A Naïve Survey on **Syntax** and **Semantics** Learning

费豪

武汉大学

All about the Natural Language Understanding

1、自然语言理解本质是结构预测

要搞清楚自然语言理解难在哪儿，先看自然语言理解任务的本质是什么。

作为人工智能关注的三大信息类型（语音、视觉、语言）之一，**自然语言文本是典型的无结构数据，由语言符号（如汉字）序列构成。**

要实现对自然语言的表意的理解，需要建立对该无结构文本背后的语义结构的预测。因此，自然语言理解的众多任务，包括并不限于中文分词、词性标注、命名实体识别、共指消解、句法分析、语义角色标注等，都是在对文本序列背后特定语义结构进行预测。

例如，中文分词就是在原本没有空格分隔的句子中增加空格或其他标识，将句子中每个词的边界标记出来，相当于添加了某些结构化语义信息到这个文本序列上。

2、自然语言理解的关键是语义表示

不过，以上NLP任务都只是在不断"逼近"对文本的理解，是对文本语义的局部表示。要实现对文本的完整理解，**需要建立更完备的语义结构表示空间**，这种更完备的语义表示经常成为上述NLP任务进行结构预测的依据。

Outline

- **Syntax Learning**
  - Word level
  - Sentence level
- **Semantics Learning**
  - Word level
  - Sentence level structure

一、Syntax Learning

1. Word (lexical) level

- 中文分词
- Part-of-Speech (POS) tagging
- Chunking
- …

- 中文分词

  - 基于词典的最大匹配分词方法

    刘源、梁南元(1986),汉语处理的基础工程——现代汉语词频统计[J]. 中文信息学报

  - 全切分路径选择方法

    张华平、刘群(2002),基于N-最短路径方法的中文词语粗分模型[J]. 中文信息学报

  - 基于字序列标注的方法
    - Tagging set: {B, I, E, S}. 然而很难利用到词级别的信息。
    - HMM,CRF,RNN

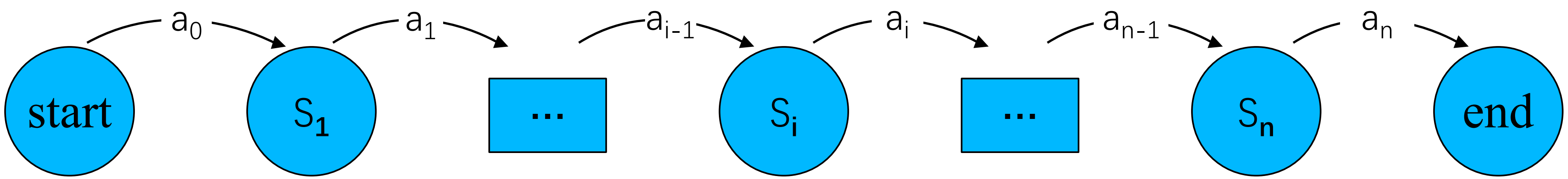
    Nianwen Xue, Susan P. Converse,2002, Combining Classifiers for Chinese Word Segmentation. SIGHAN@COLING

  - 基于转移的方法

    基于转移的方法是通过转移动作序列来进行分词，即从左往右判断一个每两个相邻的字是分还是不分。

Transition System

- State
  - Corresponds to partial results during decoding
    - start state, end state, S_i



- Actions
  - The operations that can be applied for state transition
  - Construct output incrementally

An Example :

S-SHIFT
R-REDUCE
AL-ARC-LEFT
AR-ARC-RIGHT

Zhang and Clark, 2007. Chinese Segmentation with a Word-Based Perceptron Algorithm. ACL

- 中文分词
  - 基于词典的最大匹配分词方法

    刘源、梁南元(1986),汉语处理的基础工程——现代汉语词频统计[J]. 中文信息学报

  - 全切分路径选择方法

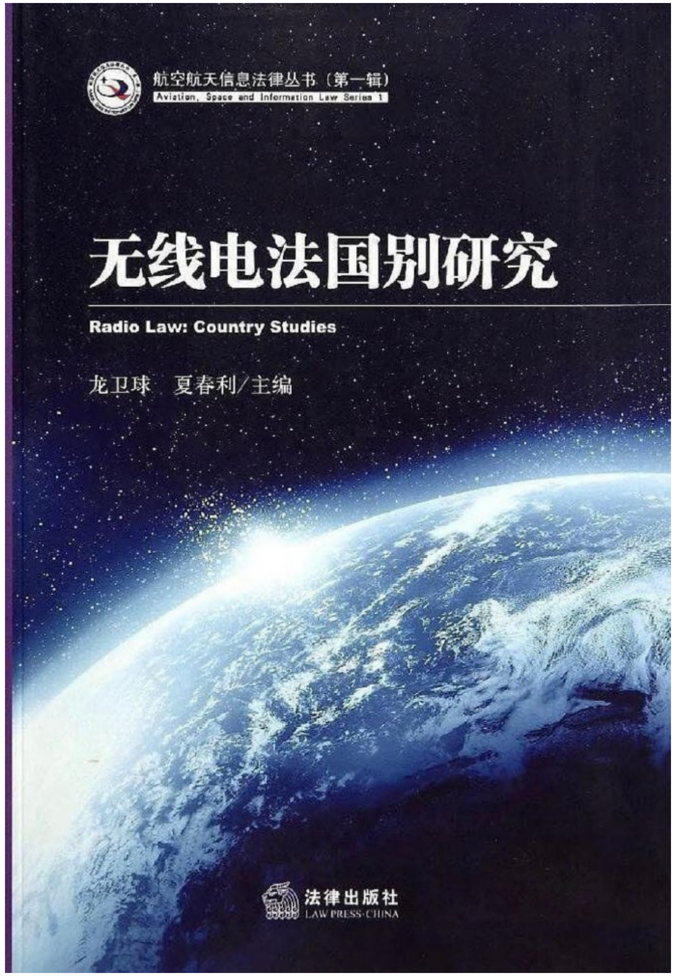
    张华平、刘群(2002),基于N-最短路径方法的中文词语粗分模型[J]. 中文信息学报

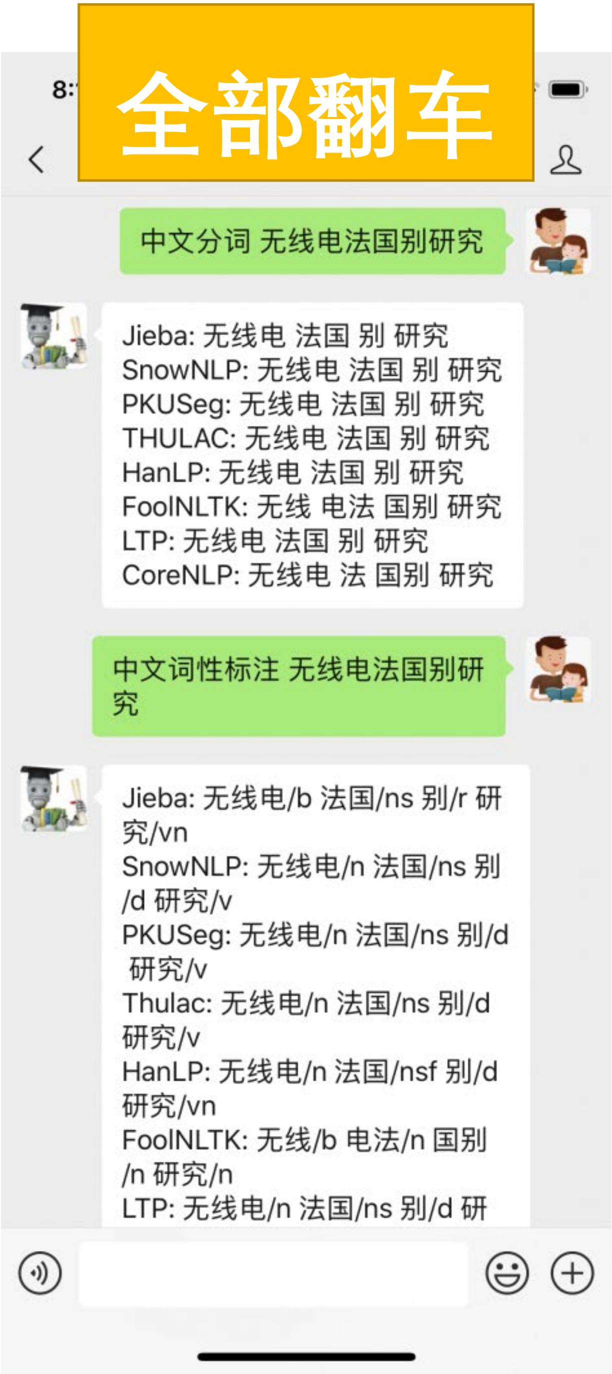
  - 基于字序列标注的方法
    - Tagging set: {B, I, E, S}
    - HMM,CRF,RNN

    Nianwen Xue, Susan P. Converse,2002, Combining Classifiers for Chinese Word Segmentation. SIGHAN@COLING

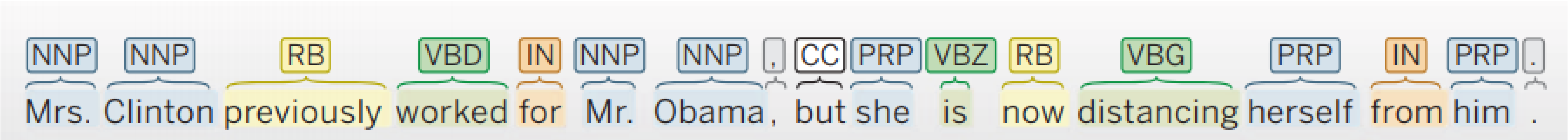
  - 基于转移的方法
  - 基于图的方法

- 中文分词

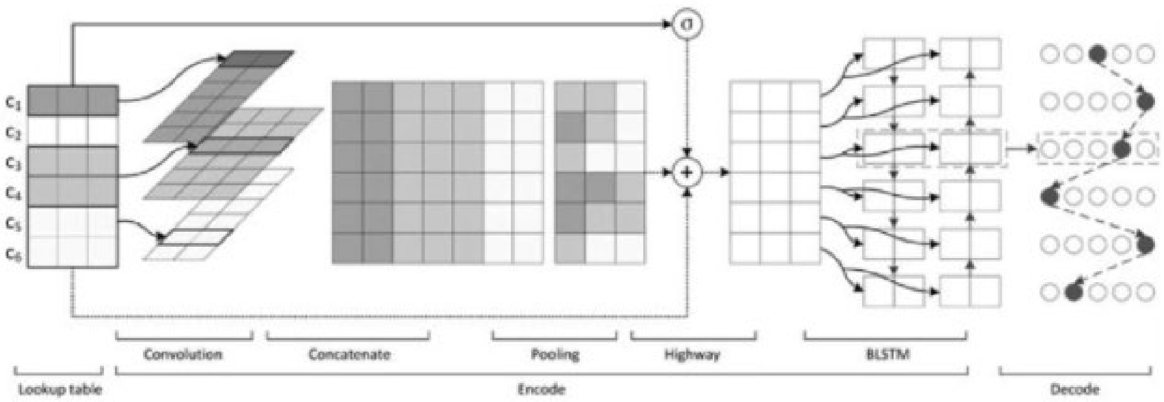

航空航天信息法律丛书（第一辑）
Aviation, Space and Information Law Series 1
无线电法国别研究
Radio Law: Country Studies
龙卫球 夏春利/主编
法律出版社
LAW PRESS·CHINA

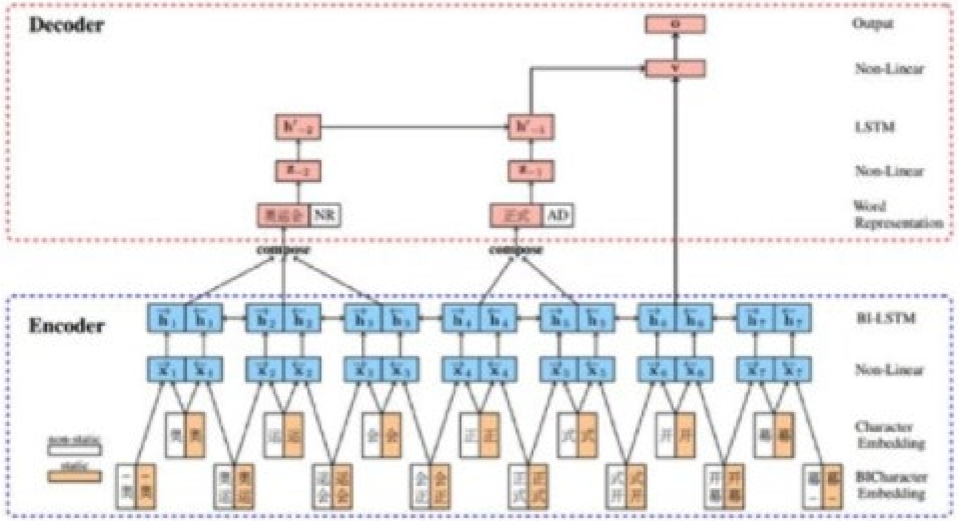

全部翻车
中文分词 无线电法国别研究
Jieba: 无线电 法国 别 研究
SnowNLP: 无线电 法国 别 研究
PKUSeg: 无线电 法国 别 研究
THULAC: 无线电 法国 别 研究
HanLP: 无线电 法国 别 研究
FoolNLTK: 无线 电法 国别 研究
LTP: 无线电 法国 别 研究
CoreNLP: 无线电 法 国别 研究
中文词性标注 无线电法国别研究
Jieba: 无线电/b 法国/ns 别/r 研究/vn
SnowNLP: 无线电/n 法国/ns 别/d 研究/v
PKUSeg: 无线电/n 法国/ns 别/d 研究/v
Thulac: 无线电/n 法国/ns 别/d 研究/v
HanLP: 无线电/n 法国/nsf 别/d 研究/vn
FoolNLTK: 无线/b 电法/n 国别/n 研究/n
LTP: 无线电/n 法国/ns 别/d 研

- POS Tagging
  - Sequential labeling
  - Transition-based
  - Graph-based



- Joint parsing
  - Joint **POS** with CWS





Xinchi Chen, Xipeng Qiu, Xuanjing Huang. 2017. A Feature-Enriched Neural Model for Joint Chinese Word Segmentation and Part-of-Speech Tagging. IJCAI. 3960-3966

Meishan Zhang, Nan Yu, Guohong Fu. 2018. A Simple and Effective Neural Model for Joint Word Segmentation and POS Tagging. IEEE/ACM Trans. Audio, Speech & Language Processing 26(9): 1528-1538

- Joint parsing
  - Joint Segmentation/Tagging/Chunking

Character-based chunking

Action: FIN W

chunk buffer	word buffer	character buffer
[NP 他/NR] [VP 到达/VV] [NP 北京/NR]	[机场/NN]	。

[NP 他/NR] [VP 到达/VV] [NP 北京/NR 机场/NN] [O 。/PU]

[He] [arrived] [Beijing airport] [.]

Chen Lyu, Yue Zhang and Donghong Ji. *Joint Word Segmentation, POS-Tagging and Syntactic Chunking.* In Proceedings of the AAAI 2016, Phoenix, Arizona, USA, February.

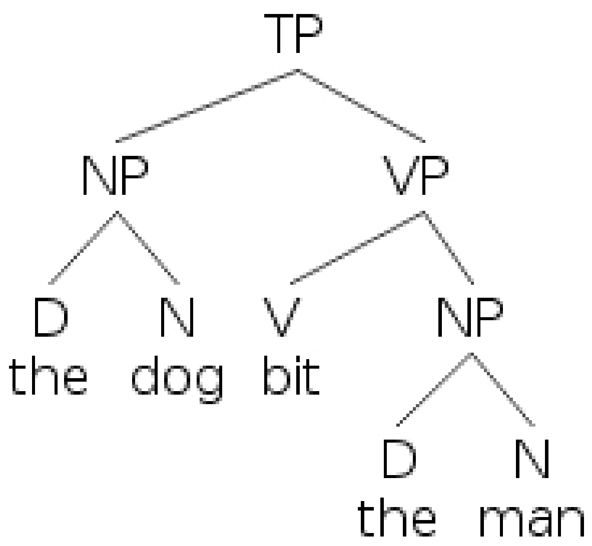
一、Syntax Learning

2. Sentence level

- 短语结构句法分析(成分句法分析)
- 依存句法分析
- 深层文法句法分析
- …

- 短语结构句法分析

  - 短语结构句法分析，（phrase structure parsing），也叫成分句法分析（**constituent** syntactic parsing）。
  - 作用是识别出句子中的短语结构以及短语之间的层次句法关系
  - 短语结构句法分析的研究基于上下文无关文法（Context Free Grammar， CFG）。



上下文无关文法可以定义为四元组<T, N, S, R>，其中 T 表示终结符的集合（即词的集合），N 表示非终结符的集合（即文法标注和词性标记的集合），S 表示充当句法树根节点的特殊非终结符，而 R 表示文法规则的集合，其中每条文法规则可以表示为 $N_i \rightarrow \gamma$，这里的γ表示由非终结符与终结符组成的一个序列(允许为空)。

◆ Supervised constituency learning

- PCFG (经典，Stanford parser)

  Dan Klein and Christopher D. Manning. 2003. Accurate Unlexicalized Parsing. Proceedings of the 41st Meeting of the Association for Computational Linguistics, pp. 423-430.

  Yoon Kim, Chris Dyer, Alexander M. Rush. 2019. Compound Probabilistic Context-Free Grammars for Grammar Induction. ACL: 2369-2385

- Transition-based

  Yue Zhang, Stephen Clark. 2009. Transition Based Parsing of the Chinese Treebank using a Global Discriminative Model. IWPT 162-171.

  Muhua Zhu, Yue Zhang, Wenliang Chen, Min Zhang, Jingbo Zhu. 2013. Fast and Accurate Shift-Reduce Constituent Parsing. ACL, 434-443

- Seq2seq

  Lemao Liu, Muhua Zhu, Shuming Shi. 2018. Improving Sequence-to-Sequence Constituency Parsing. AAAI: 4873-4880

- Self-attention parser

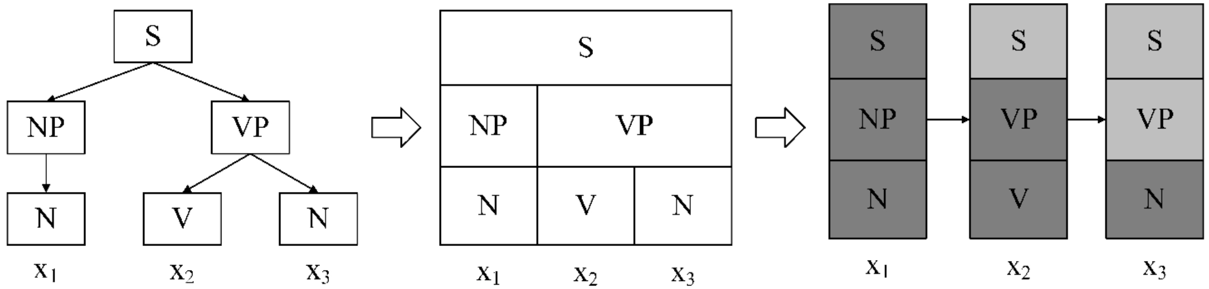
  Nikita Kitaev, Dan Klein. 2018. Constituency Parsing with a Self-Attentive Encoder. ACL: 2676-2686

◆ Unsupervised Grammar Induction

- PRPN

Yikang Shen, Zhouhan Lin, Chin-Wei Huang, and Aaron C. Courville. 2018. Neural language modeling by jointly learning syntax and lexicon. In Proceedings of the ICLR.

- On-LSTM



(a) Constituency tree (b) Block view (c) ON-LSTM cell states

Yikang Shen, Shawn Tan, Alessandro Sordoni, and Aaron C. Courville. 2018. Ordered neurons: Integrating tree structures into recurrent neural networks. CoRR, abs/1810.09536.

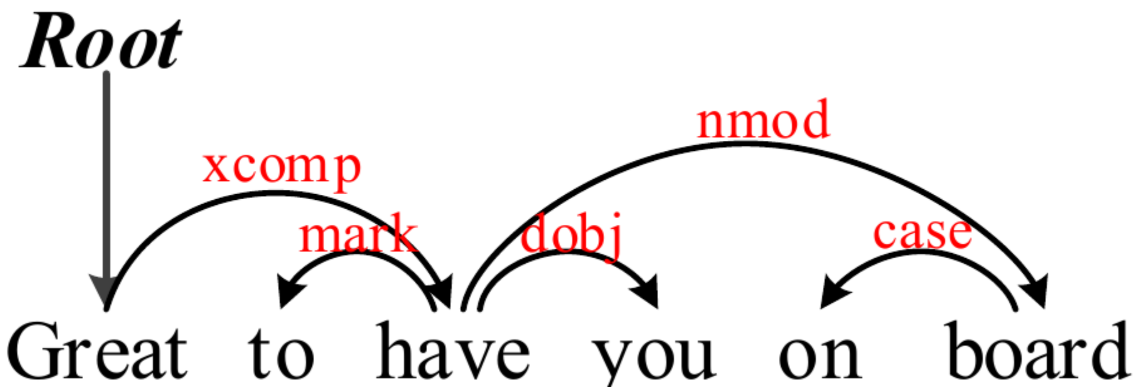
- URNNG

Yoon Kim, Alexander Rush, Lei Yu, Adhiguna Kuncoro, Chris Dyer, and Ga´bor Melis. 2019b. Unsupervised recurrent neural network grammars. In Proceedings of the NAACL, pages 1105–1117.

- DIORA

Andrew Drozdov, Patrick Verga, Mohit Yadav, Mohit Iyyer, and Andrew McCallum. 2019. Unsupervised latent tree induction with deep inside-outside recursive autoencoders. CoRR, abs/1904.02142.

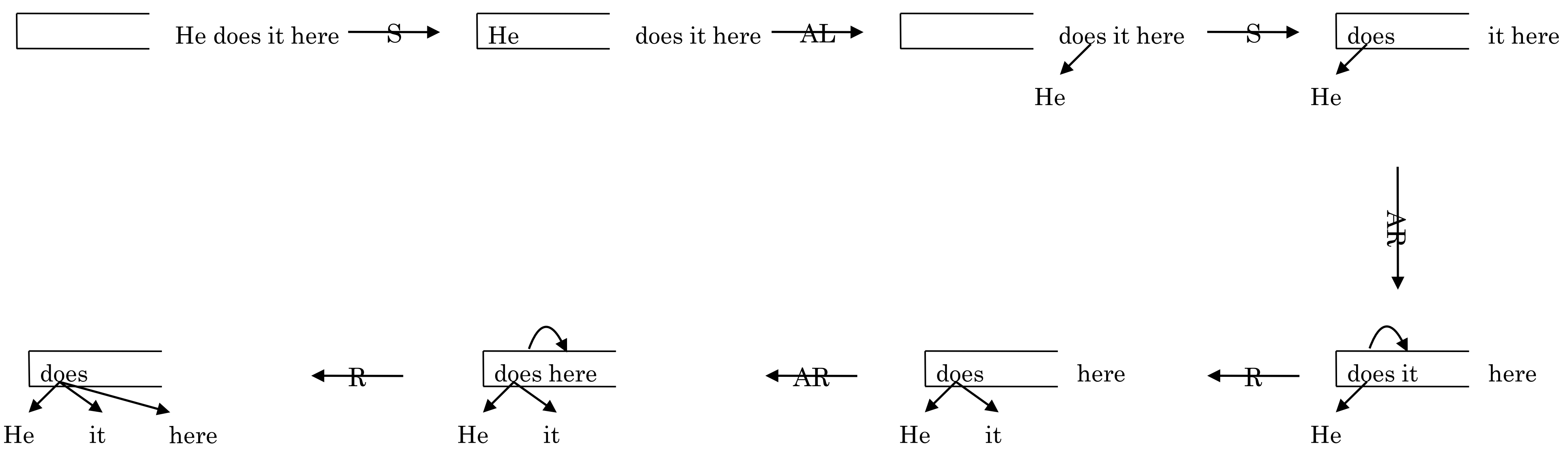
- 依存句法分析（dependency parsing）
  - 依存语法存在一个共同的基本假设：句法结构本质上包含词和词之间的依存（修饰）关系。
  - 一个依存关系连接两个词，分别是核心词（head）和依存词（dependent）。



- 依存句法分析

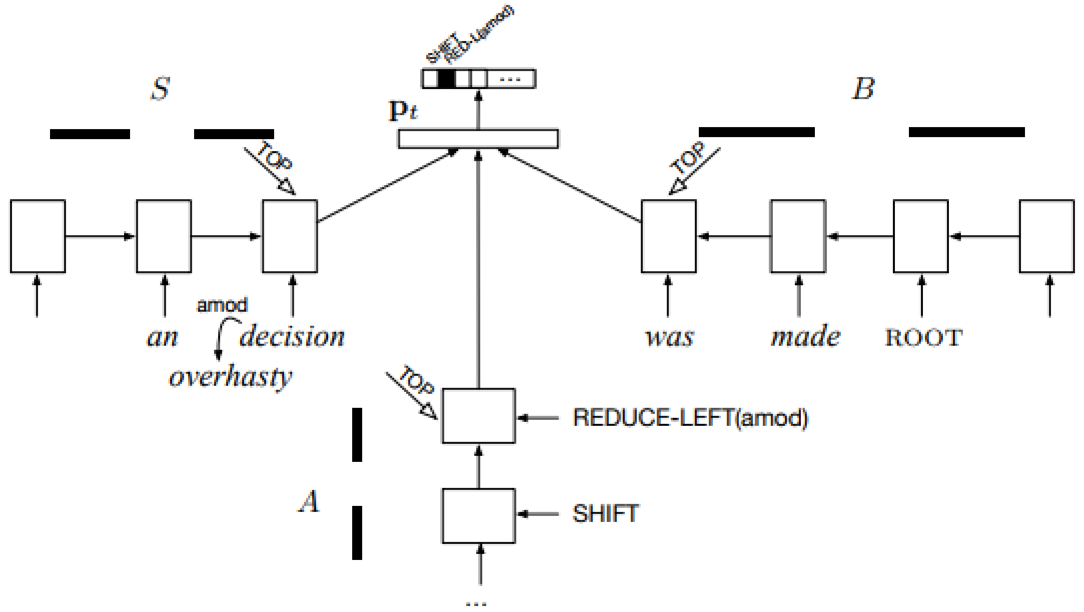
  - Transition-based

    - S-SHIFT
    - R-REDUCE
    - AL-ARC-LEFT
    - AR-ARC-RIGHT



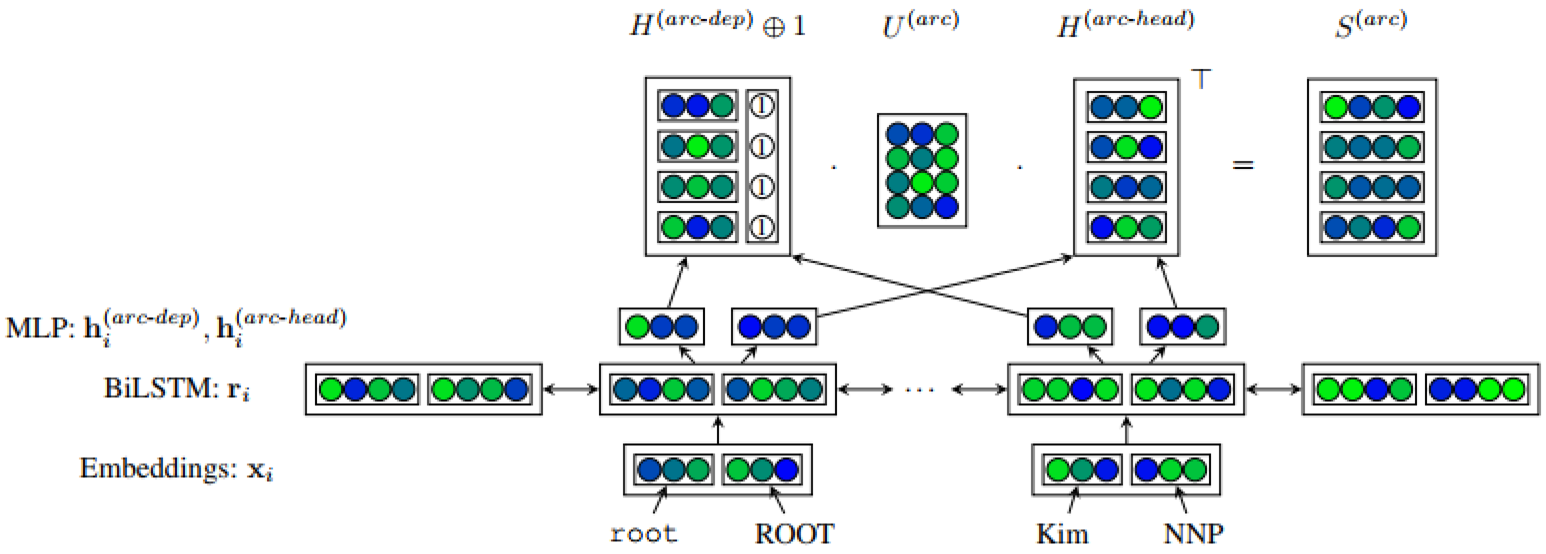
Yue Zhang and Stephen Clark. 2011. *Syntactic Processing Using the Generalized Perceptron and Beam Search*. In *Computational Linguistics*, 37(1), March.

- 依存句法分析
    - Transition-based



Chris Dyer, Miguel Ballesteros, Wang Ling, Austin Matthews, Noah A. Smith. 2015. Transition-Based Dependency Parsing with Stack Long Short-Term Memory. ACL: 334-343

- 依存句法分析
    - Graph-based

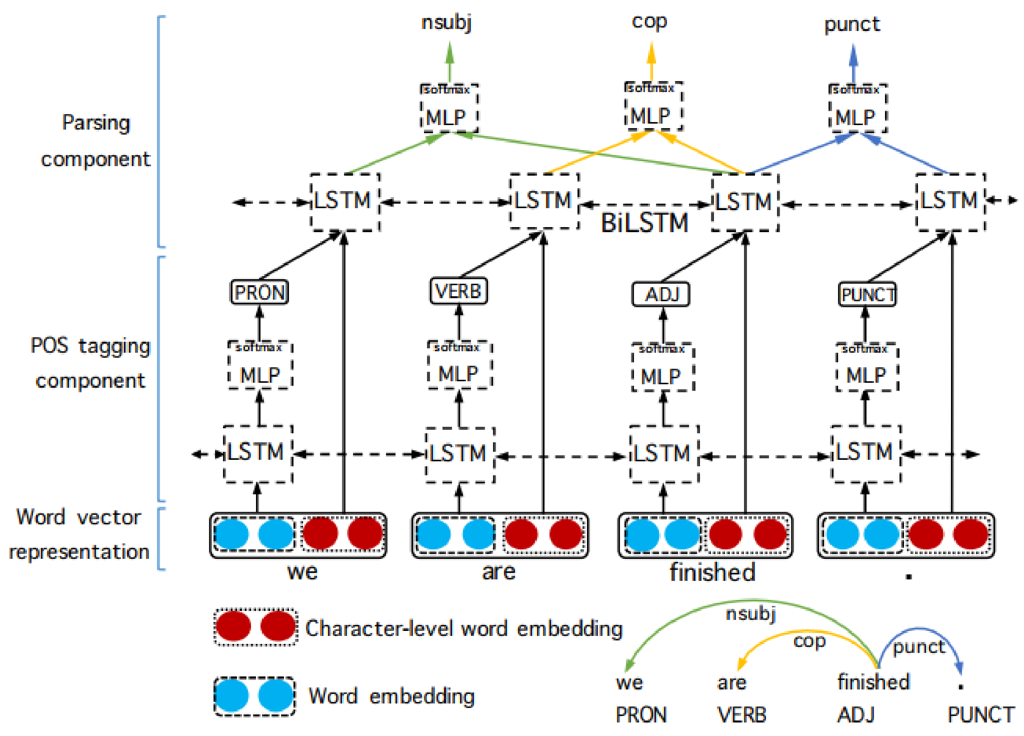


Timothy Dozat, Christopher D. Manning. 2017. Deep Biaffine Attention for Neural Dependency Parsing. ICLR.

- 依存句法分析
  - 当前，依存树的使用比句法成分结构更为广泛。
  - 普遍采用 Recursive model 方案做Encoding Dependency structure.
    - Tree-LSTM(RNN/GRU)
      - Child-Sum Tree-LSTMs for Dependency tree.
      - *N*-ary Tree-LSTMs for Constituency tree.
      - TreeLSTM + CRFs
    - Syntax aware-LSTM
    - SDP (shortest dependency path) LSTM
    - GCN

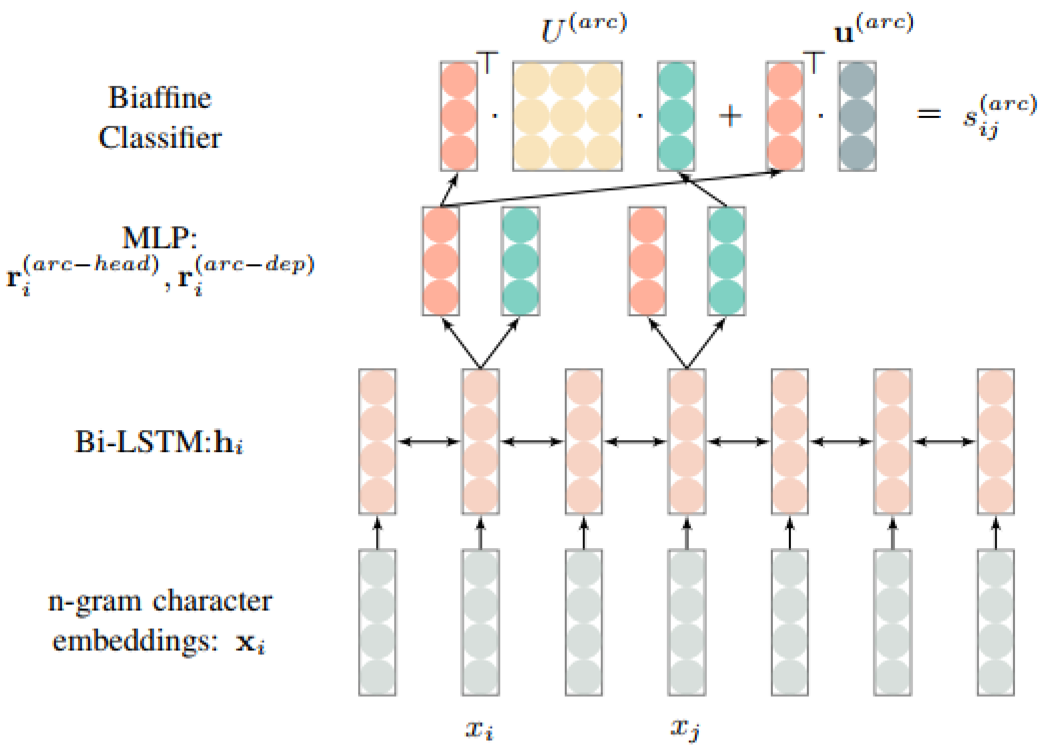
- *Kai Sheng Tai, Richard Socher, Christopher D. Manning. 2015. Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks. ACL: 1556-1566*
- *Wenya Wang, Sinno Jialin Pan, Daniel Dahlmeier, Xiaokui Xiao。 2016. Recursive Neural Conditional Random Fields for Aspect-based Sentiment Analysis. EMNLP: 616-626*
- *Feng Qian, Lei Sha, Baobao Chang, Lu Chen Liu, and Ming Zhang. 2017. Syntax aware LSTM model for semantic role labeling. In The Workshop on Struc- tured Prediction for Natural Language Processing, pages 27–32.*
- *Xu, Y.; Mou, L.; Li, G.; Chen, Y.; Peng, H.; and Jin, Z. 2015. Classifying relations via long short term memory networks along shortest dependency paths. In Proceedings of EMNLP, 1785–1794.*
- *Diego Marcheggiani and Ivan Titov. 2017. Encoding sentences with graph convolutional networks for semantic role labeling. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1506–1515, Copenhagen, Denmark.*

- Joint parsing with lexical level tasks
    - Joint POS Tagging & dependency parsing



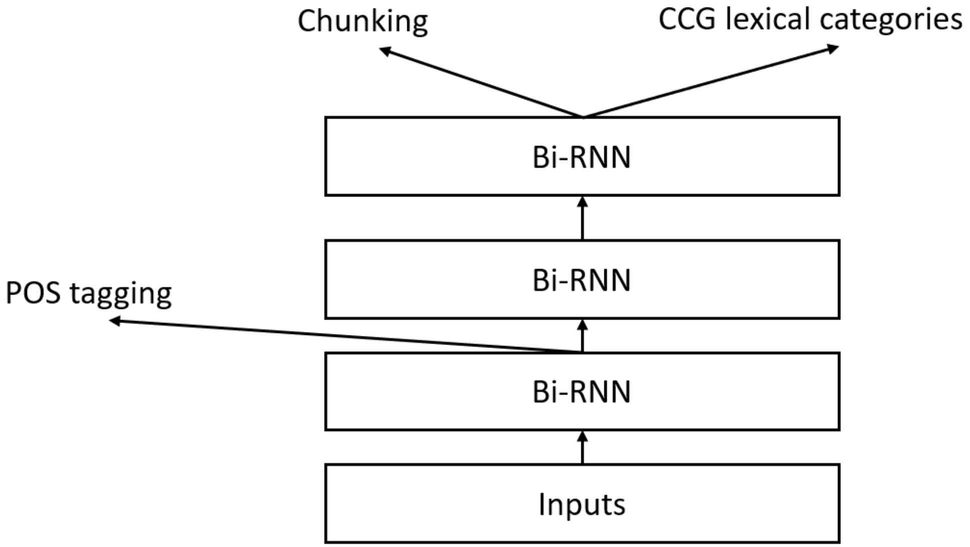
Dat Quoc Nguyen, Karin Verspoor. An Improved Neural Network Model for Joint POS Tagging and Dependency Parsing. CoNLL Shared Task (2) 2018: 81-91

- Joint parsing with lexical level tasks
  - Joint CWS & dependency parsing



Hang Yan, Xipeng Qiu, Xuanjing Huang. A Unified Model for Joint Chinese Word Segmentation and Dependency Parsing. CoRR abs/1904.04697 (2019)

- Joint parsing with lexical level tasks
  - Joint POS tagging/Chunking and CCG Super Tagging



Søgaard, Anders, and Yoav Goldberg. "Deep multi-task learning with low level tasks supervised at lower layers." Proceedings *of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Vol. 2. 2016.

- 深层文法句法分析
  - 利用深层文法，对句子进行深层的句法以及语义分析。
  - 如:
    - 词汇化树邻接文法（Lexicalized Tree Adjoining Grammar，**LTAG**）、
    - 词汇功能文法（Lexical Functional Grammar，**LFG**）、
    - 组合范畴文法（Combinatory Categorial Grammar，**CCG**）等。

- *Masashi Yoshikawa, Hiroshi Noji, Yuji Matsumoto. 2017. A* CCG Parsing with a Supertag and Dependency Factored Model. ACL: 277-287*
- *Maria Nadejde, Siva Reddy, Rico Sennrich, Tomasz Dwojak, Marcin Junczys-Dowmunt, Philipp Koehn, Alexandra Birch. 2017. Predicting Target Language CCG Supertags Improves Neural Machine Translation. WMT: 68-79*

一、Syntax Learning

*3. Discourse level

- RST修辞结构理论
- 基于逻辑结构的篇章理论
- 句群理论和复句理论
- …

二、 Semantics Learning

How to define **Semantic**？

エ̸ -1决雀予エ

二、 Semantics Learning

#NLP太难了#

难度： ※※ 两颗星

1.来到杨过曾经生活过的地方，小龙女动情地说：“我也想过过过儿过过的生活。”

2.来到儿子等校车的地方，邓超对孙俪说：“我也想等等等等等过的那辆车。”

3.赵敏说：我也想控忌忌己不想无忌。

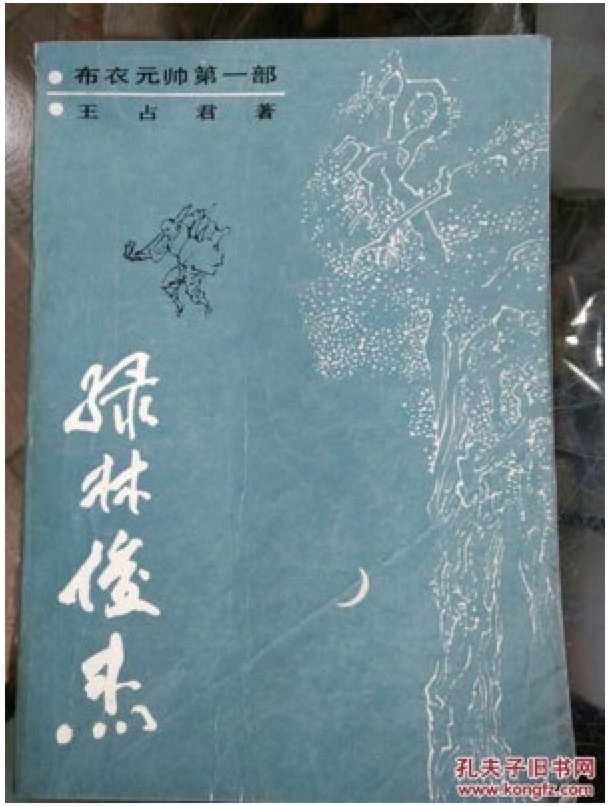
4.你也想犯范范范玮琪犯过的错吗

5.对叙打击是一次性行为?

6.《绿林俊杰》--林俊杰做错了什么？为什么要绿他

7.那辆白车是黑车，那辆车到时是白车还是黑车啊!

8.能穿多少穿多少，到底是多穿点还是少穿点，怎么感觉冬天喝夏天不一样了



二、 Semantics Learning

#NLP太难了#　　难度：※※※ 三颗星

1.写给卖豆芽的对联：长长长长长长长，长长长长长长长。
(solution:changzhangchangzhangchangchangzhangzhangchangzhangchangzhangzhangchang,
zhangchangchangzhangchangzhangchang，zhangchangzhangchangzhangchangchang)

2.季姬寂，集鸡，鸡即棘鸡。棘鸡饥叽，季姬及箕稷济鸡。鸡既济，跻姬笈，季姬忌，急咭鸡，鸡急，继圾几，季姬急，即籍箕击鸡，箕疾击几伎，伎即齑，鸡叽集几基，季姬急极屐击鸡，鸡既殛，季姬激，即记《季姬击鸡记》。

3.石室诗士施氏，嗜狮，誓食十狮。氏时时适市视狮。十时，适十狮适市。是时，适施氏适市。施氏视是十狮，恃矢势，使是十狮逝世。氏拾是十狮尸，适石室。石室湿，氏使侍拭石室。石室拭，氏始试食是十狮尸。食时，始识是十狮尸，实十石狮尸。试释是事。《施氏食狮史》

4.去商店买东西一算账1001块，小王对老板说：“一块钱算了。” 老板说好的。于是小王放下一块钱就走了，老板死命追了小王五条街又要小王付了1000，小王感慨：#自然语言理解太难了#

5.“碳碳键键能能否否定定律一”

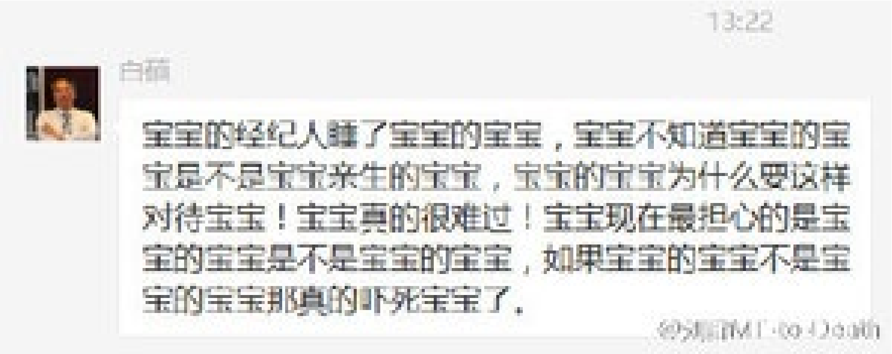
6.中文里面“大胜”和“大败”意思相同，刚发现英文里面也有类似的现象：valuable和invaluable都是表示非常有价值的意思

二、 Semantics Learning

#NLP太难了#

难度： ※※※※ **四颗星**

1.宝宝的经纪人睡了宝宝的宝宝，宝宝不知道宝宝的宝宝是不是宝宝的亲生的宝宝，宝宝的宝宝为什么要这样对待宝宝！宝宝真的很难过！宝宝现在最担心的是宝宝的宝宝是不是宝宝的宝宝，如果宝宝的宝宝不是宝宝的宝宝那真是吓死宝宝了。

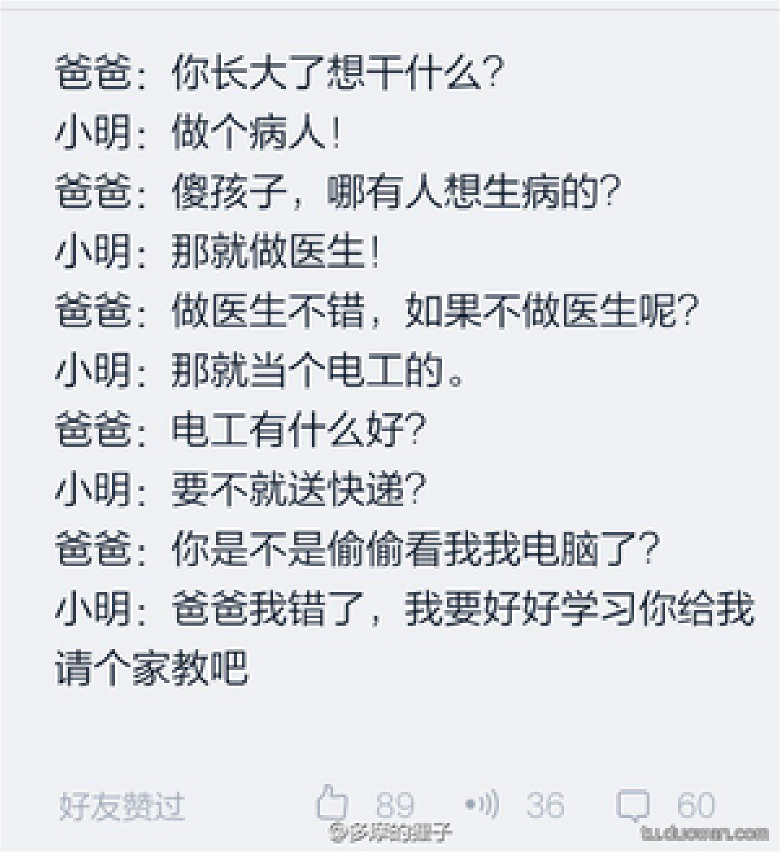


2.中不不建交是受印度的影响，中不建交不受印度的影响。

二、 Semantics Learning

#NLP太难了#

难度：※※※※※ **五颗星**



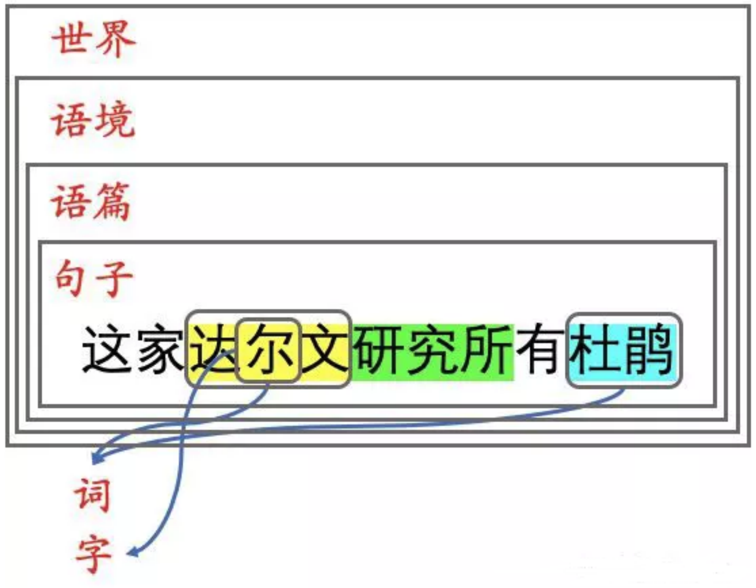
爸爸：你长大了想干什么?
小明：做个病人!
爸爸：傻孩子，哪有人想生病的?
小明：那就做医生!
爸爸：做医生不错，如果不做医生呢?
小明：那就当个电工的。
爸爸：电工有什么好?
小明：要不就送快递?
爸爸：你是不是偷偷看我我电脑了?
小明：爸爸我错了，我要好好学习你给我请个家教吧

好友赞过 89 36 60

需要配合知识图谱

二、 Semantics Learning

- 语言使用需要考虑其复杂的语境。
- 以语言的多义性为例，存在多义的语言单元，总需要其外部的复杂语境信息进行消歧：
  - 字的多义性至少需要所组成的词来消歧；
  - 词的歧义性至少需要所在的句子来消歧；
  - 句子的意思至少要放在语篇或对话语境中，甚至需要复杂的世界知识来帮助理解。



- *刘知远. 2018. 自然语言理解难在哪儿？知乎*
- *Julia Hirschberg and Christopher D. Manning. Advances in Natural Language Processing. Science, 2015.*
- *Waltz, David L . "Semantic Structures: Advances in Natural Language Processing." L. Erlbaum Associates Inc. 2014.*

二、 Semantics Learning

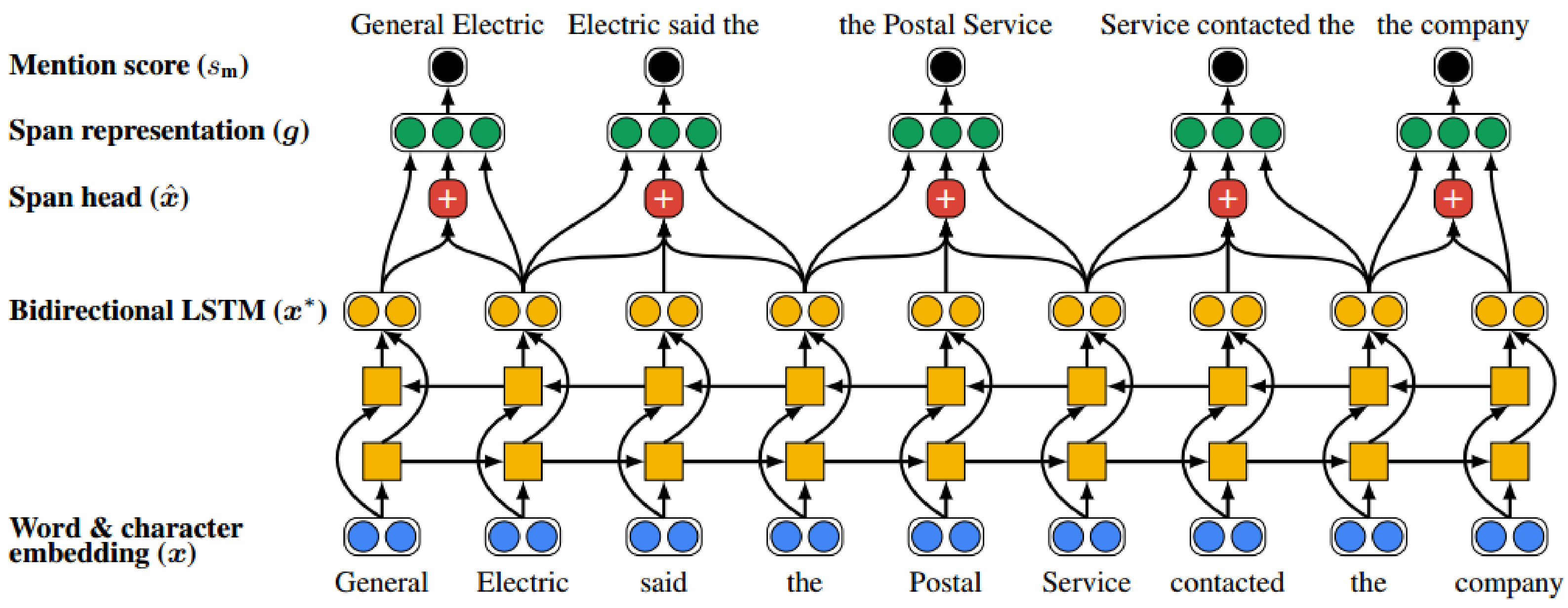
1. Word level

- 词义消歧、指代消解
- Word Representation
- Normalizarion
- …

- 词义消歧、指代消解

  - Mention-pair classifier
    - *Vincent Ng. 2010. Supervised noun phrase coreference research: The first fifteen years. In Proceedings of the 48th annual meeting of the association for computational linguistics, pages 1396–1411. Association for Computational Linguistics.*
    - *Eric Bengtson and Dan Roth. 2008. Understanding the value of features for coreference resolution. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, pages 294–303. Association for Computational Linguistics.*
  - Entity-level models
    - *Aria Haghighi and Dan Klein. 2010. Coreference resolution in a modular, entity-centered model. In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL, 385–393.*
    - *Kevin Clark and Christopher D. Manning. 2015. Entity-centric coreference resolution with model stacking. In ACL.*
    - *Kevin Clark and Christopher D. Manning. 2016a. Deep reinforcement learning for mention-ranking coreference models. In Empirical Methods on Natural Language Processing (EMNLP).*
  - Latent-tree models
    - *Eraldo Rezende Fernandes, C´ıcero Nogueira Dos Santos, and Ruy Luiz Milidiu. 2012. Latent structure perceptron with feature induction for unrestricted coreference resolution. In Joint Conference on EMNLP and CoNLL-Shared Task, pages 41–48.*
    - *Anders Bjorkelund and Jonas Kuhn. 2014. Learning structured perceptrons for coreference resolution with latent antecedents and non-local features. In ACL.*
  - Mention-ranking models
    - *Greg Durrett and Dan Klein. 2013. Easy victories and uphill battles in coreference resolution. In EMNLP.*
    - *Sam Wiseman, Alexander M. Rush, Stuart M. Shieber, and Jason Weston. 2015. Learning anaphoricity and antecedent ranking features for coreference resolution. In ACL.*
    - *Kenton Lee, Luheng He, Mike Lewis, Luke Zettlemoyer. 2017. End-to-end Neural Coreference Resolution. EMNLP: 188-197*

- 词义消歧、指代消解
  - Span-graph model



Kenton Lee, Luheng He, Mike Lewis, Luke Zettlemoyer. 2017. End-to-end Neural Coreference Resolution. EMNLP: 188-197

- Word Representation

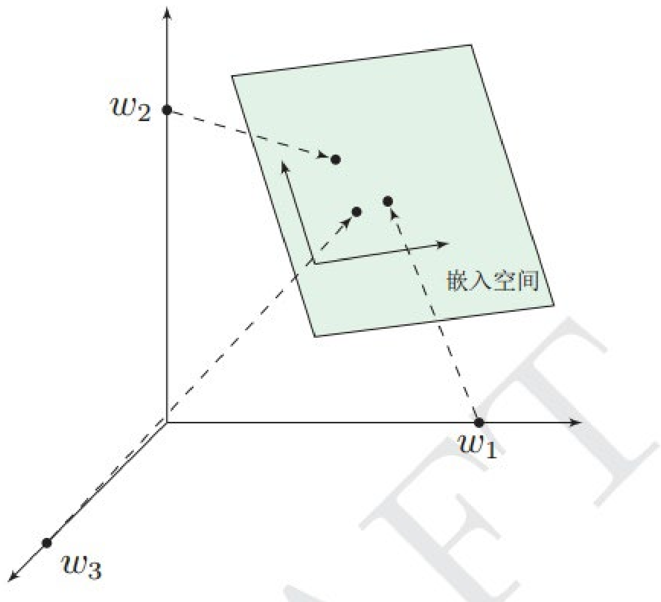
  - 词的独热表示one-hot representation

    “话筒”表示为 [0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 ...]
    “麦克”表示为 [0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 ...]

  - 词的分布式表示distributed representation

    1、将vector每一个元素由整形改为浮点型，变为整个实数范围的表示；

    2、将原来稀疏的巨大维度压缩嵌入到一个更小维度的空间。

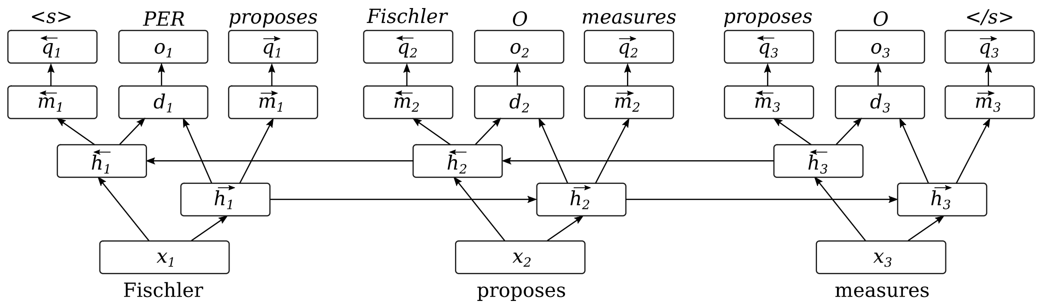


- 词的分布式表示distributed representation
  - 基于矩阵的分布表示
    - *Glove Vector*
  - 基于聚类的分布表示
    - *Word cluster*
  - 基于神经网络的分布表示，词嵌入（word embedding）

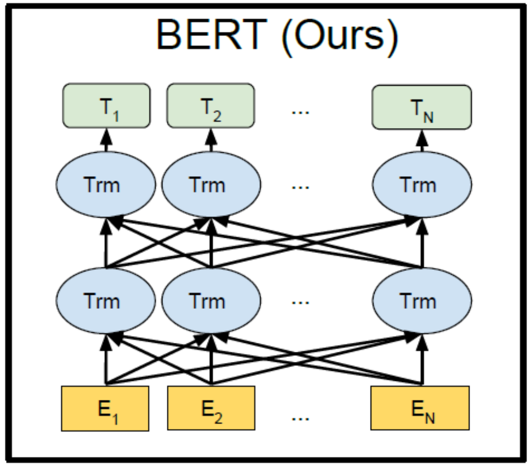
    神经网络语言模型（NNLM）的产出即为词向量。
    - *CBOW 或 Skip-gram 的产出：Word2vec*

◆ 词的分布式表示distributed representation

- 传统词向量
  - *Glove、word2vec*
  - *词级别静态固定表示，问题：一词多义无法解决*

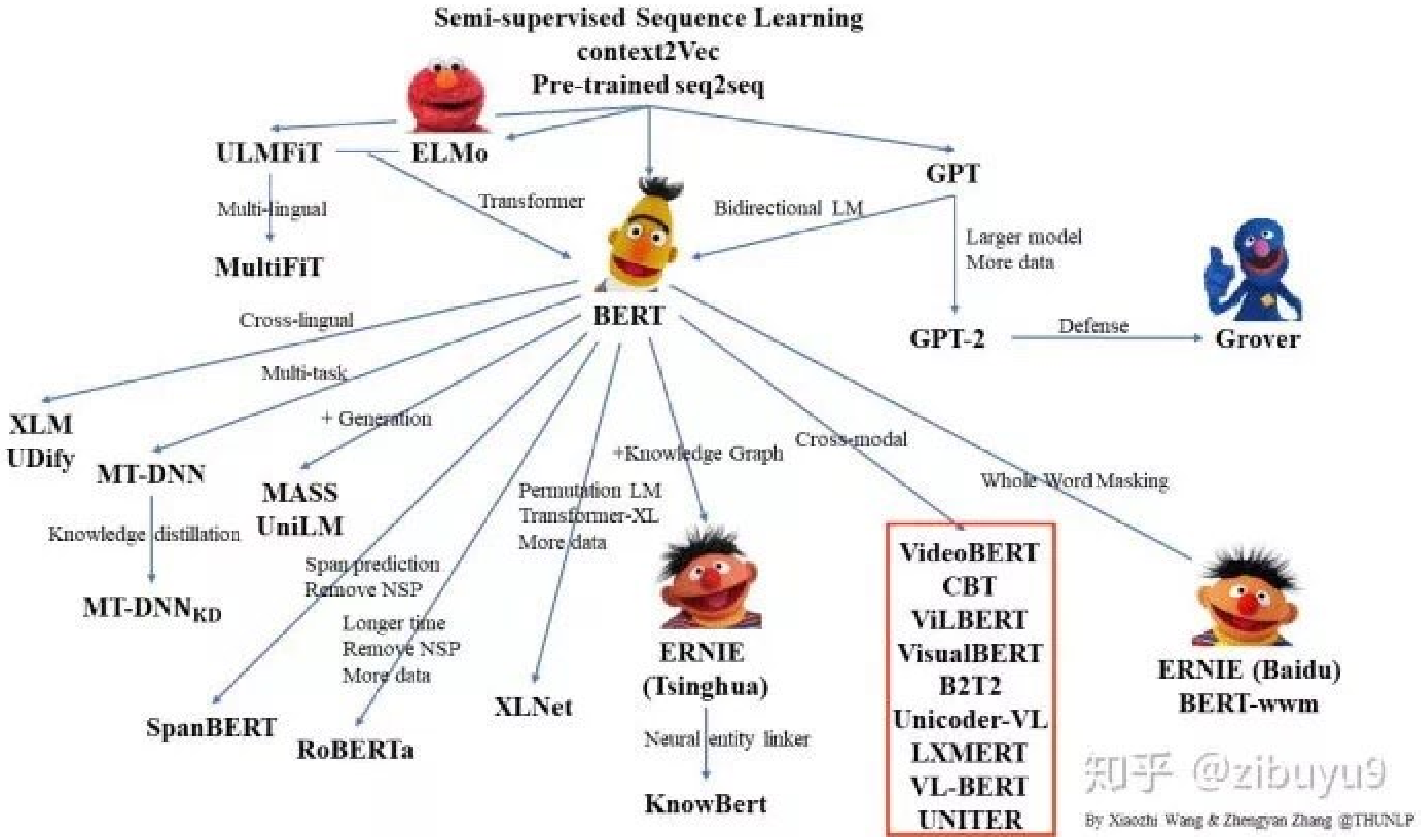


- SOTA Contextual Language Model
  - *ELMo(Embeddings from Language Models)*
  - *BERT(Pre-training of Deep Bidirectional Transformers)*
  - *GPT*
  - *XLNet*
  - *ULMFit*
  - …



- *Peters, Matthew E., et al. "Deep contextualized word representations." arXiv preprint arXiv:1802.05365 (2018).*
- *Devlin, Jacob, et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv preprint arXiv:1810.04805 (2018).*

◆ Bert family



Semi-supervised Sequence Learning
context2Vec
Pre-trained seq2seq
ULMFiT
ELMo
GPT
Transformer
Bidirectional LM
Multi-lingual
MultiFiT
Larger model
More data
BERT
GPT-2
Defense
Grover
Cross-lingual
Multi-task
XLM
UDify
+ Generation
MT-DNN
MASS
UniLM
Knowledge distillation
$MT\text{-}DNN_{KD}$
Span prediction
Remove NSP
Longer time
Remove NSP
More data
Permutation LM
Transformer-XL
More data
+Knowledge Graph
Cross-modal
Whole Word Masking
SpanBERT
RoBERTa
XLNet
ERNIE
(Tsinghua)
Neural entity linker
KnowBert
VideoBERT
CBT
ViLBERT
VisualBERT
B2T2
Unicoder-VL
LXMERT
VL-BERT
UNITER
ERNIE (Baidu)
BERT-wwm
知乎 @zibuyu9
By Xiaozhi Wang & Zhengyan Zhang @THUNLP

二、 Semantics Learning

Semantic Parsing

2. Sentence level structure

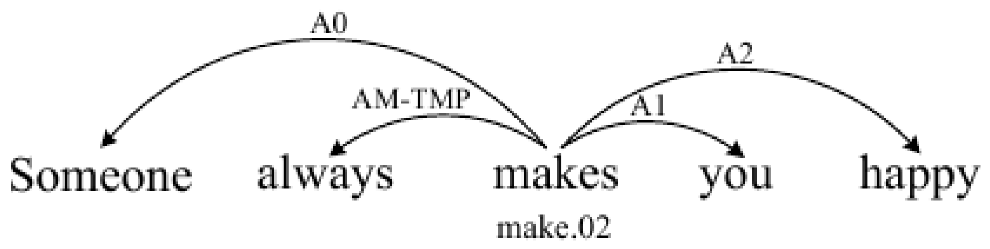
Surface form → Logic form

- Domain specific
  - Event Frame
    - VerbNet, FrameNet, PropBank semantics(SRL)
    - Universal Conceptual Cognitive Annotation parsing (UCCA)
    - Semantic Dependency Parsing (SDP)
    - Abstract mining representation (AMR)
    - Universal Decompositional Semantics (UDS)
    - Parallel Meaning Bank parsing
    - Event Extraction
  - Lambda Expression
  - Sequence-to-SQL
  - Sequence-to-math_expression
  - …
- Open domain
  - 语义槽填充 (Semantic Slot Filling) for Dialogue
  - …

时间有限，只提一点。

- Semantic Role Labeling (SRL) on PropBank

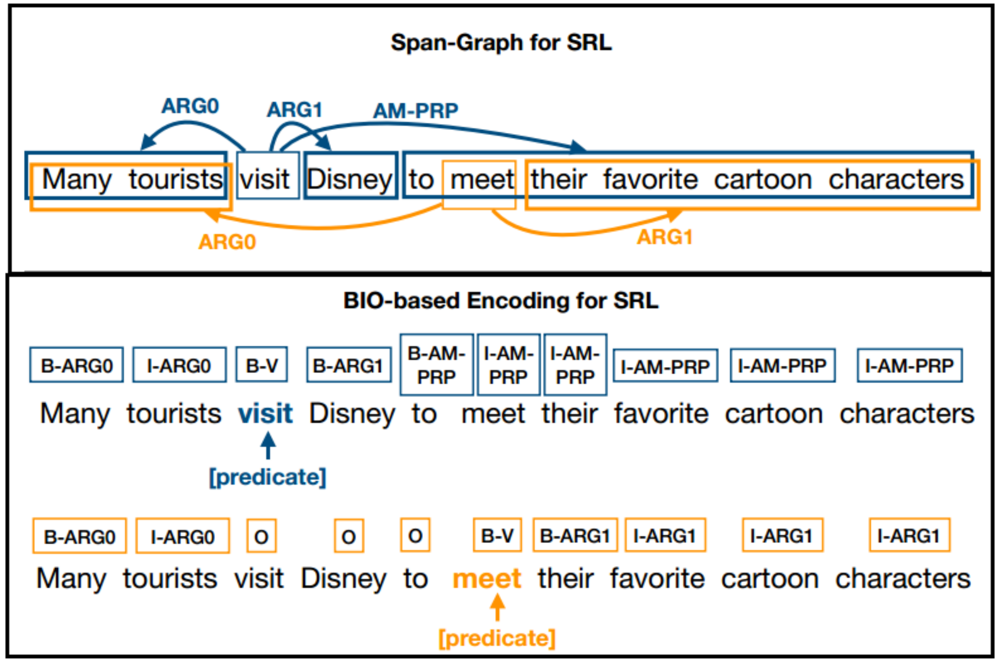
  - 找出句子中谓词(predicate)的相应语义角色(arguments role)成分，包括:
    - 核心语义角色（如施事者、受事者等）. (A0-A5 and AA)
    - 附属语义角色（如地点、时间、方式、原因等）. (AM-adj)



ARG0	agent	ARG3	starting point, benefactive, attribute
ARG1	patient	ARG4	ending point
ARG2	instrument, benefactive, attribute	ARGM	modifier

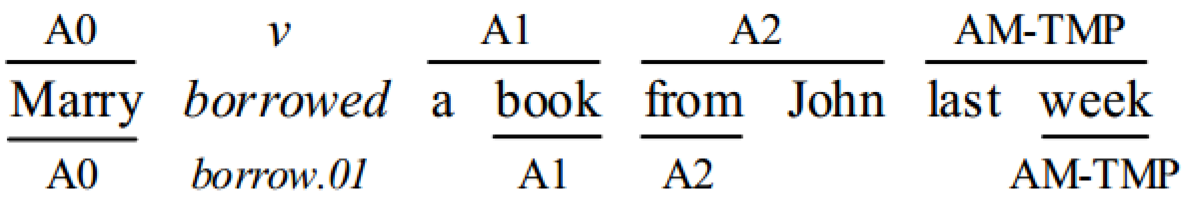
- *Martha Palmer, Dan Gildea, Paul Kingsbury, The Proposition Bank: A Corpus Annotated with Semantic Roles Computational Linguistics Journal, 31:1, 2005.*
- *Paul Kingsbury and Martha Palmer. From Treebank to PropBank. 2002. In Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002), Las Palmas, Spain.*
- *The Necessity of Parsing for Predicate Argument Recognition. Daniel Gildea and Martha Palmer. 2002. In Proceedings of ACL 2002, Philadelphia, PA.*

- Semantic Role Labeling (SRL) on PropBank
  - 玩法1
    - Argument role labeling (given predicate)
      - **BIO sequence labeling**
    - Joint(end2end): Predicate & Argument role labeling
      - **Span-graph model**
      - **Transition method**



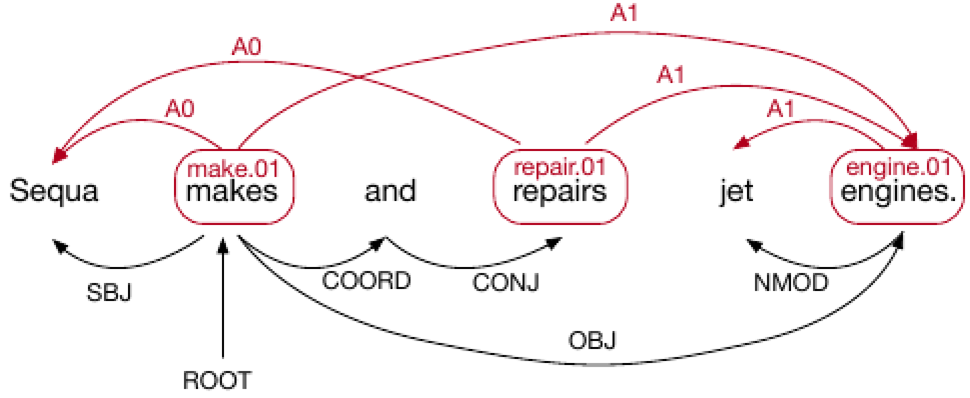
- *Luheng He, Kenton Lee, Omer Levy, Luke Zettlemoyer. 2018. Jointly Predicting Predicates and Arguments in Neural Semantic Role Labeling. ACL*
- *Dependency or Span, End-to-End Uniform Semantic Role Labeling.*
- *Luheng He, Kenton Lee, Mike Lewis, Luke Zettlemoyer. 2017. Deep Semantic Role Labeling: What Works and What's Next. ACL (1) 2017: 473-483.*
- *Zhixing Tan, Mingxuan Wang, Jun Xie, Yidong Chen, Xiaodong Shi. 2018. Deep Semantic Role Labeling With Self-Attention. AAAI 2018: 4929-4936*

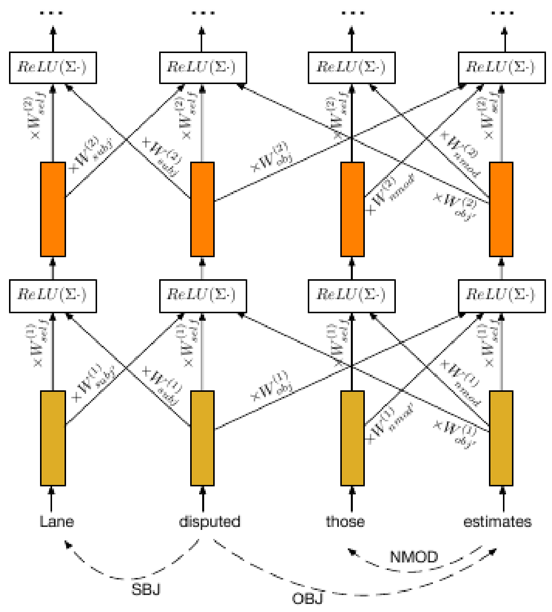
- Semantic Role Labeling (SRL) on PropBank
  - 玩法2
    - Span-based SRL: CoNLL 2005/2012
      - **Argument: Phrase span**
    - Dependency-based SRL: CoNLL 2008/2009
      - **Argument: Dependency head**

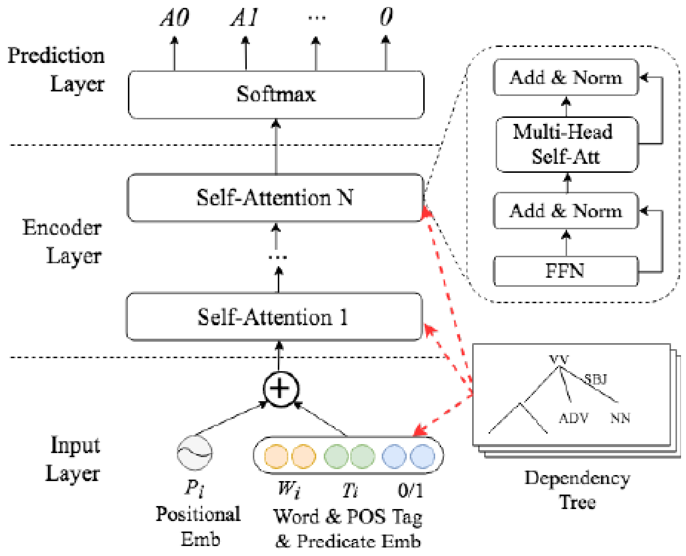


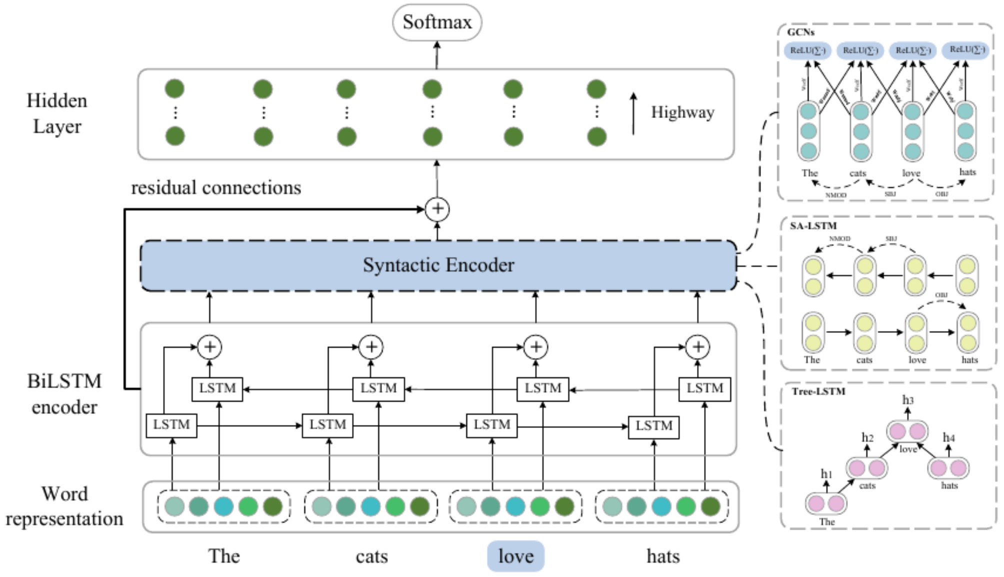
- *Zuchao Li, Shexia He, Hai Zhao, Yiqing Zhang, Zhuosheng Zhang, Xi Zhou, Xiang Zhou. 2019. Dependency or Span, End-to-End Uniform Semantic Role Labeling. AAAI 2019: 6730-6737*

- SRL using syntax information







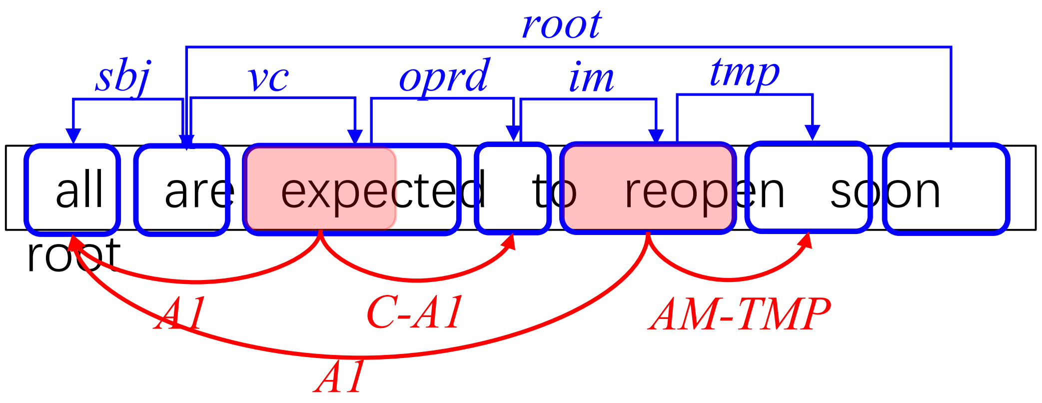


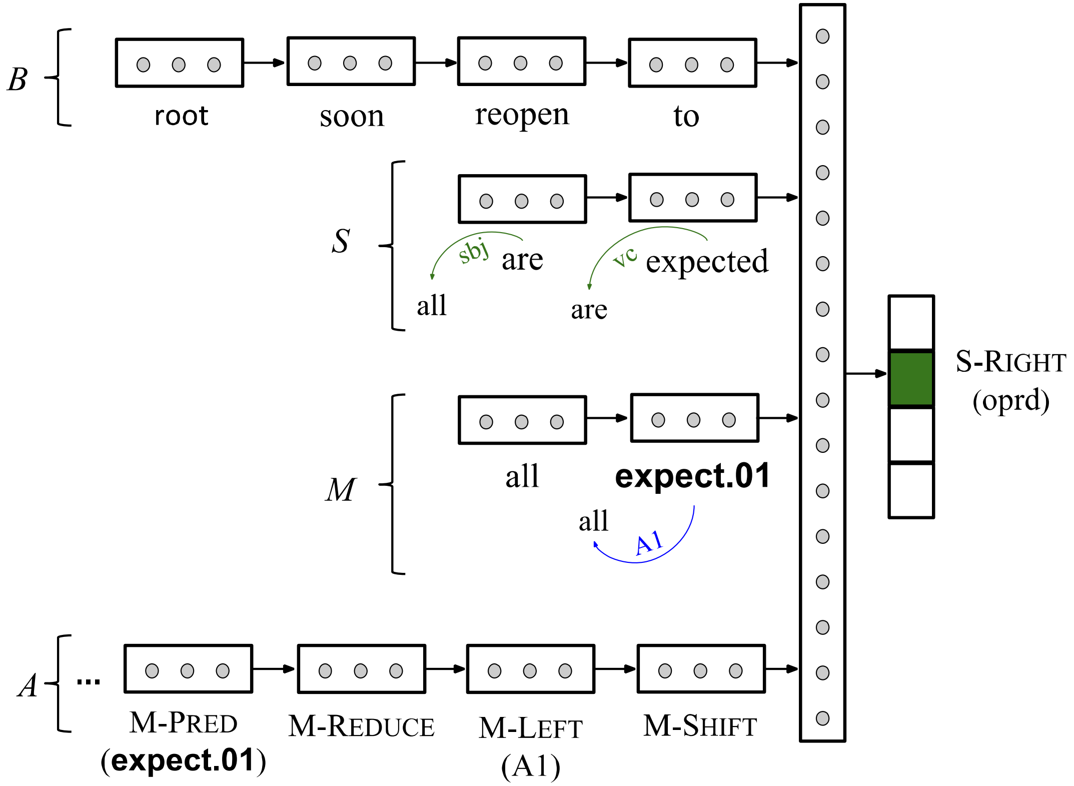
- *Diego Marcheggiani, Ivan Titov. 2017. Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling. EMNLP : 1506-1515*
- *Zuchao Li, Shexia He, Jiaxun Cai, Zhuosheng Zhang, Hai Zhao, Gongshen Liu, Linlin Li, Luo Si. 2018. A Unified Syntax-aware Framework for Semantic Role Labeling. EMNLP : 2401-2411*
- *Yue Zhang, Rui Wang, Luo Si. 2019. Syntax-Enhanced Self-Attention-Based Semantic Role Labeling. EMNLP/IJCNLP (1) 2019: 616-626*

- Joint parsing with syntactic dependency parsing

Transition Action

M-REDUCE





Stack [S]	Buffer [M]	Queue [B]
root		root

Swabha Swayamdipta, Miguel Ballesteros, Chris Dyer and Noah A. Smith. Greedy, Joint Syntactic-Semantic Parsing with Stack LSTMs In proceedings of CoNLL (CoNLL 2016).

三、 Spark & Inspiration

1. Deep inside one task, and try to be an expert.
2. Joint A with B.
    - Lexical & sentence-level parsing
    - Syntax & semantic parsing
    - With other general domain tasks
3. Cross-learning(Joint Modeling / Joint Structure / Parameter-sharing/ …)
    - Cross Task
    - Cross Lingual
    - Cross Domain
    - Cross Standard
4. Try Models
    - Neural Transition-based Models
    - Neural Graph-based Models
5. Try some Tasks where few people take place
6. Advanced and In-depth techniques for retrofitting existing methods
    - A * search? Beam search
    - Pruning while decoding
    - Decomposition and constraint optimization
7. Transform into other modeling scheme?

Thanks
QA