

# 언어지식공학 연구실

Language Knowledge Engineering Lab.

<http://lake.cbnu.ac.kr>

사람의 언어를 이해하고 처리하는 컴퓨터 개발을 목표로 인공지능의 주요 분야인 자연언어 처리와 정보검색을 통합한 언어지식 공학을 연구하고 이와 관련된 Machine Learning (Deep Learning) 기법 등을 연구한다.



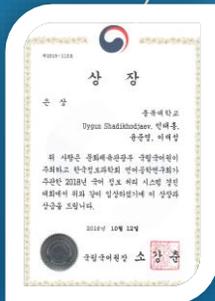
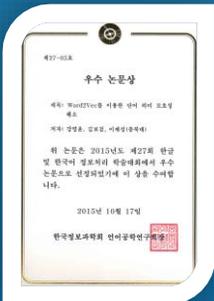
## 이재성 교수 (Prof. Jae Sung Lee)

- 1979 ~ 1983 서울대학교 컴퓨터공학과 학사
- 1983 ~ 1985 한국과학기술원 전산학과 석사
- 1995 ~ 1999 한국과학기술원 전산학과 박사
- 1985 ~ 1988 (주)큐닉스컴퓨터 과장
- 1988 ~ 1993 미국 및 한국 마이크로소프트 차장
- 1999 ~ 2000 한국전자통신연구원 팀장
- 2009 ~ 2010 교수학습지원센터장
- 2014 ~ 2016 전산정보원장
- 2017 ~ 2019 인공지능연구소장



표 1. 한국어 의존 구문 분석 결과

Team	Model	UAS	LAS
가은 최고 모델 [4]	Multi-layer Pointer networks	85.55	83.07
강원대 [5]	Multi-head attention and pointer networks	85.31	82.72
서강대 [6]	Deep bi-affine attention and stack pointer networks	85.24	83.02
충북대 [7]	Deep bi-affine attention	85.19	82.89
고려대 [8]	Stack pointer networks	84.60	82.02
NHN [9]	Deep bi-affine attention	84.43	80.55
A 시스템	Stack pointer networks	83.03	80.41
B 시스템	Arc-eager	69.22	-
C 시스템	Arc-eager	47.80	21.81



The SIGMORPHON 2019 Shared Task:

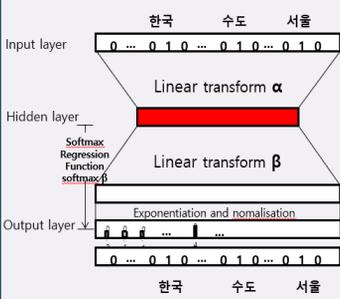
Team	Lemma Accuracy	Lemma Levenshtein	Morph Accuracy	Morph F1
CBNU[01]	94.07	0.13	88.09	91.84
CHARLES-MALTA-01	74.95	0.62	50.37	58.81
CHARLES-SAARLAND-02	95.00	0.11	93.23	96.82
CMU-02	92.20	0.17	85.06	88.97
CMU-DataAug-01	92.51	0.17	86.53	91.18
Edinburgh-01	94.20	0.13	88.93	92.89
ITU-01	94.46	0.11	86.67	90.54
NLP-CUBE-01	91.43	2.43	84.92	88.67
OHIOSTATE-01	93.43	0.17	87.42	92.51
RUG-01	93.91	0.14	90.53	94.54
RUG-02	93.06	0.15	88.80	93.22
USAL-PRAQUE-01	95.78	0.10	93.19	95.92
UNTHLTING-02	83.14	0.55	15.69	51.87
EDINBURGH-02*	97.35	0.06	93.02	95.94
CMU-Monolingual*	88.31	0.27	84.60	91.18
CMU-PolyGlot-01*	76.81	0.54	60.98	75.42
Baseline	94.17	0.13	73.16	87*

# 주요 연구 내용

## 언어의 벡터 표현방법 연구

신경망 학습 (Deep Learning)  
(단어 벡터 생성)

단어 벡터 연산



왕 - 남자 + 여자 = [ ]  
한국 - 서울 + 도쿄 = [ ]  
가다 - 갔다 + 왔다 = [ ]  
인생 - 사랑 + 미움 = ?

한국어의 특성을 반영한 단어 벡터 생성 방법 연구

## 질의응답 시스템

Q 서울은 청주의 어느 쪽에 있나요?

⇒ Q 방향<서울, 청주, ?>

Knowledge Base Source:

천안은 북쪽에 있는 서울에서 많은 인구가 유입되고 있다.  
청주에서 천안을 가려면 경부 고속도로를 타고 북쪽으로 약 40분 정도 걸린다.

Knowledge Base:

방향<서울, 천안, 북쪽>  
방향<천안, 청주, 북쪽>  
=>(추론)  
방향 <서울, 청주, 북쪽>

A. ? = 북쪽



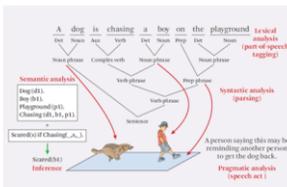
## 형태소 및 구문 분석

문장을 형태소 단위로 구분하고, 언어 구조를 파악.

언어를 분석하고 문법 규칙을 사용하여 구성 요소를 식별하고 의미를 학습.



어근, 접두사/접미사, 품사(POS: Part-of-Speech) 등을 구분.



구성 요소 간의 위계 관계를 분석하여 문장의 구조 결정.

## 기계독해 시스템

"질문-지문" 쌍을 입력으로 받아 지문에서 정답의 문구를 찾는 연구.

질문 : 1989년 6월 30일 평양축전에 대표로 파견된 인물은?

지문 : ... 1989년 3월 12일 서울지방검찰청 공안부는 임중석의 사전구속영장을 발부받았다. 같은 해 6월 30일 평양축전에 임수경을 대표로 파견하여 국가보안법위반 혐의가 추가되었다. 경찰은 12월 18일~20일 사이 서울 경희대학교에서 임중석이 성명 발표를 추진하고 있다는 첩보를 입수했다. ...

↓ 지문에서 정답을 예측하여 추출

정답 : 임수경

# 언어지식공학 연구실

Language Knowledge Engineering Lab.

<http://lake.cbnu.ac.kr>

교수 연구실 : S4-1 324호 (043-261-2252), 학생 연구실 : S4-1 302호