

Question Generation

Guiding the Growth: **Difficulty-Controllable** Question Generation
through Step-by-Step Rewriting

--2021-main-ACL

Controllable Open-ended Question Generation with A New
Question Type Ontology

--2021-main-ACL

Guiding the Growth: **Difficulty-Controllable** Question Generation through Step-by-Step Rewriting

Background: Difficulty-Controllable Question Generation(factoid questions)

QUESTION REWRITING

Q₁ : Who starred Top Gun?

Q₂ : Who starred the film directed by Tony Scott ? *(BRIDGE)*

Q₃ : Who starred a 1986 action film directed by Tony Scott? *(INTERSECTION)*

Q₄ : Who starred Rain Man and a 1986 action film directed by Tony Scott? *(INTERSECTION)*

Q₅ : Who starred a film directed by Barry Levinson and a 1986 action film directed by Tony Scott? *(BRIDGE)*

Two challenges of this task

- The definition of difficulty level
- Controllability of difficulty

Contribution

- The first work of difficulty-controllable question generation, with question difficulty defined as the inference steps to answer it.
- Propose a novel framework that achieves DCQG through step-by-step rewriting

QUESTION REWRITING

Q₁ : Who starred Top Gun?

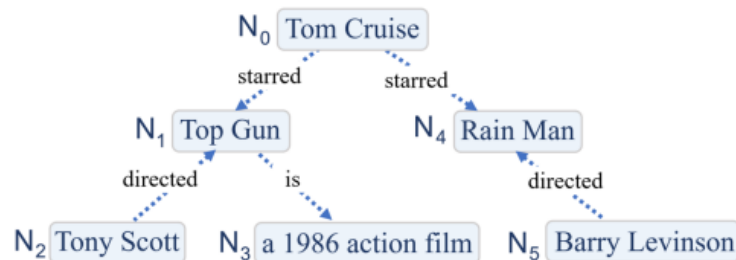
Q₂ : Who starred the film directed by Tony Scott ? *(BRIDGE)*

Q₃ : Who starred a 1986 action film directed by Tony Scott? *(INTERSECTION)*

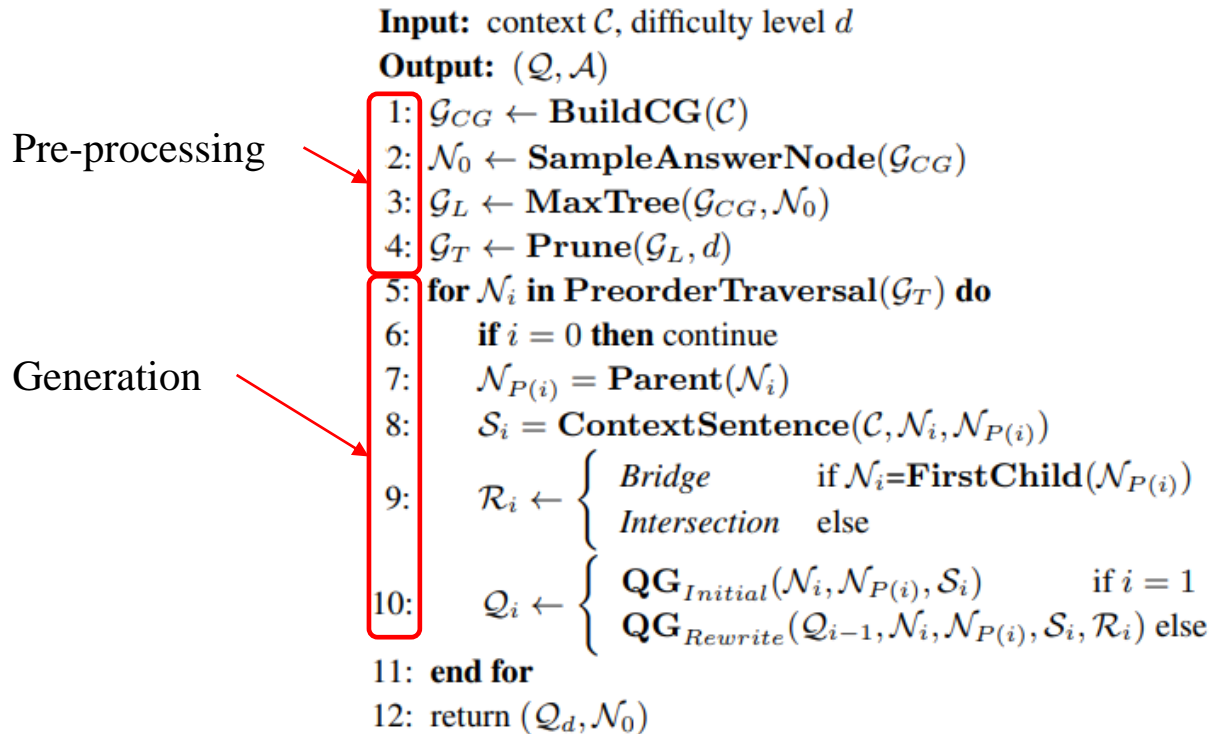
Q₄ : Who starred Rain Man and a 1986 action film directed by Tony Scott? *(INTERSECTION)*

Q₅ : Who starred a film directed by Barry Levinson and a 1986 action film directed by Tony Scott? *(BRIDGE)*

REASONING CHAIN



Method: Overview



Method: Pre-processing

- Context Graph Construction

Open information extraction_[1] \rightarrow Coreference resolution_[2]

- Reasoning Chain Selection

$\mathcal{N}_0 \leftarrow \text{SampleAnswerNode}(\mathcal{G}_{CG})$

$\mathcal{G}_L \leftarrow \text{MaxTree}(\mathcal{G}_{CG}, \mathcal{N}_0)$

$\mathcal{G}_T \leftarrow \text{Prune}(\mathcal{G}_L, d)$

1. Gabriel Stanovsky, Julian Michael, Luke Zettlemoyer, and Ido Dagan. 2018. Supervised open information extraction.

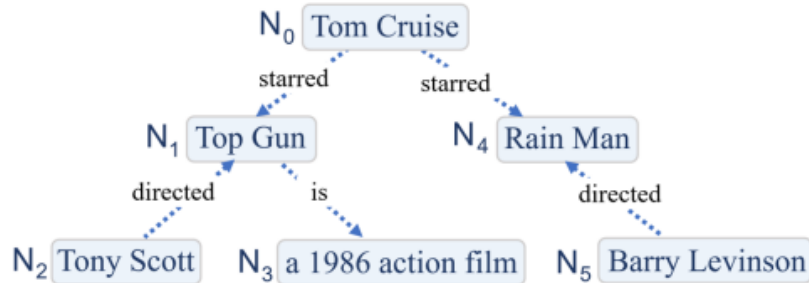
2. Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. 2017. End-to-end neural coreference resolution. I

Method: Generation

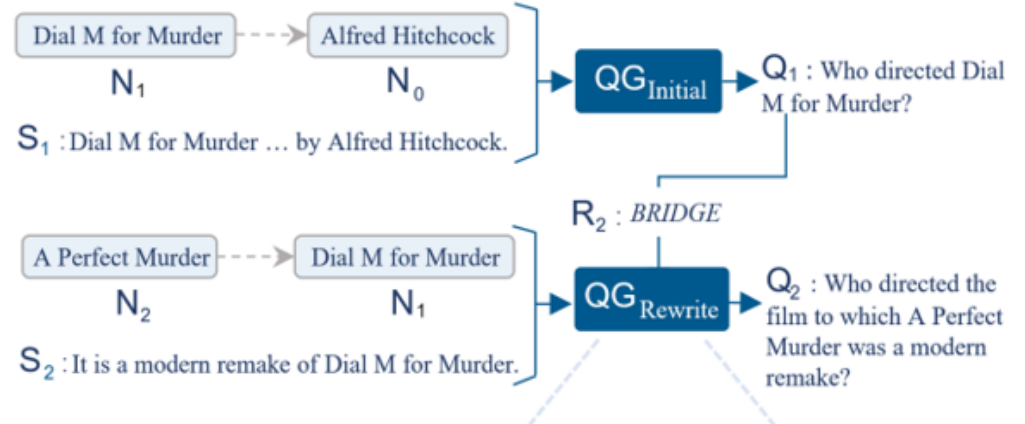
QUESTION REWRITING

- Q₁ : Who starred Top Gun?
- Q₂ : Who starred the film directed by Tony Scott ? *(BRIDGE)*
- Q₃ : Who starred a 1986 action film directed by Tony Scott? *(INTERSECTION)*
- Q₄ : Who starred Rain Man and a 1986 action film directed by Tony Scott? *(INTERSECTION)*
- Q₅ : Who starred a film directed by Barry Levinson and a 1986 action film directed by Tony Scott? *(BRIDGE)*

REASONING CHAIN



Step-by-step Question Generation:



for \mathcal{N}_i **in** $\text{PreorderTraversal}(\mathcal{G}_T)$ **do**

if $i = 0$ **then** **continue**

$\mathcal{N}_{P(i)} = \text{Parent}(\mathcal{N}_i)$

$\mathcal{S}_i = \text{ContextSentence}(\mathcal{C}, \mathcal{N}_i, \mathcal{N}_{P(i)})$

$\mathcal{R}_i \leftarrow \begin{cases} \text{Bridge} & \text{if } \mathcal{N}_i = \text{FirstChild}(\mathcal{N}_{P(i)}) \\ \text{Intersection} & \text{else} \end{cases}$

$\mathcal{Q}_i \leftarrow \begin{cases} \text{QG}_{\text{Initial}}(\mathcal{N}_i, \mathcal{N}_{P(i)}, \mathcal{S}_i) & \text{if } i = 1 \\ \text{QG}_{\text{Rewrite}}(\mathcal{Q}_{i-1}, \mathcal{N}_i, \mathcal{N}_{P(i)}, \mathcal{S}_i, \mathcal{R}_i) & \text{else} \end{cases}$

end for

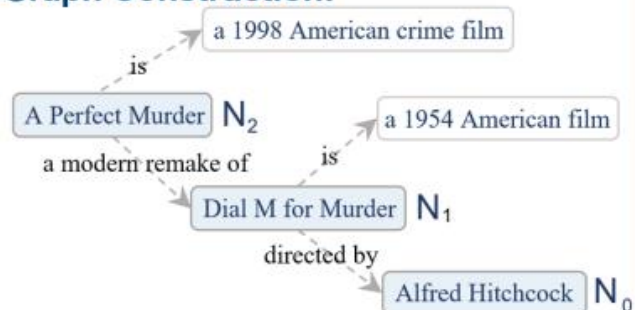
Method: Model

Context:

Paragraph A: A Perfect Murder is a 1998 American crime film. It is a modern remake of Dial M for Murder.

Paragraph B: Dial M for Murder is a 1954 American film directed by Alfred Hitchcock

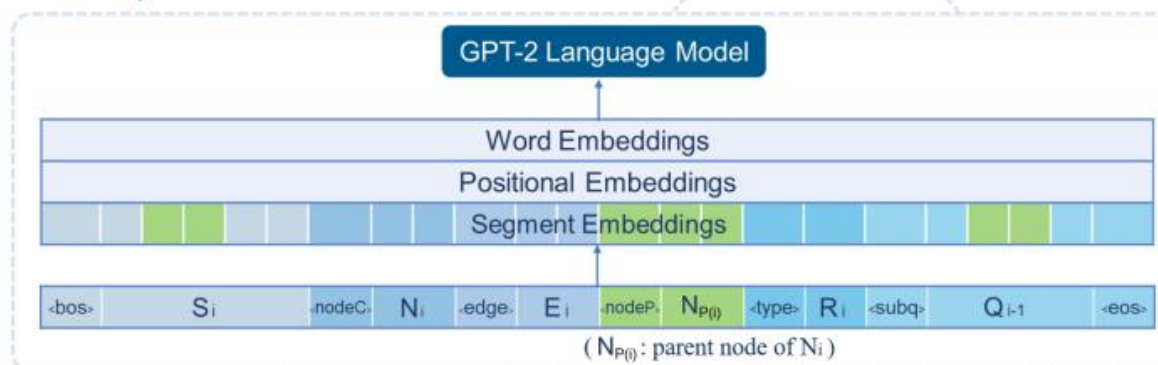
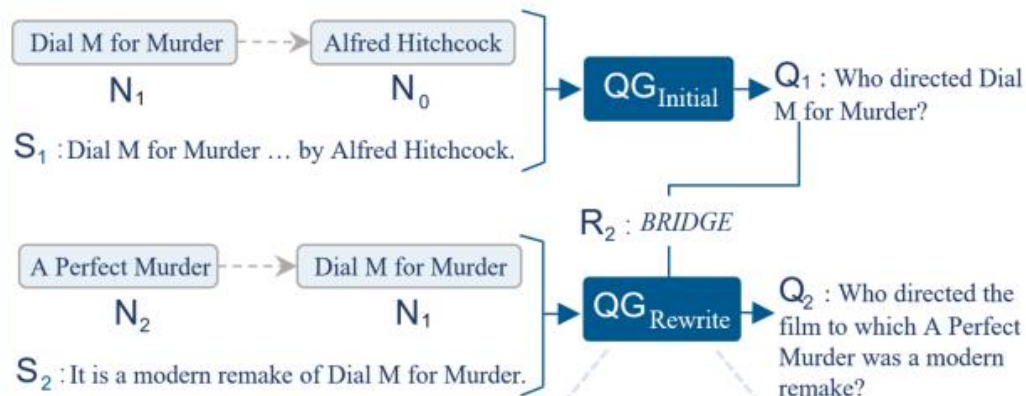
Graph Construction:



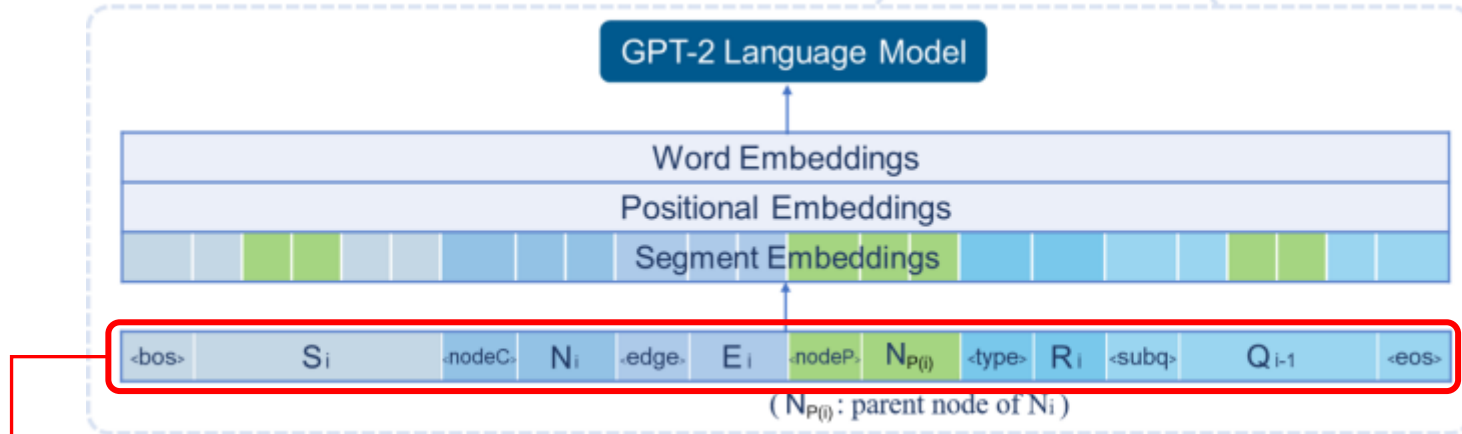
Implementation of QG_{Rewrite} :

(Implementation of QG_{Initial} is similar except without "`<type> Ri <subq> Qi-1`")

Step-by-step Question Generation:



Method: Data construction



Input: context $\mathcal{C} = \{\mathcal{P}_1, \mathcal{P}_2\}$, QA pair (Q_2, \mathcal{A}_2) , supporting facts \mathcal{F}

Output: $\mathcal{R}_1, (Q_1, \mathcal{A}_1), S_1, S_2, \{\mathcal{N}_0, \mathcal{E}_1, \mathcal{N}_1, \mathcal{E}_2, \mathcal{N}_2\}$

1: $\mathcal{R}_1 \leftarrow \text{TypeClassify}(Q_2)$

2: **if** $\mathcal{R}_1 \notin \{\text{Bridge}, \text{Intersection}\}$ **then** return

3: $\text{subq}_1, \text{subq}_2 \leftarrow \text{DecompQ}(Q_2)$

4: $\text{suba}_1, \text{suba}_2 \leftarrow \text{QA}(\text{subq}_1), \text{QA}(\text{subq}_2)$

5: $Q_1, \mathcal{A}_1 \leftarrow \begin{cases} \text{subq}_2, \text{suba}_2 & \text{if } \mathcal{A}_2 = \text{suba}_2 \\ \text{subq}_1, \text{suba}_1 & \text{else} \end{cases}$

6: $S_1, S_2 \leftarrow \begin{cases} \mathcal{F} \cap \mathcal{P}_1, \mathcal{F} \cap \mathcal{P}_2 & \text{if } Q_1 \text{ concerns } \mathcal{P}_1 \\ \mathcal{F} \cap \mathcal{P}_2, \mathcal{F} \cap \mathcal{P}_1 & \text{else} \end{cases}$

7: $\mathcal{N}_2 \leftarrow \text{FindNode}(\mathcal{A}_2)$

8: $\mathcal{N}_0, \mathcal{E}_1, \mathcal{N}_1, \mathcal{E}_2 \leftarrow \text{Match}(\text{subq}_1, \text{subq}_2)$

One-hop data: QG_{Initial}

Two-hop data: QG_{Rewrite}

Result:

Model	BLEU3	BLEU4	METEOR	CIDEr
NQG++	15.41	11.50	16.96	-
ASs2s	15.21	11.29	16.78	-
SRL-Graph	19.66	15.03	19.73	-
DP-Graph	19.87	15.23	20.10	1.40
GPT2	20.98	15.59	24.19	1.46
Ours _{2-hop}	21.07	15.26	19.99	1.48

Difficulty Level	Model	Well-formed			Concise			Answerable		Answer Matching	
		Yes	Acceptable	No	Yes	Acceptable	No	Yes	No	Yes	No
2-hop	DP-Graph	28%	41%	31%	41%	53%	6%	49%	51%	39%	61%
	GPT2	57%	34%	9%	47%	50%	3%	69%	31%	66%	34%
	Ours _{2-hop}	74%	19%	7%	67%	30%	3%	78%	22%	69%	31%
	Gold _{2-hop}	72%	22%	6%	56%	40%	4%	92%	8%	87%	13%
1-hop	Ours _{1-hop}	46%	46%	8%	65%	25%	10%	81%	19%	72%	28%
	Gold _{1-hop}	56%	39%	5%	80%	16%	4%	84%	16%	79%	21%

Controllable Open-ended Question Generation with A New **Question Type** Ontology

Background: Controllable Open-ended Question Generation

Input: It's a difficult task to undertake. Teenagers tend to identify gangs with “fitting” in. Peer pressure plays a large part in it and sometimes teenagers have problems with their own identity being part of a gang deals with those issues. It also provides a little bit of respect on the street ...

BART SAMPLING:

- How do you stop a teen from joining a gang? (PROCEDURAL)
- How do you get teenagers to stop being in gangs? (PROCEDURAL)
- How do you get teens out of gangs? (PROCEDURAL)

BART + QWORD:

- **How** do you get a teenager out of a gang? (PROCEDURAL)
- **What** is the best way to get teenagers out of gangs? (PROCEDURAL)
- **Why** do teenagers join gangs? (CAUSE)

Three challenges of this task

- Specifying the question type
- Capturing central concepts as question focus
- Encouraging the diversity of generated questions

Contribution

- **A new question type ontology**

Previous: what, how, why,...

Now: How do..., what do you think...


- **A type-aware framework**

Question type classifier

- **Use templates to improve controllability and generation diversity**

Task Formulation:

Input:

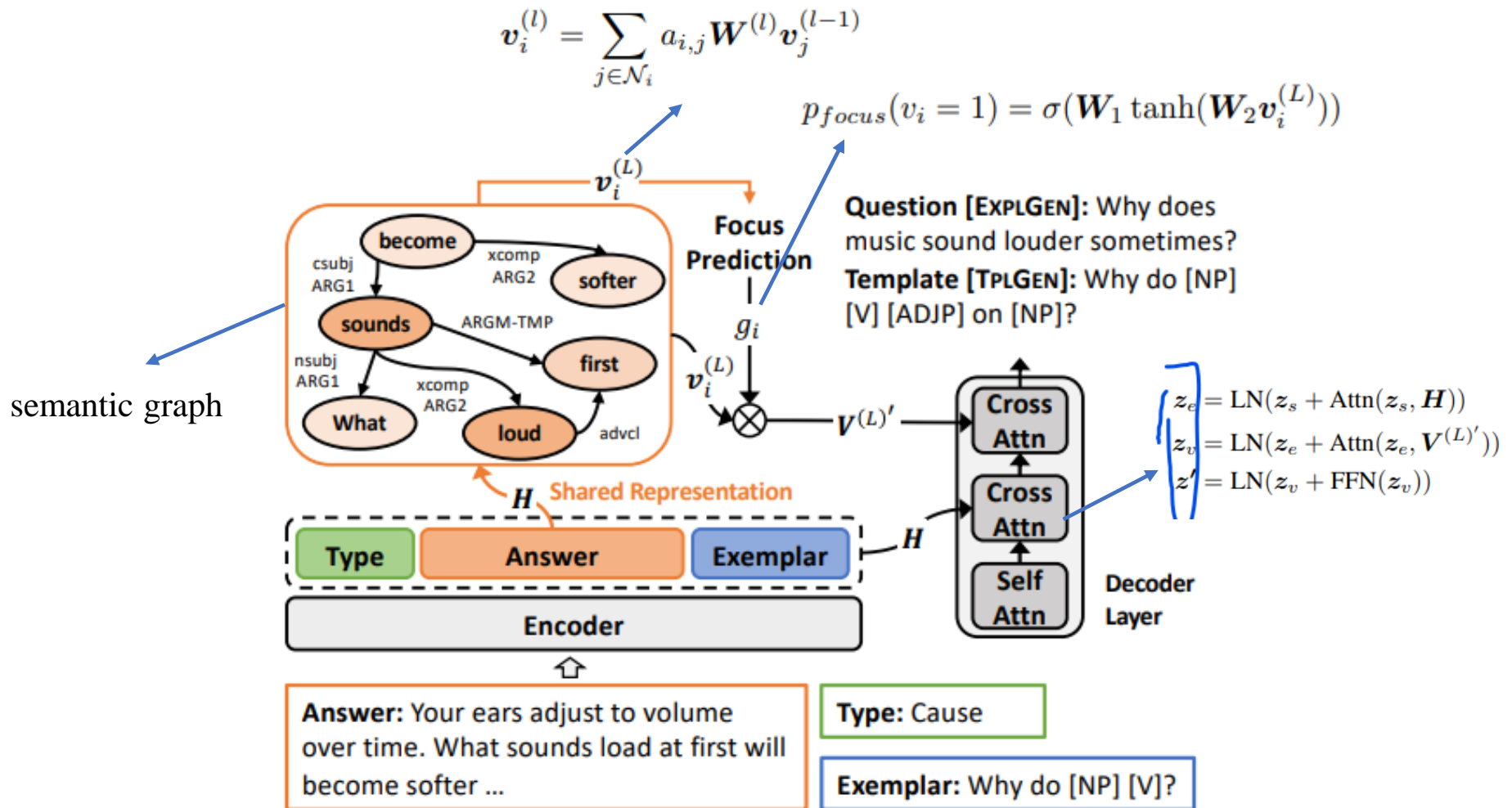
- Contextual text: $x = \{x_1, x_2, \dots, x_n\}$;
 - Question type: t ;
 - Template: x_T
- 

Output:

- Question: y

Question Type	Template Exemplars
VERIFICATION	"Is [NP] [NP]?", "Is there [NP]?", "Is [NP] [ADJP]?", "Can [NP] [V] [NP]?", "Do [NP] [V] [NP]?", "Does anyone have [NP]?", "Is it [ADJP] to [V] [NP]?"
DISJUNCTIVE	"Is [NP] [NP] or [NP]?", "Is [NP] [ADJP] or [ADJP]?", "Who is [NP] or [NP]?", "What came [ADVP] [NP] or [NP]?", "Which is [NP] or [NP]?", "What is [NP] or [NP]?"
CONCEPT	"What is [NP]?", "What does [NP] mean?", "Who is [NP]?", "Where is [NP]?", "What is the meaning of [NP]?", "What does [NP] do?", "What do you know about [NP]?", "When is [NP]?", "What is meant by [NP]?", "Where did [NP] come from?", "Which is [NP]?", "When was [NP] [V]?", "What is the definition of [NP]?", "How is [NP]?", "Does anyone know anything about [NP]?", "What happened to [NP]?"
EXTENT	"What is [NP]?", "How [OTHER] is [NP]?", "How many [OTHER] are in [NP]?", "How many [NP]?", "How much does [NP]?"
EXAMPLE	"What are [NP]?", "What is a good [NP]?", "What is the best [NP]?", "Where can I [V] [NP]?", "What are some good [NP]?", "Does anyone have [NP]?"

Method: model



Result:

Model	B-4	Yahoo		B-4	Reddit	
		MTR	R-L		MTR	R-L
DEEPQG	6.53	25.92	27.56	–	–	–
BART	21.88	38.01	39.16	19.45	35.46	37.82
BART+QWORD	22.02	38.44	39.32	19.80*	35.85	38.48*
Type-aware Models						
BART+QTYPE	22.12	38.62	39.72	19.90*	35.83	38.68*
JOINTGEN (ours)	22.56*	38.63	40.40*	20.09*	35.75	39.07*
w/o graph	22.21	38.21	39.93	19.81*	35.60	38.47*
EXPLGEN (ours)	21.74	37.52	39.70	18.67	33.28	36.74
TPLGEN (ours)	21.51	36.55	39.63	17.83	31.69	36.05

Model	Appro.	Ans.	Scp.	Top 1
REFERENCE	4.77	3.96	3.79	34.5%
BART	4.93	4.02	3.81	39.7%
BART+QWORD	4.86	4.14	3.85	40.8%
BART+QTYPE	4.92	4.23	3.94	48.7%
JOINTGEN	4.90	4.25	3.96	50.5%
TPLGEN	4.92	4.19	3.87	46.4%

(a) YAHOO

Model	Appro.	Ans.	Scp.	Top 1
REFERENCE	4.90	4.43	4.37	47.1%
BART	4.89	4.27	4.21	43.9%
BART+QWORD	4.88	4.29	4.21	46.7%
BART+QTYPE	4.88	4.39	4.26	49.6%
JOINTGEN	4.84	4.45	4.38	50.3%
TPLGEN	4.81	4.21	4.19	33.1%