 관계에 대응되는 점수가 임계값 이상이라면, 후보 개체는 입력 개체와 복수 개의 특정 관계를 가질 수 있다. 예를 들어, 입력 시퀀스가 "IBM chief Ginni Rometty puts emphasis on responsible use of data" 이고, 입력 개체가 "IBM"인 경우 후보 개체 "Ginni Rometty"는 입력 개체와 "employee" 관계일수도 있고, "top employee"관계일수도 있다.

[0070] 도 4는 일 실시예에 따른 슬롯 채우기 방법의 순서도이다.

[0071] 도 4를 참조하면, 단계(410, 420)은 도 1 내지 도 3를 참조하여 전송된 슬롯 채우기 장치(100)에 의해 수행될 수 있다. 슬롯 채우기 장치(100)는 하나 또는 그 이상의 하드웨어 모듈, 하나 또는 그 이상의 소프트웨어 모듈, 또는 이들의 다양한 조합에 의하여 구현될 수 있다.

[0072] 단계(410)에서, 슬롯 채우기 장치(100)는 입력 시퀀스와 입력 개체를 미리 학습된 제1 신경망에 입력하여 입력 개체와 특정 관계를 갖을 수 있는 하나 이상의 후보 개체를 추출한다.

[0073] 단계(420)에서, 슬롯 채우기 장치(100)는 후보 개체를 미리 학습된 제2 신경망에 입력하여 입력 개체와 후보 개체 사이의 특정 관계를 추출한다.

[0074] 제1 신경망과 상기 제2 신경망은 제1 신경망의 손실함수와 제2 신경망의 손실함수를 결합하여 생성된 새로운 손실함수를 사용하여 학습될 수 있다.

[0075] 제1 신경망의 손실 함수는 수학적 식 1과 같을 수 있다.

수학적 식 1

[0076]
$$loss_{cand} = -\log(P(Y_{cand}^*|X))$$

[0077] $loss_{cand}$ 는 제1 신경망의 손실 함수, X 는 입력 시퀀스, Y_{cand}^* 는 제1 신경망의 출력 입력 시퀀스에 대응되는 태그 시퀀스일 수 있다.

[0078] 제2 신경망의 손실 함수는 수학적 식 2와 같을 수 있다.

수학적 식 2

[0079]
$$loss_{rerank} = \begin{cases} \sum_{i=0}^n cross_entropy(Y_{relation}^{i*}, reranker(X, e_1, e_2^i)), & if\ n = 0 \\ 0, & else \end{cases}$$

[0080] $loss_{rerank}$ 는 제2 신경망의 손실 함수, e_1 은 입력 개체, e_2^i 는 i 번째 후보 개체일 수 있다. 예를 들어, e_1 은 "Barack Obama", e_2^1 는 "President", e_2^2 는 "Michelle Obama"일 수 있다. $Y_{relation}^*$ 은 입력 개체와 후보 개체 사이의 특정 관계일 수 있고, $Y_{relation}^*$ 은 수학적 식 3과 같을 수 있다.

수학적 식 3

[0081]
$$Y_{relation}^* = (e_2^1 = President, Y_{relation}^{1*} = title), (e_2^2 = Michelle Obama, Y_{relation}^{2*} = spouse)$$

[0082] 제1 신경망의 손실함수와 제2 신경망의 손실함수를 결합하여 생성된 새로운 손실함수는 수학적 식 4와 같을 수 있다.

수학적 식 4

[0083]
$$loss_{joint} = \alpha * loss_{cand} + (1 - \alpha)loss_{rerank}$$

[0084] 수학적 식 4와 같이 두 모듈의 손실 함수를 결합하여 전체 슬롯 채우기 문제에 대해 최적화할 수 있는 새로운 손실 함수를 사용할 수 있다.

[0086] 도 5는 일 실시예에 따른 슬롯 채우기 방법을 설명하기 위한 도면이다.

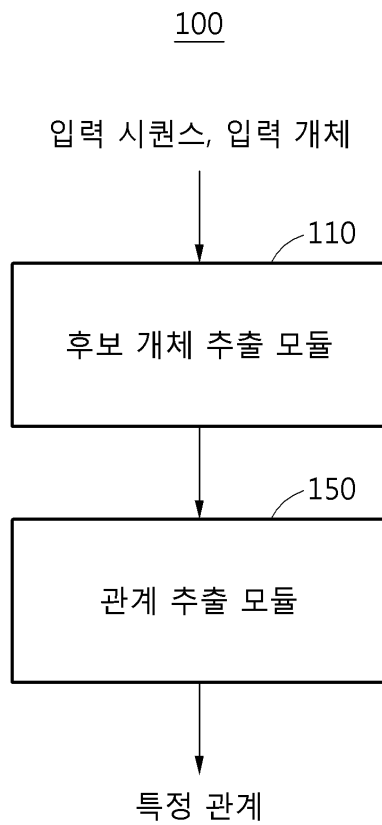
- [0087] 도 5를 참조하면, 일 실시예에 따른 슬롯 채우기 장치(500)의 후보 개체 추출 모듈(510)에 입력 시퀀스 "President Barack Obama and First Lady Michelle Obama welcome Donald Trump at White House"와 입력 개체는 "Barack Obama"이 입력될 수 있고, 후보 개체 추출 모듈(510)은 태그 시퀀스 "B-cand E1 E1 O O O B-cand I-cand O O O O O O"를 출력할 수 있다. 태그 시퀀스 "B-cand E1 E1 O O O B-cand I-cand O O O O O O"에 기초하여 "President", "Michelle Obama"를 후보 개체로 추출할 수 있다. 후보 개체 추출 모듈(510)은 양방향 LSTM-CRF 모델을 포함할 수 있다.
- [0088] 관계 예측 모듈(550)은 후보 개체 "President", "Michelle Obama"와 입력 개체 "Barack Obama" 사이의 특정 관계를 추출할 수 있다. 예를 들어, "president"는 입력 개체 "Barack Obama"와 "title"의 관계에 있고, 후보 개체 "Michelle Obama"는 입력 개체 "Barack Obama"와 "spouse"의 관계에 있음을 추출할 수 있다. 또한, 후보 개체 "president"는 입력 개체 "Barack Obama"와 "title"의 관계에 있음과 함께 0.87의 점수를 가짐을, 마찬가지로 "Michelle Obama"는 입력 개체 "Barack Obama"와 "spouse"의 관계에 있음과 함께 0.46의 점수를 가짐을 추출할 수 있다. 관계 예측 모듈(550)은 구간적 합성곱 신경망(Piecewise Convolutional Neural Network: PCNN)을 포함할 수 있다.
- [0090] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0091] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0092] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0093] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태

로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0094] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

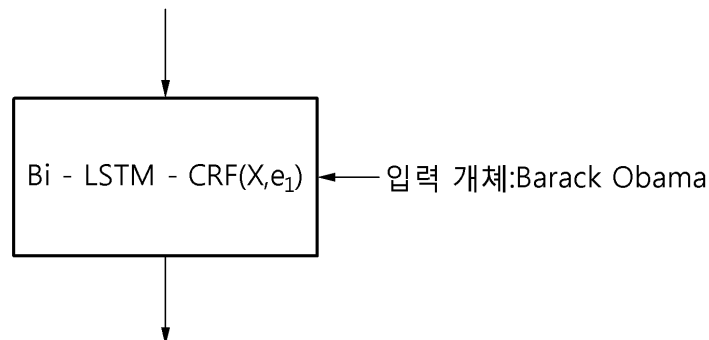
도면1



도면2

X:입력 시퀀스

President Barack Obama and First Lady Michelle Obama welcome Donald Trump at White House.



Y:태그 시퀀스

President Barack Obama and First Lady Michelle Obama welcome Donald Trump at White House.

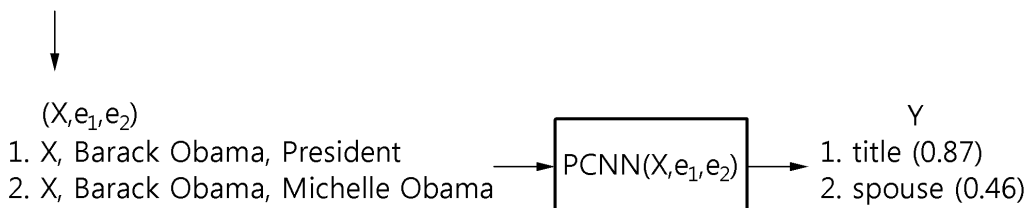
B-cand E1 E1 O O O B-cand I-cand O O O O O O

도면3

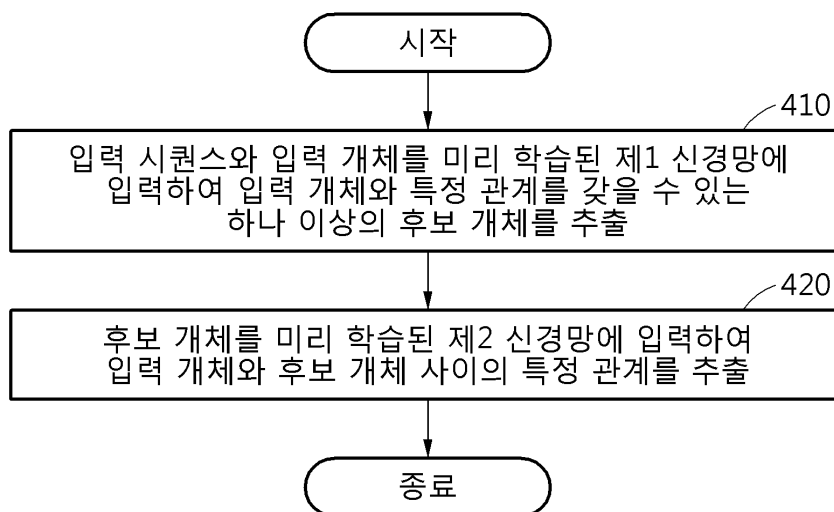
Y:태그 시퀀스

President Barack Obama and First Lady Michelle Obama welcome Donald Trump at White House.

B-cand E1 E1 O O O B-cand I-cand O O O O O O



도면4



도면5

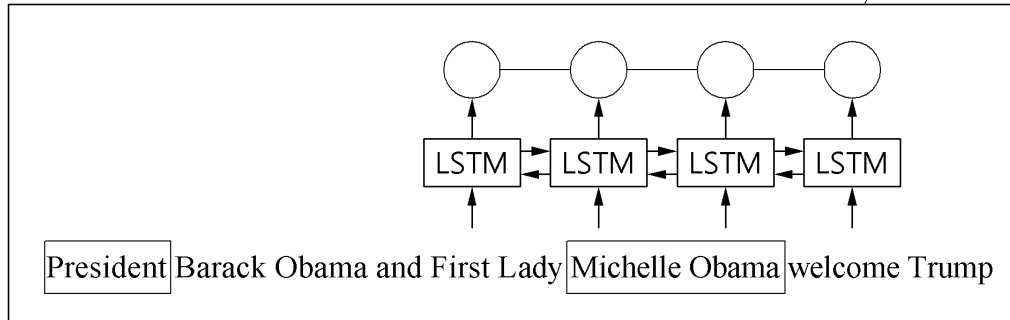
500

입력 시퀀스

President Barack Obama and First Lady Michelle Obama welcome Trump



510



550

