

순환 구조 인코더를 활용한 상식 기반 QA 모델의 개선 방법

정지원¹ 이소영² 박호건¹²¹성균관대학교 인공지능학과, ²성균관대학교 전기전자컴퓨터공학과
jwjw9603@g.skku.edu, lsy7451@g.skku.edu, hogunpark@g.skku.edu

Improving Commonsense-based QA Model through a Cycle-Encoder

Jiwon Jeong¹ Soyoung Lee² Hogun Park¹²¹Department of Artificial Intelligence, Sungkyunkwan University²Department of Electrical and Computer Engineering, Sungkyunkwan University

요약

질의응답 시스템은 인간이 제시하는 질문에 대하여 스스로 답변을 제공하고자 하고, 나아가 인간 수준의 추론 능력을 가지는 것에 목표를 두고 있다. 최근 딥러닝 기술의 발전으로 언어 모델 기반 질의응답 시스템은 문맥이 주어질 시 많은 벤치마크 데이터 세트에서 인간에 준하는 답변 선택 능력을 보여주고 있으며, 지식 그래프를 함께 사용시, 문맥이 없는 상식 기반 질의응답에서도 큰 성능 향상을 관찰할 수 있다. 그러나 대부분의 지식 그래프 기반 질의응답 시스템 연구들은 추론을 위하여 질문/선택지를 지식 그래프와 통합하는 과정에서 많은 순환 구조(Cycle)를 생성하게 되는데, 여기서 발생하는 영향에 대해서는 연구자들이 아직 거의 주목하지 않았다. 일반적으로 논리 구조가 순환적이면 논리적 결함을 유발시킨다. 따라서 순환 구조를 고려하지 않고 학습 시, 신경망의 학습이 실패하거나 논리 오류가 발생할 수 있다. 본 논문에서는 상식 추론 과정에서 그래프 신경망의 입력으로 사용하는 하위 그래프 내 순환 구조를 식별할 수 있는 인코더를 사용하여, 기존의 QA 모델을 개선하고자 한다. 제안한 방법은 2가지 상식 기반 QA 데이터 세트에서 최대 4.4%의 성능 향상을 보인다.

1. 서론

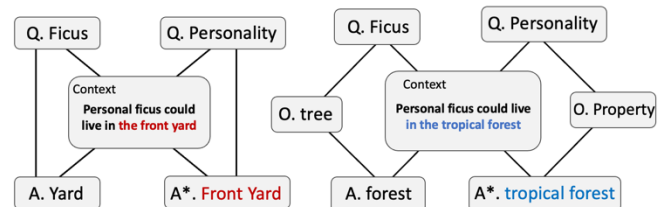
최근 등장한 대규모 사전 학습 언어 모델(Pretrained language Model)은 연구자들로 하여금 인공지능이 인간과 유사한 복잡한 추론 문제에 대해 대답하는 것을 기대할 수 있게 만들었다[1]. Question Answering System(QA)은 이러한 추론 연구 중 인간이 제기한 질문에 기계가 얼마나 자연스러운 대답을 생성하는지 알아보기 위한 연구로 크게 정보기반 질의응답 시스템과 지식 그래프 기반 질의응답 시스템으로 나뉜다.

지식 그래프(Knowledge graph) 기반 질의응답 시스템은 지식 그래프로 표현된 지식들을 사용하여 질의응답 시스템 학습에 사용하는데, 여기서 지식 그래프는 객체(Entity)들을 나타내는 노드(Node)와 객체 사이의 관계(Relation)를 의미하는 에지(Edge)로 구성된 삼중 그래프 데이터(Head relation tail triplet) 데이터베이스를 의미한다.

이러한 지식 그래프 기반 질의응답 시스템(KG-QA)의 대표적인 모델로 QA-GNN[2]와 Graph Soft Counter(GSC)[3]가 있는데, QA-GNN은 언어 모델을 통해 추출된 컨텍스트(Context)를 지식 그래프에 추가적인 노드로 사용하여 언어 모델과 지식 그래프의 joint reasoning을 도입하였으며 GSC는 QA-GNN의 그래프 신경망을 단순화하여 복잡도를 낮추는 방식을 제안하였다. 그러나 이러한 컨텍스트의 도입에는 질문, 답변 후보 노드와 양방향으로 연결이 이뤄지면서 필연적으로 순환 그래프가 발생하게 된다.

Q. Where could a personal ficus live?

A. Front yard B. Cabin in the wood C. California D. conservatory E. tropical forest



Graph A

Graph B

그림 1. 컨텍스트 노드가 도입된 그래프 구조

그림 1은 컨텍스트 노드의 도입으로 순환 그래프가 생성된 KG-QA 데이터 구조이다. 주어진 질문과 5개의 답안에 따라 각 답안과 질문을 연결한 5개의 하위 그래프들을 생성하게 된다. 생성된 하위 그래프들 중 두 개의 하위 그래프, Graph A와 Graph B를 보면 공통적으로 질문 벡터 Q 와 답안 벡터 A 에 연결된 Context 노드로 구성되어 있다. 그러나 Graph B에는 지식 그래프를 통해 연결된 그 외 노드(Other node) O 가 존재하기 때문에 사각형의 모습이 된다. 이러한 데이터 구조를 학습하기 위해 기존 모델들은 그래프 신경망(GNN)을 사용하는데 그래프 신경망은 그래프의 구조를 학습할 때 삼각형과 사각형을 구분하지 못하는 문제가 있다[7]. 이것은 QA 그래프 구조에서 노드의 값들을 알 수 없을 때, 추가적인

Other node의 존재 유무가 주는 영향력을 배제하게 만든다. 때문에 우리는 이러한 그래프 구조에 특성을 더 잘 담아낼 수 있는 순환 추론 인코더를 사용하여 KG-QA 시스템의 향상된 표현력을 이끌어 내고자 하였다. 우리는 대표적인 두가지 질의응답 시스템 벤치마크에서 제안 모델의 우수성을 입증한다.

2. 본론

2.1 제안 모델 : 순환 구조(Cycle) 인코더 기반

질의응답 시스템

본 논문은 질의응답 시스템 중 상식을 기반으로 다루는 상식 추론 시스템을 다루며 그래프의 순환에 문제점을 제기하고 이를 해결하기 위한 방법을 제시한다. 순환 구조란 경로 중 시작 노드와 끝 노드가 동일한 경우를 나타내며 순환 구조 그래프는 순환 구조를 가진 그래프를 지칭한다. 순환 구조가 질의응답에서 중요한 이유는 아래와 같다.

- (1) 기존 지식 그래프 기반 질의응답 시스템은 노드 간의 관계에만 집중하고 그래프의 전체적인 위상, 구조(특히, 순환)를 고려하지 않았다.
- (2) 순환 구조를 포함한 서로 다른 그래프는 그래프 신경망으로 구분할 수 없는 문제가 있고, 기존의 지식 그래프 기반 QA 모델들은 이러한 그래프 신경망을 그대로 사용하고 있다[4, 5].
- (3) 순환 구조가 추론에 사용 시, 순환 추론(논법)에 해당하는 상황이 발생할 수 있다. 순환 추론이란 화자가 어떤 주장을 펼침에 있어 근거로 그 주장을 다시 사용하는 논리 오류에 해당한다.

본 논문에서는 위와 같은 문제를 해결하기 위해, Fast-IO-GNN [7]에서의 순환 구조 개수를 준 지도 그래프 학습에 사용한 아이디어에 착안하여, 순환 구조의 길이를 사용하는 인코더를 통하여 기존의 QA 모델을 개선하고자 한다.

알고리즘 1은 질의응답 모델 중 하나인 GSC에 순환 구조 개수 벡터를 적용한 방법을 의사 코드로 표현하였다. adj 는 에지 인덱스(Edge index)를 나타낸 것으로 에지의 방향을 기준으로 출발 노드와 도착 노드가 두 개의 차원으로 구성되어 있다. $node_type$ 은 노드의 타입을 의미하는 벡터이고, A 는 하위 그래프의 인접 행렬을 의미한다. $cycle_feature$ 은 식 1, 식 2를 통해 구한 순환 구조 개수 벡터 $x_{v,3}^+$ 를 의미한다 (알고리즘1, 5번째 줄). 순환 구조 개수(Cycle count) 벡터 $x_{v,n}^+ \in \mathbb{R}^{1 \times K}$ 는 시작 노드 v 가 n 번의 이동으로 자기 자신 노드로 돌아올 수 있는 경우의 수로 구성되어있는 K 차원 벡터이다. 즉, 그래프 내에서 n 번의 이동으로 자기 자신에게 돌아오는 경우의 수를 측정한 값으로 이루어져 있다. 순환 구조 개수 벡터의 크기(K)는 QA-GNN, GSC 모델에 따라 다르며 실험을 통해 각각 30, 5차원으로 지정하였다.

$$x_{v,3}^+ = \text{Diag}(A^3)[v] \quad (1)$$

$$x_v = \text{MLP}\left(\text{CONCAT}(x_v, x_{v,3}^+)\right) \quad (2)$$

식 1을 통해 구한 우리의 순환 구조 개수 벡터 $x_{v,3}^+$ 는 길이 30에 해당하는 순환, 즉 3번 이동으로 자기 자신으로 되돌아오는 경우의 수를 요소(Element)값으로 하고 있는데, 기존 연구의 순환 구조 개수 벡터 같은 경우 벡터의 i 번째 자리에 i 길이 순환의 경우의 수가 값으로 계산되었으나 지식 그래프의 특성 상 삼각형과 사각형을 구분하는 3번째 요소의 값이 가장 큰 영향을 주어 우리는 계산상 편의를 위하여 $i = 3$ 인 값만 가지는 순환 구조 개수 벡터를 생성하고 기존 노드 feature에 연결(Concatenation)하는 방식을 취한다. 이에 따라 순환 구조 개수 벡터는 동일한 값을 가지는 K 차원의 벡터가 된다.

이러한 순환 구조 개수 벡터와 노드 초깃값(GSC는 0으로 초깃값 설정)을 연결(Concatenation)한 값을 두 개의 레이어로 이뤄진 Fully connected layer를 통해 1차원 크기의 새로운 벡터로 반환한 것을 우리는 노드 임베딩($node_emb$)이라고 한다. 에지 임베딩($edge_emb$)이란, 에지 인코더($edge_encoder$)를 통해 노드 타입($node_type$)과 에지 타입($edge_type$) 그리고 에지 인덱스(adj)를 입력으로 1차원 벡터를 반환한 값을 의미한다. 에지 인코더는 두 개의 레이어로 이루어진 Fully connected layer이다. GSC 모델의 학습 과정을 요약하자면 다음과 같다.

- (1) 출발 노드 임베딩($node_emb$)과 에지의 임베딩($edge_emb$)값을 더한다. (알고리즘1, 12번째 줄)
- (2) 더한 값을 도착 노드에 전파(Scatter)하고 (1)의 과정을 레이어 수만큼 반복한다. (알고리즘1, 14번째 줄)

위와 같이 그래프에 속한 노드에 순환 구조 개수 벡터를 추가하는 방법은 다양한 그래프 신경망 기반 QA 모델에 적용할 수 있다. 예를 들어, QA-GNN 모델은 내부 Message Passing 과정의 변경없이, 노드 임베딩 초깃값에 순환 구조 개수 벡터를 연결하고 두 개의 레이어로 이루어진 fully connected layer를 통과하도록 구성하여 모델을 개선할 수 있다.

Algorithm 1 PyTorch-style code of our Cycle encoder-based GSC model

```

1: # adj : edge index with shape 2 x N_edge
2: # edge_type : edge type with shape 1 x N_edge
3: # node_type : node type with shape 1 x N_node
4: # A : adj_matrix used for cycle_feature
5: # cycle_feature : A3.Diagonal() with shape N_node x 5
6: # num_gsc_layers : iteration of message passing

7: edge_emb = edge.encoder(adj, edge_type, node_type)
8: node_emb = torch.concat([torch.zeros(N_node, 1), cycle_feature], -1)
9: node_emb = 2-layer MLP(node_emb)

10: for i_layer in range(num_gsc_layers) do
11:     # propagate from node to edge
12:     edge_emb += node_emb[adj[0]]
13:     # aggregate from edge to node
14:     node_emb = scatter(inputs, adj[1])
15: end for

```

알고리즘 1. 순환 구조 개수 벡터를 적용한 GSC Message Passing 과정

3. 실험

3.1 데이터 세트

본 연구는 두 가지 모델을 두 가지 데이터 세트 CommonsenseQA[8], OpenBookQA[9]로 평가를 진행했다. CommonsenseQA는 오지선다의 문제로 문맥 없이 상식에 의존하여 질의에 대한 정답을 찾는 문제이다. 공식적인 모델 평가는 이주에 한 번씩 리더보드를 통해 진행되므로 본 실험은 공식적인 리더보드가 아닌 [10]에서 진행한 데이터 분할을 통해 In-House로 실험을 진행했다. OpenBookQA는 사지선다의 문제로 기초적인 과학 상식을 요구하는 문제이다. [11] 방식을 사용하여 데이터 분할을 진행 후 In-House로 실험을 진행했다. CommonsenseQA와 OpenBookQA를 위해, 본 논문은 ConceptNet[12] 지식 그래프를 활용했다. 지식 그래프의 전처리 방식은 기존 모델 QA-GNN, GSC에서 사용한 방식과 동일하다.

3.3 실험 결과

본 실험은 CommonsenseQA와 OpenBookQA 데이터 세트를 사용하여 실험을 진행했다. 사전 학습 언어 모델은 RoBERTa-large[13]를 사용했으며 QA-GNN, GSC 두 모델에 순환 구조 개수 벡터를 적용하였다. 각 모델들의 성능은 해당 논문의 성능 결과를 참고하였다.

표 1. CommonsenseQA 에 대한 결과

Methods	1Hdev-ACC%	1Htest-ACC%
KagNet	73.47%	69.01%
RN	74.57%	69.08%
MHGRN	74.45%	71.11%
QA-GNN	76.54%	73.41%
GSC	79.11%	74.48%
QA-GNN(+cycle feature)	78.05%	74.62%
GSC(+cycle feature)	79.36%	75.58%

표 1은 CommonsenseQA 데이터 세트의 결과이며 성능 지표는 정확도(Accuracy)를 사용하였다. QA-GNN과 GSC 두 모델에 순환 구조 개수 벡터를 적용했을 경우 QA-GNN은 검증 데이터 세트는 1.51%, 테스트 데이터 세트는 1.69%의 성능 향상이 있었다.

표 2. OpenBookQA 에 대한 결과

Methods	1Htest-ACC%
RoBERTa-large	64.80%
RN	65.20%
MHGRN	66.85%
QA-GNN	67.80%
GSC	70.33%
QA-GNN(+cycle feature)	72.20%
GSC(+cycle feature)	72.40%

표 2는 OpenBookQA 데이터 세트의 결과이며 표1과 같이 성능 지표로 정확도(Accuracy)를 적용하였다. 기존 OpenBookQA 데이터 세트를 실험한 논문에서 테스트 데이터 세트에 대한 정확도만 표시하였으므로 본 논문도 테스트 데이터 세트에 대한 결과만 기록하였다. QA-GNN은 기존 모델 대비 4.40%, GSC는 2.07%의 성능이 향상되었다. 비교적 CommonsenseQA

데이터 세트에 비해 성능 향상이 컸다.

4. 결론 및 향후 연구

본 연구에서는 지식 그래프와 사전 학습 언어 모델을 질의응답 중 상식 추론 과정에 사용하는 경우 그래프의 순환 구조의 영향에 주목하고, 이에서 오는 문제를 해결하는 질의 응답 시스템을 제안하였다. 실험 결과 각 데이터 별로 평균 0.92%, 3.24% 성능 향상이 있었다. 향후 다양한 데이터, 모델에 대한 검증과 순환 그래프의 영향에 대한 구체적인 분석과 순환 추론에 더욱 적합한 그래프 신경망을 구성하는 것이 필요하다.

감사의 글

이 성과는 2021년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2021R1C1C1005407). 또한, 이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2019-0-00421, 인공지능대학원지원(성균관대학교))과 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신산업진흥원의 지원을 받아 수행된 헬스케어 AI융합 연구개발 사업을 통해 나온 결과물임(N0.S0254-22-1001).

[참고 문헌]

- [1] WU et al. Graph neural networks for natural language processing: A survey. arXiv:2106.06090, 2021.
- [2] YASUNAGA et al. QA-GNN: Reasoning with language models and knowledge graphs for question answering. In: NAACL. 2021.
- [3] WANG et al. GNN is a counter? revisiting GNN for question answering. In: ICLR. 2022.
- [4] XU et al. How powerful are graph neural networks?. In: ICLR. 2019.
- [5] ZHANG and LI. Nested graph neural networks. In: NeurIPS 2021.
- [6] GILMER et al. Neural message passing for quantum chemistry. In: ICML. 2017.
- [7] YOU et al. Identity-aware graph neural networks. In: AAAI. 2021.
- [8] TALMOR et al. Commonsenseqa: A question answering challenge targeting commonsense knowledge. In: NAACL. 2019.
- [9] MIHAYLOV et al. Can a suit of armor conduct electricity? a new dataset for open book question answering. In: EMNLP. 2018.
- [10] LIN et al. Kagnet: Knowledge-aware graph networks for commonsense reasoning. In: EMNLP. 2019.
- [11] Mihaylov and Todor. Knowledgeable reader: Enhancing cloze-style reading comprehension with external commonsense knowledge. In: ACL. 2018.
- [12] Speere et al. Conceptnet 5.5: An open multilingual graph of general knowledge. In: AAAI 2017.
- [13] Liu et al. Roberta: A robustly optimized bert pretraining approach. arXiv:1907.11692, 2019.