

# KagNet

## Knowledge-Aware Graph Networks

BaekTree  
자여너학땅

# What to come: KagNet

1. Contribution: Abstract & Conclusion
2. Introduction
  - Problem Statement
  - Knowledge Graph
3. Reasoning Flow: 그래서 뭘 어떻게?
4. Explainability

# 1. Contribution: Abstract and Conclusion

- Problem
  - **Commonsense** Inference
  - Particularly, Ask question with some answer candidates
  - N 지선다 문제
- 방법
  - utilize **external** and structured **commonsense knowledge graphs**
- 성과
  - **explainable** inference
  - SOTA(A little bit... :/)
  - outperforms BERT and GPT in Commonsense QA
    - BERT랑 GPT는 그냥 데이터만 많이 쑤서 넣음

## 2. Introduction: Problem Definition

- Q
  - “Where do adults use glue sticks?”
- A
  - {classroom(X), office (O), desk drawer (X)}
- BERT or GPT?
  - FAR LESS than human performance
  - lacking transparency and interpretability

Model	10(%) of IHtrain		50(%) of IHtrain		100(%) of IHtrain	
	IHdev-Acc.(%)	IHtest-Acc.(%)	IHdev-Acc.(%)	IHtest-Acc.(%)	IHdev-Acc.(%)	IHtest-Acc.(%)
Random guess	20.0	20.0	20.0	20.0	20.0	20.0
GPT-FINETUNING	27.55	26.51	32.46	31.28	47.35	45.58
GPT-KAGNET	28.13	<b>26.98</b>	33.72	<b>32.33</b>	48.95	<b>46.79</b>
BERT-BASE-FINETUNING	30.11	29.78	38.66	36.83	53.48	53.26
BERT-BASE-KAGNET	31.05	<b>30.94</b>	40.32	<b>39.01</b>	55.57	<b>56.19</b>
BERT-LARGE-FINETUNING	35.71	32.88	55.45	49.88	60.61	55.84
BERT-LARGE-KAGNET	36.82	<b>33.91</b>	58.73	<b>51.13</b>	62.35	<b>57.16</b>
Human Performance	-	<b>88.9</b>	-	<b>88.9</b>	-	<b>88.9</b>

Table 1: Comparisons with large pre-trained language model fine-tuning with different amount of training data.

근데 사실 KagNet도 사람에 비해서는 엄청 형편 없음.... Explainable하다는 것에 의의를 뒀야...

## 2. Introduction: Knowledge Graph

- External Knowledge
  - 외부 지식 창고를 사용하자!
  - 미리 만들어둔다!
  - Graph 자료구조 사용
  - Vertices: Concept
  - Edge: Relation
- Q -> Semantic Space -> A
  - 모델이 input을 받아서, Semantically 문제 이해해서 정답을 맞춘다(BERT or GPT).
- Q -> Symbolic Space -> A
  - 그 과정에서 여기에 외부 지식 창고의 논리 관계를 참고해서 정답을 맞춘!
- 문법/문장 구조와 논리 관계를 둘 다 쓴다!

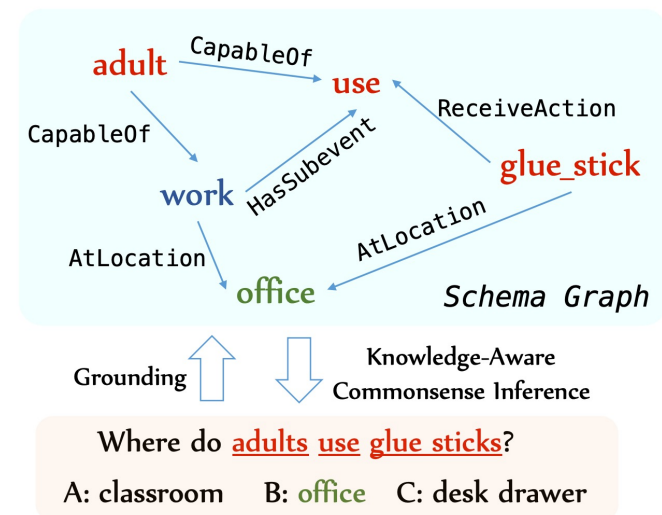
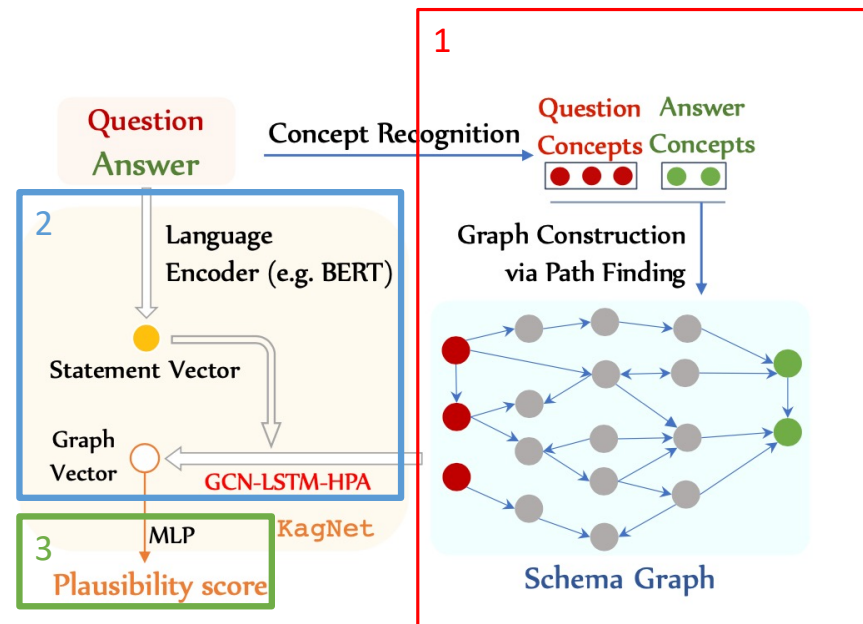


Figure 1: An example of using **external commonsense knowledge** (symbolic space) for inference in natural language commonsense questions (semantic space).

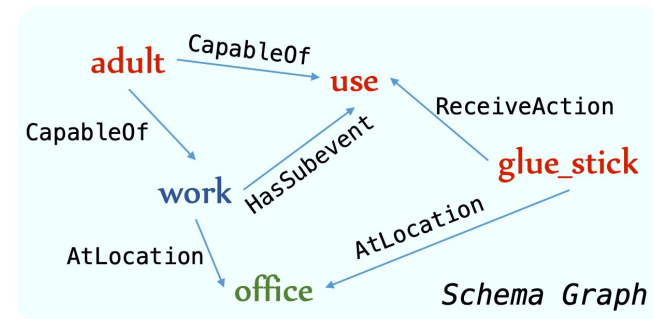
### 3. Reasoning Flow: 그래서 뭘 어떻게?

1. Input에 해당하는 subgraph 찾기!(Concept Recognition)
  1. 그리고 가공하기
2. subgraph의 concept과 relation을 분석하기!
  1. 논리 관계 파악
  2. 문장/문법 구조 파악
3. 정답 맞추기!



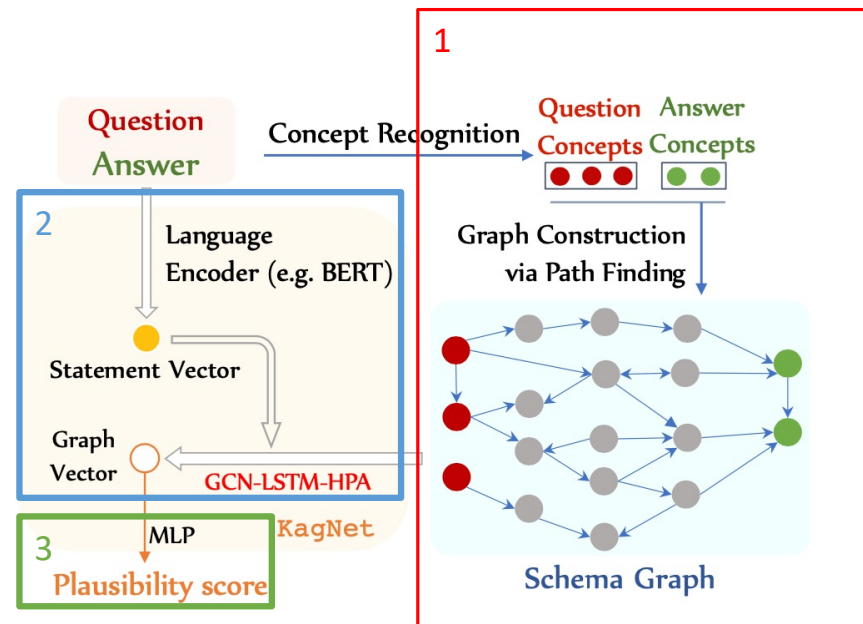
# 3.1. Input에 해당하는 subgraph 찾기!

1. input
  1. Where do adults use glue sticks **SEP** classroom
  2. Where do adults use glue sticks **SEP** office
  3. Where do adults use glue sticks **SEP** desk
  4. Where do adults use glue sticks **SEP** drawer
2. find vertices(concept)
  1. ~~n-gram match~~
3. find edges that link vertices
  1. ~~select maximum edge length k~~
4. 가공하기(그래프 -> 벡터화)
  1. ~~KGE method으로 graph -> vector~~
  2. ~~concept을 연결하는 relation 중 의미 없는 것 제거(pruning)~~



### 3. Reasoning Flow: 이제 분석해보자!

1. ~~Input에 해당하는 subgraph 찾기!(Concept Recognition)~~
  1. ~~그리고 가공하기~~
2. subgraph의 concept과 relation을 분석하기!
  1. 논리 관계 파악
  2. 문장/문법 구조 파악
  3. 필요한 정보에만 집중
3. 정답 맞추기!

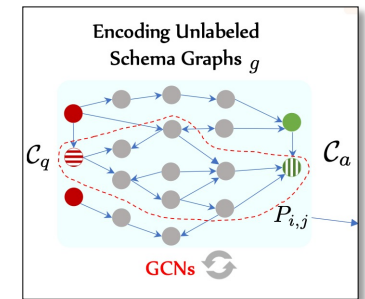
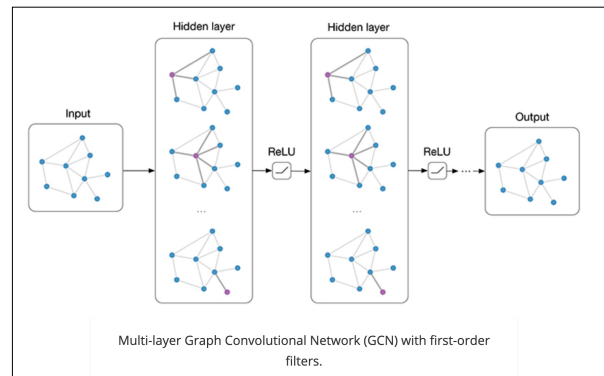




## 3.2.1 논리 관계 파악

- 개념간 관계를 강화하자!
  - 주변 개념들에서 공통 feature 파악
  - Graph에서 Convolution Net (GCN)
  - ~~각 concept마다 문맥 고려한 새로운 vector으로 바꿈!~~
  - ~~$h(l) \rightarrow h(l+1)$~~
  - ~~k번째 인접 concept 고려~~
- 개념간 연결 관계를 파악하자!
- 문맥 파악 및 필요한 정보에만 집중

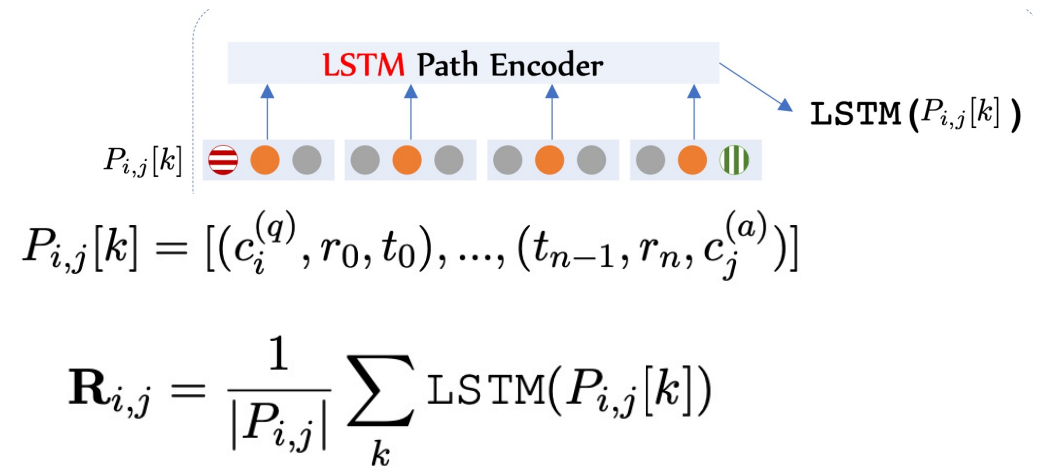
$$h_i^{(l+1)} = \sigma(W_{self}^{(l)} h_i^{(l)} + \sum_{j \in N_i} \frac{1}{|N_i|} W^{(l)} h_j^{(l)})$$



<https://tkipf.github.io/graph-convolutional-networks/>

## 3.2.1 논리 관계 파악

- 개념간 관계 강화 하자!
  - 주변 개념들에서 공통 feature 파악
  - Graph에서 Convolution Net (GCN)
- 개념간 연결 관계를 파악하자!
  - 논리 관계는 논리의 연결이다.
  - A -> B
  - B -> C
  - 그러면 A -> C
  - 따라서 시계열 데이터와 유사
  - LSTM
  - 연결 거리가 k일 때 관계 파악한 결과 벡터 도출
  - 직관: 개념들의 평균적인 논리 연결 관계에 대한 feature representation
- 문맥 파악 및 필요한 정보에만 집중



## 3.2. 2&3 문맥 파악 및 필요한 정보에만 집중

- 개념간 관계 강화!
  - 주변 개념들에서 공통 feature 파악
  - Graph에서 Convolution Net (GCN)
- 개념간 연결 관계를 파악하자!
  - 논리 관계는 논리의 연결이다.
  - 시계열 데이터와 유사
  - LSTM
- 쿼리: Q와 A를 연결하는 semantic feature T(2. 문맥 파악)
- 필요한 정보에만 집중(3. 집중)
  - attention
  - 어떤 논리 연결에 집중? a
  - 어떤 concept 연결에 집중? b

$$\mathbf{T}_{i,j} = \text{MLP}([\mathbf{s}; \mathbf{c}_q^{(i)}; \mathbf{c}_a^{(j)}])$$

$$\alpha_{(i,j,k)} = \mathbf{T}_{i,j} \mathbf{W}_1 \text{LSTM}(P_{i,j}[k]),$$

$$\hat{\alpha}_{(i,j,\cdot)} = \text{SoftMax}(\alpha_{(i,j,\cdot)}),$$

$$\hat{\mathbf{R}}_{i,j} = \sum_k \hat{\alpha}_{(i,j,k)} \cdot \text{LSTM}(P_{i,j}[k]).$$

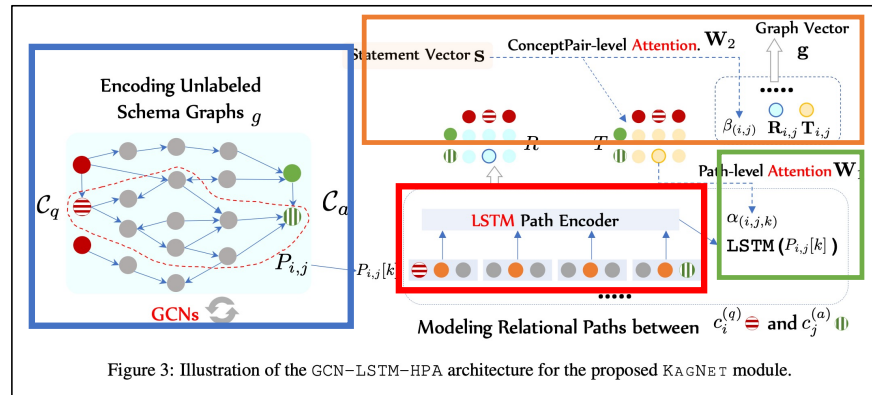
$$\beta_{(i,j)} = \mathbf{s} \mathbf{W}_2 \mathbf{T}_{i,j}$$

$$\hat{\beta}_{(\cdot,\cdot)} = \text{SoftMax}(\beta_{(\cdot,\cdot)})$$

$$\hat{\mathbf{g}} = \sum_{i,j} \hat{\beta}_{(i,j)} [\hat{\mathbf{R}}_{i,j}; \mathbf{T}_{i,j}]$$

## 3.2. Revisit: Concept과 Relation을 분석하기!

- 개념과 관계를 추출하자!
  - 주변 concept 파악
  - 연결 관계를 강화
  - GCN
- 개념간 연결 관계를 파악하자!
  - 논리 관계는 논리의 연결이다.
  - A -> B
  - B -> C
  - 그러면 A -> C
  - 따라서 시계열 데이터와 유사
  - LSTM
- 필요한 정보에만 집중
  - Attention on 논리 연결 관계
  - Attention on 문법/문장 구조
- Combine all!
- 최종 output: graph vector,  $\hat{g}$



$$\mathbf{T}_{i,j} = \text{MLP}([\mathbf{s}; \mathbf{c}_q^{(i)}; \mathbf{c}_a^{(j)}])$$

$$\begin{aligned} \alpha_{(i,j,k)} &= \mathbf{T}_{i,j} \mathbf{W}_1 \text{LSTM}(P_{i,j}[k]), \\ \hat{\alpha}_{(i,j,\cdot)} &= \text{SoftMax}(\alpha_{(i,j,\cdot)}), \\ \hat{\mathbf{R}}_{i,j} &= \sum_k \hat{\alpha}_{(i,j,k)} \cdot \text{LSTM}(P_{i,j}[k]). \end{aligned}$$

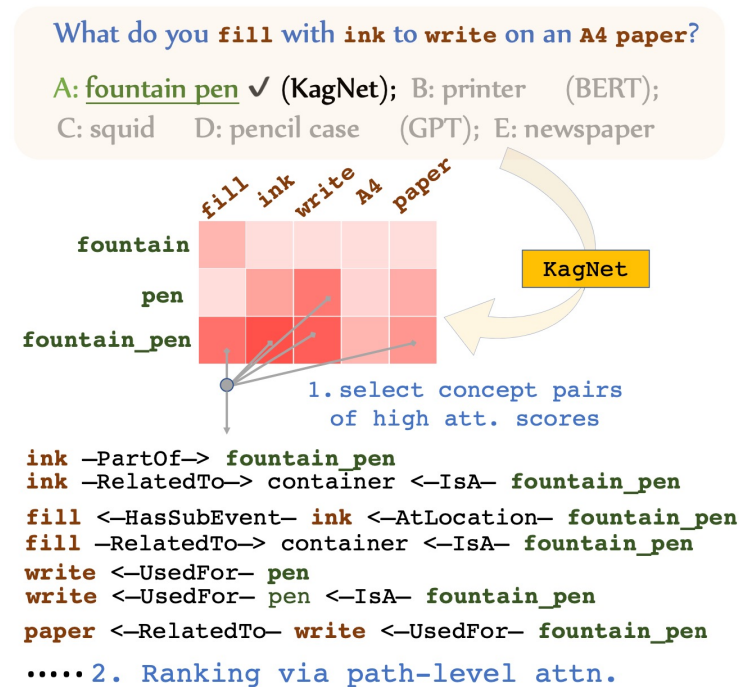
$$\begin{aligned} \beta_{(i,j)} &= \mathbf{s} \mathbf{W}_2 \mathbf{T}_{i,j} \\ \hat{\beta}_{(\cdot,\cdot)} &= \text{SoftMax}(\beta_{(\cdot,\cdot)}) \\ \hat{\mathbf{g}} &= \sum_{i,j} \hat{\beta}_{(i,j)} [\hat{\mathbf{R}}_{i,j}; \mathbf{T}_{i,j}] \end{aligned}$$

# 정답 맞추기

- 분석 output: graph vector
- for each input
  1. *Where do adults use glue sticks SEP*  
classroom
  2. *Where do adults use glue sticks SEP*  
office
  3. *Where do adults use glue sticks SEP*  
desk
  4. *Where do adults use glue sticks SEP*  
drawer
- sigmoid for each input -> graph vector
- get the max probability choice

# Explainability

- Q
  - What do you fill with ink to write on an A4 paper?
- A
  - fountain pen
- Attention Matrix
  - Concepts are focusing on fountain\_pen
  - Why pick this answer?
    - because they are related!
    - fill <-> ink, paper<-> write, ...
    - ink <-> fountain\_pen, write <-> fountain\_pen, ...



# Revisit: Contribution: Abstract and Conclusion

- Problem
  - Commonsense Inference
  - Particularly, Question and Answering Task
- 방법
  - utilize external and structured commonsense knowledge graphs
- 성과
  - explainable inference
  - SOTA(A little bit... :/)
  - outperforms BERT and GPT in Commonsense QA
    - BERT랑 GPT는 그냥 데이터만 많이 쑤셔넣음

