Relation Extraction

Motivation for relation extraction, application of relation extraction, relation types, classic approaches for relation extraction, evaluation methods.

Ruixin Yang, ryang9@ncsu.edu
PhD Student in Dr. Samatova's research lab,
Department of Computer Science
North Carolina State University

Why Relation Extraction?

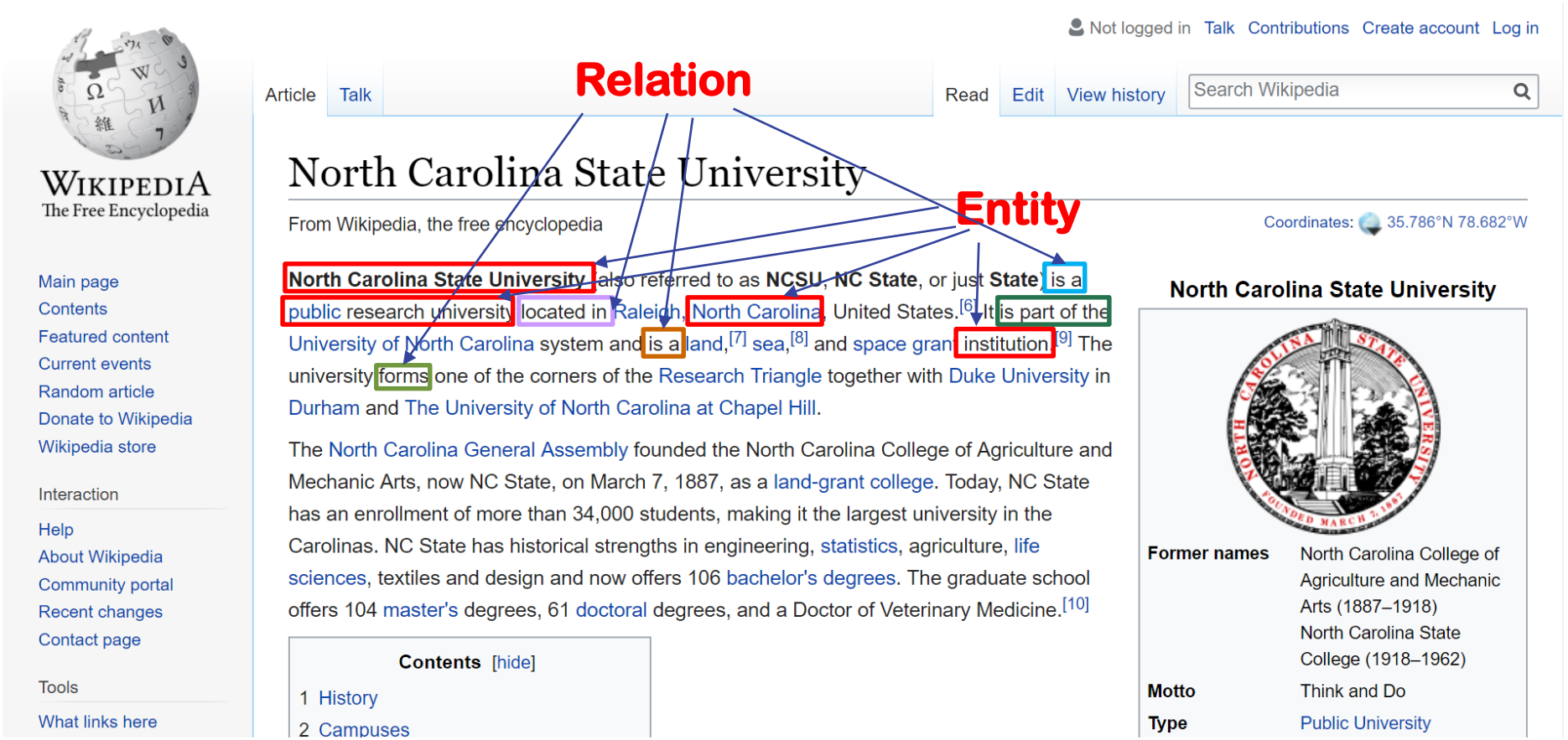
Motivation

- The world is now like a graph --- almost everything (people, organization, event and even gene, disease, protein etc.) is closed connected with each other.
- A particular set of relations between two or more entities in text can be foundation of querying answering, automated reasoning and knowledge graph construction.
- Once named entities have been identified in unstructured text data, to effectively manage, search, and mine the data, we then need to extract relation information.

Applications of relation extraction

- Question Answering (e.g., "*When was Washington born? ", " Where is key west?"*)
- Biotext Mining (e.g., "*Protein X binds with Protein Y*", "*Gene X with mutation Y leads to malignancy Z*")

Overview of Relation Extraction



Input:

- **A collection of entities.**

Output:

- **A set of relations between given entities.**

Relation types:

- **Geospatial**
- **Affiliations**
- **Ownership**
- **Employment**
- **Subsidiary**

⋮

Src: https://en.wikipedia.org/wiki/North_Carolina_State_University

Classic Approaches for Relation Extraction

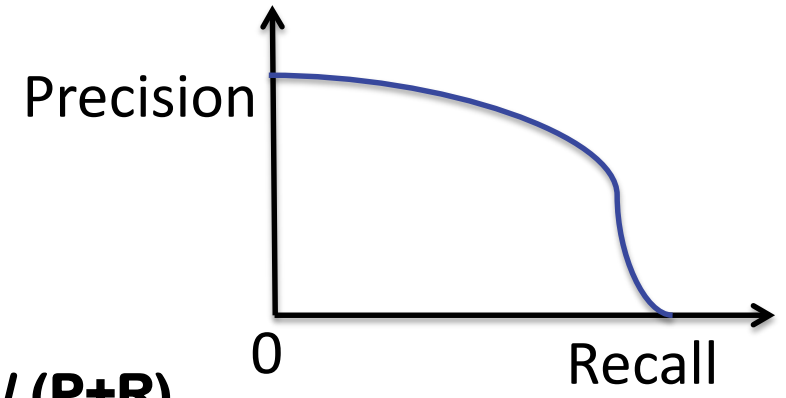
- Supervised methods (early approaches: Kambhatla et al. 2004, Zhou et al. 2005)
  - Idea: Using hand-labelled data where entities and the relations between them were annotated (e.g., MUC [3], ACE [4]).
  - Pros: Can achieve high accuracy of we have lots of hand-labeled training data.
  - Cons: Labeling is expensive and cannot generalize to different relations.

- Unsupervised methods (Banko et al. 2007, Yao et al. 2012)
  - Idea: Trying to find entity pairs that are related in a similar way by a clustering approach with no training data and predetermined list of relations.
  - Pros: Can result in very general relations and relation instances.
  - Cons: it is not possible to predict what relations will be found, and human judgement is needed to determine whether a found relation is informative.

- Distant Supervised methods (Surdeanu et al. 2012, Zeng et al. 2015)
  - Idea: using a database of relations to get lots of noisy training examples.
  - Pros: Has advantages of both supervised and unsupervised approaches.
  - Cons: Need more computation resources (GPU, TPU) when conducting deep Neural Network.

Performance Evaluation

- **Evaluation of Supervised Methods**
  - **Precision (P) = $\frac{\text{\# of correct answers given by system}}{\text{\# of answers given by system}}$**
  - **Recall (R) = $\frac{\text{\# of correct answers given by system}}{\text{total \# of possible correct answers in text}}$**
  - **F-measure (combination of Precision/Recall) = 2 * P* R/ (P+R)**



- **Evaluation of Unsupervised and Distant Supervised methods**
  - **Corpus-level task (similar as Supervised evaluation): Still using Precision, Recall and F-measure to evaluate.**
  - **Discovery new patterns and relations task: Small sample drawn randomly from the output is treated as a representative of the output and manually checked for actual relations. Then, the approximate estimate of the P, R F are calculated (Human evaluation).**

Recent Works of Relation Extraction (state of the art)

Attention-Based Bidirectional Long Short-Term Memory Networks (ACL 16) [9]

This model utilizes neural attention mechanism with Bidirectional Long Short-Term Memory Networks(BLSTM) [10] to capture the most important semantic information in a sentence, without utilizing any features derived from lexical resources or NLP systems.

Neural Relation Extraction with Selective Attention over Instances (ACL 16) [11]

The model employs convolutional neural networks (CNN) to embed the semantics of sentences and builds sentence-level attention over multiple instances, which is expected to dynamically reduce the weights of those noisy instances.

Neural Relation Extraction with Multi-lingual Attention (ACL 17) [12]

This model introduces a multi-lingual neural relation extraction framework which employs monolingual attention to utilize the information within mono-lingual texts. And cross-lingual attention was proposed to consider the information consistency and complementarity among cross-lingual text.

Recent Works of Relation Extraction (state of the art)

Deep Residual Learning for Weakly-Supervised Relation Extraction (EMNLP 2017) [13]

This work designs a novel convolutional neural network (CNN) with residual learning, and investigates its impacts on the task of distantly supervised noisy relation extraction. In contradictory to popular beliefs that ResNet only works well for very deep networks, this model performs well even using 9 layers of CNNs.

Cross-Sentence N-ary Relation Extraction with Graph LSTMs (TACL 17)[14]

The model uses a general relation extraction framework based on graph long short-term memory networks (graph LSTMs), which can be easily extended to cross-sentence n-ary relation extraction. The graph formulation provides a unifying way to explore different LSTM approaches and incorporate various intra-sentential and inter-sentential dependencies, such as sequential, syntactic, and dis-course relations.

Joint extractions of entities and relations based on a novel tagging scheme (ACL 17) [15]

This work proposes a novel tagging scheme converts the joint extraction task to a tagging problem and utilizes a different end-to-end models to extract entities and their relations directly, without identifying entities and relations separately.

Relation Extraction Software

Model	Link	Language
ReVerb (ACL 11)	Washington.reveb	Java
Stanford Relation Extractor	Stanford.nlp	Java
Ollie (EMNLP 12)	Washington.ollie	Java
Open IE (ACL 15)	Stanford.ie	Python
Natural Language Toolkit	nltk	Python
MITIE	MIT.ie	Python/R/Java/C/C++
Att-BiLSTM (ACL 16)	Tsinghua.nlp	Python
NRE_SAI (ACL 16)	Tsinghua.nre	C++
NRE_MA (ACL 17)	Tsinghua.nre	Python
ResCNN (EMNLP 2017)	Taiwan.res	Python
CS_N-ary (TACL 2017)	Johns_Hopkins.nary	Python
Joint_E (ACL 17)	CAS.joint	Python
GATE	Sheffield.gate	Open source software

References

[1] Kambhatla, Nanda. "Combining lexical, syntactic, and semantic features with maximum entropy models for extracting relations." Proceedings of the ACL 2004 on Interactive poster and demonstration sessions. Association for Computational Linguistics, 2004.

[2] Zhou, et al. "Exploring various knowledge in relation extraction." Proceedings of the 43rd annual meeting on association for computational linguistics. Association for Computational Linguistics, 2005.

[3] http://www.itl.nist.gov/iaui/894.02/related_projects/muc/

[4] https://catalog.ldc.upenn.edu/ldc2006t06

[5] Banko, Michele, et al. "Open information extraction from the web." IJCAI. Vol. 7. 2007.

[6] Yao, Limin, Sebastian Riedel, and Andrew McCallum. "Unsupervised relation discovery with sense disambiguation." Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1. Association for Computational Linguistics, 2012.

[7] Surdeanu, Mihai, et al. "Multi-instance multi-label learning for relation extraction." Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning. Association for Computational Linguistics, 2012.

[8] Zeng, Daojian, et al. "Distant supervision for relation extraction via piecewise convolutional neural networks." Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015.

9

References

[9] Zhou, Peng, et al. "Attention-based bidirectional long short-term memory networks for relation classification." Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Vol. 2. 2016.

[10] Zhang, Shu, et al. "Bidirectional long short-term memory networks for relation classification." Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation. 2015.

[11] Lin, Yankai, et al. "Neural relation extraction with selective attention over instances." Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vol. 1. 2016.

[12] Lin, Yankai, Zhiyuan Liu, and Maosong Sun. "Neural Relation Extraction with Multi-lingual Attention." Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vol. 1. 2017.

[13] Huang, Yi Yao, and William Yang Wang. "Deep Residual Learning for Weakly-Supervised Relation Extraction." arXiv preprint arXiv:1707.08866 (2017).

[14] Peng, Nanyun, et al. "Cross-sentence n-ary relation extraction with graph lstms." arXiv preprint arXiv:1708.03743 (2017).

[15] Zheng, Suncong, et al. "Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme." arXiv preprint arXiv:1706.05075 (2017).

10