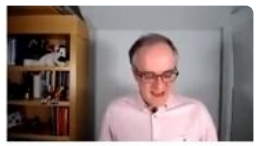
Natural Language Processing with Deep Learning

CS224N/Ling284

Christopher Manning

Lecture 4: Dependency Parsing

Stanford



Lecture Plan

Syntactic Structure and Dependency parsing

1. Syntactic Structure: Consistency and Dependency (25 mins)
2. Dependency Grammar and Treebanks (15 mins)
3. Transition-based dependency parsing (15 mins)
4. Neural dependency parsing (20 mins)

Reminders/comments:

In Assignment 3, out on Tuesday, you build a neural dependency parser using PyTorch

Start installing and learning PyTorch (Ass 3 has scaffolding)

Come to the PyTorch tutorial, Friday 10am (under the Zoom tab, not a Webinar)

Final project discussions – **come meet with us**; focus of Thursday class in week 4

Stanford

1. Two views of linguistic structure: Constituency = phrase structure grammar = context-free grammars (CFGs)

Phrase structure organizes words into nested constituents

Starting unit: words

the, cat, cuddly, by, door

Words combine into phrases

the cuddly cat, by the door

Phrases can combine into bigger phrases

the cuddly cat by the door

Stanford

3

사람들이 문장의 구조에 대해 어떻게 생각하는지, 어떻게 사람 언어에서 의미를 전달하는지 설명하기 위해 구조를 위에 얹는지

words : 사람 언어의 의미에서 가장 중요한 부분

그만큼 문장의 구조도 중요함

단어로 문장이 어떻게 구성되는지 생각할 때, 단어들이 문장에서 어떤 부분을 차지하는지에 따라 이름을 붙임

1. Two views of linguistic structure: Constituency = phrase structure grammar = context-free grammars (CFGs)

Phrase structure organizes words into nested constituents

Starting unit: words

the, cat, cuddly, by, door

Det N Adj P N

a
the
this
that
every

Words combine into phrases

the cuddly cat, by the door

Phrases can combine into bigger phrases

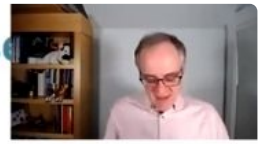
the cuddly cat by the door

Stanford

3

명사, 전치사, 형용사, 한정사(a, the, this, that, every) 등

context-free grammars (CFGs)



1. Two views of linguistic structure: Constituency = phrase structure grammar = context-free grammars (CFGs)

Phrase structure organizes words into nested constituents

Starting unit: words

the, cat, cuddly, by, door

Det N Adj P N

a, the, this, that, every

Words combine into phrases

NP [the cuddly cat,] [by [the door]$_{NP}$]$_{PP}$

Prep

Phrases can combine into bigger phrases

the cuddly cat by the door

NP NP PP NP

Stanford

3

조금 더 큰 단위인 phrase로 나눌 수 있음

noun phrase, preposition phrase 등, 작은 phrase들을 더 큰 phrase로 묶을 수도 있음

non-terminal symbol : 얼마든지 rule을 선택하고 위치를 바꿀 수 있는 요소

noun phrase, preposition phrase는 context free grammar에서 non-terminals

Two views of linguistic structure: Constituency = phrase structure grammar = context-free grammars (CFGs)

Grammar

Phrase structure organizes words into nested constituents.

the	cat
a	dog
large	in a crate
barking	on the table
cuddly	by the door
large	barking

NP → Det (Adj)* N (PP)

PP → P NP

VP → V PP

S → NP VP

Lexicon

N → dog

N → cat

Det → a

Det → the

P → in, on, by

V → talk, walked

V P

talk to

walked behind

the cat walked behind the dog

Stanford

6

<Grammar>

NP -> Det (Adj)* N (PP)

** (___) : 선택사항, (___)* : 선택사항이면서 0번이상 출현

PP -> P NP

VP -> V PP

S -> NP VP

<Lexicon>

N -> dog, cat

Det -> a, the

P -> in, on, by

V -> talk, walked

Two views of linguistic structure: Dependency structure

- Dependency structure shows which words depend on (modify, attach to, or are arguments of) which other words.

Look in the large crate in the kitchen by the door

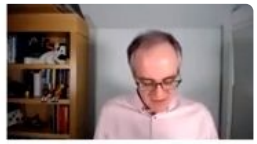
Stanford

7

dependency grammar

문장이 있을 때, 각 단어에 대해 어떤 단어가 그것을 수식하는지 설명하는 것이 목표

Two views of linguistic structure: Dependency structure



- Dependency structure shows which words depend on (modify, attach to, or are arguments of) which other words.

Stanford

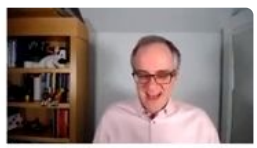
7

example)

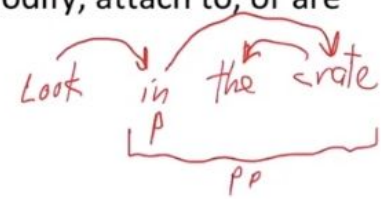
large가 crate 수식, the가 crate 수식, the가 kitchen 수식, the가 door 수식 -> 이것을 단어 위에서 화살표로 표기

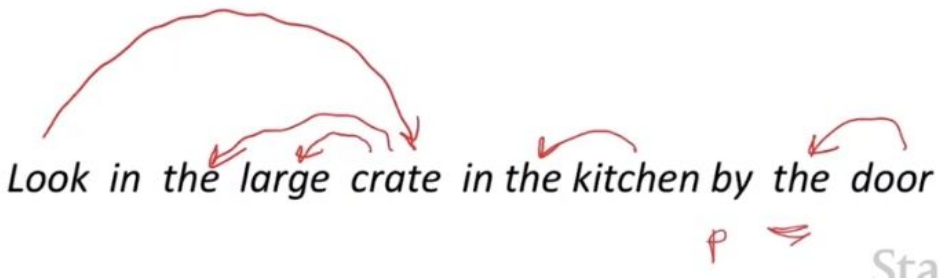
in the large crate가 look을 수식

Two views of linguistic structure: Dependency structure



- Dependency structure shows which words depend on (modify, attach to, or are arguments of) which other words.





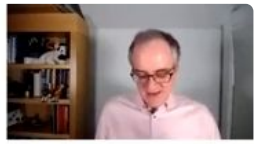
Stanford

7

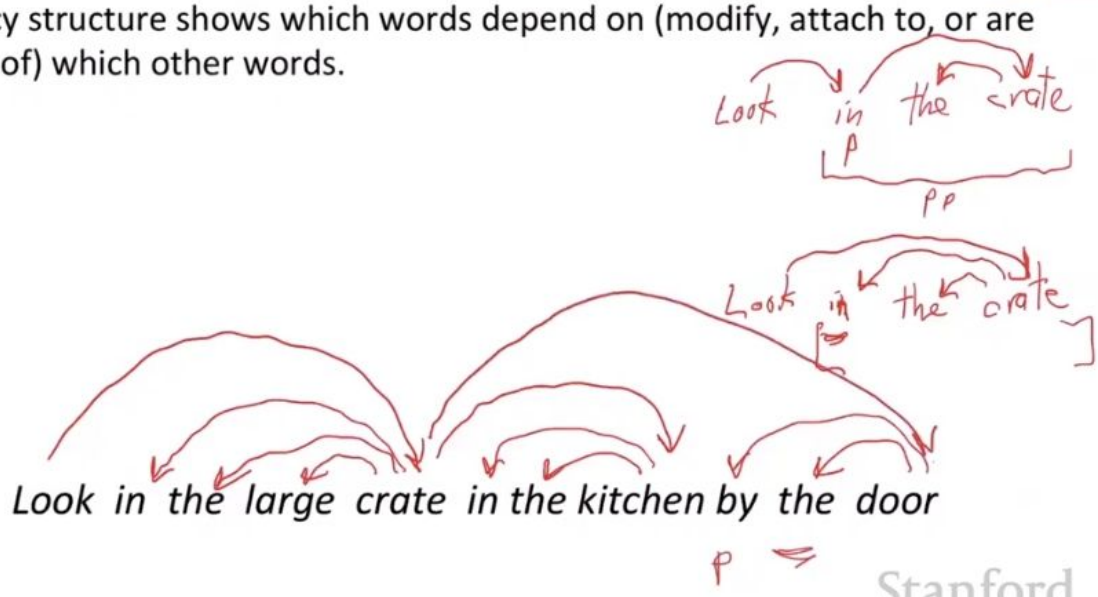
the는 crate에 의존, 'look in' 'in the crate' 각각 의존관계

전치사는 prepositional phrase의 맨 앞에서 이끔

Two views of linguistic structure: Dependency structure



- Dependency structure shows which words depend on (modify, attach to, or are arguments of) which other words.



Stanford

7

영어를 기본으로 하는 다른 외국어들은 이런 전치사 역할을 하기 위한 다양한 새로운 마크들이 있음

in the kitchen은 crate를 수식, by the door은 crate를 수식

Why do we need sentence structure?

Humans communicate complex ideas by composing words together into bigger units to convey complex meanings

Listeners need to work out what modifies [attaches to] what

A model needs to understand sentence structure in order to be able to interpret language correctly

Stanford

9

Prepositional phrase attachment ambiguity

San Jose cops kill man with knife

Close

Text Paper Translate Listen

San Jose cops kill man with knife

Ex-college football player, 23, shot 9 times allegedly charged police at fiancee's home

By Hamed Aleaziz and Vivian Ho

A man fatally shot by San Jose police officers while allegedly charging at them with a knife was a 23-year-old former football player at De Anza College in Cupertino who was distraught and depressed, his family said Thursday.

Police officials said two officers opened fire Wednesday afternoon on Phillip Watkins outside his fiancee's home because they feared for their lives. The officers had been drawn to the home, officials said, by a 911 call reporting an armed home invasion that, it turned out, had been made by Watkins himself.

But the mother of Watkins' fiancee, who also lives in the home on the 1300 block of Sherman Street, said she witnessed the shooting and described it as excessive. Faye Buchanan said the confrontation happened shortly after she called a suicide intervention hotline in hopes of getting Watkins medical help.

Watkins' 911 call came in at 5:01 p.m., said Sgt. Heather Randol, a San Jose police spokeswoman. "The caller stated there was a male breaking into his home armed with a knife," Randol said. "The caller also stated he was locked in an upstairs bedroom with his children and requested help from police."

She said Watkins was on the sidewalk in front of the home when two officers got there. He was holding a knife with a 4-inch blade and ran toward the officers in a threatening manner, Randol said.

"Both officers ordered the suspect to stop and drop the knife," Randol said. "The suspect continued to charge the officers with the knife in his hand. Both officers, fearing for their safety and defense of their life, fired at the suspect."

On the police radio, one officer said, "We have a male with a knife. He's walking toward us."

"Shots fired! Shots fired!" an officer said moments later.

A short time later, an officer reported, "Male is down. Knife's still in hand."

Buchanan said she had been prompted to call the

Shoot continues on D8

Back Continue

Stanford

자연어에서 찾을 수 있는 몇 가지 모호성 예시

다양한 언어들에서 모호성을 찾을 수 있는데, 구문구조가 부분적으로 언어의 세부사항에 따라 달라지는 모호성은 무엇일까?

다른 언어들은 다른 구문구조를 가지고 있고, 다른 단어 배열을 가지고 있고, 다른 형식을 가진 단어들(case markings, 격표시)을 각자 다른 양만큼 가지고 있음 -> 이렇게 세부사항이 달라짐에 따라 각자 다른 모호성을 가지고 있음

위 슬라이드의 예시는 영어에서 가장 흔한 모호성

"San Jose cops kill man with knife"

1) San Jose cops가 칼로 한 남자를 죽였음

2) San Jose cops가 칼을 든 남자를 죽였음

Prepositional phrase attachment ambiguity

San Jose cops kill man with knife

Close

Text Paper Translate Listen

San Jose cops kill man with knife

Ex-college football player, 23, shot 9 times allegedly charged police at fiancee's home

By Hamed Aleaziz and Vivian Ho

A man fatally shot by San Jose police officers while allegedly charging at them with a knife was a 23-year-old former football player at De Anza College in Cupertino who was distraught and depressed, his family said Thursday.

Police officials said two officers opened fire Wednesday afternoon on Phillip Watkins outside his fiancee's home because they feared for their lives. The officers had been drawn to the home, officials said, by a 911 call reporting an armed home invasion that, it turned out, had been made by Watkins himself.

But the mother of Watkins' fiancee, who also lives in the home on the 1300 block of Sherman Street, said she witnessed the shooting and described it as excessive. Faye Buchanan said the confrontation happened shortly after she called a suicide intervention hotline in hopes of getting Watkins medical help.

Watkins' 911 call came in at 5:01 p.m., said Sgt. Heather Randol, a San Jose police spokeswoman. "The caller stated there was a male breaking into his home armed with a knife," Randol said. "The caller also stated he was locked in an upstairs bedroom with his children and requested help from police."

She said Watkins was on the sidewalk in front of the home when two officers got there. He was holding a knife with a 4-inch blade and ran toward the officers in a threatening manner, Randol said.

"Both officers ordered the suspect to stop and drop the knife," Randol said. "The suspect continued to charge the officers with the knife in his hand. Both officers, fearing for their safety and defense of their life, fired at the suspect."

On the police radio, one officer said, "We have a male with a knife. He's walking toward us."

"Shots fired! Shots fired!" an officer said moments later.

A short time later, an officer reported, "Male is down. Knife's still in hand."

Buchanan said she had been prompted to call the

Shoot continues on D8

Back Continue

Stanford

10

1)

San Jose cops가 kill의 주어(subject)

man이 kill의 목적어(object)

knife는 kill의 도구(oblique modifier)

Prepositional phrase attachment ambiguity

Stanford

10

2)

a man with a knife & San Jose cops

San Jose cops가 kill의 주어(subject)

man이 kill의 목적어(object)

knife가 man의 명사수식어(noun modifer)

Prepositional phrase attachment ambiguity

Stanford

10

이렇게 전치사구가 문장 맨 뒤에 오면, 그것이 직전에 나오는 명사구를 수식하는 것인지, 그 전에 오는 동사를 수식하는 것인지 애매함

영어에서는 전치사구가 모든 문장에서 모호하게 붙기 때문에 이러한 모호성을 매번 마주하지만, 사람들은 잘 알아채지 못함

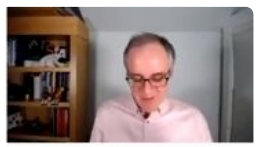
사람의 뇌는 가능한 해석들을 모두 고려하고, 문맥에 적절한 해석을 골라내는 것을 잘 하기 때문

자연어에서 다른 언어는 다른 모호성을 가짐

위의 예시로 든 영어에서의 전치사구 모호성은 중국어에서는 안 나타남

중국어에서 동사를 수식하는 전치사구는 동사 앞에 위치하기 때문

그러나 중국어에서는 또 다른 모호성이 있을 것임



Prepositional phrase attachment ambiguity

Stanford

10

또 다른 예시

"Scientists count whales from space"

Prepositional phrase attachment ambiguity

Scientists count whales from space

✓

Scientists count whales from space

✗

1) 과학자들이 우주에서 고래를 관찰함 (from space가 count를 수식)

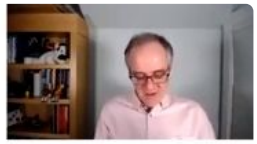
2) 과학자들이 우주의 고래를 관찰함 (from space가 whales를 수식)

PP attachment ambiguities multiply

- A key parsing decision is how we 'attach' various constituents
  - PPs, adverbial or participial phrases, infinitives, coordinations,

The board approved [its acquisition] [by Royal Trustco Ltd.]
[of Toronto]
[for $27 a share]
[at its monthly meeting].

PP attachment ambiguities multiply



- A key parsing decision is how we 'attach' various constituents
  - PPs, adverbial or participial phrases, infinitives, coordinations,

The board approved [its acquisition] [by Royal Trustco Ltd.]
[of Toronto]
[for $27 a share]
[at its monthly meeting].

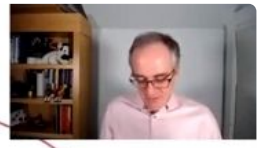
Stanford

PP attachment ambiguities multiply

- A key parsing decision is how we 'attach' various constituents
  - PPs, adverbial or participial phrases, infinitives, coordinations,

The board approved [its acquisition] [by Royal Trustco Ltd.]
[of Toronto]
[for $27 a share]
[at its monthly meeting].

Stanford



PP attachment ambiguities multiply

- A key parsing decision is how we 'attach' various constituents
  - PPs, adverbial or participial phrases, infinitives, coordinations,

The board approved [its acquisition] [by Royal Trustco Ltd.]
[of Toronto]
[for $27 a share]
[at its monthly meeting].

- Catalan numbers: $C_n = (2n)!/[(n+1)!n!]$
- An exponentially growing series, which arises in many tree-like contexts:
  - E.g., the number of possible triangulations of a polygon with $n+2$ sides
    - Turns up in triangulation of probabilistic graphical models (CS228)....

Stanford

12

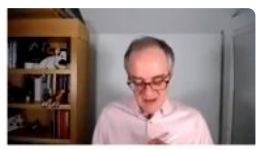
문장의 끝에 k개의 전치사구가 있으면, 문장은 Catalan numbers 만큼 parse 할 수 있음

Catalan numbers는 기하급수적으로 늘어나는 급수

이러한 prepositional phrase attachment ambiguity와 같은 구조에 대해 알아둘 것은, 문장의 구조상 이러한 모호성을 해결할 수 있는 것은 아무것도 없다는 것임

프로그래밍 언어를 다루는 컴파일러에서는 문법이 모호성을 가지지 않도록 만들어짐, 만약 모호성이 생긴다면 그것을 특정한 방법으로 parse 하도록 하는 기본 규칙이 있음

그러나 사람 언어는 그렇지 않음



Coordination scope ambiguity

[Shuttle veteran] and [longtime NASA executive Fred Gregory] appointed to board

2 people

[Shuttle veteran and longtime NASA executive Fred Gregory] appointed to board

1 person

Stanford

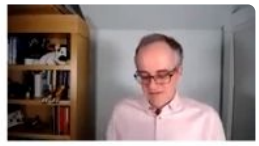
14

"Shuttle veteran and longtime NASA executive Fred Gregory appointed to board"

1) 'Shuttle veteran', 'longtime NASA executive Fred Gregory' -> 2 people

2) 'Shuttle veteran and longtime NASA executive' = Fred Gregory -> 1 person

Coordination scope ambiguity



THE NEWS-GAZETTE

NATION / WO

PRESIDENT'S FIRST PHYSICAL

Doctor: No heart, cognitive issues

But Trump needs to reduce his cholesterol, lose weight

By JILL COLVIN

6-foot-3 president weighed in at 239 pounds — three pounds heavier than he was in September 2016, the last time Trump revealed his weight to the public.

Trump's blood pressure was 122 over 74, and his

with no medical issues." Trump has no heart disease and no family history of it.

The 71-year-old president performed "exceedingly well" on cognitive screening, which is not

White H reporter

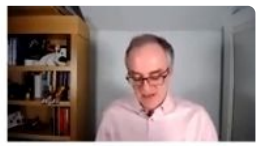
15

기사제목이 짧게 함축해야 해서 모호한 문장이 많음

1) 'No heart' and 'cognitive issues'

2) 'heart' or 'cognitive' issues

Adjectival/Adverbial Modifier Ambiguity



numbers, including some that featured a bucket and bells brigade of performers who beat rhythms on buckets and trash cans with drums sticks and hammer mallets. PHOTO BY JENNIFER STULTZ

MENTORING DAY

Students get first hand job experience

By Gale Rose
grose@pratttribune.com

Eager students invaded businesses all over Pratt Tuesday, October 24 as they looked for future job opportunities on Disability Mentoring Day.

The 97 students from 12 schools fanned out across Pratt and got first hand experience what it would be like to work at those 40 businesses. They asked questions and got some hands on experience with various operations.

Paola Luna of Pratt High School, Gina Patton of Kingman High School and America Fernandez of St. John chose the Main Street Small Animal Veterinarian Clinic for their business. Students got a tour of the facility, learned what happens in an examination, got to handle various animals and watched a snake eat a mouse.

Luna said she was interested in animal health and wanted to know more about caring for hurt animals. Patton likes all kinds of animals and said she learned a lot from the experience. Watching the snake eat the mouse impressed her the most.

Fernandez wants to become a veterinarian and enjoyed learning everything that veterinarians

SEE MENTORING, 6

ng Meyer
ty Commissioner

- Hospital Pharmacist for 41 years
- 4 years Commissioner for Pratt Planning and Zoning Board of Appeals
- 3 years Pratt City Commission
- Graduate of Pratt High School and KU School of Pharmacy
- Past Member and President of Civic Groups and Organizations
- Experience and Knowledge of Financial Responsibility and Budgeting
- Supports Family Values, Education, and Business Growth
- Common Sense Approach for the Sustained Progress of Pratt

16

또 다른 모호성 예시 (adjectival, adverbial)

1)

first가 hand 수식

job은 experience 수식하는 복합명사

first hand가 experience 수식

first hand job experience가 get의 목적어

students가 get의 주어

Adjectival/Adverbial Modifier Ambiguity

numbers, including some that featured a bucket and bells brigade of performers who beat rhythms on buckets and trash cans with drums sticks and hammer mallets. PHOTO BY JENNIFER STULTZ

MENTORING DAY

Students get first hand job experience

By Gale Rose
grose@pratttribune.com

Eager students invaded businesses all over Pratt Tuesday, October 24 as they looked for future job opportunities on Disability Mentoring Day.

The 97 students from 12 schools fanned out across Pratt and got first hand experience what it would be like to work at those 10 businesses. They asked questions and got some hands on experience with various operations.

Paola Luna of Pratt High School, Gina Patton of Kingman High School and America Fernandez of St. John chose the Main Street Small Animal Veterinarian Clinic for their business. Students got a tour of the facility, learned what happens in an examination, got to handle various animals and watched a snake eat a mouse.

Luna said she was interested in animal health and wanted to know more about caring for hurt animals. Patton likes all kinds of animals and said she learned a lot from the experience. Watching the snake eat the mouse impressed her the most.

Fernandez wants to become a veterinarian and enjoyed learning everything that veterinarians

SEE MENTORING, 6

ug Meyer
ty Commissioner

- Hospital Pharmacist for 41 years
- 4 years Commissioner for Pratt Planning and Zoning Board of Appeals
- 3 years Pratt City Commission
- Graduate of Pratt High School and KU School of Pharmacy
- Past Member and President of Civic Groups and Organizations
- Experience and Knowledge of Financial Responsibility and Budgeting
- Supports Family Values, Education, and Business Growth
- Common Sense Approach for the Sustained Progress of Pratt

Stanford

16

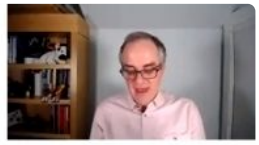
2)

hand가 job 수식

first가 experience 수식

job은 experience 수식

Verb Phrase (VP) attachment ambiguity



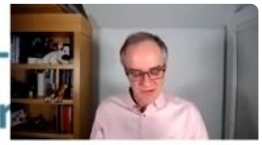
17

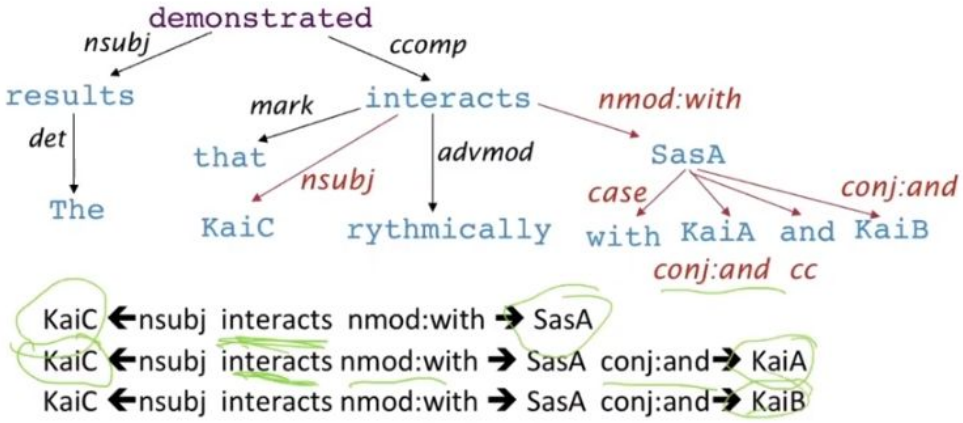
Verb Phrase (VP) attachment ambiguity

17

'to be used for Olympics beach volleyball'이 'Rio beach'를 수식할 수도, 'Mutilated body'를 수식할 수도 있음

Dependency paths help extract semantic interpretation – simple practical example: extracting protein-protein inter





[Erkan et al. EMNLP 07, Fundel et al. 2007, etc.]

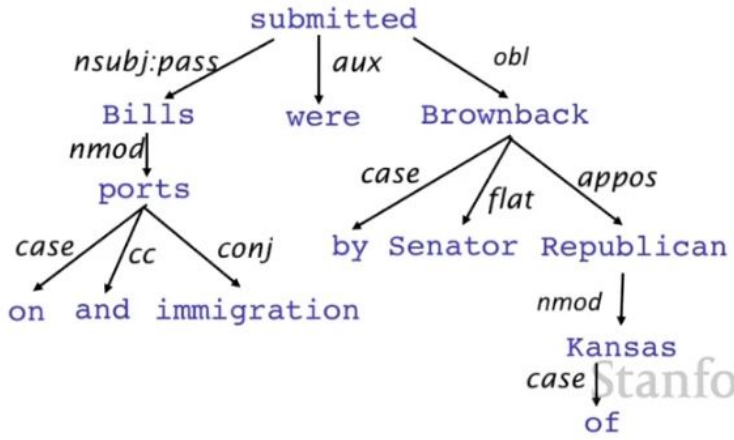
Stanford

18

Dependency Grammar and Dependency Structure

Dependency syntax postulates that syntactic structure consists of relations between lexical items, normally binary asymmetric relations ("arrows") called dependencies

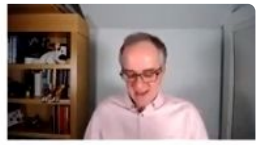
The arrows are commonly typed with the name of grammatical relations (subject, prepositional object, apposition, etc.)



Stanford

20

화살표는 문법적 관계를 칭함

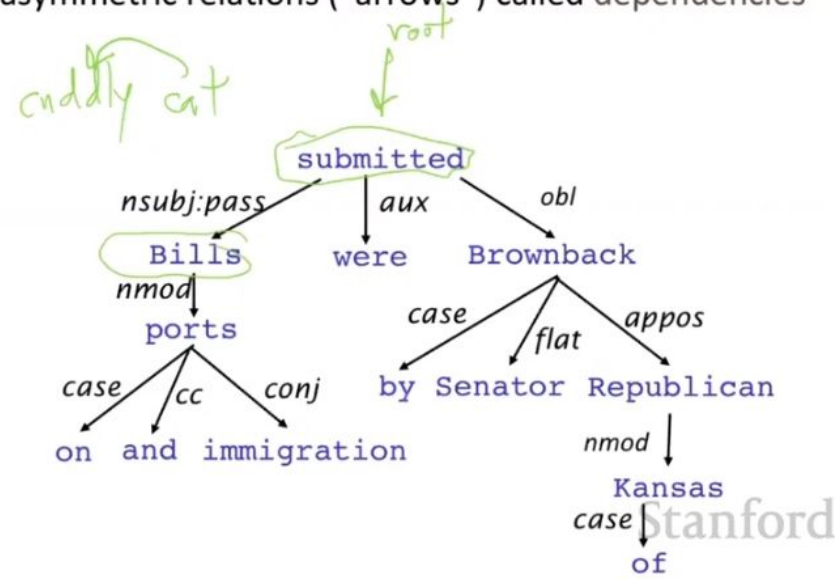


Dependency Grammar and Dependency Structure

Dependency syntax postulates that syntactic structure consists of relations between lexical items, normally binary asymmetric relations ("arrows") called dependencies

An arrow connects a head (governor, superior, regent) with a dependent (modifier, inferior, subordinate)

Usually, dependencies form a tree (a connected, acyclic, single-root graph)



화살표는 head와 dependent를 연결함

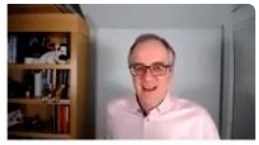
이 그래프는 single-root tree로 사이클 없음

Pāṇini's grammar (c. 5th century BCE)

Gallery: http://wellcomeimages.org/indexplus/image/L0032691.html
CC BY 4.0 File:Birch bark MS from Kashmir of the Rupavatra Wellcome L0032691.jpg

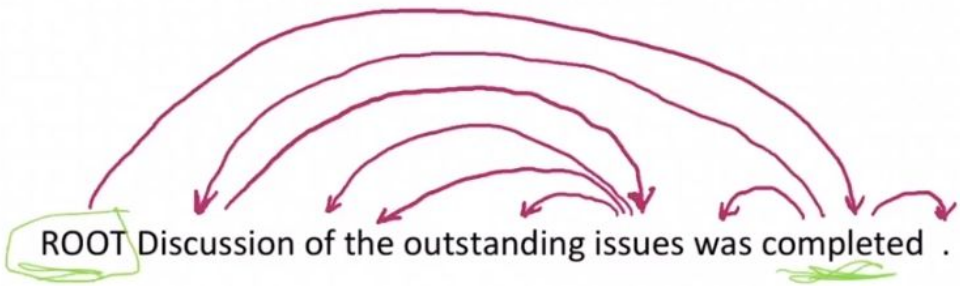
dependency grammar는 오랜 역사를 가지고 있음

Dependency Grammar/Parsing History



- The idea of dependency structure goes back a long way
  - To Pāṇini's grammar (c. 5th century BCE)
  - Basic approach of 1st millennium Arabic grammarians
- Constituency/context-free grammar is a new-fangled invention
  - 20th century invention (R.S. Wells, 1947; then Chomsky 1953, etc.)
- Modern dependency work is often sourced to Lucien Tesnière (1959)
  - Was dominant approach in "East" in 20th Century (Russia, China, …)
    - Good for free-er word order, inflected languages like Russian (or Latin!)
- Used in some of the earliest parsers in NLP, even in the US:
  - David Hays, one of the founders of U.S. computational linguistics, built early (first?) dependency parser (Hays 1962) and published on dependency grammar in *Language*

Dependency Grammar and Dependency Structure



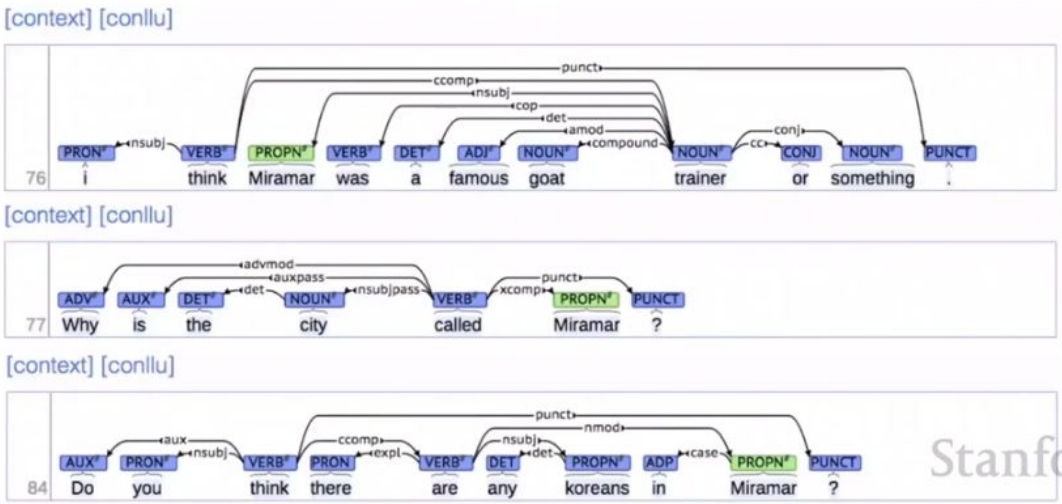
- Some people draw the arrows one way; some the other way!
  - Tesnière had them point from head to dependent – we follow that convention
- We usually add a fake ROOT so every word is a dependent of precisely 1 other node

화살표 방향 둘다 해도 되는데 from the head to the dependent 이걸로 많이 씀

맨 앞에 fake ROOT를 놓고 시작하면 편함, 모든 단어가 completed에서 시작되는 것이 보임

The rise of annotated data & Universal Dependencies tree

Brown corpus (1967; PoS tagged 1979); Lancaster-IBM Treebank (starting late 1980s);
Marcus et al. 1993, The Penn Treebank, *Computational Linguistics;*
Universal Dependencies: http://universaldependencies.org/



Stanford

25

데이터를 어디서 얻어오는지

treebank : 손으로 parse, dependency 구조 직접 분석한 것 모음 (80년대 후반-90년대)

그러나 treebank에 관심 없음, 문법을 적용할 수 있는 효율적인 parser를 구축하는 것이 더 관심, 문법을 직접 정리

머신러닝 기술 들어오면서 다시 treebank 중요도 높아짐

The rise of annotated data

Starting off, building a treebank seems a lot slower and less useful than writing a grammar (by hand)

But a treebank gives us many things

- Reusability of the labor
  - Many parsers, part-of-speech taggers, etc. can be built on it
  - Valuable resource for linguistics
- Broad coverage, not just a few intuitions
- Frequencies and distributional information
- A way to evaluate NLP systems

Stanford

26

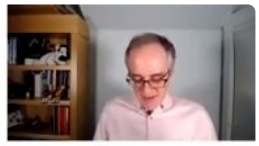
treebank는 재사용성이 높음

커버리지가 넓음

통계량을 얻을 수 있음

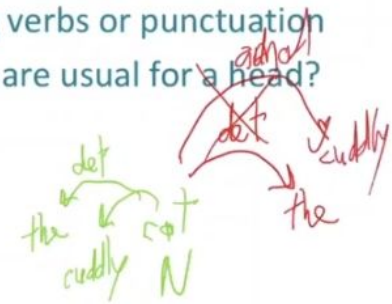
NLP 시스템을 평가하는 방법임

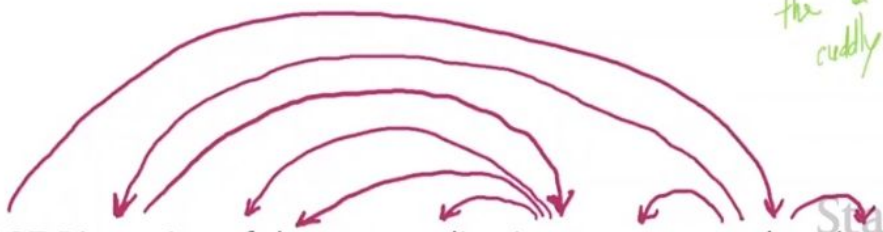
Dependency Conditioning Preferences



What are the sources of information for dependency parsing?

1. Bilexical affinities — The dependency [discussion → issues] is plausible
2. Dependency distance — Most dependencies are between nearby words
3. Intervening material — Dependencies rarely span intervening verbs or punctuation
4. Valency of heads — How many dependents on which side are usual for a head?





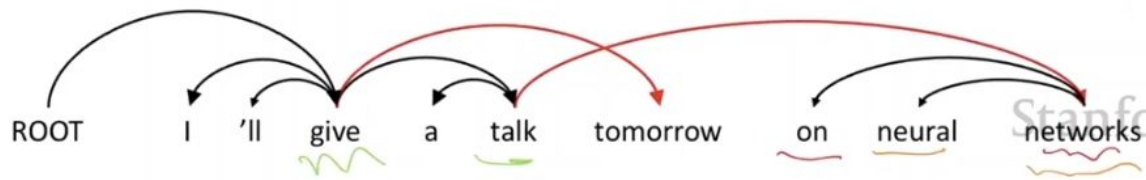
Stanford

27

Dependency Parsing

- A sentence is parsed by choosing for each word what other word (including ROOT) it is a dependent of

- Usually some constraints:
  - Only one word is a dependent of ROOT
  - Don't want cycles A → B, B → A
- This makes the dependencies a tree
- Final issue is whether arrows can cross (be non-projective) or not



Stanford

28

non-projective sentence ???

I'll give a talk on neural networks tomorrow

Projectivity

- Definition of a projective parse: There are no crossing dependency arcs when the words are laid out in their linear order, with all arcs above the words
- Dependencies corresponding to a CFG tree must be projective
  - I.e., by forming dependencies by taking 1 child of each category as head
- Most syntactic structure is projective like this, but dependency theory normally does allow non-projective structures to account for displaced constituents
  - You can't easily get the semantics of certain constructions right without these nonprojective dependencies

???

dependency parsers 구축하는 방법

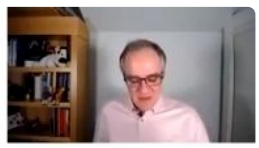
1) Dynamic Programming -> O(n^3) : 작은 부분으로 쪼개서 문제해결, 상향식

2) Graph Algorithms

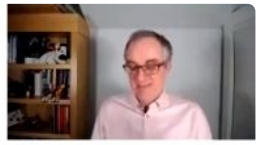
3) Constraint Satisfaction

4) Transition-based parsing (deterministic dependency parsing) : shift reduce parsing

Greedy transition-based parsing [Nivre 2003]



- A simple form of greedy discriminative dependency parser
- The parser does a sequence of bottom-up actions
  - Roughly like "shift" or "reduce" in a shift-reduce parser, but the "reduce" actions are specialized to create dependencies with head on left or right
- The parser has:
  - a stack σ, written with top to the right
    - which starts with the ROOT symbol
  - a buffer β, written with top to the left
    - which starts with the input sentence
  - a set of dependency arcs A
    - which starts off empty
  - a set of actions



Basic transition-based dependency parser

Start: $\sigma = [ROOT], \beta = w_1, \ldots, w_n, A = \emptyset$

1. Shift $\sigma, w_i|\beta, A \rightarrow \sigma|w_i, \beta, A$
2. Left-Arc$_r$ $\sigma|w_i|w_j, \beta, A \rightarrow \sigma|w_j, \beta, A \cup \{r(w_j, w_i)\}$
3. Right-Arc$_r$ $\sigma|w_i|w_j, \beta, A \rightarrow \sigma|w_i, \beta, A \cup \{r(w_i, w_j)\}$

Finish: $\sigma = [w], \beta = \emptyset$

relation

Stanford

32

1) 초기화

stack (sigma) : 'ROOT', buffer (beta) : parse하려는 문장의 단어들, dependency arcs (A)

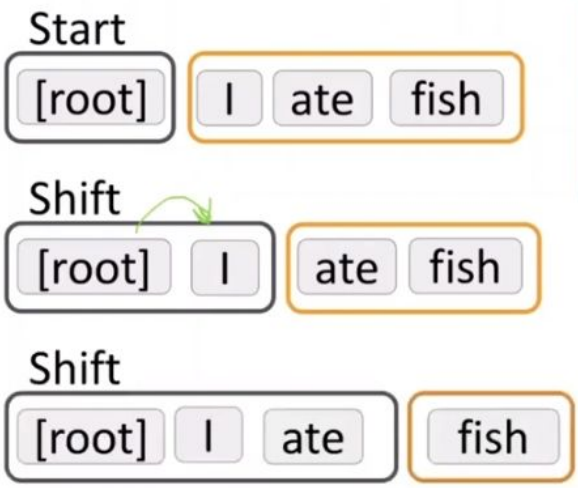
2) shift : 다음 단어를 스택에 넣기

3) 스택에서 단어 두 개 가져와서 의존(수식)관계 파악하기(Left or Right Arc), A에 두 단어의 문법적 관계와 함께 (j->i or i->j) 라는 화살표 추가하고 수식받은 단어는 스택에서 사라짐

Arc-standard transition-based parser

(there are other transition schemes ...)

Analysis of "I ate fish"



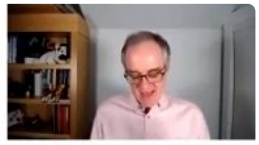
Start: $\sigma = [ROOT], \beta = w_1, \ldots, w_n, A = \emptyset$
1. Shift $\sigma, w_i|\beta, A \rightarrow \sigma|w_i, \beta, A$
2. Left-Arc$_r$ $\sigma|w_i|w_j, \beta, A \rightarrow \sigma|w_j, \beta, A \cup \{r(w_j, w_i)\}$
3. Right-Arc$_r$ $\sigma|w_i|w_j, \beta, A \rightarrow \sigma|w_i, \beta, A \cup \{r(w_i, w_j)\}$

Finish: $\sigma = [w], \beta = \emptyset$

Stanford

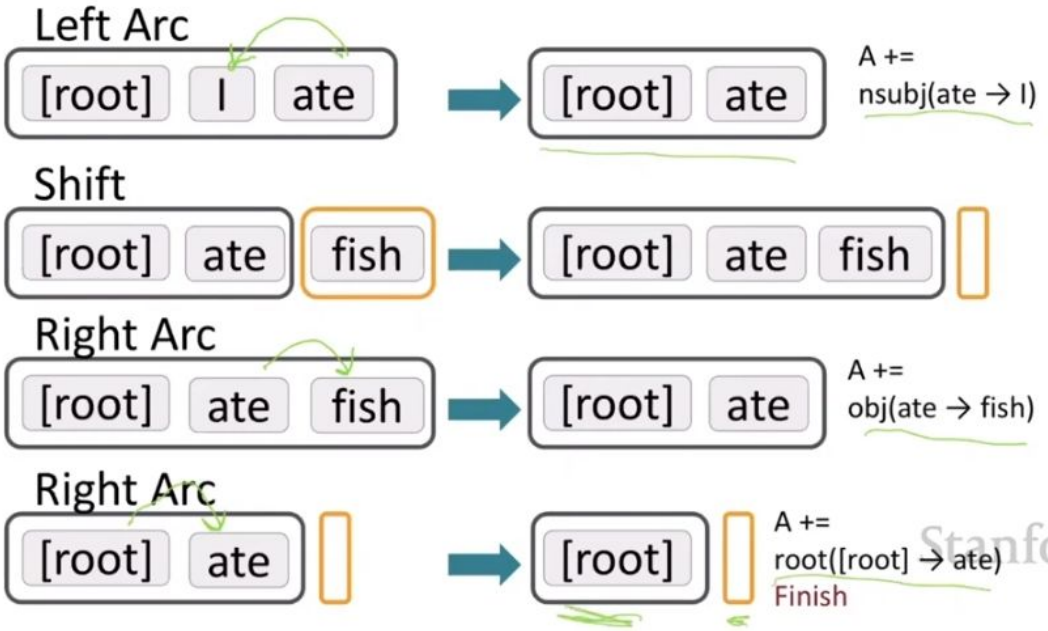
33

I 가 head 아님 -> shift 한 번 더



Arc-standard transition-based parser

Analysis of "I ate fish"



ate가 I 수식 -> A에 nsubj(ate->I) 추가하고 스택에서 I 제거

ate가 fish 수식 -> A에 obj(ate->fish) 추가하고 스택에서 fish 제거

ate가 전체 문장의 head -> A에 root(root->ate) 추가하고 스택에서 ate 제거

스택에 root만 남고 buffer가 비었으므로 종료

MaltParser [Nivre and Hall 2005]

- We have left to explain how we choose the next action 🤷
  - Answer: Stand back, I know machine learning!
- Each action is predicted by a discriminative classifier (e.g., softmax classifier) over each legal move
  - Max of 3 untyped choices; max of $|R| \times 2 + 1$ when typed
  - Features: top of stack word, POS; first in buffer word, POS; etc.
- There is NO search (in the simplest form)
  - But you can profitably do a beam search if you wish (slower but better): You keep *k* good parse prefixes at each time step
- The model's accuracy is *fractionally* below the state of the art in dependency parsing, but
- It provides **very fast linear time parsing**, with high accuracy – great for parsing the web

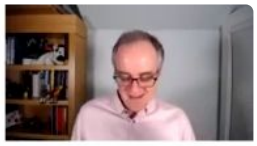
다음 액션이 뭔지 (shift, left or right arc) 어떻게 결정?

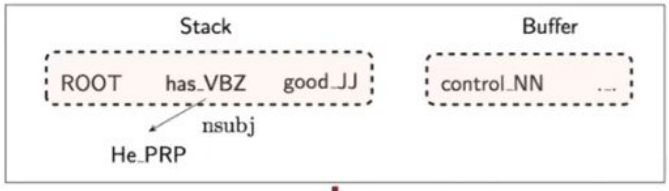
naively explore(아무거나 해보기)

머신러닝 classifier 통해서 결정 -> 시간복잡도가 3차에서 1차로

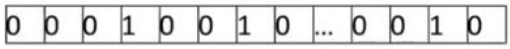
beam search

Conventional Feature Representation





binary, sparse
dim =10^6–10^7

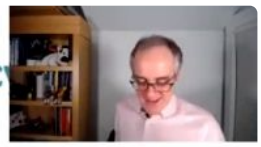


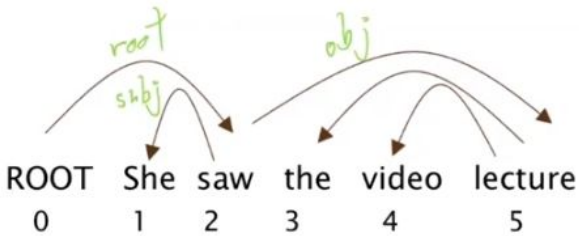
Indicator features

$$s1.w = \text{good} \wedge s1.t = \text{JJ}$$
$$s2.w = \text{has} \wedge s2.t = \text{VBZ} \wedge s1.w = \text{good}$$
$$lc(s_2).t = \text{PRP} \wedge s_2.t = \text{VBZ} \wedge s_1.t = \text{JJ}$$
$$lc(s_2).w = \text{He} \wedge lc(s_2).l = \text{nsubj} \wedge s_2.w = \text{has}$$

Stanford

Evaluation of Dependency Parsing: (labeled) dependency accurac





Acc = $\frac{\text{\# correct deps}}{\text{\# of deps}}$

UAS = 4 / 5 = 80%
LAS = 2 / 5 = 40%

Gold			
1	2	She	nsubj
2	0	saw	root
3	5	the	det
4	5	video	nn
5	2	lecture	obj

Parsed			
1	2	She	nsubj
2	0	saw	root
3	4	the	det
4	5	video	nsubj
5	2	lecture	ccomp

Stanford