

# Relation Extraction

By 邢俊劼



# Content

- Overview
- Methods & Examples
- Neural Network Approach



# Content

- Overview
- *Methods & Examples*
- *Neural Network Approach*



# What's the purpose

- Machine reading...
- Free text -> structured data





# Information extraction

- IE = extracting information from text
- The super-task of relation extraction, including:
  - Extract entities (NER)
  - Extract relations between entities (RE)
  - Extract events (such as Weibo public opinion analysis)



# Named Entity Recognition

- The task:
  - find the entities, classify them by type
  - also can be treated as a sequence tagging prob

[Eric PER] is leading [National Business Union ORG], and his office is located in [Shanghai LOC].



# RE example

- CHICAGO (AP) — Citing high fuel prices, United Airlines said Friday it has increased fares by \$6 per round trip on flights to some cities also served by lower-cost carriers. American Airlines, a unit of AMR, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increase took effect Thursday night and applies to most routes where it competes against discount carriers, such as Chicago to Dallas and Atlanta and Denver to San Francisco, Los Angeles and New York.

Subject	Relation	Object
American Airlines	subsidiary	AMR
Tim Wagner	employee	American Airlines
United Airlines	subsidiary	UAL



# Relations are unlimited

- Dozens in ACE 2003
- Thousand in Freebase, 23m entities
- types: traditional, geographical, disease breakout, protein interaction, word sense relation...
- Note that, word sense relation is quite important for NLP applications: QA, ChatBot, Summarization...;
- hyponymy, antonymy, meronymy...



# Content

- Overview
- Methods & Examples
- Neural Network Approach



# 5 easy methods

- pattern, rule based
- bootstrapping
- supervised
- distant supervised
- unsupervised



# 5 easy methods

- pattern, rule based
- bootstrapping
- supervised
- distant supervised
- unsupervised



# pattern, rule based

- For <company> appoints <person> <position>
- <people> works at <company>
- <hypernym>, such as <hyponym>
- NYU Proteus system (1997)
- Hearst: Automatic Acquisition of Hyponyms ~om Large Text Corpora



# pattern, rule based

- Agar is a substance prepared from a mixture of red algae, such as Gelidium, for laboratory or industrial use.
- What is Gelidium? How do you know that?
- Other pattern Hearst uses:
  - Y such as X ((, X)\* (, and/or) X)      such Y as X...
  - X... or other Y      X... and other Y
  - Y including X...      Y, especially X...



# pattern, rule based

## Problems:

- It requires hand-built patterns for each relation:
  - hard to write; hard to maintain
  - too many patterns to cover, may zillions
  - domain-dependent
- The accuracy is not satisfying:
  - Hearst: 66% acc on hyponym extraction



# 5 easy methods

- pattern, rule based
- bootstrapping
- supervised
- distant supervised
- unsupervised



# Bootstrapping

- When annotated text is not enough,
- But we have:
  - some seeds of the relation or some patterns works well
  - lots of unannotated text
- We can use bootstrapping



# Bootstrapping

- Example: DIPRE, extract (author, book) relation
- Starts with:

Author	Book
Isaac Asimov	The Robots of Dawn
David Brin	Startide Rising
James Gleick	Chaos: Making a New Science
Charles Dickens	Great Expectations
William Shakespeare	The Comedy of Errors

- Learn patterns and use patterns to get more instances and patterns.
- After iteration: 15,000 pairs with 95% acc.



# Bootstrapping

## Problems:

- Requires seeds for each relation
  - and the result is sensitive to the seeds
- Semantic drift at each iteration
- Precision not high
- No probabilistic interpretation (no confident)



# 5 easy methods

- pattern, rule based
- bootstrapping
- supervised
- distant supervised
- unsupervised



# Supervised

- Supervised methods requires:
  - Define output labels: true/false, relations types
  - Collect labeled data
  - Define features
  - Choose a classifier



# Supervised

ACE 2008:

Type	Subtype
ART (artifact)	User-Owner-Inventor-Manufacturer
GEN-AFF (General affiliation)	Citizen-Resident-Religion-Ethnicity, Org-Location
METONYMY*	<i>None</i>
ORG-AFF (Org-affiliation)	Employment, Founder, Ownership, Student-Alum, Sports-Affiliation, Investor-Shareholder, Membership
PART-WHOLE (part-to-whole)	Artifact, Geographical, Subsidiary
PER-SOC* (person-social)	Business, Family, Lasting-Personal
PHYS* (physical)	Located, Near



# Supervised

- Futures:
  - Bags of words, bigram before, between after entities
  - Distance between entities
  - Type of entity
  - Dependency-tree path
  - ...
- Classifiers:
  - SVM
  - logistic regression
  - naive Bayes
  - ...



# Supervised

- It can achieve high accuracy
- Problem:
  - Labeling data is expensive and time-consuming
  - It doesn't generalize to different relations, all the features are sensitive to the relation type and corpora



# 5 easy methods

- pattern, rule based
- bootstrapping
- supervised
- distant supervised
- unsupervised



# Distant supervise

- If two entities belong to a certain relation, any sentence containing those two entities is likely to express that relation
- use a database of relations to get lots of noisy training examples



# Distant supervise

- The advantages:
  - leverage rich, reliable hand-created knowledge
  - leverage unlimited amounts of text data
  - not sensitive to corpus
- Mintz, Mike, et al. "Distant supervision for relation extraction without labeled data." Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2-Volume 2. Association for Computational Linguistics, 2009.



# 5 easy methods

- pattern, rule based
- bootstrapping
- supervised
- distant supervised
- unsupervised



# unsupervised

- If two paths tend to occur in similar contexts, the meanings of the paths tend to be similar.
- Example: “X solves Y”:
  - Y is solved by X                      X resolves Y
  - X finds a solution to Y              X tries to solve Y
  - X deals with Y                        Y is resolved by X
  - X addresses Y                        ...



# Content

- Overview
- *Methods & Examples*
- Neural Network Approach



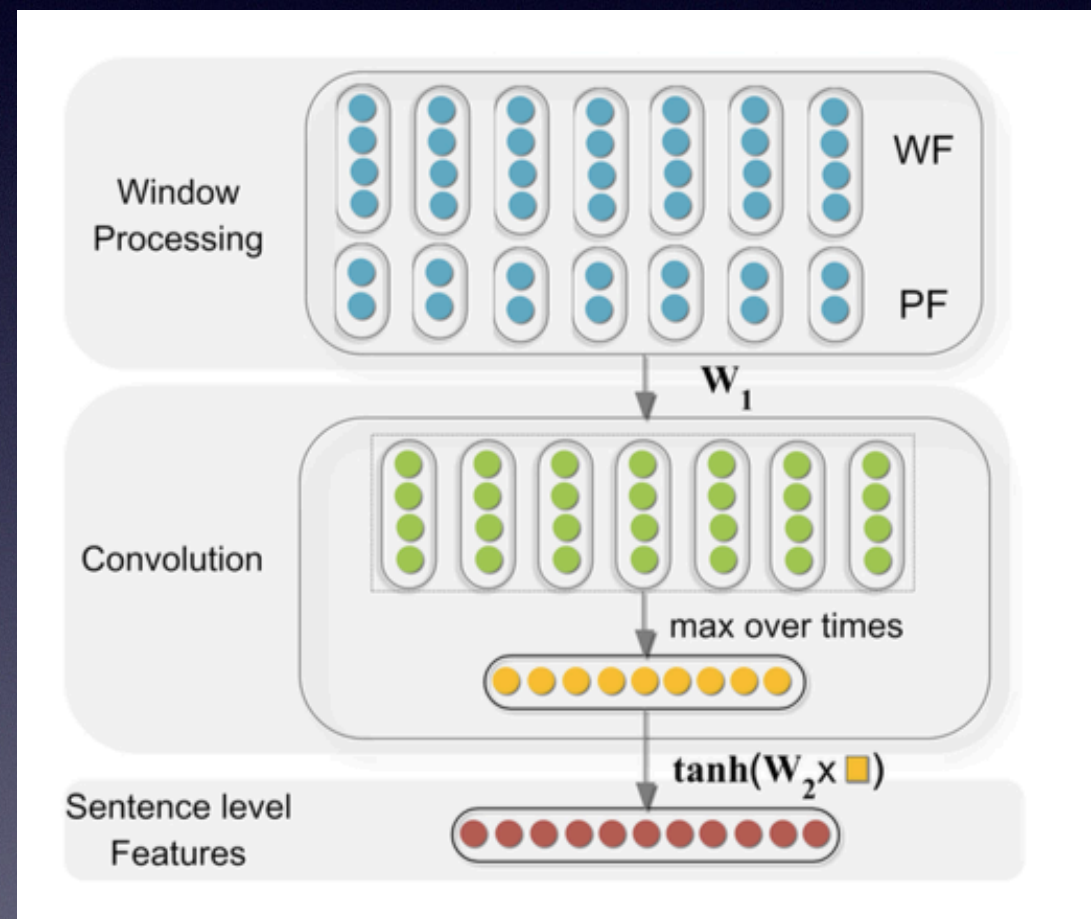
# Neural Network Approach

- Since 2013, researchers starts to use neural network to deal with the task.



# Neural Network Approach

Uses max-pooling to get sentence feature

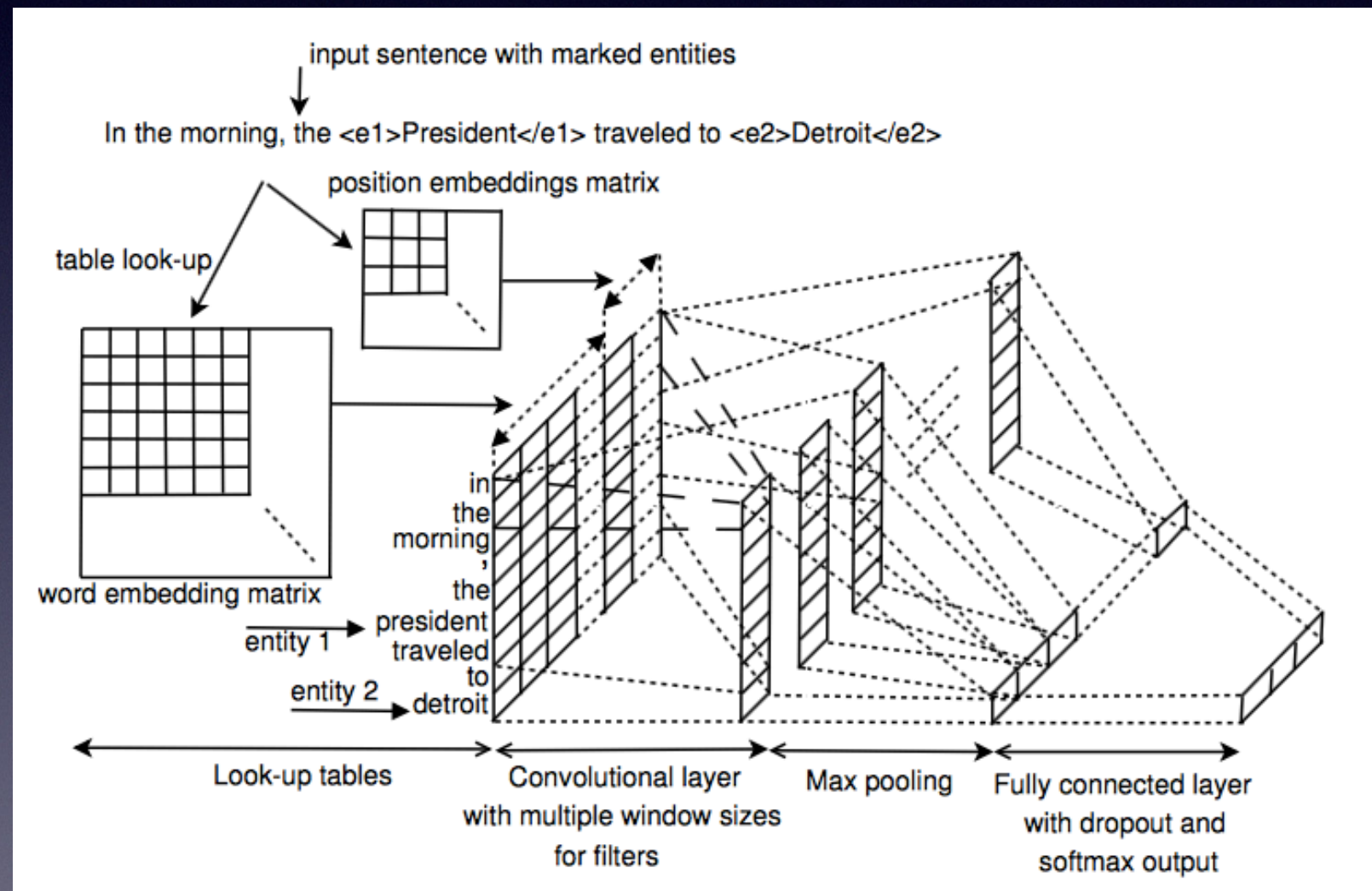


Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, Jun Zhao, et al. 2014. Relation classification via convolutional deep neural network. In COLING. pages 2335–2344.



# Neural Network Approach

- use multi-sized kernels in CNN





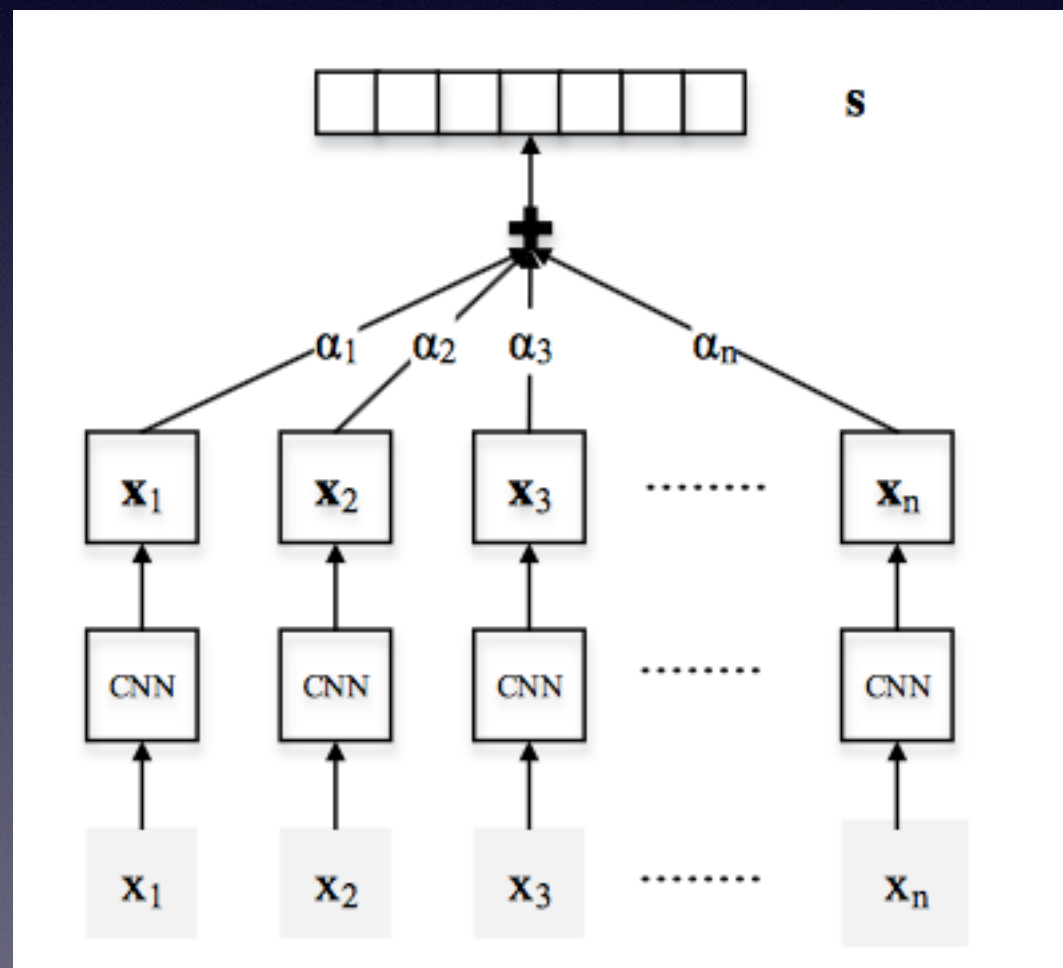
# Neural Network Approach

- multi-instance:
- If two entities belong to a certain relation, ***at least one*** sentence containing those two entities is likely to express that relation
- The assumption of distant supervise is too strong, so it is a relaxed assumption



# Neural Network Approach

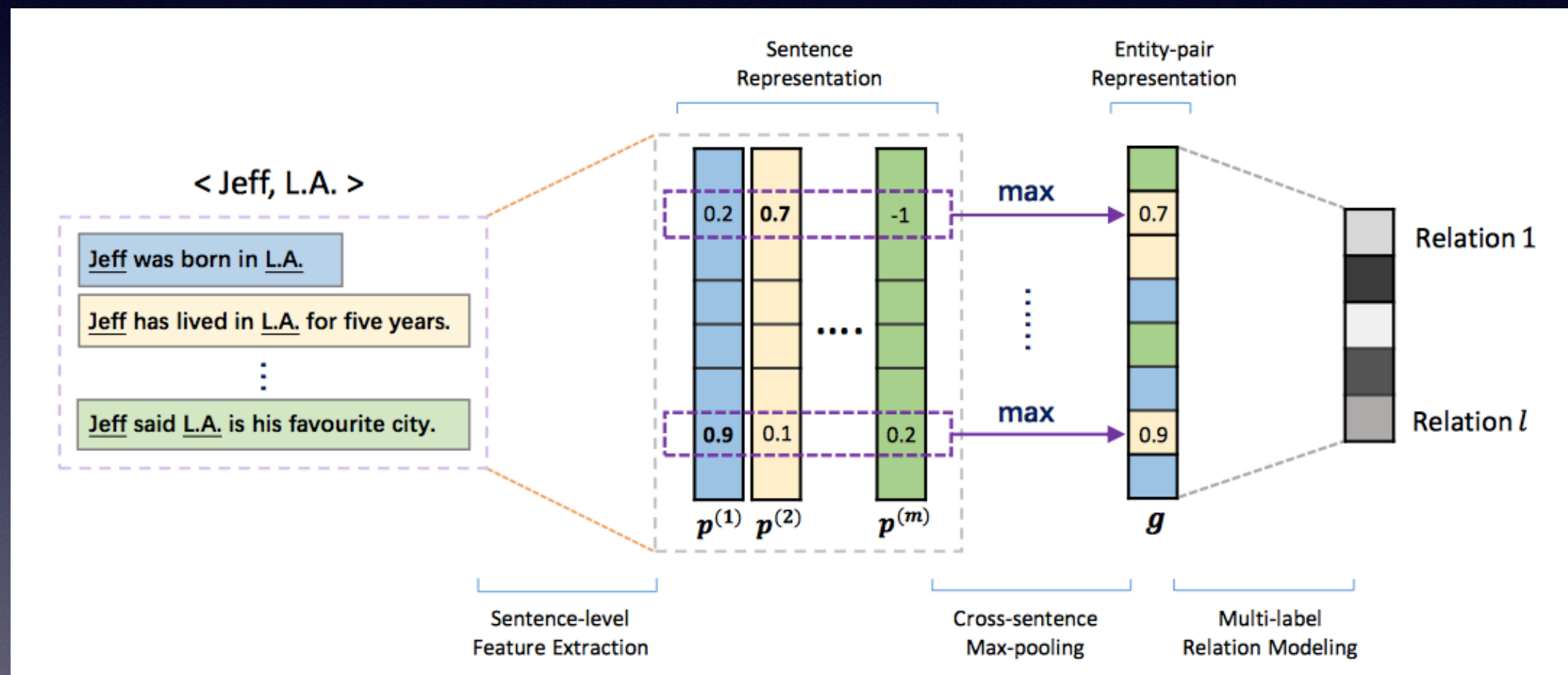
- Use attention to get as much as information from the bags of instances.





# Neural Network Approach

- Use sigmoid as the last layer

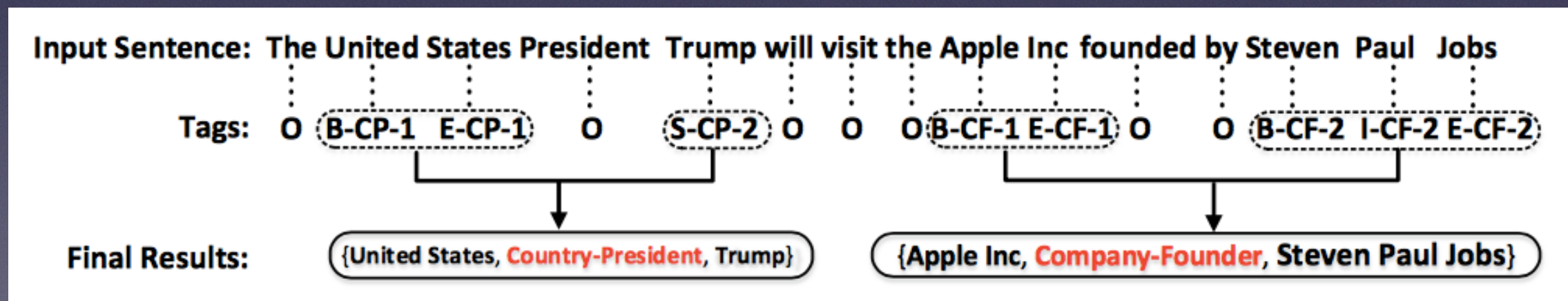
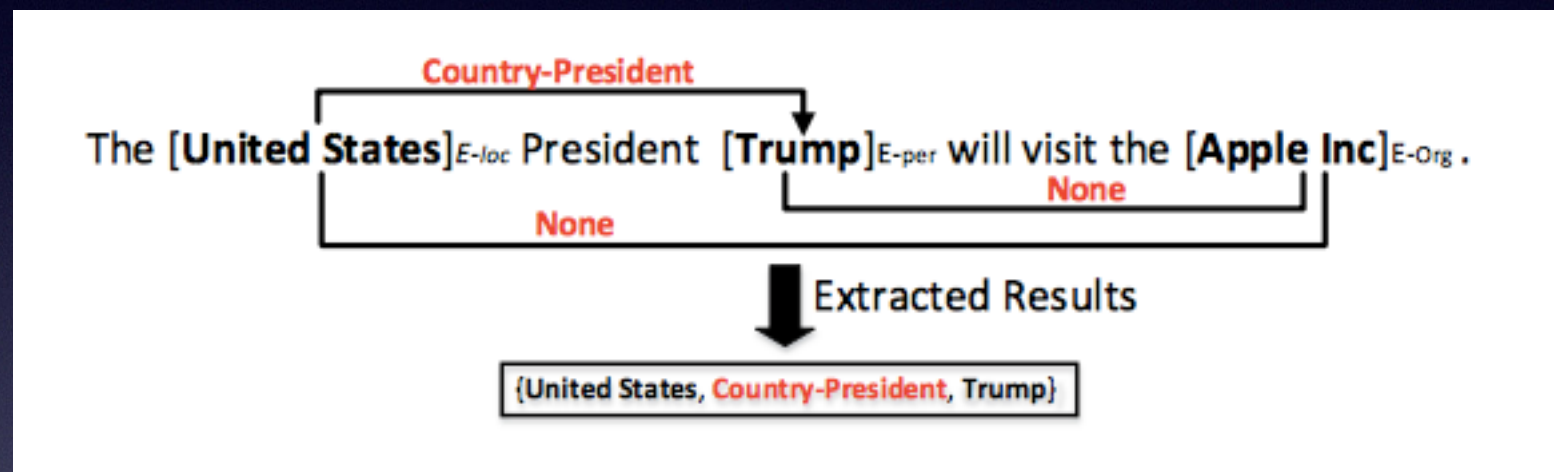


Jiang, Xiaotian, et al. "Relation Extraction with Multi-instance Multi-label Convolutional Neural Networks." *COLING*. 2016.



# Neural Network Approach

- New tagging schema to joint extraction.





# Neural Network Approach

- Katiyar, Arzoo, and Claire Cardie. "Going out on a limb: Joint Extraction of Entity Mentions and Relations without Dependency Trees." *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vol. 1. 2017.
- Ye, Hai, Wenhan Chao, and Zhunchen Luo. "Jointly Extracting Relations with Class Ties via Effective Deep Ranking." *arXiv preprint arXiv:1612.07602* (2016).
- Luo, Bingfeng, et al. "Learning with noise: enhance distantly supervised relation extraction with dynamic transition matrix." *arXiv preprint arXiv:1705.03995* (2017).
- ...



# Thanks