# 순환 신경망 병렬화를 사용한 의존 구문 분석 및 의미역 결정 통합 모델

박성식○, 김학수
강원대학교 컴퓨터 정보통신공학과
{a163912, nlpdrkim}@kangwon.ac.kr

# Joint Model for Dependency Parser and Semantic Role Labeling using Recurrent Neural Network Parallelism

Seong Sik Park○, Hark Soo Kim
Kangwon National University, Department of Computer and Communications Engineering

**요 약**

의존 구문 분석은 문장을 구성하는 성분들 간의 의존 관계를 분석하고 문장의 구조적 정보를 얻기 위한 기술이다. 의미역 결정은 문장에서 서술어에 해당하는 어절을 찾고 해당 서술어의 논항들을 찾는 자연어 처리의 한 분야이다. 두 기술은 서로 밀접한 상관관계가 존재하며 기존 연구들은 이 상관관계를 이용하기 위해 의존 구문 분석의 결과를 의미역 결정의 자질로써 사용한다. 그러나 이런 방법은 의미역 결정 모델의 오류가 의존 구문 분석에 역전파 되지 않으므로 두 기술의 상관관계를 효과적으로 사용한다고 보기 어렵다. 본 논문은 포인터 네트워크 기반의 의존 구문 분석 모델과 병렬화 순환 신경망 기반의 의미역 결정 모델을 멀티 태스크 방식으로 학습시키는 통합 모델을 제안한다. 제안 모델은 의존 구문 분석 및 의미역 결정 말뭉치인 UProbBank를 실험에 사용하여 의존 구문 분석에서 UAS 0.9327, 의미역 결정에서 PIC F1 0.9952, AIC F1 0.7312의 성능 보였다.

주제어: 의존 구문 분석, 의미역 결정, 통합 모델, 멀티 태스크 학습

## 1. 서론

의존 구문 분석(Dependency Parsing)은 단어나 어절, 구처럼 문장을 구성하는 성분들 간의 의존 관계를 분석하여 문장의 구조적 정보를 파악하기 위한 기술이다. 의미역 결정(Semantic Role Labeling)은 문장에서 서술어(Predicate)를 찾아내고 문장의 각 구성 성분이 해당 서술어에 대해 어떤 역할을 하는지 분류하는 자연어 처리의 한 분야이다. 구성 성분이 가질 수 있는 역할에는 대표적으로 “누가, 무엇을, 왜” 등이 있으며 이런 역할을 갖는 구성 성분들을 논항(Argument)이라 한다. 일반적으로 서술어와 논항 사이에는 직접적 또는 간접적으로 의존 관계가 존재하기 때문에 의존 구문 분석과 의미역 결정 간에는 높은 상관관계가 존재한다고 볼 수 있다. 최근 대부분의 의미역 결정 연구는 심층 학습(Deep Learning)을 기반으로 한 의존 구문 분석 모델과 의미역 결정 모델을 독립적으로 설계하고 의존 구문 분석 모델의 결과를 의미역 결정 모델의 자질로 사용하는 파이프라인(Pipeline) 방식으로 연구를 진행한다[1-3]. 그러나 이 방법은 의존 구문 분석 모델과 의미역 결정 모델 간 역전파(Backpropagation)를 통한 학습이 진행되지 않기 때문에 두 모델의 상관관계를 효과적으로 사용하지 못한다는 단점이 존재한다. 본 논문은 이 문제점을 해결하기 위해 멀티 태스크 학습(Multi-task Learning)으로 구문 정보와 의미역 정보를 동시에 학습하는 통합 모델을 제안한다.

## 2. 관련 연구

최근 대부분의 의존 구문 분석 연구는 순환 신경망을 응용하는 방법을 중심으로 진행되고 있다. [4]는 포인터 네트워크[5]를 기반으로 의존 관계와 의존 관계 레이블(Label)을 동시에 학습하는 모델을 제안했다. 포인터 네트워크는 각 어절의 중심어 위치를 찾아내는데 좋은 성능을 보였으며 의존 관계 레이블을 계산하는데도 높은 성능을 보였다. [6]은 의존 구문 분석에서 포인터 네트워크만으로 문장의 구조적인 정보를 파악하는데 한계가 있다고 판단하고 멀티헤드 어텐션(Multi-head Attention)[7]을 활용한 자가 주의집중 (Self Attention)을 통해 의존 구문 분석의 성능 향상을 보였다. 의미역 결정 또한 순환 신경망을 사용하는 방법들이 많이 연구되고 있다. [8]은 중심어가 후위에 위치하는 한국어의 특징에 따라 역방향 LSTM CRFs(Long Short-Term Memory, Conditional Random Fields)를 의미역 결정에 적용했다. [3]은 문자열 기반 양방향 LSTM CRFs에 구문 정보 자질을 활용하는 방법으로 의미역 결정에 좋은 성능을 보였다. 구문 정보는 의미역 결정에 있어서 효과적인 자질로 사용될 수 있으나 구문 정보가 없는 데이터를 사용할 시 구문 분석이 선행되어야 한다. 또한 미리 습득한 구문 정보를 사용하기 때문에 의미역 결정에서 생긴 오류 정보를 구문 분석 모델에서 학습 할 수 없다는 단점이 존재한다. 본 논문의 제안 모델은 의

존 구문 분석과 의미역 결정 모델을 멀티 태스크 학습시켜 이 문제를 보완하고자 한다.

## 3. 의존 구문 분석 및 의미역 결정 통합 모델

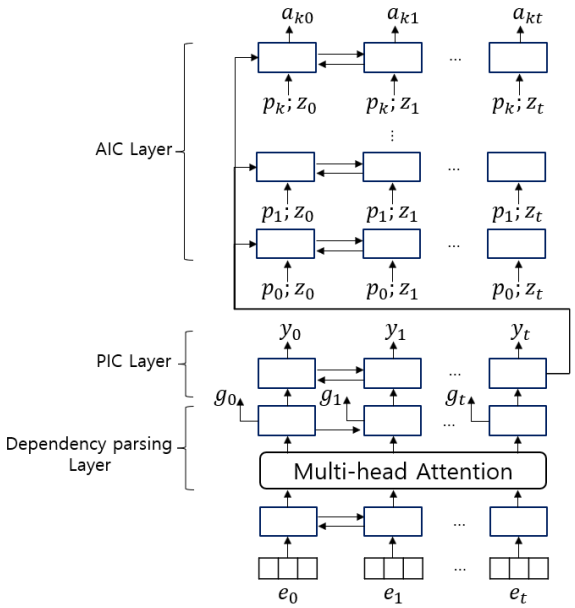


그림 1. 의존 구문 분석 및 의미역 결정 통합 모델 구조도

제안 모델의 전체 구조도는 그림 1과 같다. 제안 모델은 각 어절의 중심어 위치를 출력하는 포인터 네트워크 기반의 구문 분석 계층(Dependency Parsing Layer)과 양방향 순환 신경망 기반의 서술어 인식 및 분류 계층(PIC layer, Predicate Identification and Classification)과 순환 신경망 병렬화를 통한 논항 인식 및 분류 계층(AIC layer, Argument Identification and Classification)으로 구성된다.

### 3.1 포인터 네트워크 기반 의존 구문 분석 계층

의존 구문 분석 계층은 [6]의 모델을 기반으로 설계한다. 의존 구문 분석 계층의 입력은 어절 단위 임베딩(Embedding)을 사용한다. 어절을 구성하는 형태소들의 임베딩과 음절들의 임베딩을 각각 합성곱 신경망(Convolutional Neural Network)에 입력하여 하나의 어절을 표현하는 임베딩 $e_t$를 생성한다[9]. 형태소 임베딩은 사전 학습된 단어 임베딩과 임의 초기화(Random Initialize)한 품사 태그 임베딩을 연결(Concatenated)해 사용하며 음절 임베딩은 임의 초기화 임베딩을 사용한다. 각 어절의 임베딩은 순환 신경망의 일종인 양방향 GRU(Gated Recurrent Unit)[10]를 통해 인코딩(Encoding)된다. 이후 인코딩된 값들은 멀티헤드 어텐션을 통해 자가 주의집중 가중치가 반영된 문맥 벡터(Context Vector) 계산에 사용되고 문맥 벡터들은 포인터 네트워크에 입력되어 각 어절의 중심어 위치 $g_t$를 계산하는데 사용된다.

### 3.2 서술어 인식 및 분류 계층과 논항 인식 및 분류 계층

서술어 인식 및 분류 계층의 입력은 어절 임베딩, 인코딩 계층 출력, 자가 주의집중 계층의 출력을 연결한 벡터다. 서술어 인식 및 분류 계층은 양방향 GRU를 이용한 이진 분류를 통해 해당 어절이 서술어에 해당하는지를 나타내는 결과 값 $y_t$를 출력한다. 논항 인식 및 분류 계층에서는 양방향 GRU를 서술어의 수만큼 병렬로 수행한다. 병렬 수행 방법은 [11]의 방법을 응용한 것으로 [11]에서는 병렬 순환 신경망 사이의 가중치가 공유되지 않으며 최종적으로 각 순환 신경망의 출력을 합쳐서 하나의 예측 열을 출력한다. 그러나 제안 모델은 병렬 GRU의 가중치를 공유하도록 설계하고 GRU의 출력 각각을 서술어 하나에 대한 예측 열로 간주하여 서술어 수만큼 정답 열을 출력 할 수 있도록 변경했다. 각 병렬 GRU의 입력에서 $p_k$는 $k$번째 서술어에 해당하는 어절의 임베딩이며, $z_t$는 문장을 구성하는 각 어절의 임베딩, 인코더 계층 출력, 자가 주의집중 계층 출력, 의존 구문 분석 계층의 출력을 모두 연결한 벡터다. 논항 인식 및 분류 계층의 출력은 $k$개 서술어 각각의 논항 예측 열 $a_t$이다.

## 4. 실험

### 4.1 실험 환경

표 1. 모델 파라미터

| 모델 파라미터 | 값 |
|---|---|
| 형태소 임베딩 차원 수 | 100 |
| 품사 태그 임베딩 차원 수 | 32 |
| 음절 임베딩 차원 수 | 50 |
| 형태소 CNN 필터 크기 | 2, 3, 4 |
| 형태소 CNN 필터 개수 | 150 |
| 음절 CNN 필터 크기 | 2, 3, 5 |
| 음절 CNN 필터 개수 | 50 |
| 인코더 최대 길이 | 50 |
| GRU 최대 병렬화 횟수 | 10 |
| 드랍 아웃 | 0.2 |
| 학습률 | 0.001 |

본 논문은 실험을 위해 의존 관계 및 의미역 말뭉치인 UProbBank[12]를 사용한다. 말뭉치에는 의존 관계와 서술어, 의미역이 태깅(Tagging)된 문장이 약 14만개가 존재하며 이 중 너무 길거나(51어절 이상) 서술어가 너무 많이 존재(11개 이상)하는 문장을 제외한 132,384 문장을 8:2의 비율로 나누어 학습 및 평가 데이터로 사용한다. 말뭉치에 의존 관계는 태깅되어 있지만 의존 관계 레이블은 태깅되어 있지 않기 때문에 의존 구문 분석의 성능 척도로 UAS(Unlabeled Attachment Score)만 사용한

다. 서술어 및 논항 분류의 성능 척도는 재현율(Recall)과 정밀도(Precision)의 조화 평균인 F1 점수를 사용한다. 형태소 임베딩은 20GB의 뉴스기사 데이터로 학습시킨 GloVe[13]를 사용했다. 모델에 사용된 파라미터는 표 1과 같다.

### 4.2 실험 결과

표 2. 의존 구문 분석 성능 비교

| 모델 | UAS |
|---|---|
| DP only | 0.9285 |
| DP + SRL | **0.9327** |

표 3. 의미역 결정 성능 비교

| 모델 | PIC F1 | AIC F1 |
|---|---|---|
| SRL only | 0.9949 | **0.7312** |
| DP + SRL | **0.9952** | 0.7255 |

표 2는 통합 모델의 의존 구문 분석 성능과 의존 구문 분석 단일 모델의 성능 비교를 보여준다. 의존 구문 분석 단일 모델은 통합 모델에서 서술어 인식 및 분류 계층과 논항 분류 및 분류 계층을 제거한 모델 구조다. 표 2에서 통합 모델의 의존 구문 분석 성능이 단일 모델의 성능보다 0.42%p 높은 것을 확인 할 수 있다. 이는 의미역 결정과 의존 구문 분석 간의 상관관계로 인해 의미역 결정 계층의 학습 정보가 의존 구문 분석 계층의 학습에도 도움을 주기 때문이라 판단된다. 표 3은 통합 모델과 의미역 결정 단일 모델의 성능 비교를 보여준다. 의미역 결정 단일 모델은 통합 모델에서 의존 구문 분석 계층을 제거한 모델이다. 모델의 입력은 구문 분석 정보를 제외한 어절 임베딩만 주었다. 실험 결과 서술어 인식 및 분류 성능에서 0.03%p의 성능 향상이 있고 논항 인식 및 분류 성능은 0.6%p 떨어짐을 확인 할 수 있다. 논항 인식 및 분류의 성능이 하락한 이유는 완전한 구문 정보가 아니라 중심어 위치 정보만을 사용했기 때문이라고 판단된다. [3]은 자질로 구문 레이블 정보만을 사용해 논항 인식 및 분류의 성능 향상을 보였다. 표 3의 실험 결과를 보면 중심어 위치 정보는 서술어 인식 및 분류에는 도움이 됐지만, 논항 인식 및 분류에 있어서는 노이즈(Noise)로 작용됐다고 판단된다. 따라서 구문 레이블 정보가 포함된 데이터 셋을 학습에 사용하고 그 정보를 의미역 결정에 이용 할 수 있다면 논항 인식 및 분류의 성능 저하를 보완 할 수 있을 것으로 기대된다.

## 5. 결론

본 논문은 의존 구문 분석과 의미역 결정의 상관관계를 효과적으로 활용 할 수 있는 의존 구문 분석 및 의미역 결정 통합 모델을 제안하였다. 실험 결과 의존 구문 분석, 서술어 인식 및 분류는 각각 학습했을 때보다 멀티 태스크 방식으로 동시에 학습 했을 때 더 좋은 성능을 보였으나 서술어 인식 및 분류는 성능의 저하를 보였다. 성능의 저하를 보인 이유는 의존 관계 레이블 정보의 부재로 인한 것이라고 판단된다. 완전한 의존 구문 분석과 의미역 결정 정답이 함께 태깅되어 있는 데이터 셋을 구하거나 구축하는 것은 많은 노력을 필요로 하는 작업이다. 따라서 향후 연구로 의존 구문 분석 데이터 셋과 의미역 결정 데이터 셋을 각각 습득하고 의존 구문 분석 데이터로 구문 분석 계층을 먼저 학습 시키고 그 가중치를 전달받아 의미역 결정 계층을 학습하는 전이 학습(Transfer learning) 기반 통합 모델을 실험할 예정이다. 또한 기존 의미역 결정 연구에서 주로 사용되는 말뭉치인 Korean PropBank[14]를 실험에 사용하고 기존 연구들과 정량적 성능 비교를 수행할 예정이다.

## 감사의 글

이 논문은 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2016R1A2B4007732)

## 참고문헌

[1] M.Roth and M. Lapata, “Neural Semantic Role Labeling with Dependency Path Embeddings”, Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL), pp.1192-1202, 2016

[2] 박광현, 나승훈, “Layer Nomalized LSTM CRF를 이용한 한국어 의미역 결정”, 제 29회 한글 및 한국어 정보처리 학술대회 논문집 (HCLT 2017), pp.163-166, 2017

[3] 박광현, 나승훈, “문자 기반 LSTM CRF를 이용한 한국어 의미역 결정”, *2017년 한국컴퓨터종합학술대회 논문집 (KCC 2017)*, pp.1817-1819, 2017

[4] 박천음, 이창기, “포인터 네트워크를 이용한 한국어 의존 구문 분석”, *정보과학회 논문지*, 제44권, 제8호, pp.822-831, 2017.8

[5] O. Vinyals, M. Fortunato and N. Jaitly, “Pointer Networks”, *Neural Information Processing Systems (NIPS)*, pp.2674-2682, 2015

[6] 박성식, 오신혁, 김홍진, 김시형, 김학수, “ELMo와 멀티헤드 어텐션을 이용한 한국어 의존 구문 분석”, *제 30회 한글 및 한국어 정보처리 학술대회 논문집 (HCLT 2018)*, pp.8-12, 2018

[7] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser and I. Polosukhin, “Attention Is All You Need”, *Neural Information Processing Systems (NIPS)*, pp.5998-6008, 2017

[8] 배장성, 이창기, 임수종, “Backward LSTM CRF를 이용한 한국어 의미역 결정”, *제 27회 한글 및 한국*

*어 정보처리 학술대회 논문집 (HCLT 2015)*, pp.194-197, 2015

[9] 박성식, 김학수, “주의 집중 방법을 이용한 개체명 인식”, 2018년 한국컴퓨터종합학술대회 논문집 (KCC 2018), pp.678-680, 2018.6

[10] J. Chung, C. Gulcehre, K. Cho and Y. Bengio, “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling”, *arXiv:1412.3555*, 2014

[11] M. Bouaziz, M. Morchid, R. Dufour, G. Linares and R. De Mori, “Paralell Long Short-Term Memory for Multi-Stream Classification”, *arXiv:1702.03402v1*, 2017

[12] 김완수, 옥철영, “한국어 격틀사전 기반 의미역 반자동 부착 도구”, *제 26회 한글 및 한국어 정보처리 학술대회 논문집 (HCLT 2014)*, pp.251-254, 2014

[13] J. Pennington, R. Socher and C. D. Manning. “GloVe: Global Vectors for Word Representation”, *2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp.1532-1543, 2014

[14] Martha Palmer, Shijong Ryu, Jinyoung Choi, Sinwon Yoon and Yeongmi Jeon, Korean Propbank [Online]. Available: http://catalog.ldc.upenn.edu/LDC2006T03