# CCKS2019 回顾

胡聪 2019/09/03

- ATT前沿技术讲习班
- Advanced work
- School work

# Natural Language Interface to Knowledge Graph (our experience)

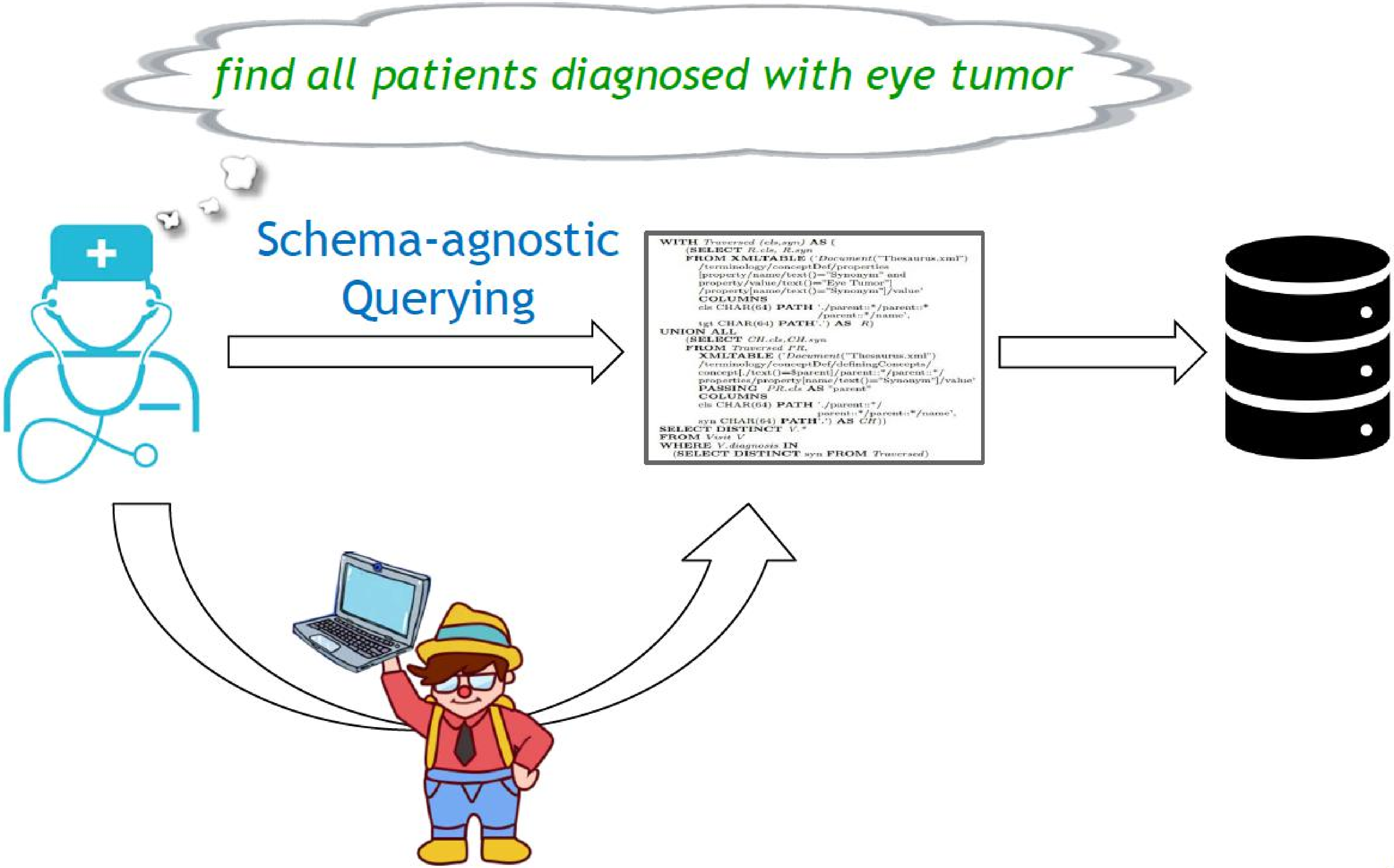
Xifeng Yan

**Dissertation slides** from Semih Yavuz, Izzeddin Gur, and Yu Su

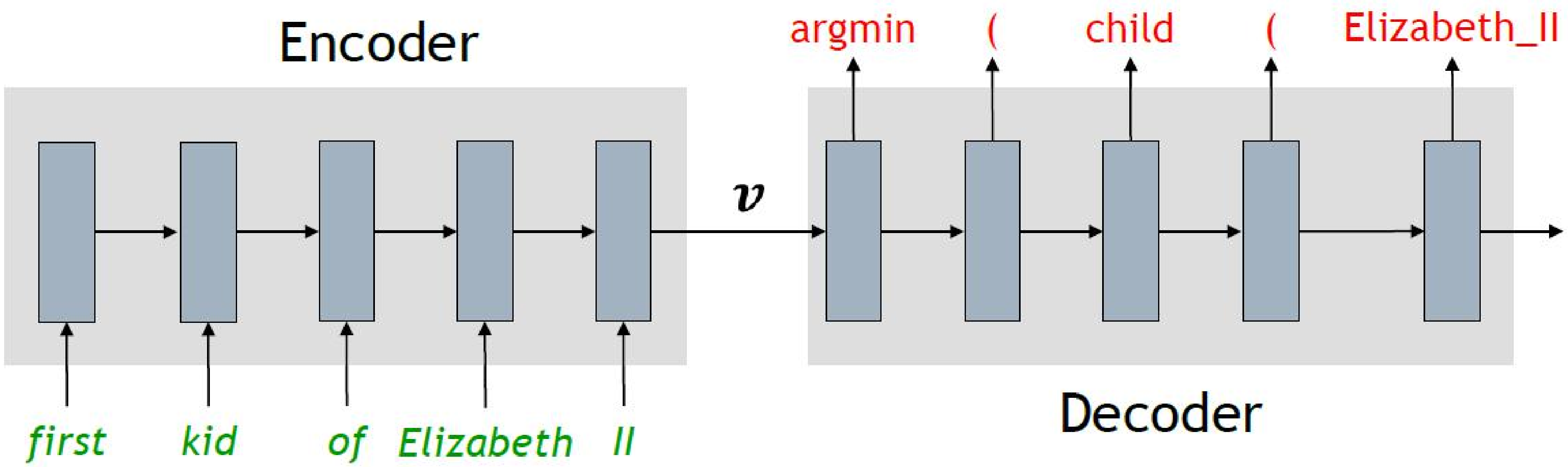
Department of Computer Science

University of California, Santa Barbara

# In Pursue of Efficiency



# Neural Natural Language Interface




*eldest kid of Elizabeth II*

*Her Majesty's first child*

*Queen Elizabeth's firstborn*

$$p(y_1 = \text{"argmin"} \mid \boldsymbol{v})$$

$$\downarrow$$

$$p(y_2 = \text{"("} \mid \boldsymbol{v}, \boldsymbol{y_1})$$

$$\downarrow$$

$$p(y_3 = \text{"child"} \mid \boldsymbol{v}, \boldsymbol{y_{1:2}})$$

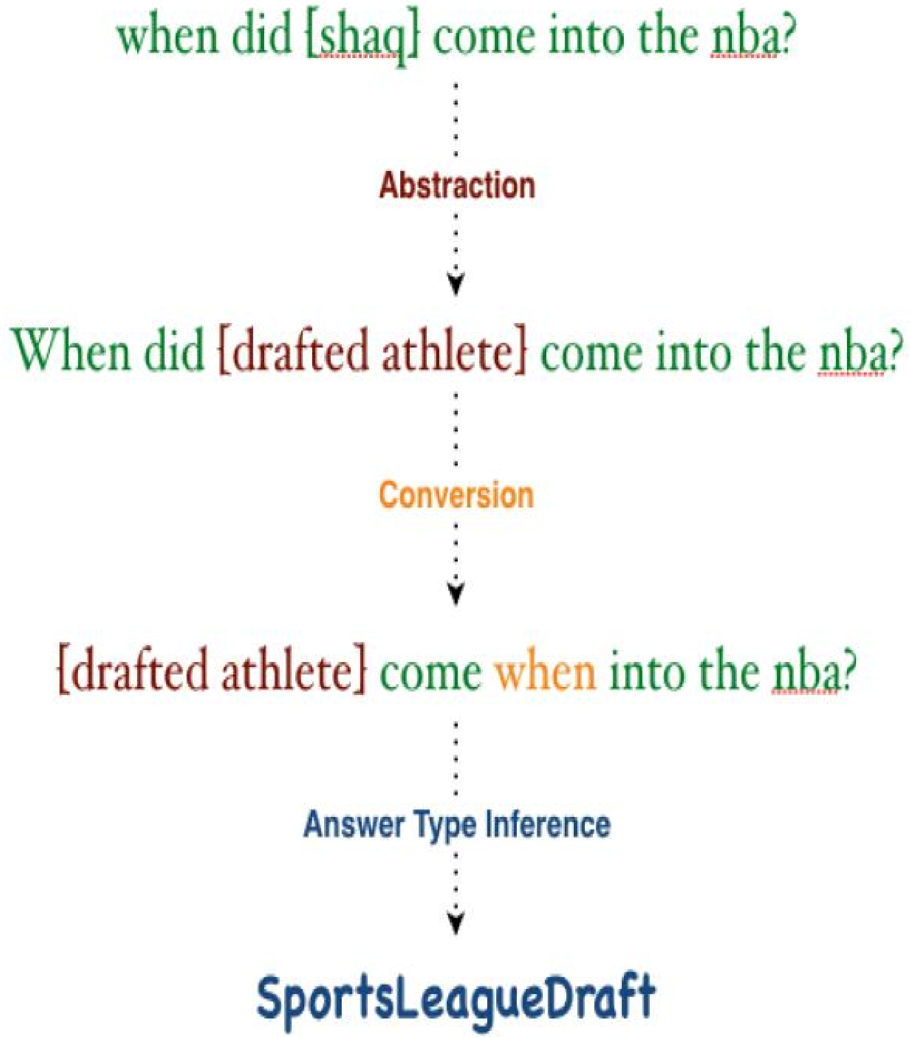
$$\downarrow$$

$$\ldots$$

[Dong & Lapata 2016], [Jia & Liang 2016], [Mei et al., 2016]

1、what if the correct answer type is enforced

去猜测question对应的answer type，使用answer type信息

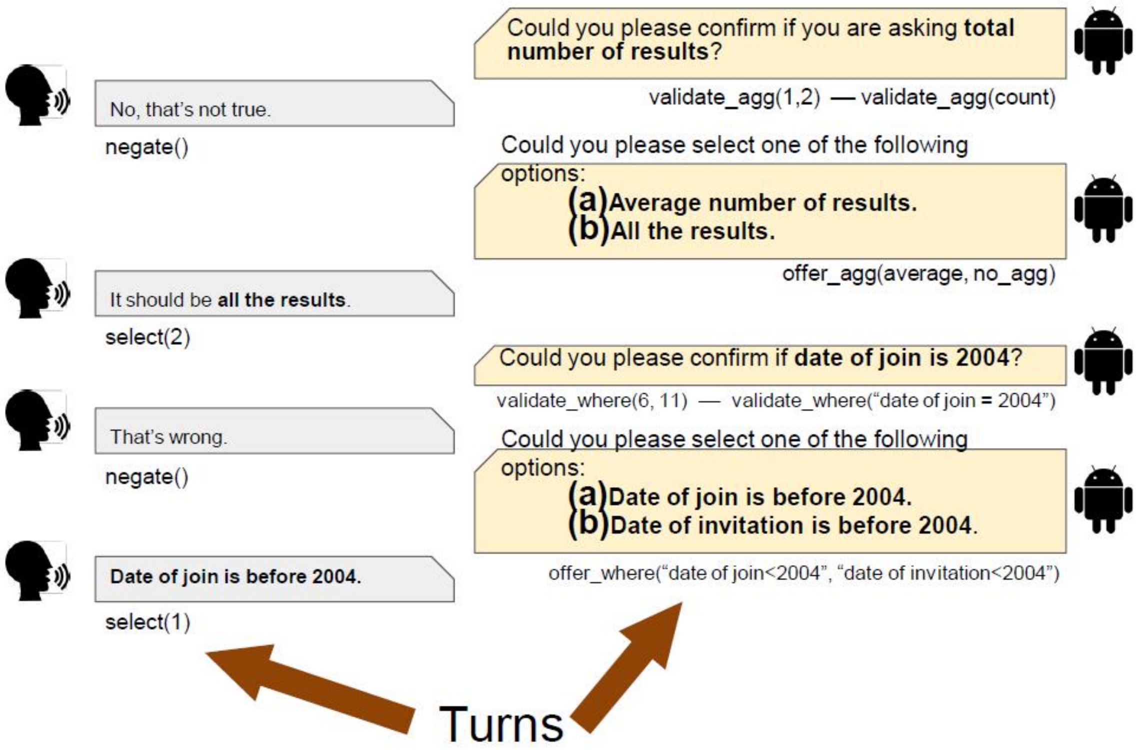



| Ranking | F1 | # Improved Qs |
|---|---|---|
| AgendaIL | 49.7 | - |
| w/ Oracle Types@10 | 57.3 | +234 |
| w/ Oracle Types@20 | 58.7 | +282 |
| w/ Oracle Types@50 | 60.1 | +331 |
| w/ Oracle Types@All | 60.5 | +345 |

Improving semantic parsing via answer type inference.[S Yavuz et al. EMNLP 2016]

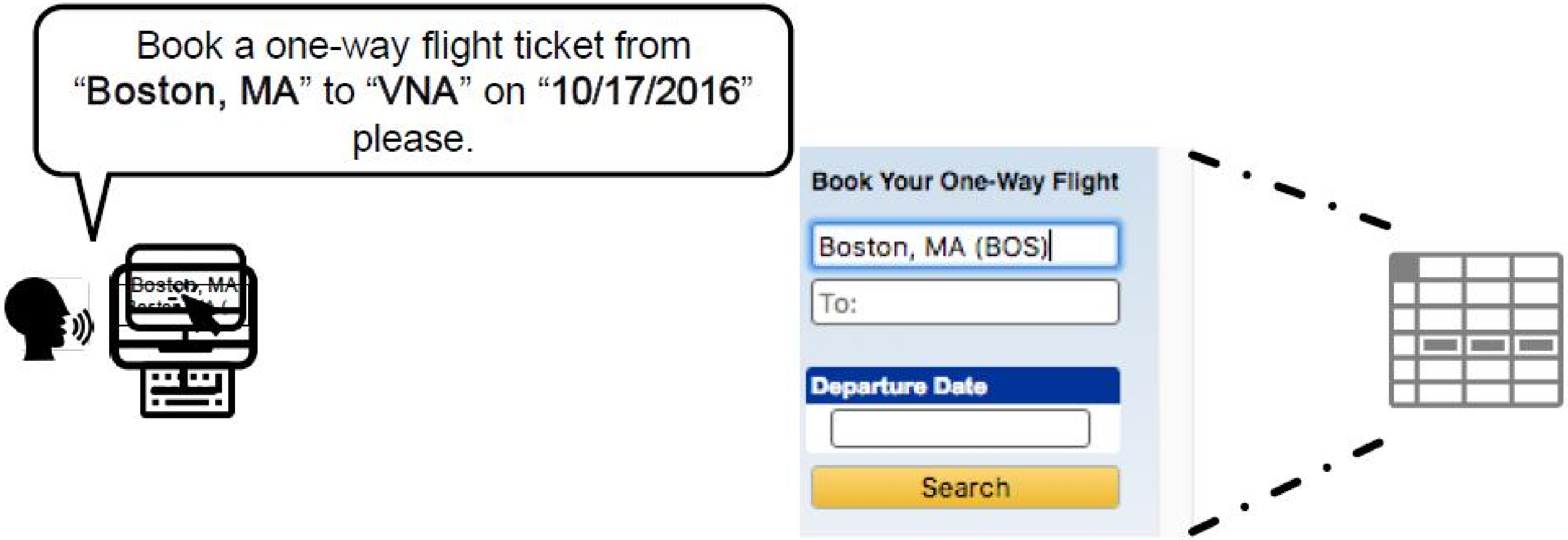
# SQL Generation via Task-Oriented Dialogue

| | |
|---|---|
| **User Question :** | *What are the countries that joined NATO before 2004 ?* |
| **GT SQL Query :** | SELECT country WHERE date of join < 2004 |
| **Candidate SQL Query :** | SELECT **count** ( country ) WHERE date of join = 2004 |




Yaghmazadeh, Navid, et al. "SQLizer: query synthesis from natural language." Proceedings of the ACM on Programming Languages 1.OOPSLA(2017):1-26.
Izzeddin Gur (UCSB), Semih Yavuz (UCSB), Yu Su (OSU), Xifeng Yan (UCSB) , DialSQL: Dialogue-Based Structural Query Generation (ACL'18)

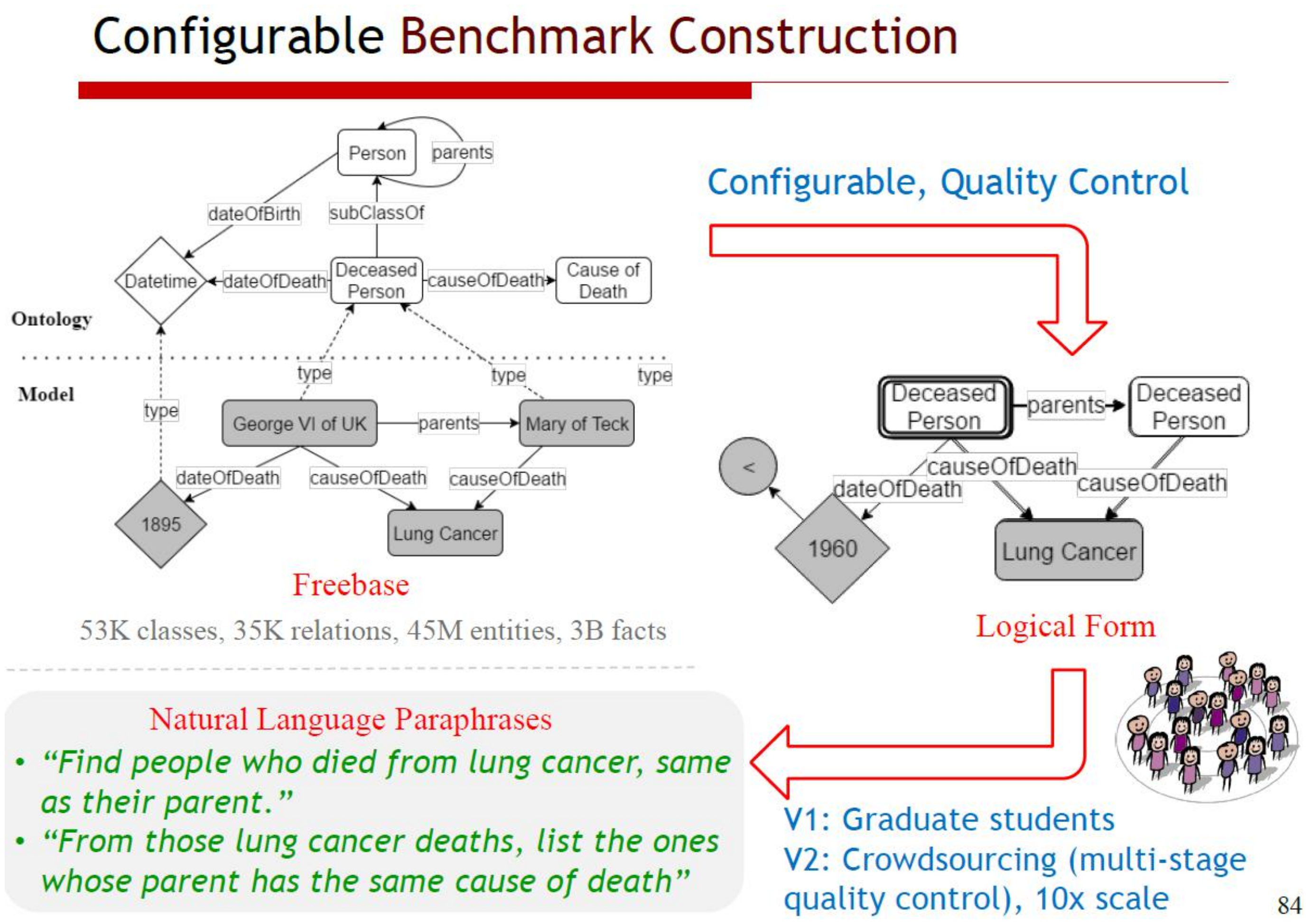
# Learning from User Interfaces that connected to knowledge graph and databases




Web Interfaces to KB/DBs

* Izzeddin Gur (UCSB), Ulrich Rueckert (Google AI), Aleksandra Faust (Google AI), Dilek Hakkani-Tur (Amazon Alexa), *Learning to Navigate the Web (ICLR '19)*
* *Izzeddin Gur (UCSB),* Xifeng Yan (UCSB), *Learning Hierarchical Policies for Navigating Web Pages (preprint '19)*

*"I want to build an NLI for my domain, but I don't have any user and training data"*




Su, Yu, et al. "On generating characteristic-rich question sets for qa evaluation." [EMNLP 2016]

# SUMMARY

| | |
|---|---|
| **1** | **Enforce answer type** |
| 2 | Task-oriented Dialog |
| 3 | User-interface/ web interface |
| 4 | Generate NLI data from KGs |

## Future research

1、**how to verify the correctness of answer**
2、conversation between human conversation and computer \how to combine QA and human interactive
3、complex query
4、common sense knowledge

# From Data to Model Programing: Injecting Structured Priors for Knowledge Extraction

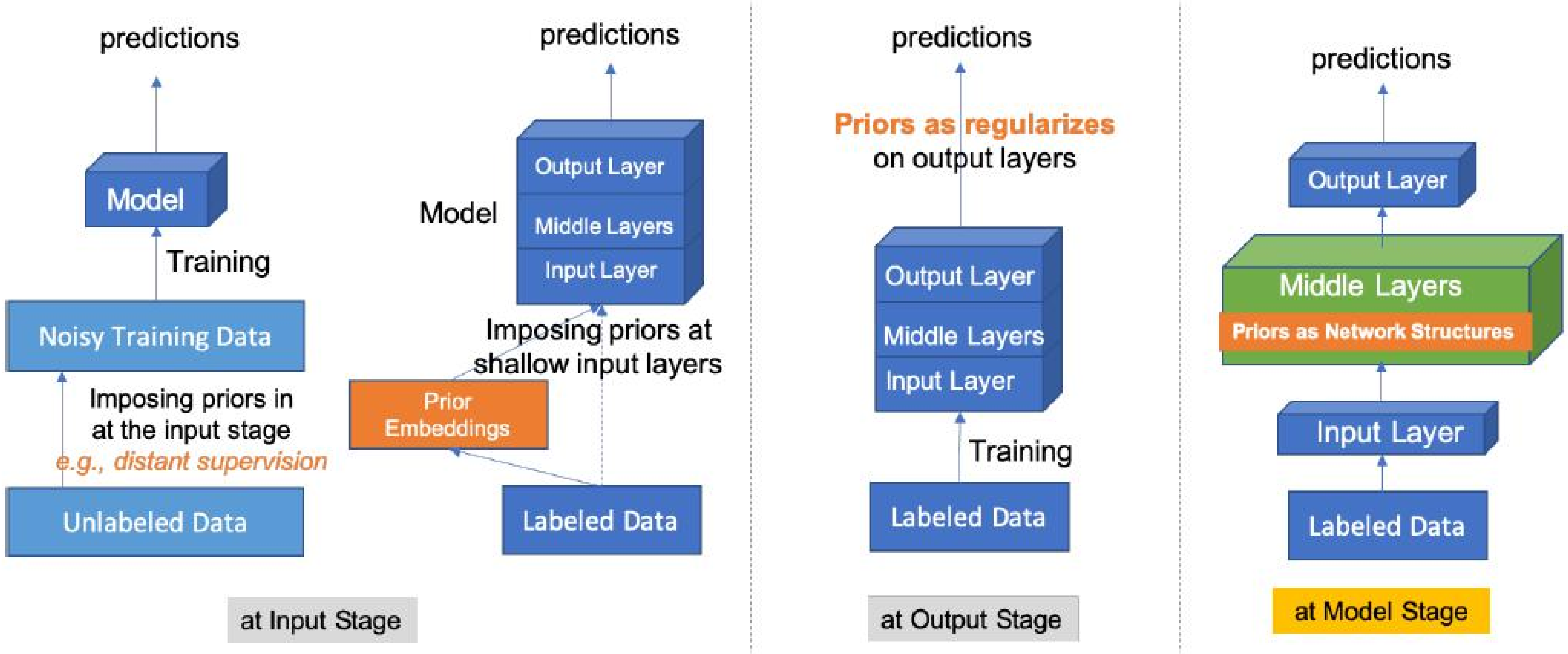
Xiang Ren

*Department of Computer Science, USC*

*USC Information Science Institute*

*USC Machine Learning Center*

# Structured Prior Knowledge




*Learning named entity tagger from domain dictionary* (Shang et al., EMNLP 2018)

*Neural rule grounding* (Zhou et al., 2019)

*KagNet: Learning to Answer Commonsense Questions with Knowledge-aware Graph Networks* (Lin et al., 2019)

# Learning Named Entity Tagger using *Domain-Specific Dictionary*

EMNLP 2018

*Joint work with Jingbo Shang, Lucas Liu, Xiaotao Gu*

# AutoNER: “Tie-or-Break” Schema

- **Label the relationship of two consecutive tokens:**
  - **Tie**, when the two tokens are matched to the same entity
  - **Unknown**, if at least one of the tokens belongs to an *out-of-dictionary phrase*
  - **Break**, otherwise.

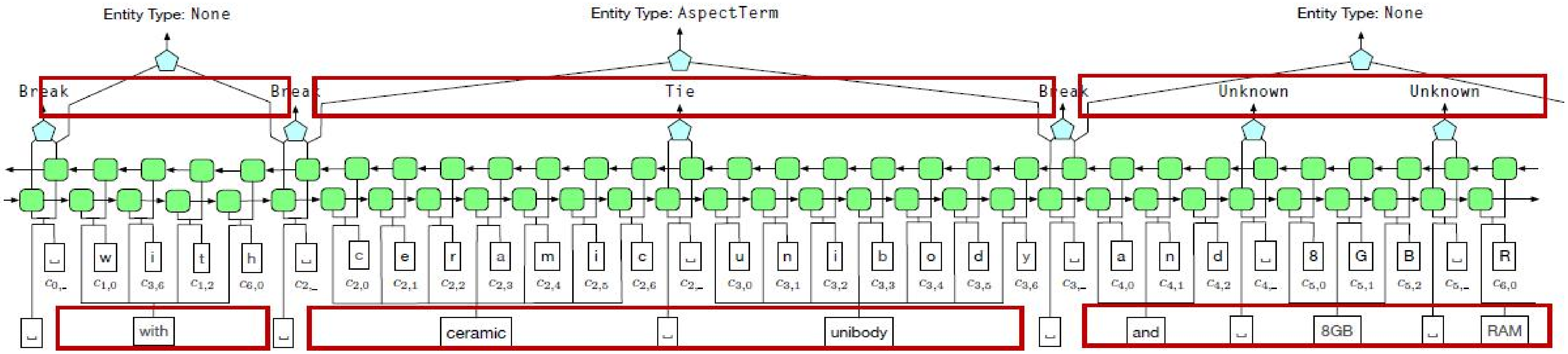
| | *Today is **Wednesday*** | *Today is Wednesday.* |
|---|---|---|
| **BIOES** | O O S-PER | O O O |
| **“Tie-or-Break”** | Break Break | Break Break |

| | *Ceramic body and 8GB RAM* | *Ceramic body and 8GB RAM* |
|---|---|---|
| **BIOES** | B-ASP E-ASP O O O | B-ASP E-ASP O O O |
| **“Tie-or-Break”** | Tie Break Break Break | Tie Break Break Unknown |

Shang J, Liu L, Ren X, et al. Learning named entity tagger using domain-specific dictionary[J]. arXiv preprint arXiv:1809.03599, 2018.

# AutoNER: Multi-task Prediction of Entity *Spans* & *Types*

- char-BiLSTM for learning contextualized representation
- 1st classification layer – “tie” or “break”
- *candidate entity spans* – merge token(s) between two “break”s

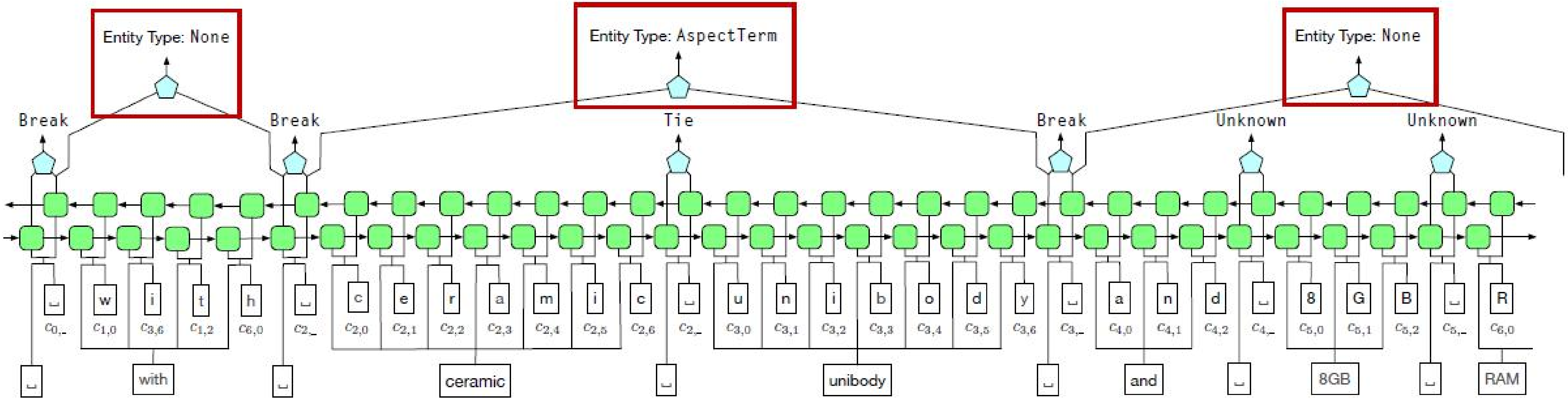



(Shang et al., EMNLP 2018)

# AutoNER: Multi-task Prediction of Entity *Spans* & *Types*

- 2nd classification layer – determine entity types

multi-class cross-entropy



# Results on Biomedical Domain

- BC5CDR NER dataset: **chemical & disease**
- Fuzzy-LSTM-CRF: models tokens with "unknown" label
- AutoNER: *close to model trained on clean labeled data*

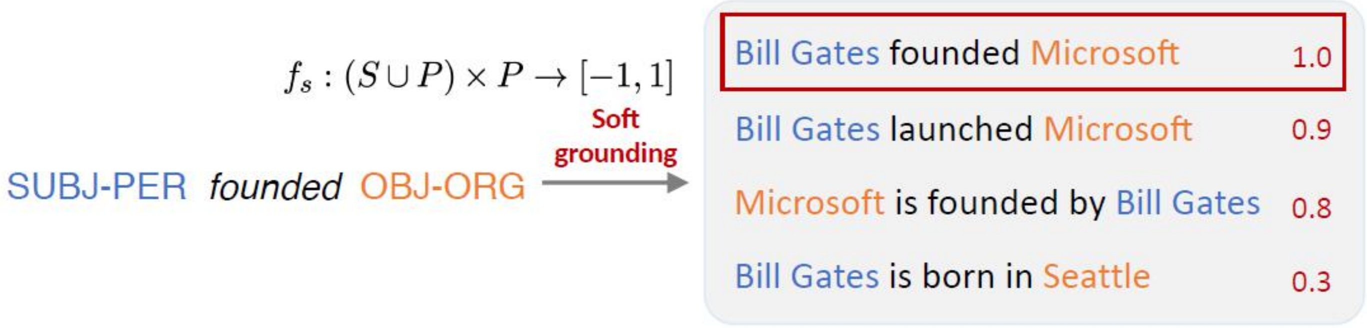
| Method | Precision | Recall | F1 |
|---|---|---|---|
| Dictionary Matching (DM)* | 93.93 | 58.35 | 71.98 |
| Fuzzy-LSTM-CRF<br>(DM + label cleaning & augmentation) | 88.27 | 76.75 | 82.11 |
| **AutoNER** | 88.96 | 81.00 | **84.80** |
| LM-LSTM-CRF on gold-standard | 88.84 | 85.16 | <u>86.96</u> |

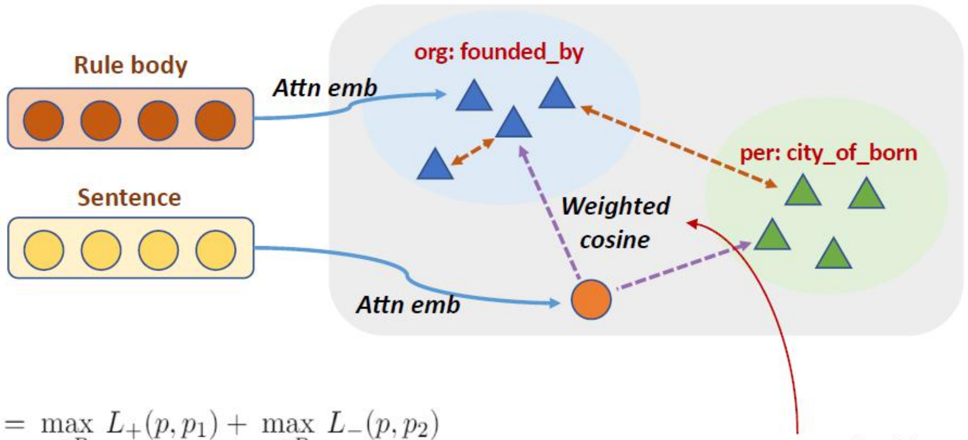
*CTD Chemical and Disease vocabularies: 322,882 Chemical and Disease entity names.

# *Neural Rule Grounding* for Low-Resource Relation Extraction

Joint work with Wenxuan Zhou & Hunter Lin, *under submission*

$$f_s(s,p)$$



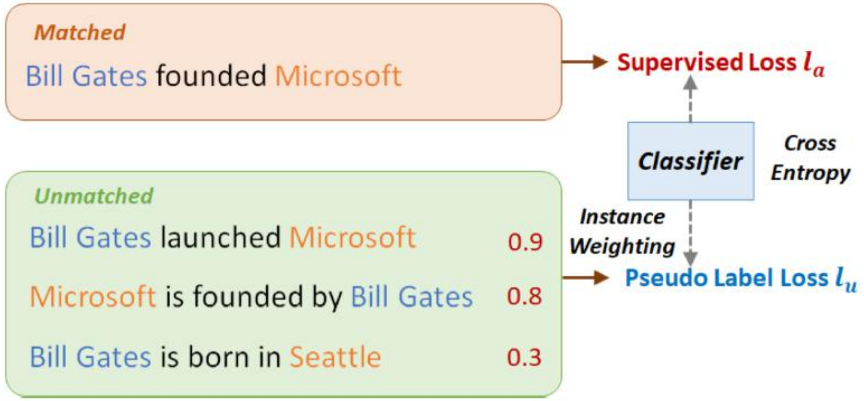

$$l_{sim} = \max_{p_1 \in P_+} L_+(p, p_1) + \max_{p_2 \in P_-} L_-(p, p_2)$$

$$L_+ = (\tau_+ - f(p, p_1))_+^2$$

$$L_- = (f(p, p_2) - \tau_-)_+^2$$

$$f_s(W_1, W_2) = \frac{z_1^T D^T D z_2}{\|z_1 D\| \|z_2 D\|}$$

- Perfect matching → score = 1
- Other cases → score = ?




$$u_s = f(s, p^*)$$

$$w_s = \frac{\exp(\theta u_s)}{\sum_{i=1}^{N_b} \exp(\theta u_i)}$$

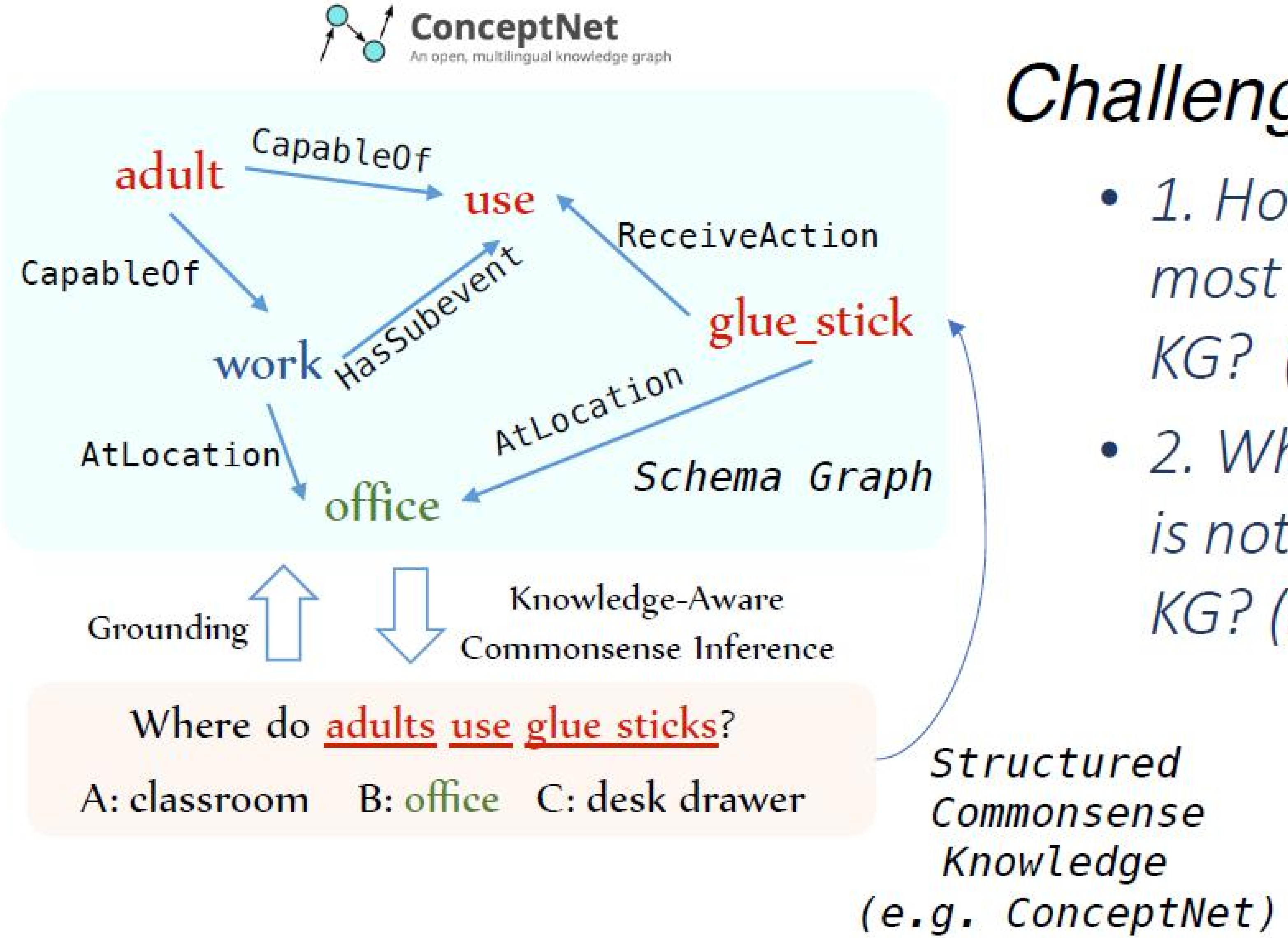
$$l_u = -\sum_{i=1}^{n} w_i \cdot \log p(r'_s | s)$$

Assign each unmatched sentence a pseudo label and weight by soft matching.

# **KagNet**: Learning to Answer Commonsense Questions with *Knowledge-aware Graph Networks*

*Joint work with Bill Lin & Jamin Chen, under submission*

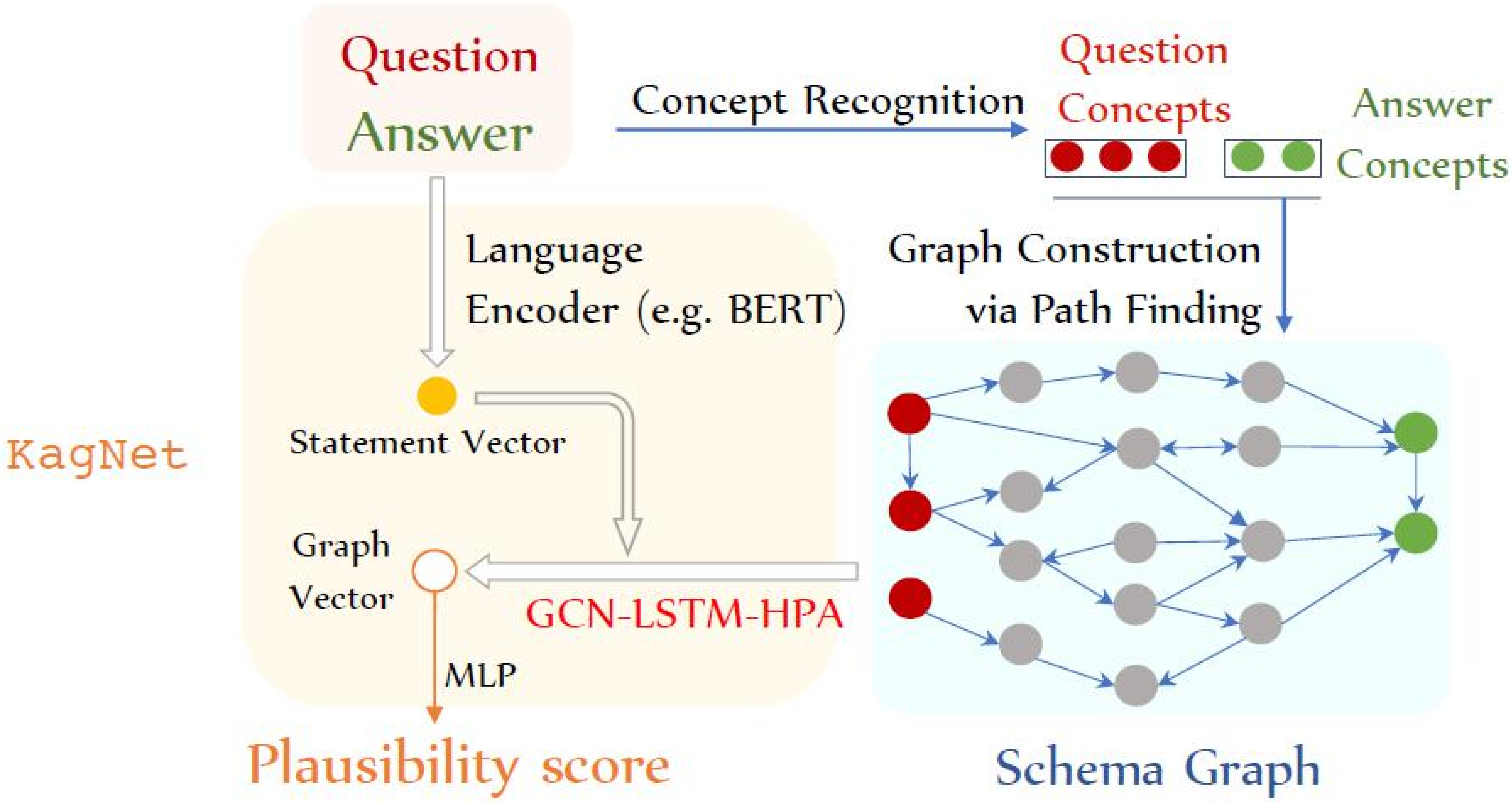
# Our Idea: Imposing External Knowledge



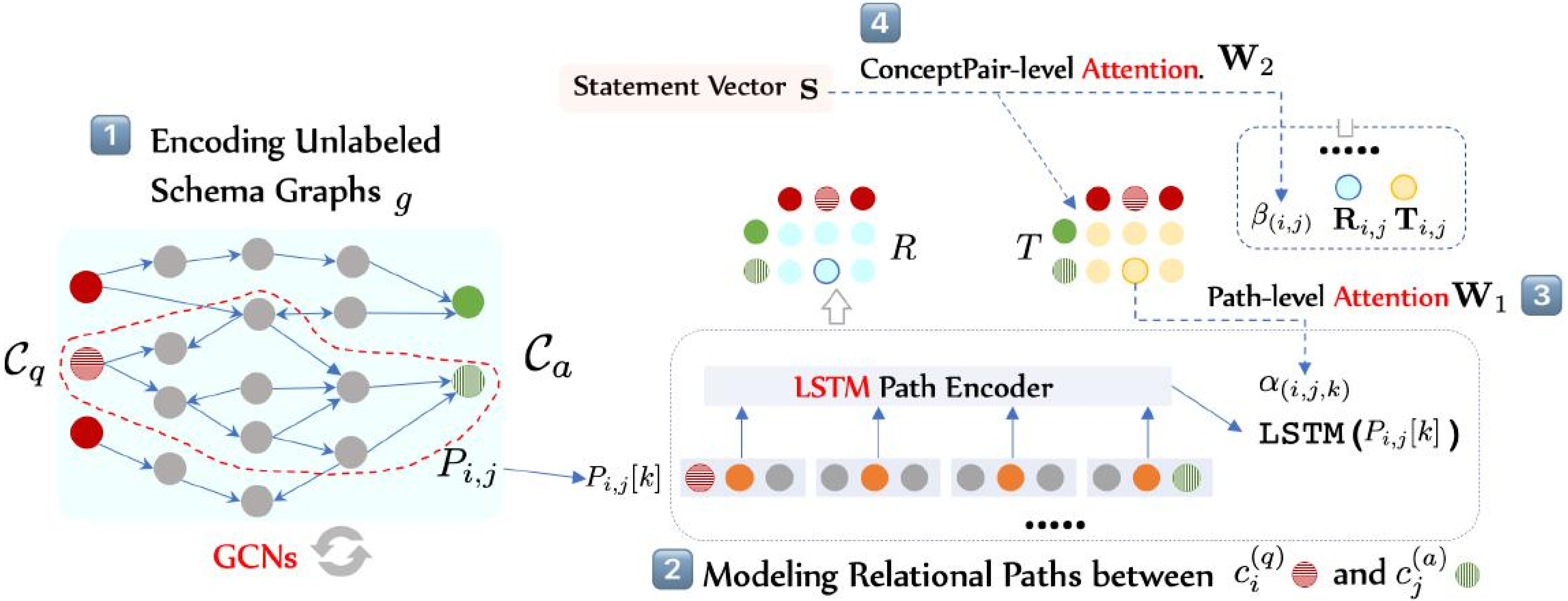

## *Challenges:*

- *1. How can we find the most relevant paths in KG? (noisy)*
- *2. What if the best path is not existent in the KG? (incomplete)*

# KagNet: Knowledge-Aware Graph Networks



# The GCN-LSTM-HPA Architecture



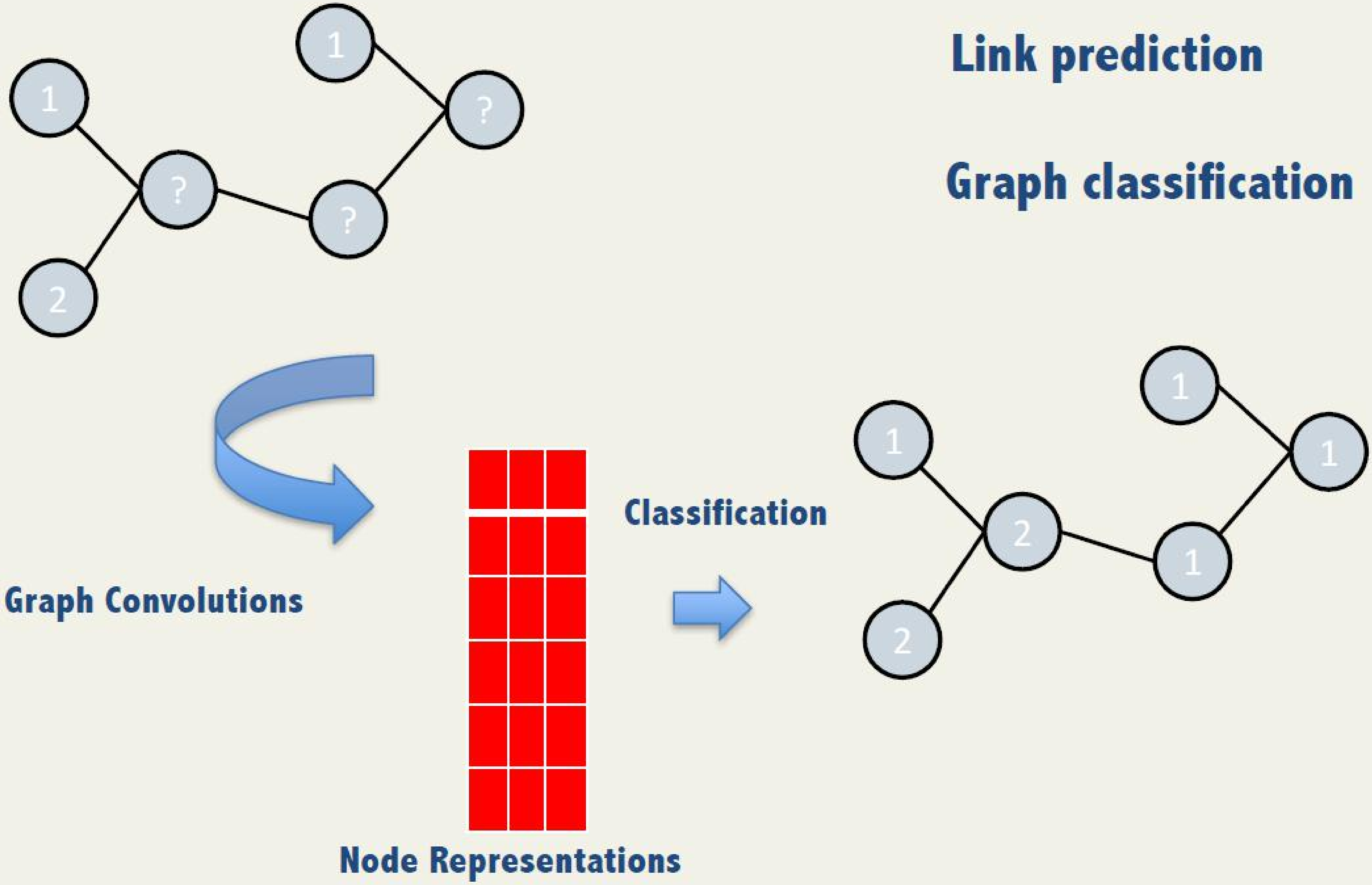
# Deep Learning on Graphs

**Jiliang Tang**

**Data Science and Engineering Lab**

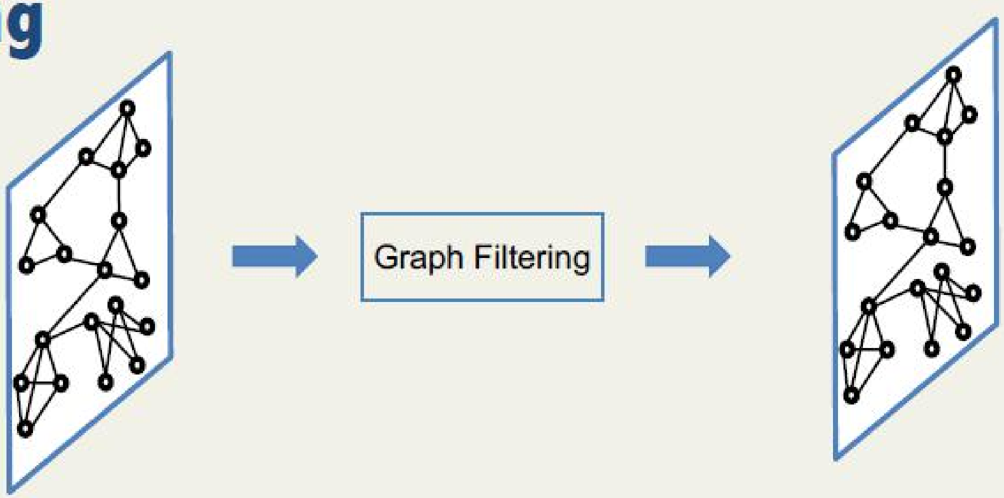
**Michigan State University**

https://www.cse.msu.edu/~tangjili/, tangjili@msu.edu


Link prediction
Graph classification
1
1
?
?
?
2
Graph Convolutions
Node Representations
Classification
1
1
2
1
1
2

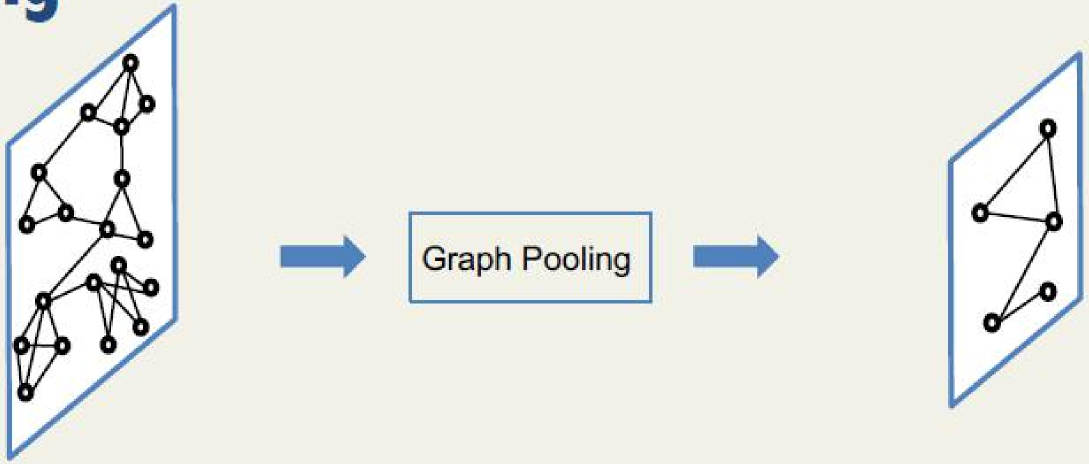
# Two Main Operations in GNN

- **Graph Filtering**



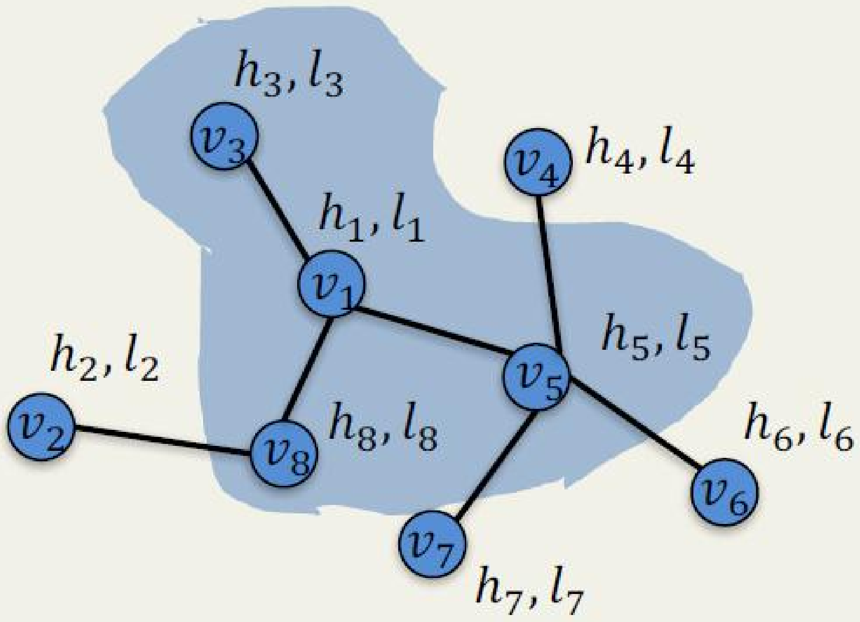

$A \in \{0,1\}^{n \times n}, X \in R^{n \times d}$ $\quad$ $A \in \{0,1\}^{n \times n}, X_f \in R^{n \times d_{new}}$

- **Graph Pooling**




$A \in \{0,1\}^{n \times n}, X \in R^{n \times d}$ $\quad$ $A_p \in \{0,1\}^{n_p \times n_p}, X_p \in R^{n_p \times d_{new}}, n_p < n$

# Graph Filtering in the First GNN Paper




$h_i$: The "state" of node $v_i$ (The features that are updated during the procedure).

$l_i$: The "label" of node $v_i$ (The input features that are fixed during the procedure).

$$h_i^{(k+1)} = \sum_{v_j \in N(v_i)} f\left(l_i, h_j^{(k)}, l_j\right), \qquad \forall\, v_i \in V.$$

$N(v_i)$: Neighbors of the node $v_i$.

$f(\cdot)$: Feedforward neural network.

**Adjacency Matrix**

$A[i,j] = 1$ if $v_i$ adjacent to $v_j$

$A[i,j] = 0$ otherwise

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

$$f(H^{(l)}, A) = \sigma\left(AH^{(l)}W^{(l)}\right)$$

$$f(H^{(l)}, A) = \sigma\left(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right)$$

H: eigen matrix (特征矩阵)

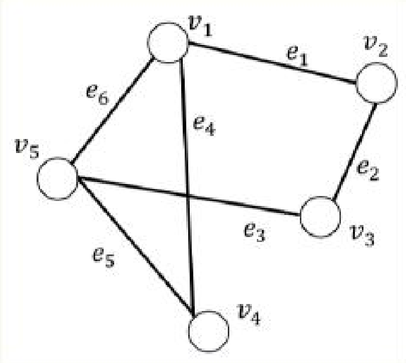
W: parameters

$\widetilde{A}$：A+I , I是单位矩阵

# Spectral Graph Theory

**Graph**

$G = \{V, E\}$




**Adjacency Matrix**

$A[i, j] = 1$ if $v_i$ adjacent to $v_j$

$A[i, j] = 0$ otherwise

$$\mathrm{D} = diag(degree(v_1), \ldots, degree(v_N))$$

**Laplacian Matrix**

$\mathrm{D} = diag(degree(v_1), \ldots, degree(v_N))$

$L = D - A$

$$\mathbf{L} = \begin{pmatrix} 3 & -1 & 0 & -1 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & 0 & -1 \\ -1 & 0 & 0 & 2 & -1 \\ -1 & 0 & -1 & -1 & 3 \end{pmatrix}$$

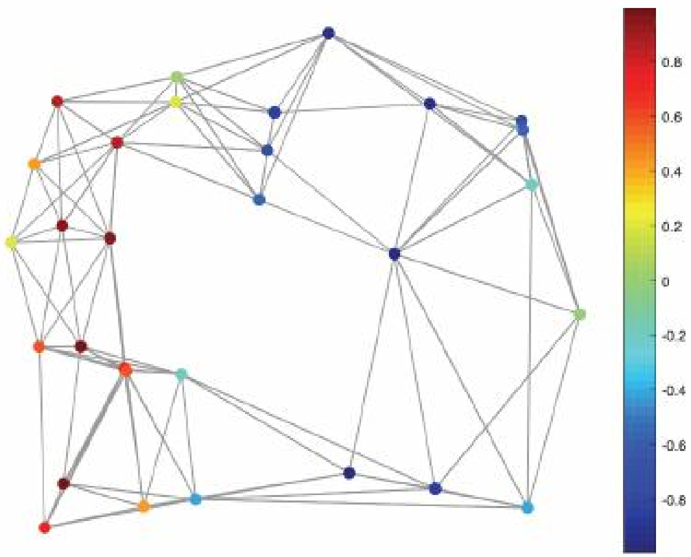
$$L = \underbrace{\begin{bmatrix} | & & | \\ \chi_0 & \cdots & \chi_{N\text{-}1} \\ | & & | \end{bmatrix}}_{\chi} \underbrace{\begin{bmatrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_{N-1} \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} \text{---} & \chi_0 & \text{---} \\ & \cdots & \\ \text{---} & \chi_{N\text{-}1} & \text{---} \end{bmatrix}}_{\chi^T}$$

$$0 = \lambda_0 < \lambda_1 \leq \cdots \leq \lambda_{N-1}$$

# Graph Fourier Transform

$$L = U \Lambda U^T$$

$L$: The Laplacian Matrix $\quad$ $U$: The eigenvectors of $L$ $\quad$ $\Lambda$ :The eigenvalues of $L$

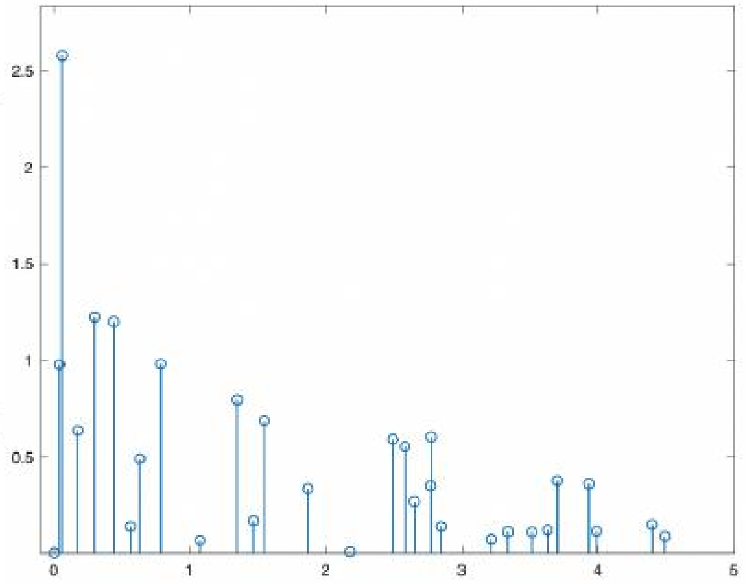



$\hat{x} = U^T x$

Graph Fourier Transform

$x = U\hat{x}$

Inverse Graph Fourier Transform




Spatial domain: $x$

Spectral domain: $\hat{x}$

# Graph Spectral Filtering for 1-D Graph Signal $f$

Recall:

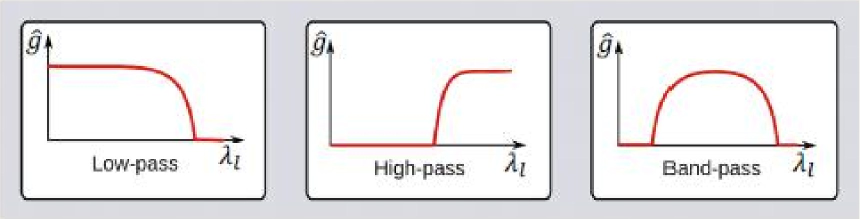
$$GFT: \hat{f} = U^T f \qquad\qquad IGFT: f = U\hat{f}$$

Filter the graph signal:

$$f \xrightarrow{GFT} U^T f \xrightarrow[\text{Filtering}]{\hat{g}(\Lambda)} \hat{g}(\Lambda)U^T f \xrightarrow{IGFT} \underbrace{U\hat{g}(\Lambda)U^T}_{\hat{g}(L)} f$$

$$\hat{g}(\Lambda) = \begin{bmatrix} \hat{g}(\lambda_0) & & 0 \\ & \ddots & \\ 0 & & \hat{g}(\lambda_{N-1}) \end{bmatrix}$$

Example:



# Recommendation Applications

- **Collaborative Filtering**
  - Graph Convolutional Neural Networks for Web-Scale Recommender Systems (KDD'18)
  - Graph Convolutional Matrix Completion (KDD'18 Deep Learning Day )
  - Neural Graph Collaborative Filtering (SIGIR'19)

- **Collaborative Filtering with Side Information**
  - Knowledge-graph-aware Recommendation
    - Knowledge Graph Convolutional Networks for Recommender Systems with Label Smoothness Regularization (KDD'19 and WWW'19)
    - KGAT: Knowledge Graph Attention Network for Recommendation (KDD'19)
  - Social Recommendation
    - Graph Neural Network for Social Recommendation (WWW'19)
- …

# Knowledge Graphs

| | |
|---|---|
| R-GCN: A First Attempt to Extend GCN to KG | https://arxiv.org/pdf/1703.06103.pdf |
| WGCN for Knowledge Graph Completion | https://arxiv.org/pdf/1811.04441.pdf |
| Estimating Node Importance Using GNN | https://arxiv.org/pdf/1905.08865.pdf |
| Cross-lingual Knowledge Graph Alignment | https://www.aclweb.org/anthology/D18-1032 |
| Zero-shot Learning using GNN on Knowledge Graph | https://arxiv.org/pdf/1803.08035.pdf |

ExplainAble Recommendation and Search (EARS)

Yongfeng Zhang
Rutgers University
yongfeng.zhang@rutgers.edu

## Explainable Recommendation Approaches

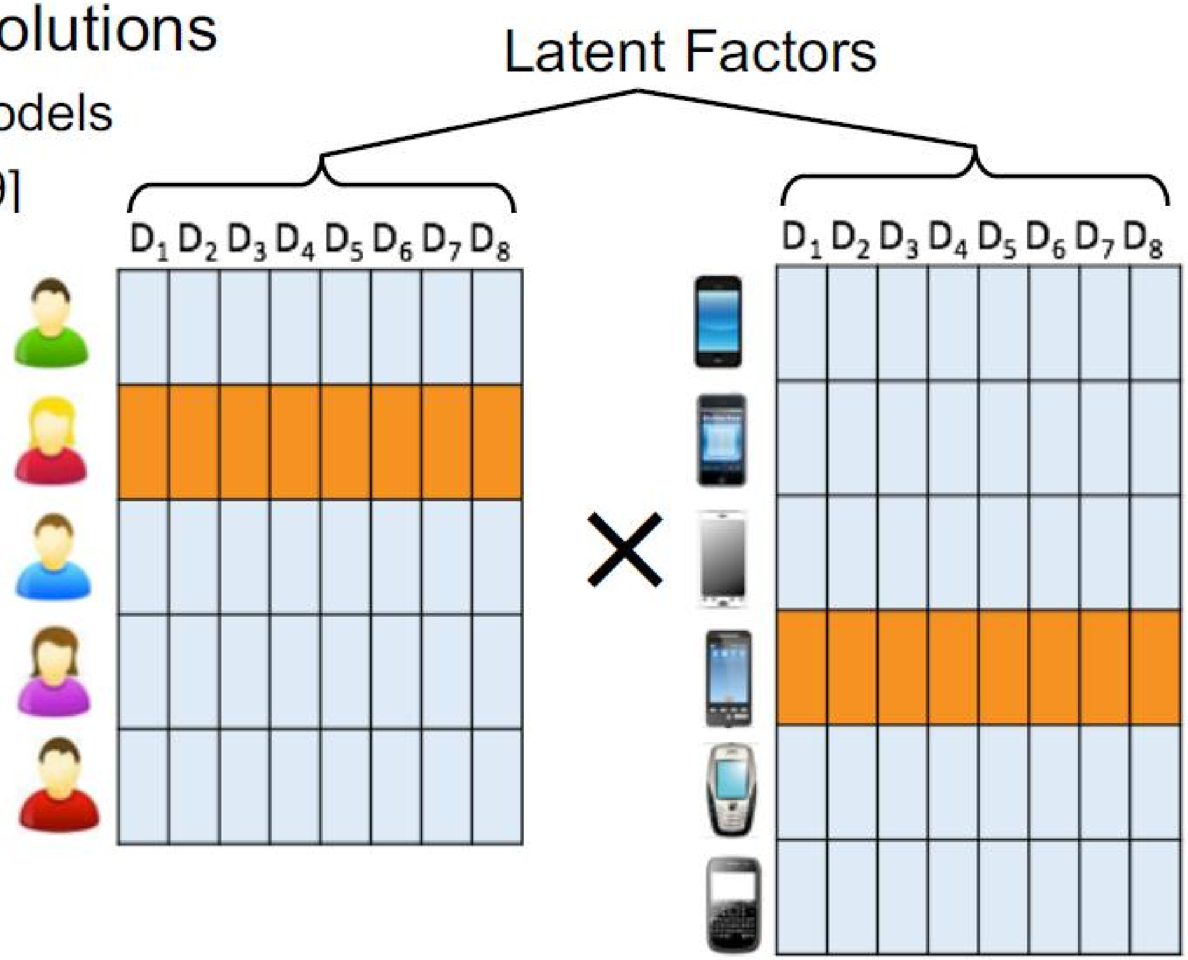
- Explainable Recommendation based on Matrix Factorization
- Explainable Recommendation based on Deep Learning
- Knowledge Graph Reasoning Approaches
- Post-hoc and Model-agnostic Approaches
- Others

# Matrix Factorization for Recommendation

- One key idea of winning solutions
  - Also called Latent Factor Models
  - [Koren et al. Computer'2009]


Original Matrix

$\approx$

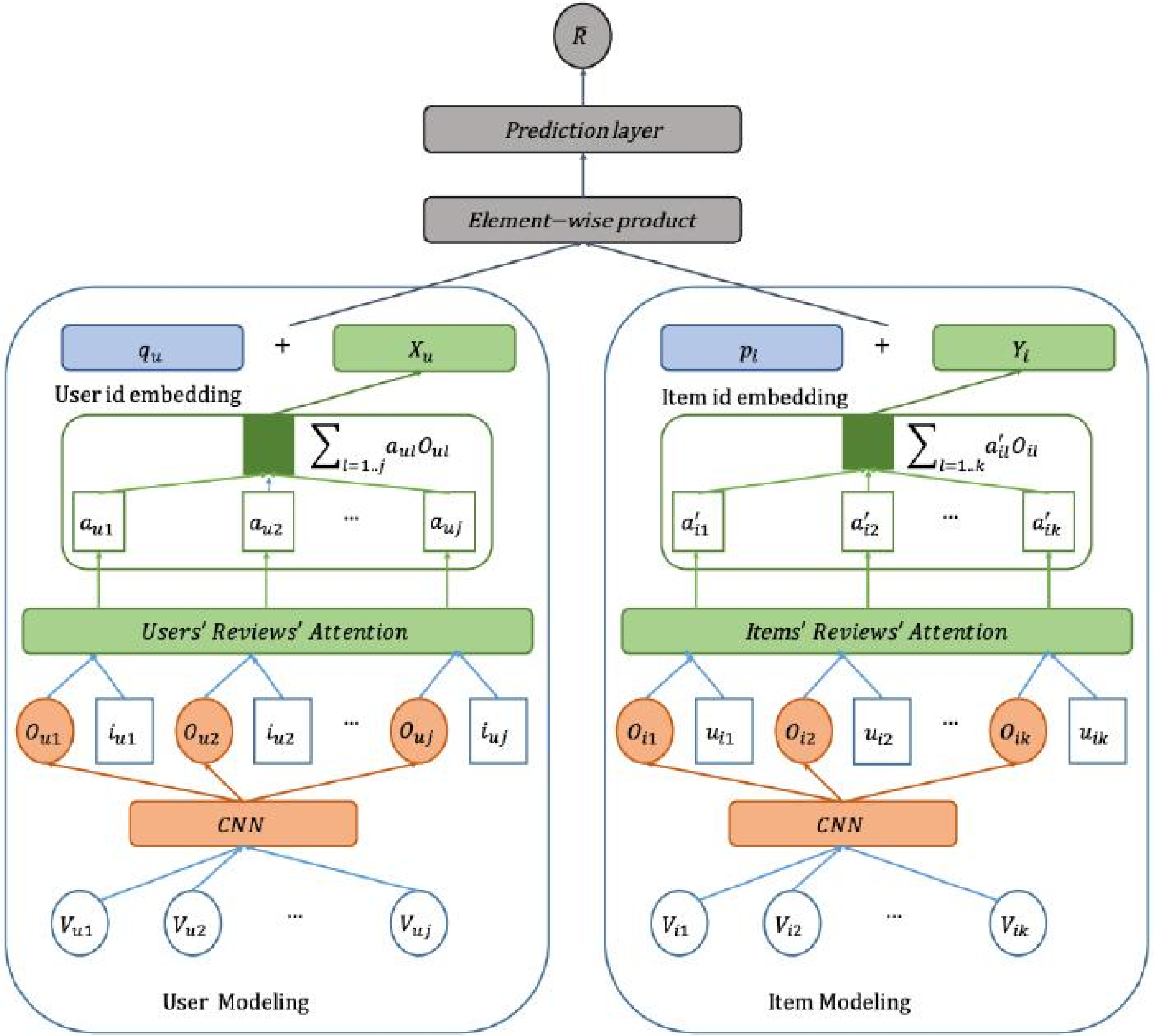



$$\hat{r}_{ui} = p_u^T q_i$$

$$\min \sum_{(u,i)\in R} \underbrace{(r_{ui} - p_u^T q_i)^2}_{\text{Goodness of fit}} + \underbrace{\lambda_1 \sum_u \| p_u \|^2 + \lambda_2 \sum_i \| q_i \|^2}_{\text{Regularization}}$$

# Review-Level Attentive Explanation

- Attentively select useful reviews as explanation [Chen et al. WWW'2018]




Reviews written by the user

Reviews of the item

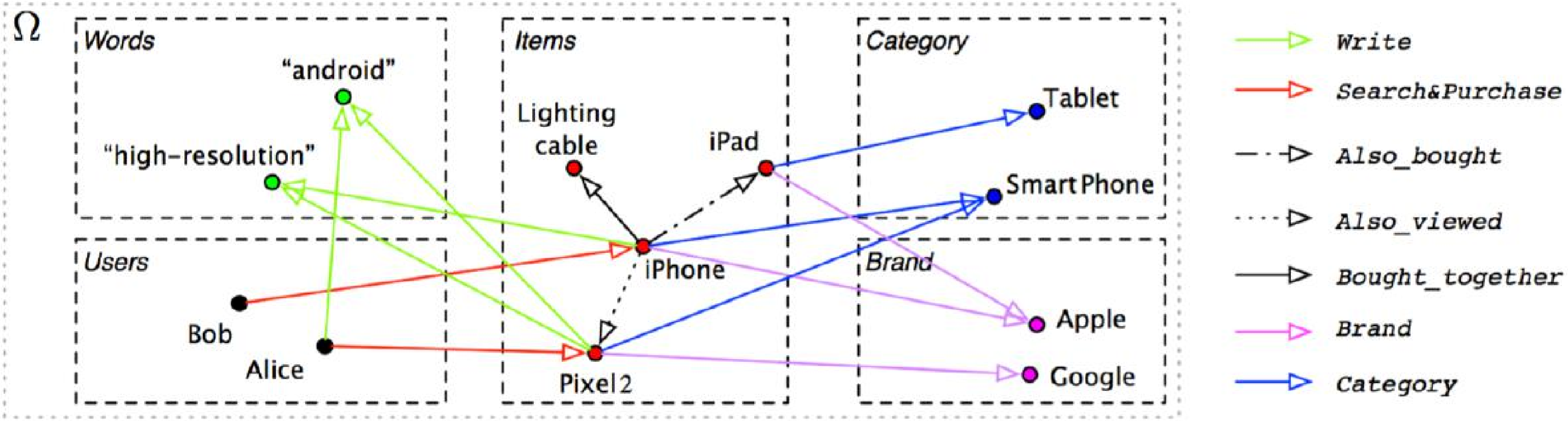
$$a^*_{il} = h^T ReLU(W_O O_{il} + W_u u_{il} + b_1) + b_2$$

$$a_{il} = \frac{exp(a^*_{il})}{\sum_{l=0}^{k} exp(a^*_{il})} \qquad O_i = \sum_{l=1,\ldots k} a_{il} O_{il}$$

Attention mechanism learns the importance of each review

# Translational KG Embedding for Recommendation

- Learning heterogeneous knowledge base embeddings for explainable recommendation [Ai et al. Alg'2018]




$$\boldsymbol{e_t} = trans(e_h, r) = \boldsymbol{e_h} + \boldsymbol{r}$$

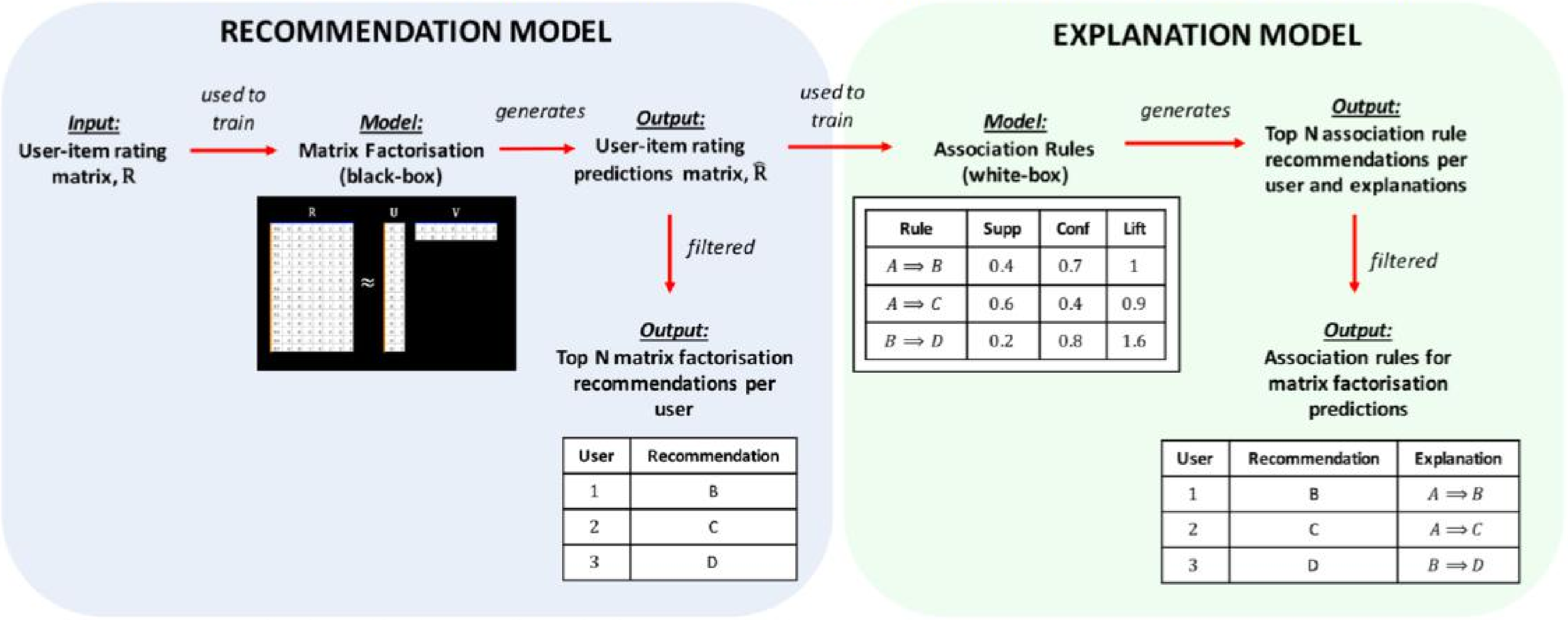
$$P(e_t | trans(e_h, r)) = \frac{\exp(\boldsymbol{e_t} \cdot trans(e_h, r))}{\sum_{e'_t \in E_t} \exp(\boldsymbol{e'_t} \cdot trans(e_h, r))}$$

$$\mathcal{L}(S) = \sum_{(e_h, e_t, r) \in S} \log \sigma(\boldsymbol{e_t} \cdot trans(e_h, r)) + k \cdot \mathbb{E}_{e'_t \sim P_t}[\log \sigma(-\boldsymbol{e'_t} \cdot trans(e_h, r))]$$

Recommendation:
Calculate $e_{user}$+$r_{purchase}$
Find top-K nearest item entity

# Post-hoc Explanation based on Association Rule Mining

- Explanation Mining: Post Hoc Interpretability of Latent Factor Models for Recommendation Systems [Peake et al. KDD'2018]




Recommendation list by a black-box model (e.g., latent factor model)

"Unexplainable Items"

Extract associate rules X->Y based on the completed matrix R. (For each user, take top-D highly predicted items as a transaction

X in training data, Y not in training data. Rank items according to some interestingness score (support/confidence/lift).

"Explainable Items" (because you liked X)

# Factorization-based Approaches

From latent factors to explicit factors

- **EFM**: Explicit factor models for explainable recommendation [Zhang et al. SIGIR'2014]
- **L2RF**: Learning to rank features for recommendation over multiple categories [Chen et al. SIGIR'2016]
- **MTER**: Explainable recommendation via multi-task learning in opinionated text data [Wang et al. SIGIR'2018]

# Explainable Recommendation with Deep Models

- Explainable Deep Models over Text
  - Based on Attention Mechanism
    - Word-level Attention [Seo et al. RecSys'2017]
    - Review-level Attention [Chen et al. WWW'2018]
    - Item-level Attention [Chen et al. WSDM'2018]
  - Based on Textual Explanation Generation
    - Sequence-to-Sequence Models with LSTM [Li et al. SIGIR'2017]
    - Generative Adversarial Networks (GAN) [Lu et al. RecSys'2018]
- Explainable Deep Models over Image
  - Based on Attention Mechanism
    - Image Region-of-Interest Explanation [Chen et al. SIGIR'2019]

# Explainable Recommendation based on KGs

Mostly based on Explanation Path between User and Item Entities

Embedding Learning Approaches

- Learn some kind of user and item representations from KG
- Recommendation based on the similarity between user-item entity
  - Translational KG Embedding for Rec and Explanation [Ai et al. Alg'2018]
  - Propagating User Preferences on the Knowledge Graph [Wang et al. CIKM'2018]
  - Learning Path Embedding for Recommendation [Wang et al. AAAI'2019]
  - Jointly Learning Explainable Rules for Recommendation [Ma et al. WWW'2019]

Symbolic Reasoning Approaches

- Recommendation based on path reasoning beginning from user entity
  - Reinforcement KG Reasoning for Explainable Recommendation [Xian et al. SIGIR'2019]

# Post-hoc and Model-Agnostic Explanation

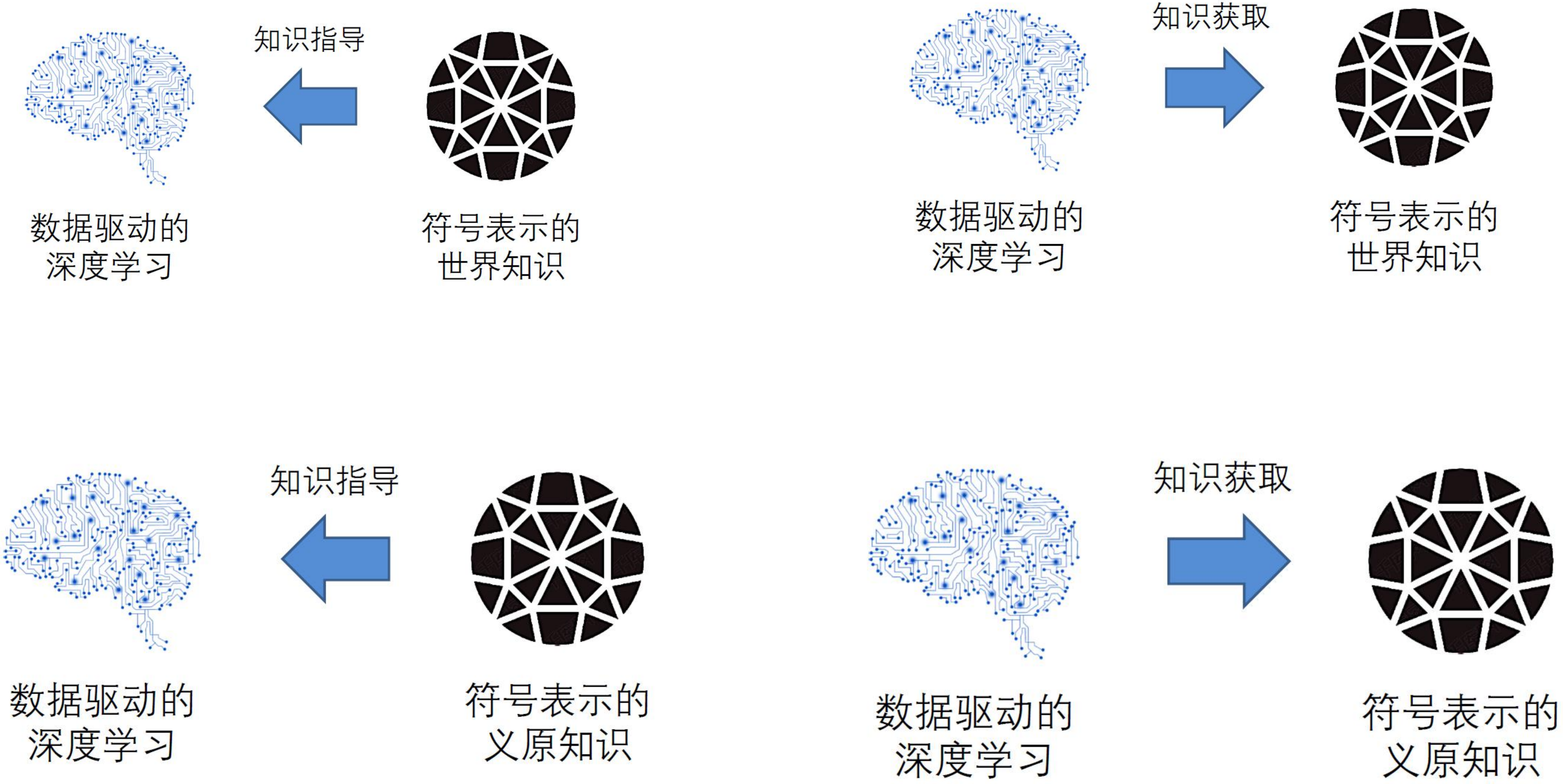
- Provide explanation for a (possibly unexplainable) model

- Mining-based Approach
  - Explanation Mining: Post Hoc Interpretability of Latent Factor Models for Recommendation Systems [Peake et al. KDD'2018]
- Learning-based Approach
  - A Reinforcement Learning Framework for Explainable Recommendation [Wang et al. ICDM'2018]

# 知识指导的自然语言处理

清华大学自然语言处理实验室

刘知远 林衍凯


知识指导
数据驱动的
深度学习
符号表示的
世界知识
知识获取
数据驱动的
深度学习
符号表示的
世界知识
知识指导
数据驱动的
深度学习
符号表示的
义原知识
知识获取
数据驱动的
深度学习
符号表示的
义原知识

# 知识指导

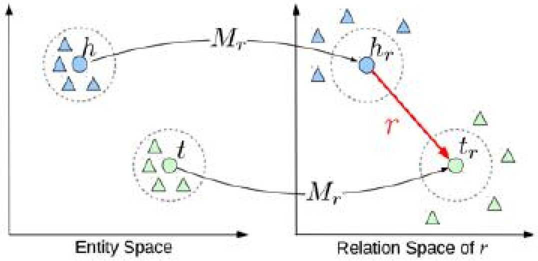
- 基于知识图谱的知识表示学习
- 知识指导的实体对齐
- 知识指导的实体细粒度分类
- 知识知道的神经网络排序
- 知识指导的预训练模型

- Zhengyan Zhang, Xu Han, Zhiyuan Liu, Xin Jiang, Maosong Sun, Qun Liu. **ERNIE: Enhanced Language Representation with Informative Entities**. ACL 2019.
- Zhenghao Liu, Chenyan Xiong, Maosong Sun, Zhiyuan Liu. **Entity-Duet Neural Ranking: Understanding the Role of Knowledge Graph Semantics in Neural Information Retrieval**. ACL 2018.
- Ji Xin, Yankai Lin, Zhiyuan Liu, Maosong Sun. **Improving Neural Fine-Grained Entity Typing with Knowledge Attention**. AAAI 2018.
- Hao Zhu, Ruobing Xie, Zhiyuan Liu, Maosong Sun. **Iterative Entity Alignment via Joint Knowledge Embeddings**. IJCAI 2017.
- Yankai Lin, Zhiyuan Liu, Maosong Sun. **Knowledge Representation Learning with Entities, Attributes and Relations**. IJCAI 2016.

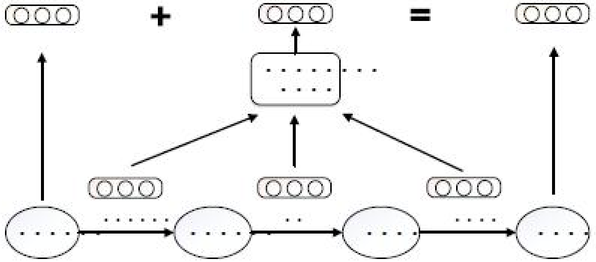
Yu Sun,Shuohuan Wang,Yukun Li,Shikun Feng.ERNIE: Enhanced Representation through Knowledge Integration(百度)

# 1、基于知识图谱的知识表示学习

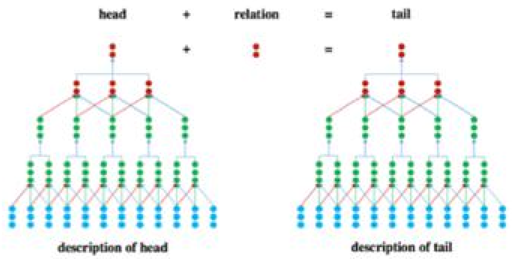
- 利用知识图谱和实体描述、类别和图像等外部信息，实现高效知识表示学习



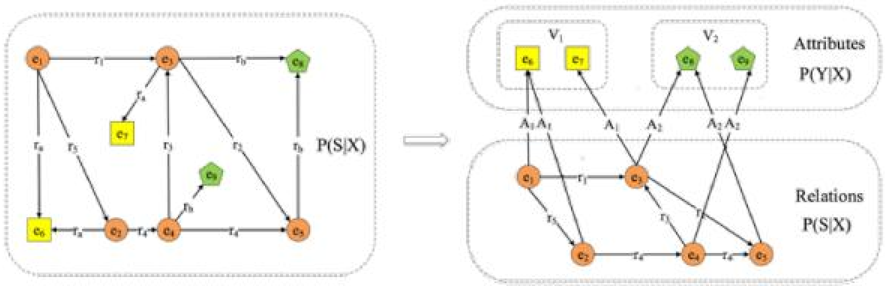

考虑复杂关系类型的知识表示
TransR (AAAI 2015)



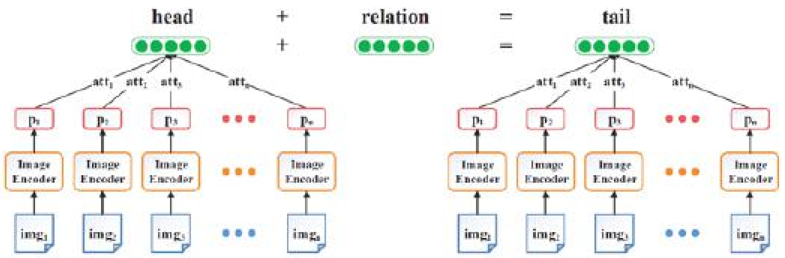
考虑关系路径的知识表示
PTransE (EMNLP 2015)




考虑实体描述信息的知识表示
DKRL (AAAI 2016)




综合考虑实体、属性与关系的知识表示
KR-EAR (IJCAI 2016)




考虑实体图像信息的知识表示
IKRL (IJCAI 2017)

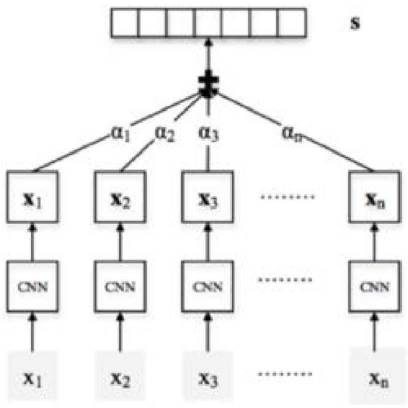
# 知识获取

- 考虑噪音过滤的关系抽取
- 多语言关系抽取
- 层次关系抽取
- 文档级别关系抽取

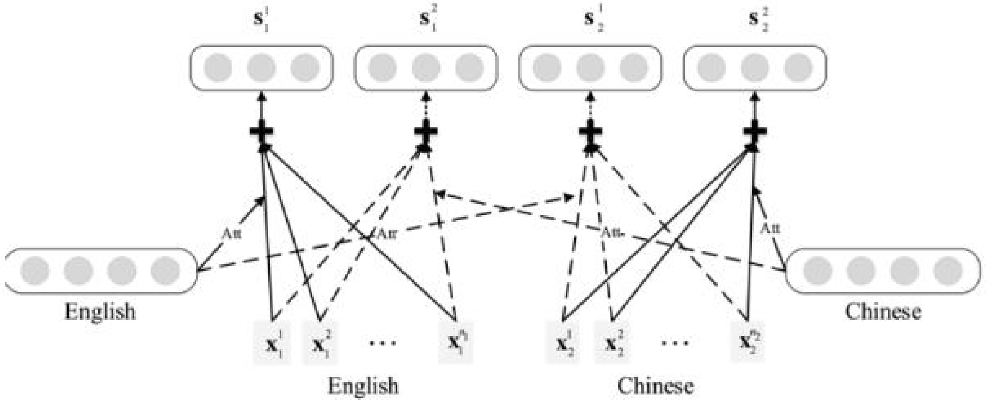
- Xu Han, Pengfei Yu, Zhiyuan Liu, Maosong Sun, Peng Li. **Hierarchical Relation Extraction with Coarse-to-Fine Grained Attention**. EMNLP 2018.
- Xiaozhi Wang, Xu Han, Yankai Lin, Zhiyuan Liu, Maosong Sun. **Adversarial Multi-lingual Neural Relation Extraction**. COLING 2018.
- Xu Han, Zhiyuan Liu, Maosong Sun. **Neural Knowledge Acquisition via Mutual Attention between Knowledge Graph and Text**. AAAI 2018.
- Wenyuan Zeng, Yankai Lin, Zhiyuan Liu, Maosong Sun. **Incorporating Relation Paths in Neural Relation Extraction**. EMNLP 2017.
- Yankai Lin, Zhiyuan Liu, Maosong Sun. **Neural Relation Extraction with Multi-lingual Attention**. ACL 2017.
- Yankai Lin, Shiqi Shen, Zhiyuan Liu, Huanbo Luan, Maosong Sun. **Neural Relation Extraction with Selective Attention over Instances**. ACL 2016.

# 1、关系抽取技术

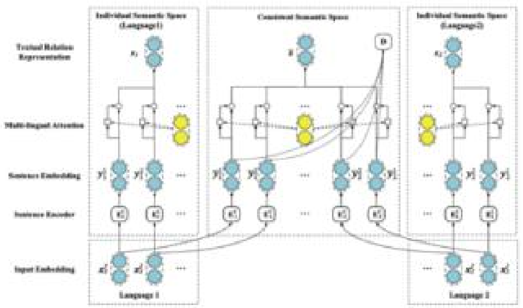
- 提出选择注意力机制自动降噪并整合多源信息



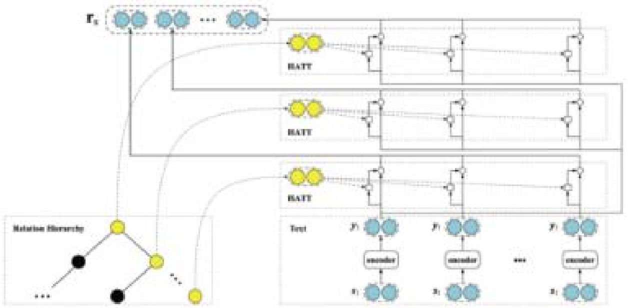

基于句级注意力的远程监督
神经网络关系抽取(ACL 2016)



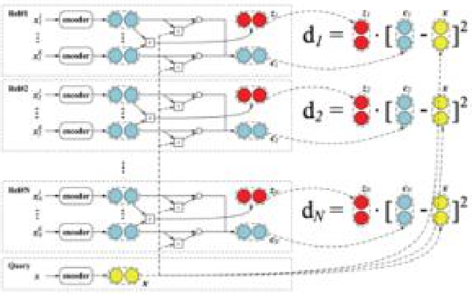

基于跨语言注意力的
神经网络关系抽取(ACL 2017)



基于对抗注意力的神经网络
关系抽取(COLING 2018)



基于层次注意力的神经网络
关系抽取(EMNLP 2018)



基于混合注意力的
少次关系抽取(AAAI 2019)

# 1、文档级别关系抽取数据集DocRED

https://github.com/thunlp/DocRED

- 为了支撑文档级别关系抽取的研究，我们构建了大规模数据集DocRED
- 规模超越了已有的所有关系抽取数据集

| Dataset | # Doc | # Word | # Sent | # Ent | # Rel | # Inst | # Fact |
|---|---|---|---|---|---|---|---|
| SemEval-2010 Task 8 | - | 205k | 10,717 | 21,434 | 9 | 8,853 | 8,383 |
| ACE 2003-2004 | - | 297k | 12,783 | 46,108 | 24 | 16,771 | 16,536 |
| TACRED | - | 1,823k | 53,791 | 152,527 | 41 | 21,773 | 5,976 |
| FewRel | - | 1,397k | 56,109 | 72,124 | 100 | 70,000 | 55,803 |
| BC5CDR | 1,500 | 282k | 11,089 | 29,271 | 1 | 3,116 | 2,434 |
| DocRED (Human-Annotated) | 5,053 | 1,002k | 40,348 | 129,767 | 96 | 71,046 | 62,762 |
| DocRED (Distantly Supervised) | 101,997 | 21,368k | 870,605 | 2,771,037 | 96 | 1,678,864 | 760,917 |

- 解决DocRED需要模型具有模式识别、逻辑推理、指代推理、常识推理等多方位能力

| Reasoning Types | % | Examples |
|---|---|---|
| Pattern recognition | 38.9 | [1] ***Me Musical Nephews*** is a 1942 one-reel animated cartoon directed by Seymour Kneitel and animated by Tom Johnson and George Germanetti. [2] Jack Mercer and Jack Ward wrote the script. ...<br>**Relation:** `publication_date` **Supporting Evidence:** 1 |
| Logical reasoning | 26.6 | [1] "Nisei" is the ninth episode of the third season of the American science fiction television series The X-Files. ... [3] It was directed by David Nutter, and written by Chris Carter, Frank Spotnitz and Howard Gordon. ... [8] The show centers on FBI special agents ***Fox Mulder*** (David Duchovny) and Dana Scully (Gillian Anderson) who work on cases linked to the paranormal, called X-Files. ...<br>**Relation:** `creator` **Supporting Evidence:** 1, 3, 8 |
| Coreference reasoning | 17.6 | [1] ***Dwight Tillery*** is an American politician of the Democratic Party who is active in local politics of Cincinnati, Ohio. ... [3] He also holds a law degree from the University of Michigan Law School. [4] ***Tillery*** served as mayor of Cincinnati from 1991 to 1993.<br>**Relation:** `educated_at` **Supporting Evidence:** 1, 3 |
| Common-sense reasoning | 16.6 | [1] ***William Busac*** (1020-1076), son of William I, Count of Eu, and his wife Lesceline. ... [4] ***William*** appealed to King Henry I of France, who gave him in marriage Adelaide, the heiress of the county of Soissons. [5] Adelaide was daughter of Renaud I, Count of Soissons, and Grand Master of the Hotel de France. ... [7] ***William*** and Adelaide had four children: ...<br>**Relation:** `spouse` **Supporting Evidence:** 4, 7 |

# 1、文档级别关系抽取数据集DocRED

- 目前已有的模型在DocRED上效果均较为薄弱

| Model | Dev Ign F1 | Dev Ign AUC | Dev F1 | Dev AUC | Test Ign F1 | Test Ign AUC | Test F1 | Test AUC |
|---|---|---|---|---|---|---|---|---|
| Supervised Setting | | | | | | | | |
| CNN | 44.55 | 41.18 | 47.57 | 45.24 | 43.04 | 40.68 | 46.35 | 44.70 |
| LSTM | 47.15 | 44.99 | 49.99 | 49.08 | 46.14 | 44.25 | 49.53 | 48.98 |
| BiLSTM | 48.31 | 46.75 | 51.16 | 50.98 | 46.77 | 45.83 | 50.09 | 50.21 |
| Context-Aware | **48.46** | **46.78** | **51.57** | **51.07** | **47.87** | **46.27** | **51.26** | **50.71** |
| Weakly Supervised Setting | | | | | | | | |
| CNN | 35.70 | 27.47 | 46.40 | 43.90 | 35.41 | 26.59 | 46.38 | 43.05 |
| LSTM | 39.79 | 27.74 | 51.92 | 46.97 | 36.64 | 26.97 | 51.33 | 46.30 |
| BiLSTM | **41.95** | **28.63** | 54.21 | **48.54** | 38.23 | **27.59** | 53.38 | 46.90 |
| Context-Aware | 41.75 | 27.88 | **54.26** | 47.74 | **39.10** | 27.14 | **53.75** | **46.91** |

- 在抽样数据上与人类水平存在极大差距，亟需进一步研究

| Method | RE P | RE R | RE F1 | RE+Sup P | RE+Sup R | RE+Sup F1 |
|---|---|---|---|---|---|---|
| Model | 55.6 | 52.6 | 54.1 | 46.4 | 43.1 | 44.7 |
| Human | **89.7** | **86.3** | **88.0** | **71.2** | **75.8** | **73.4** |

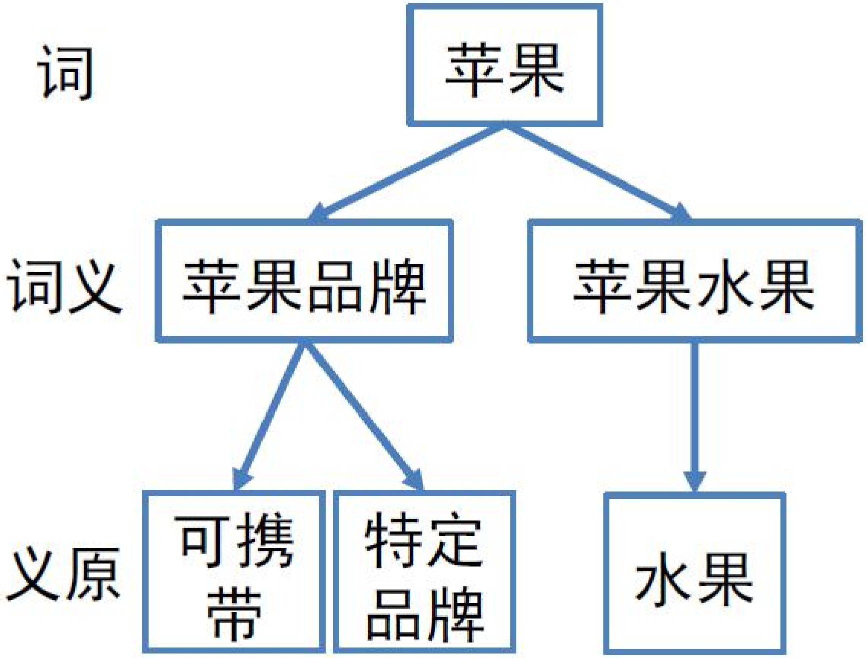
# 义原知识

- 融合义原知识的词义表示学习
- 融合义原知识的神经语言模型
- 基于语义表示学习的义原推荐
- 考虑内部汉字信息的义原推荐
- 融合HowNet义原标注的词典扩展

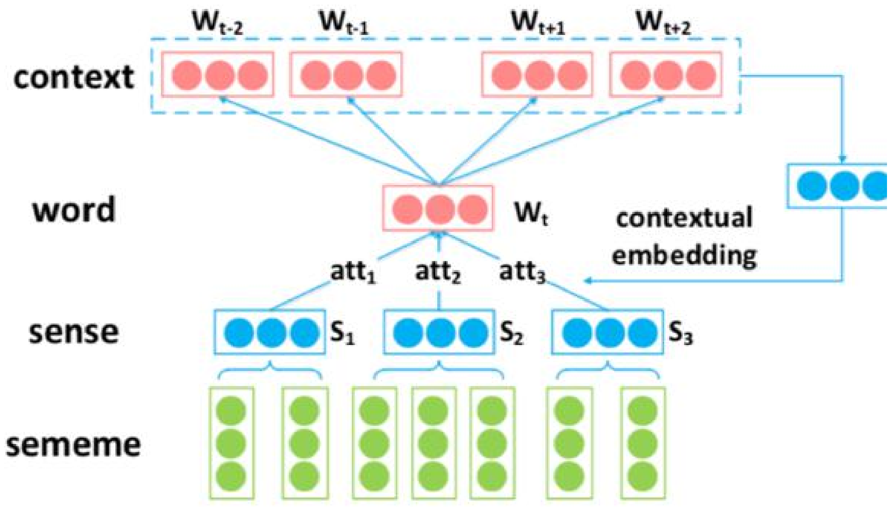

- Fanchao Qi, Junjie Huang, Chenghao Yang, Zhiyuan Liu, Xiao Chen, Qun Liu, Maosong Sun. **Modeling Semantic Compositionality with Sememe Knowledge**. ACL 2019.
- Yihong Gu, Jun Yan, Hao Zhu, Zhiyuan Liu, Ruobing Xie, Maosong Sun, Fen Lin and Leyu Lin. **Language Modeling with Sparse Product of Sememe Experts**. EMNLP 2018.
- Fanchao Qi, Yankai Lin, Maosong Sun, Hao Zhu, Ruobing Xie, Zhiyuan Liu. **Cross-lingual Lexical Sememe Prediction**. EMNLP 2018.
- Huiming Jin, Hao Zhu, Zhiyuan Liu, Ruobing Xie, Maosong Sun, Fen Lin, Leyu Lin. **Incorporating Chinese Characters of Words for Lexical Sememe Prediction**. ACL 2018.
- Xiangkai Zeng, Cheng Yang, Cunchao Tu, Zhiyuan Liu, Maosong Sun. **Chinese LIWC Lexicon Expansion via Hierarchical Classification of Word Embeddings with Sememe Attention**. AAAI 2018.
- Ruobing Xie, Xingchi Yuan, Zhiyuan Liu, Maosong Sun. **Lexical Sememe Prediction via Word Embeddings and Matrix Factorization**. IJCAI 2017.
- Yilin Niu, Ruobing Xie, Zhiyuan Liu, Maosong Sun. **Improved Word Representation Learning with Sememes**. ACL 2017.

# 融合义原知识的词义表示学习

- 考虑HowNet的词义-义原标注信息，提升词义表示性能




HowNet词义-义原标注示例




义原-词义-词汇的联合表示学习模型

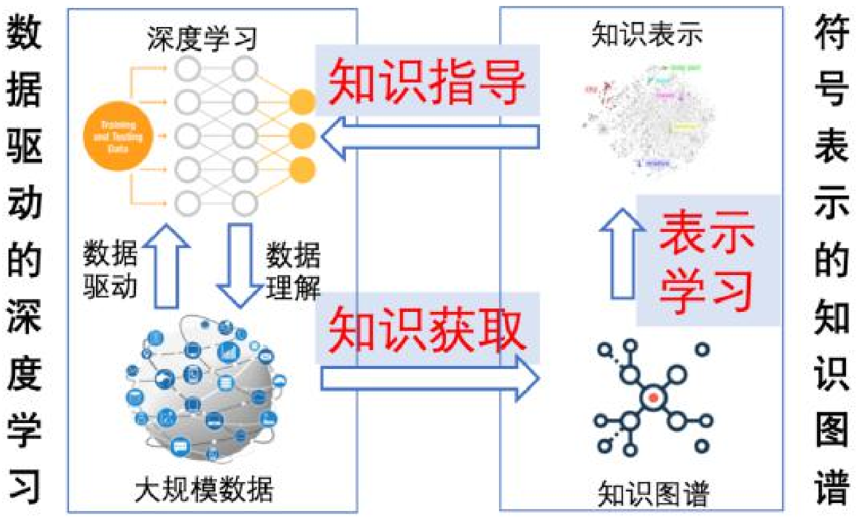
Yilin Niu, Ruobing Xie, Zhiyuan Liu, Maosong Sun. Improved Word Representation Learning with Sememes. ACL 2017.

# 总结展望

- 义原语言知识突破词汇屏障，对语言理解极具重要意义，具有极佳融合深度学习的特性
- 世界知识对于富知识文本深度理解具有重要意义，知识表示学习是目前较好的解决方案

https://github.com/thunlp

大规模文本中获取知识




五个更加

1. 更加全面的知识类型
2. 更加复杂的知识结构
3. 更加有效的知识获取
4. 更加强大的知识指导
5. 更加精深的知识推理

# Advanced Work

- 融合义原信息的pretrained Language Model (清华)
- 融合知识的的pretrained Language Model
- 知识驱动的低资源NLP
- Learning to Hash for Network/KG (From Embedding to Hashing)
- Summarization with Embeddings
- Active Search(主动搜索)

# Advanced Work Reference

## 低资源 NLP & KG

- Knowledge-Augmented Language Model and Its Application to Unsupervised Named-Entity Recognition (NAACL2019)
- Description-Based Zero-shot Fine-Grained Entity Typing (NAACL19)
- Zero-Shot Entity Linking by Reading Entity Descriptions (ACL2019)
- Zero-Shot Cross-Lingual Opinion Target Extraction (NAACL2019)
- Multi-Level Matching and Aggregation Network for Few-Shot Relation Classification (ACL19)
- Exploiting Entity BIO Tag Embeddings and Multi-task Learning for Relation Extraction with Imbalanced Data (ACL2019)
- Massively Multilingual Transfer for NER (ACL2019)

## Learing to hash for network/KG

Lian D, Zheng K, Zheng V W, et al. ***High-order Proximity Preserving Information Network Hashing***. KDD 2018: 1744-1753.

Shen X, Pan S, Liu W, et al. ***Discrete Network Embedding***. IJCAI 2018: 3549-3555.

Yang H, Pan S, Zhang P, et al. ***Binarized Attributed Network Embedding***. ICDM 2018: 1476-1481.

Wang Q, Wang S, Gong M, et al. ***Feature Hashing for Network Representation Learning***. IJCAI 2018: 2812-2818.

Misra V, Bhatia S. Bernoulli embeddings for graphs. AAAI 2018.

---

Kishimoto K, Hayashi K, Akai G, et al. ***Binarized Knowledge Graph Embeddings***. ECIR 2019: 181-196.

# Advanced Work Reference

**融合知识pretrained LM:**

[1] Yu Sun et al. ERNIE: Enhanced Representation through kNowledge IntEgration.
[2] Zhengyan Zhang et al. ERNIE: Enhanced Language Representation with Informative Entities (ACL19)

**Active search:**

Klyuchnikov N, Mottin D, Koutrika G, et al. *Figuring out the User in a Few Steps: Bayesian Multifidelity Active Search with Cokriging*. KDD. 2019: 686-695.

Kratzwald B, Feuerriegel S. *Learning from On-Line User Feedback in Neural Question Answering on the Web*. The Web Conference. 2019: 906-916.

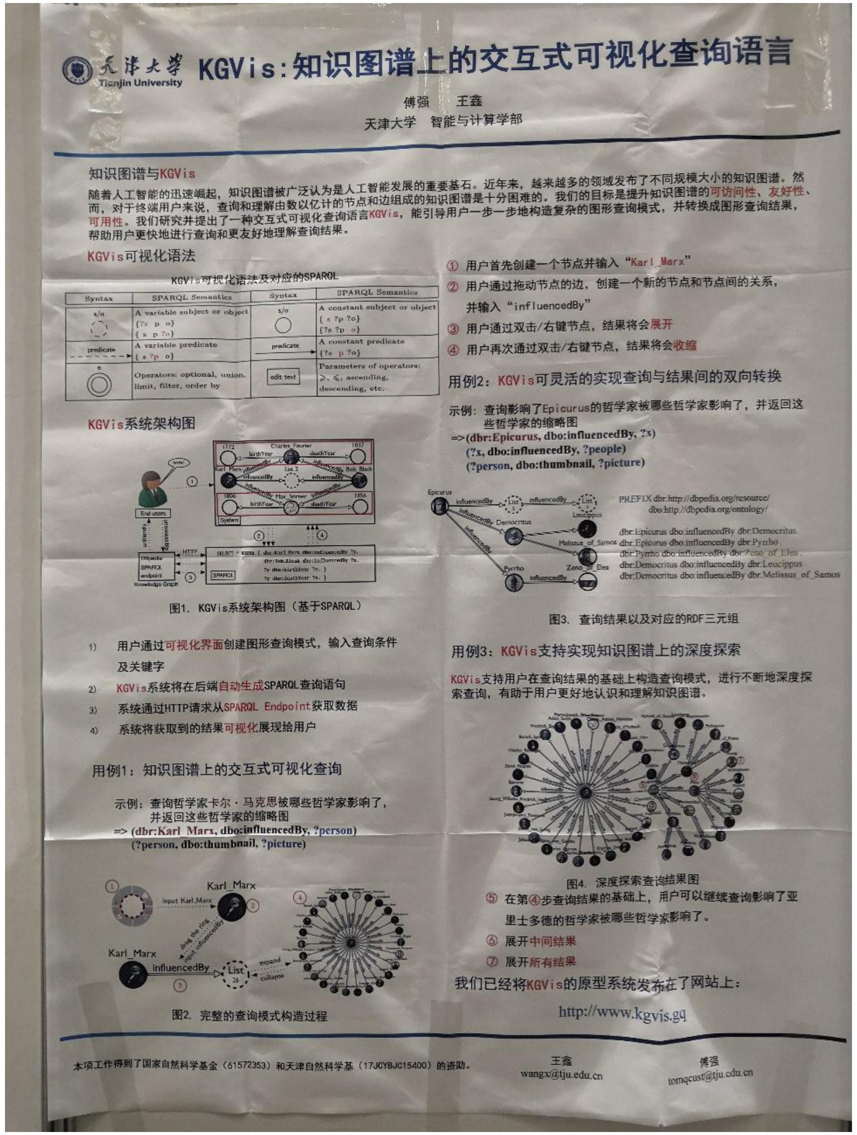
Canal G, Massimino A, Davenport M, et al. *Active Embedding Search via Noisy Paired Comparisons*. ICML. 2019: 902-911.

Rabbany R, Bayani D, Dubrawski A. *Active Search of Connections for Case Building and Combating Human Trafficking*. KDD. 2018: 2120-2129.

**Summarization with Embeddings:**

Jin D, Rossi R A, Koh E, et al. *Latent Network Summarization: Bridging Network Embedding and Summarization*. KDD 2019: to appear.

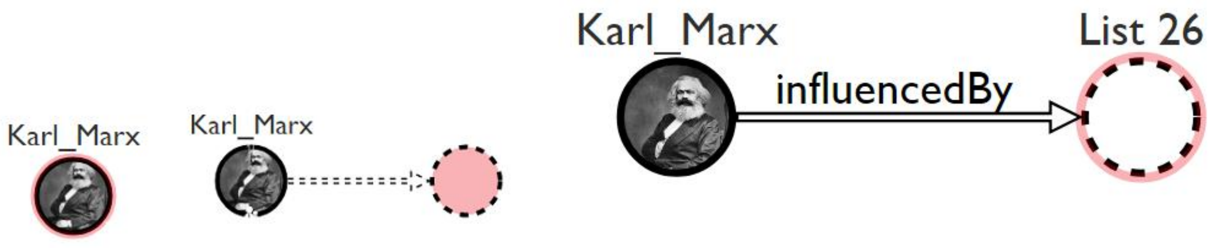
# school work

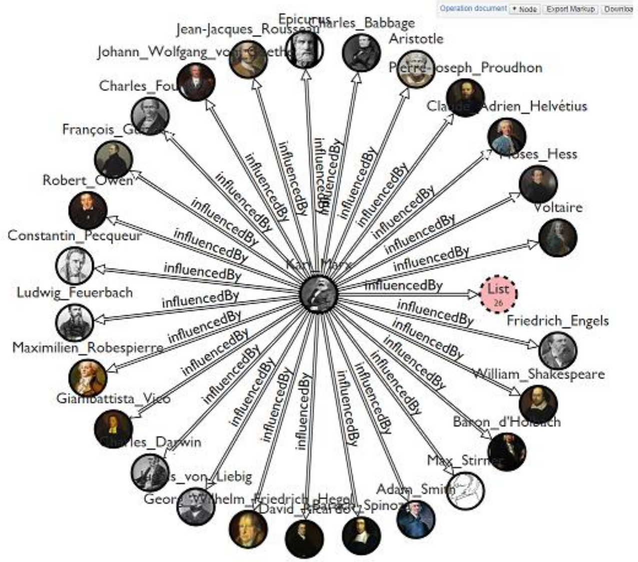



1、Task one

Drag the border of Karl_Marx, then double click the edge, and input predicate 'influencedBy'.

After this you will get a list of result. you can choose expand result while you right click the result ring.



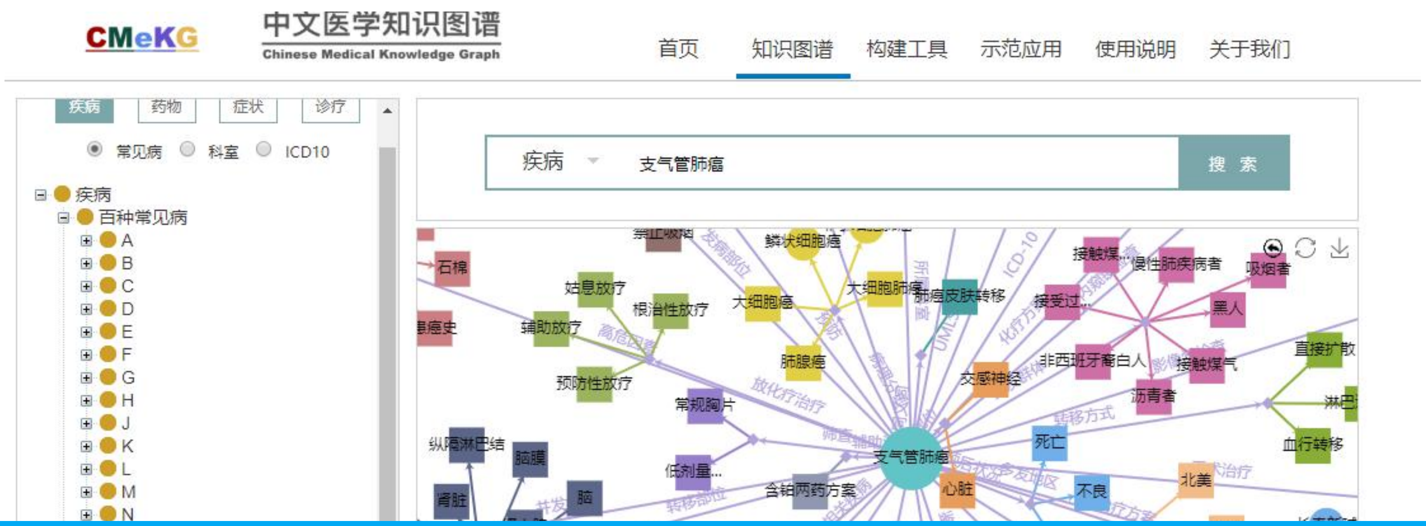

天津大学

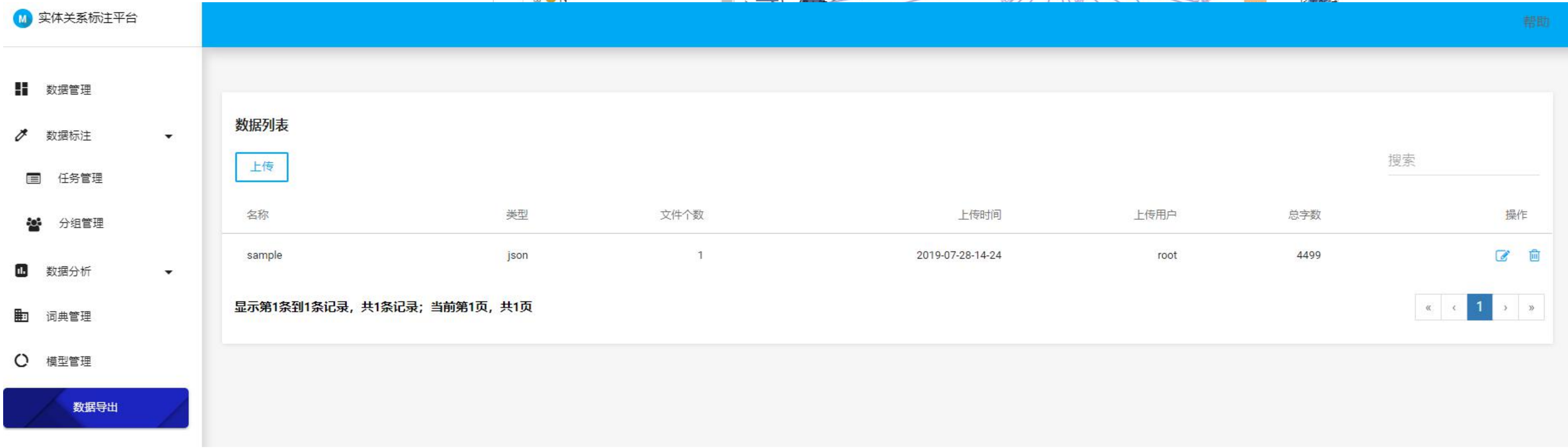
http://www.kgvis.gq/

# school work

北京大学、郑州大学

http://zstp.pcl.ac.cn:8002/



# 知识图谱发展趋势和挑战

- 知识驱动和数据驱动相结合的知识表示和推理
- 高质量大规模知识获取算法（多模态多语言知识获取，常识知识获取，机器阅读理解等）
- 集群体智能的、开放的、融合的结构化知识资源
- 知识图谱技术平台化（评测），把数据转化为知识
- 知识图谱应用的多样化，建立智能知识服务的创新应用