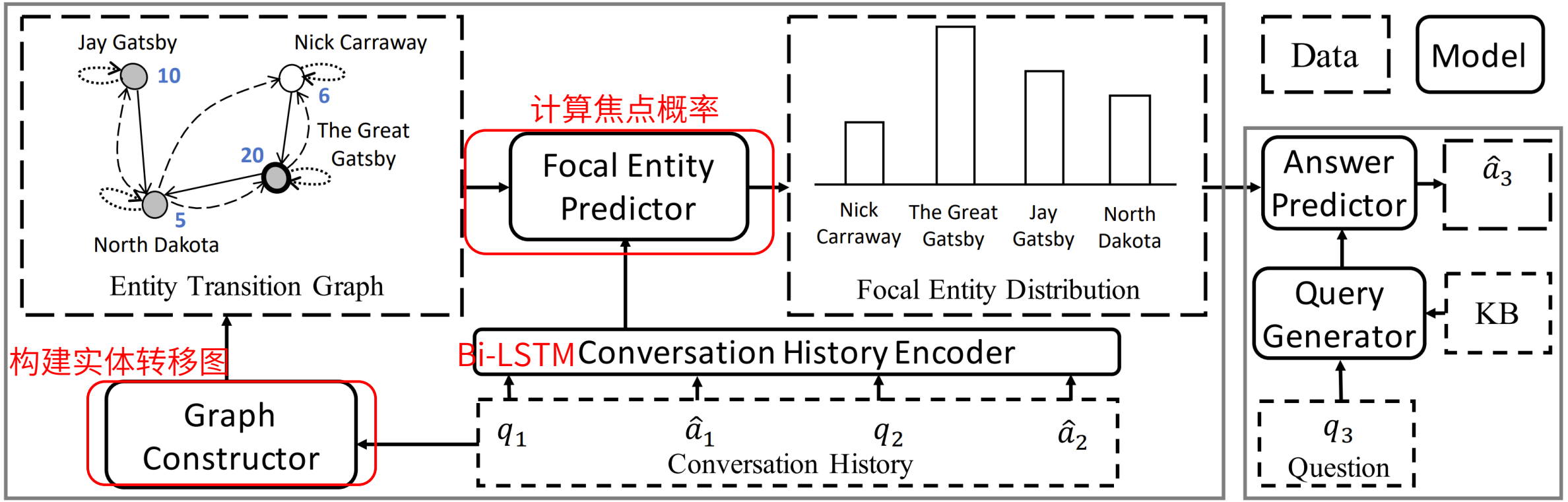
Conversational KBQA 例子：

Q1: What novel has the character named Nick Carraway?	
R1: The Great Gatsby	
Q2: Where is Jay Gatsby born?	The Great Gatsby 焦点
R2: North Dakota	
Q3: What is the name of the author?	The Great Gatsby
R3: F. Scott Fitzgerald	
Q4: What's his first novel?	F. Scott Fitzgerald
R4: This Side of Paradise	
Q5: Who was his child?	F. Scott Fitzgerald
R5: Frances Scott Fitzgerald	

Modeling Transitions of Focal Entities for Conversational Knowledge Base Question Answering. ACL2020: 3288-3297

论文概括：利用对话的焦点来实现对话答案的预测

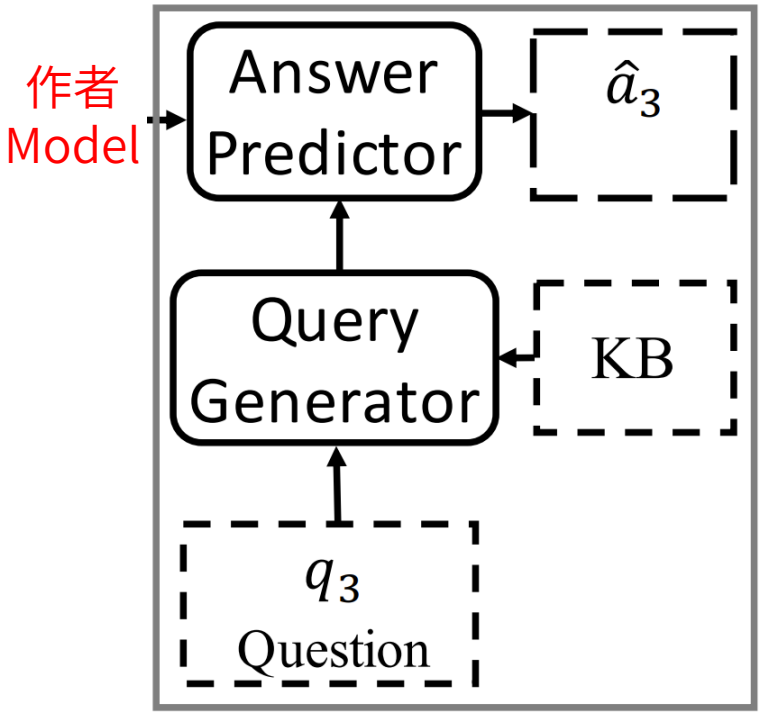
Q1: What novel has the character named Nick Carraway?	
R1: The Great Gatsby	
Q2: Where is Jay Gatsby born?	The Great Gatsby
R2: North Dakota	
Q3: What is the name of the author?	The Great Gatsby
R3: F. Scott Fitzgerald	
Q4: What's his first novel?	F. Scott Fitzgerald
R4: This Side of Paradise	
Q5: Who was his child?	F. Scott Fitzgerald
R5: Frances Scott Fitzgerald	



Modeling Transitions of Focal Entities for Conversational Knowledge Base Question Answering. ACL2020: 3288-3297

Single-turn KBQA

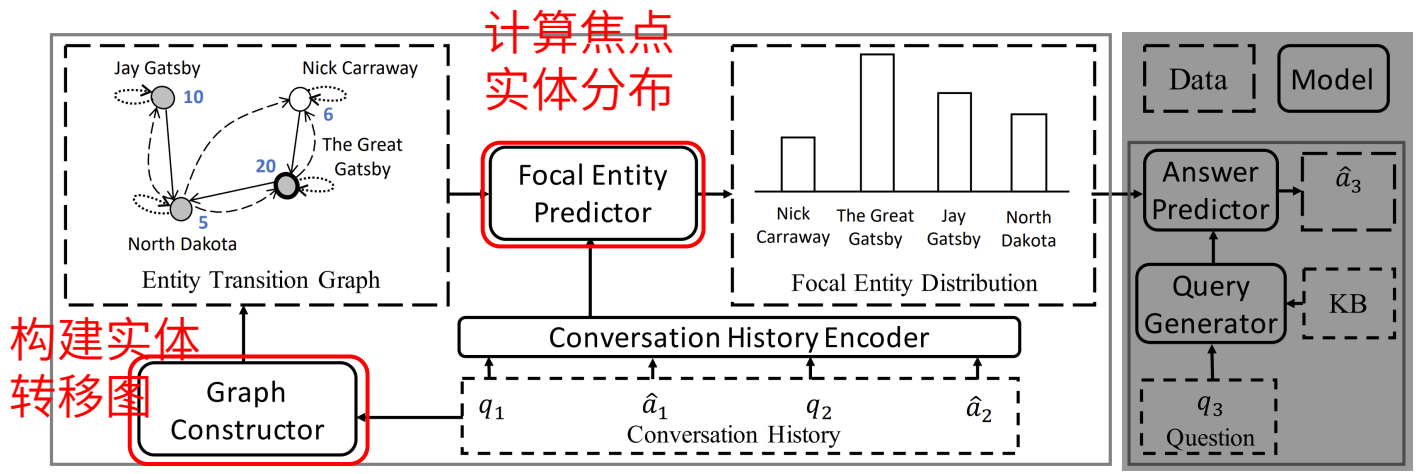
- Query Generator: 生成 candidate query graphs
- Answer Predictor: 通过 query graphs 来预测答案



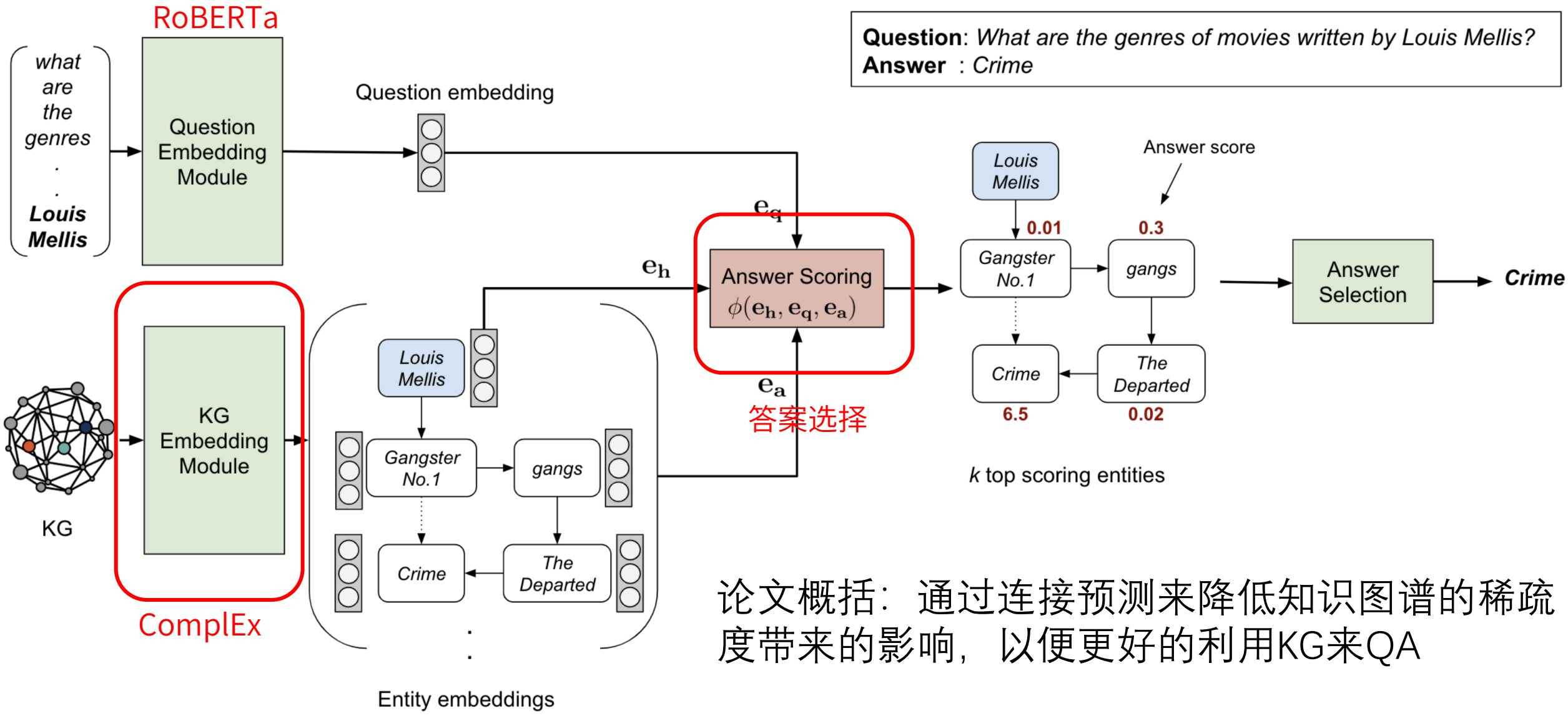
Modeling Transitions of Focal Entities for Conversational Knowledge Base Question Answering. ACL2020: 3288-3297

Author Method

- Graph Constructor: 利用一些规则在上一次的实体图上面继续更新
- Focal Entity Predictor: 通过GCN来求得实体概率分布



Modeling Transitions of Focal Entities for Conversational Knowledge Base Question Answering. ACL2020: 3288-3297

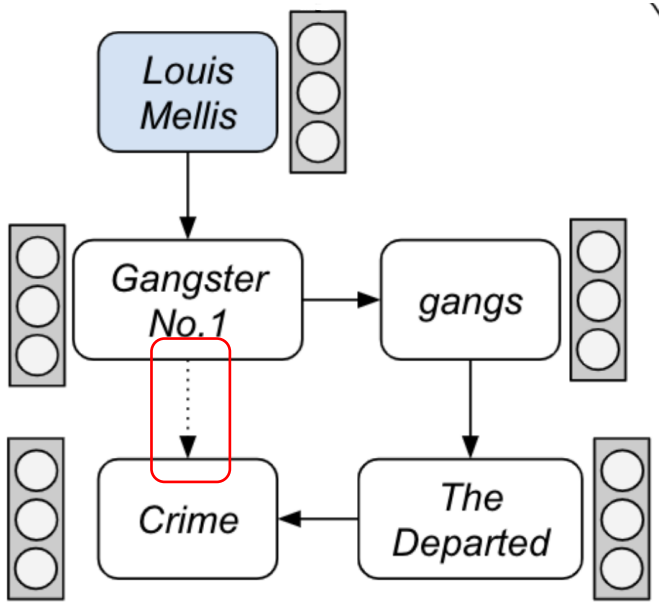


论文概括：通过连接预测来降低知识图谱的稀疏度带来的影响，以便更好的利用KG来QA

Improving Multi-hop Question Answering over Knowledge Graphs using Knowledge Base Embeddings. ACL2020: 7579-7589

KG Embedding Module

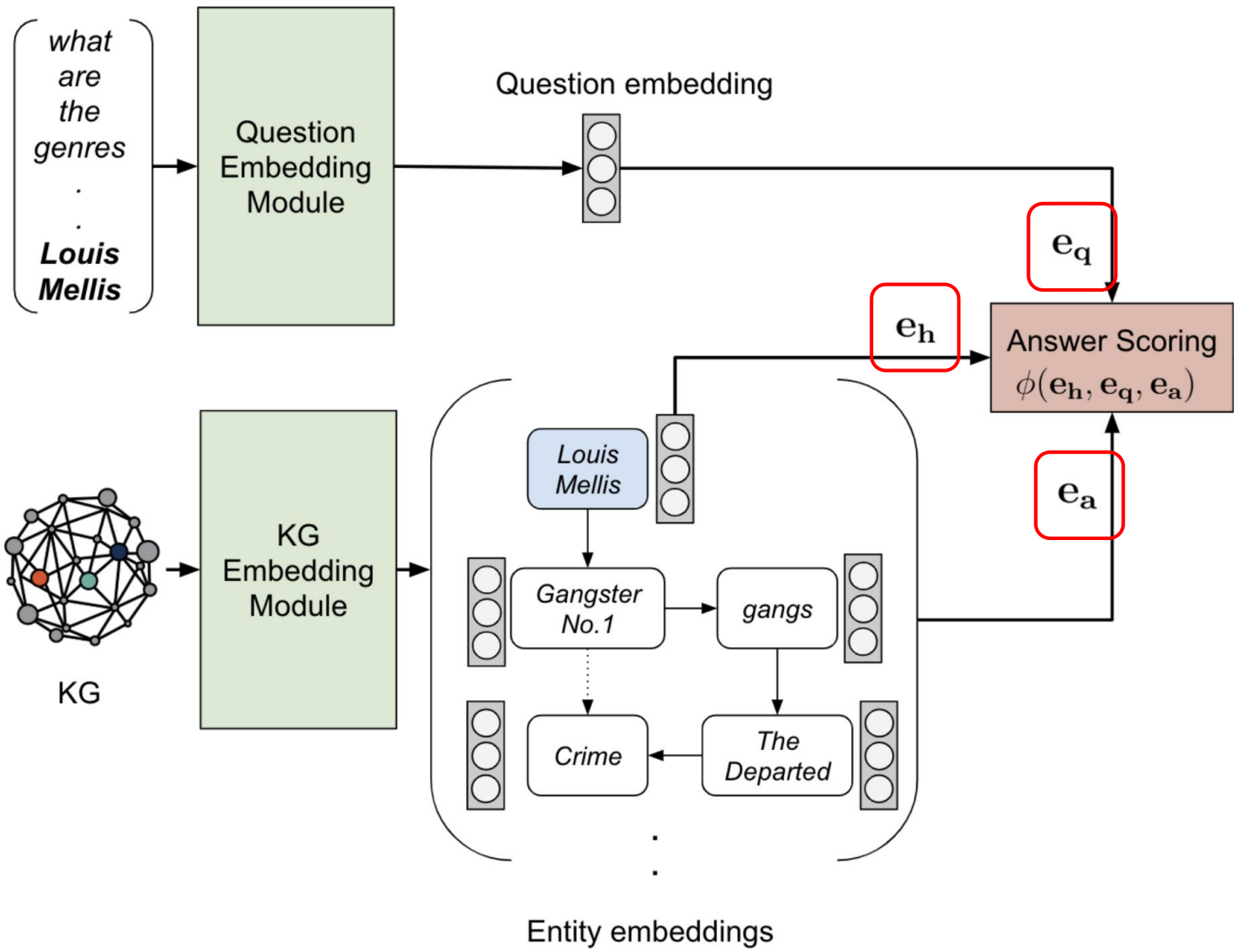
- 使用ComplEx算法对整个知识库中的实体和边嵌入到固定大小的向量（复数）
- 打分函数： $\emptyset(h,r,t) = RE(< e_h, e_r, e_t >), \emptyset(h,r,t) > 0$代表实体h, t之间有关系r
- 作用：弥补了稀疏KG的缺陷



Improving Multi-hop Question Answering over Knowledge Graphs using Knowledge Base Embeddings. ACL2020: 7579-7589

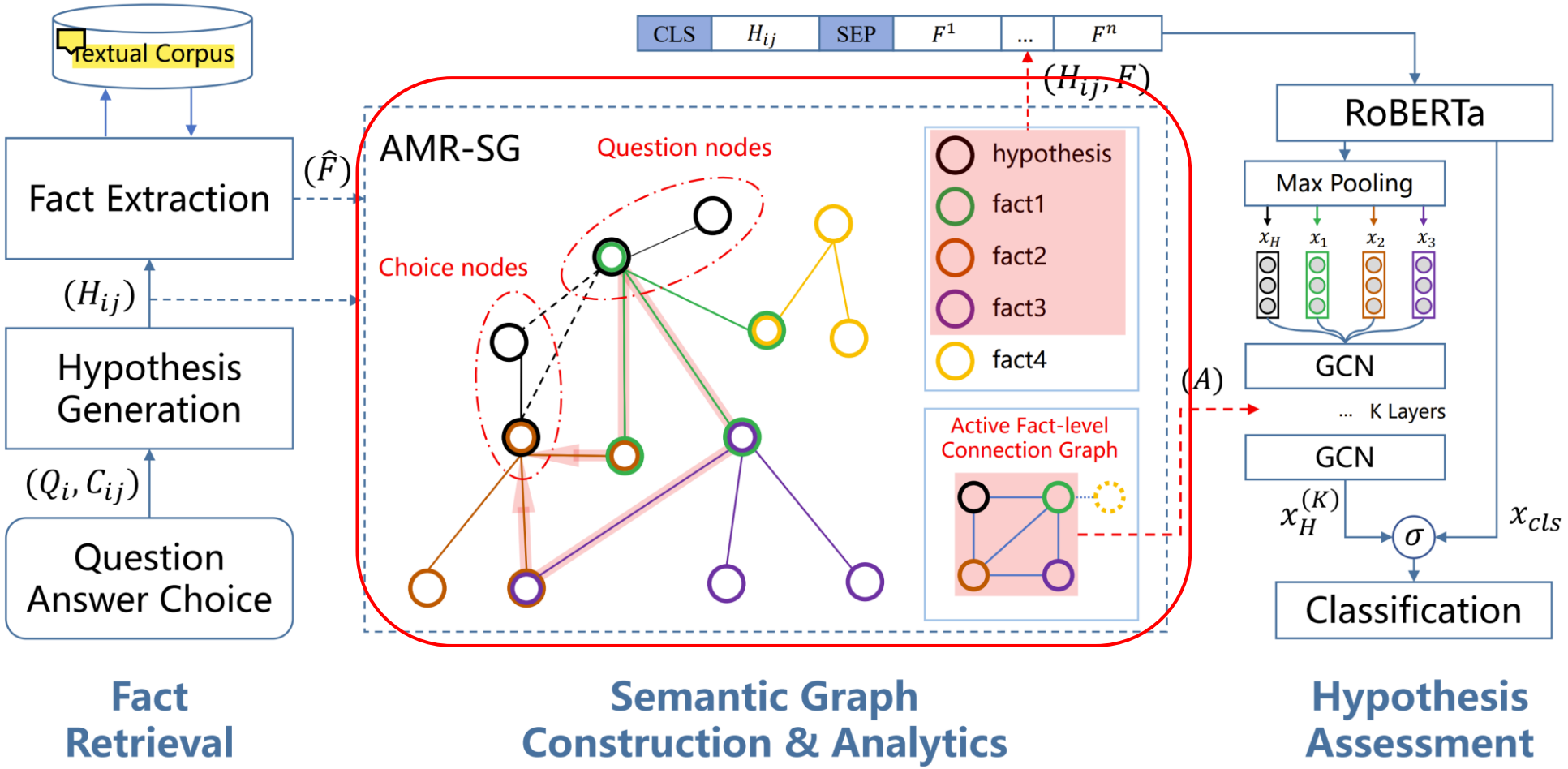
Answer Scoring

- $\text{argmax}(\emptyset(\mathrm{e}_h, e_q, e_a))$



Improving Multi-hop Question Answering over Knowledge Graphs using Knowledge Base Embeddings. ACL2020: 7579-7589

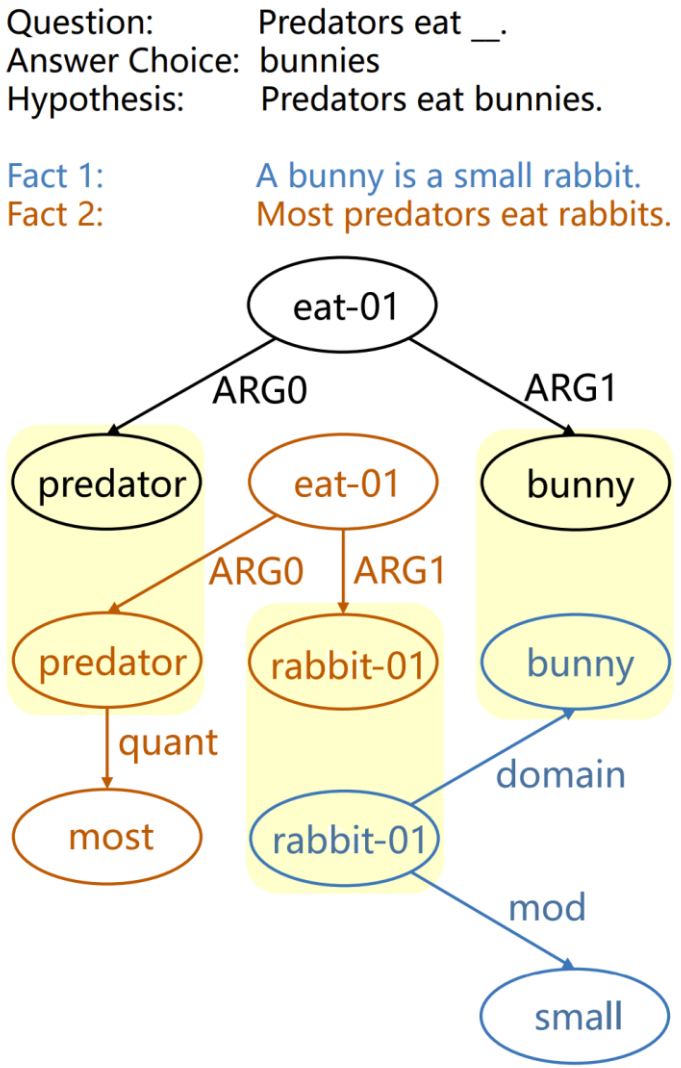
论文概括：设计让问题-选择的推理过程中经过一些事实节点来提升效果



Dynamic Semantic Graph Construction and Reasoning for Explainable Multi-hop Science Question Answering. ACL2021Findings: 1044-1056

AMR

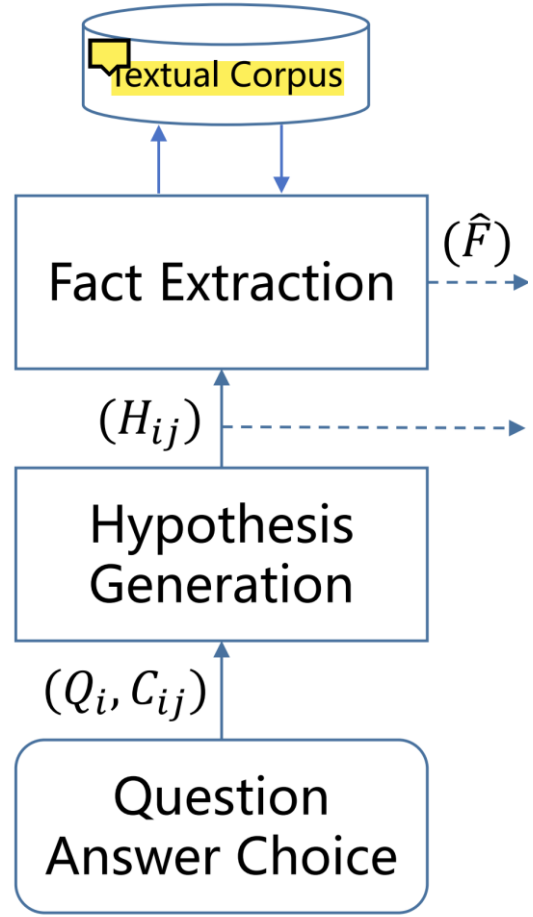
- 作用: 将句子的意义表示为一个有根、有方向的图
- 本文并不关注边的意义



Dynamic Semantic Graph Construction and Reasoning for Explainable Multi-hop Science Question Answering. ACL2021Findings: 1044-1056

Fact Retrieval

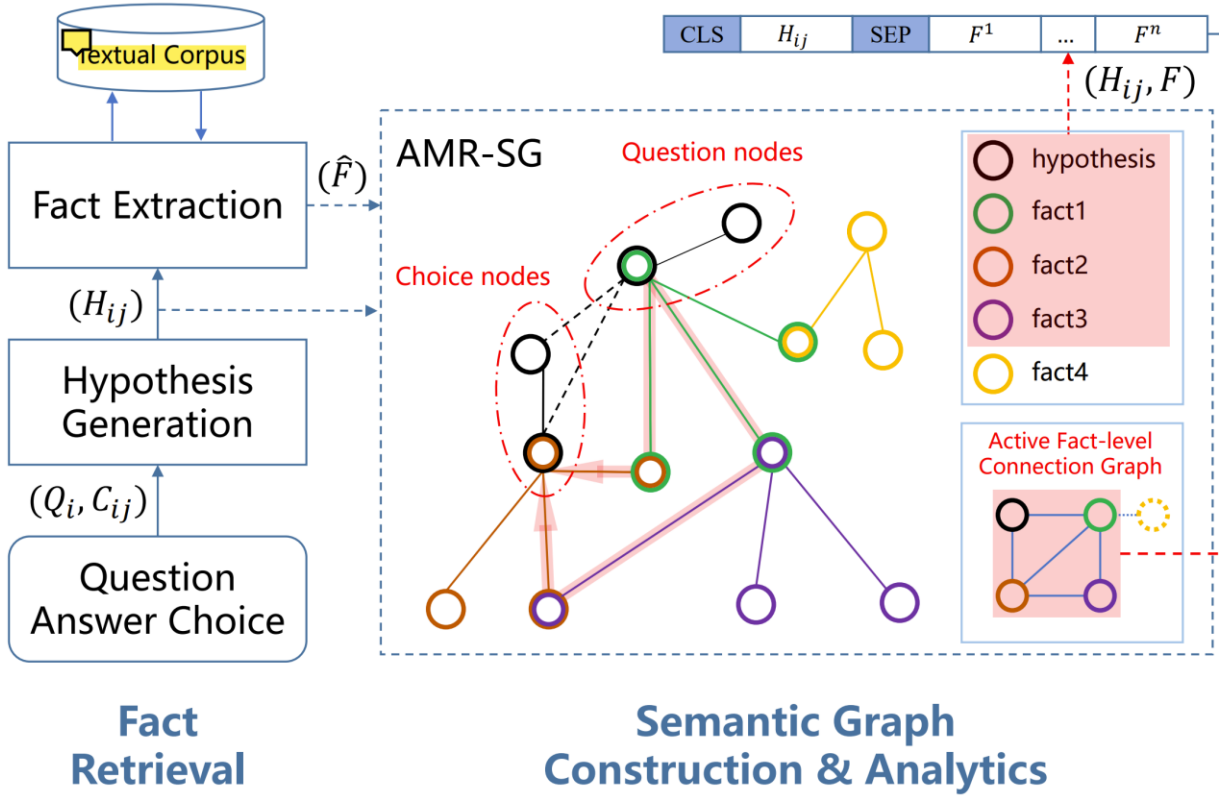
- Fact Extraction：从语料库中提取与问题答案相关的Fact
- Hypothesis Generation：不是简单的拼接问题答案，而是按照模板生成一个Hypothesis，这样可以包含更少的无意义的词，有一个更优秀的语法结构



Dynamic Semantic Graph Construction and Reasoning for Explainable Multi-hop Science Question Answering. ACL2021Findings: 1044-1056

Semantic Graph Construction & Analytics

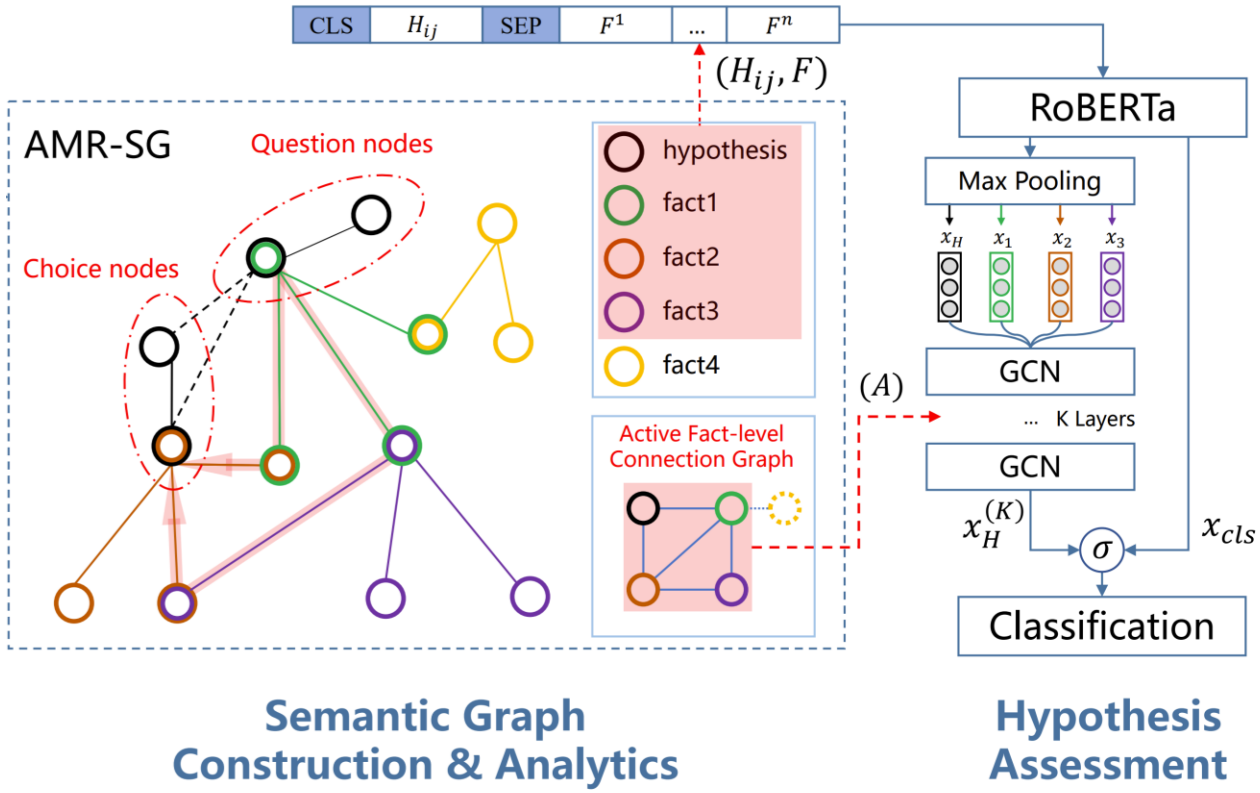
- AMR-SG：将Fact，Hypothesis放入图中；切断Choice与Question之间的联系；将相同意义的Share one node；寻找从Q到C的所有Path以及最短路
- AFCG：将被贯穿经过的Fact与Hypothesis作为节点，连接原来Share One node的Fact或Hypothesis。



Dynamic Semantic Graph Construction and Reasoning for Explainable Multi-hop Science Question Answering. ACL2021Findings: 1044-1056

Hypothesis Assessment

- 刚刚AFCG图进行GCN后得到的向量与Roberta的cls一起推断答案



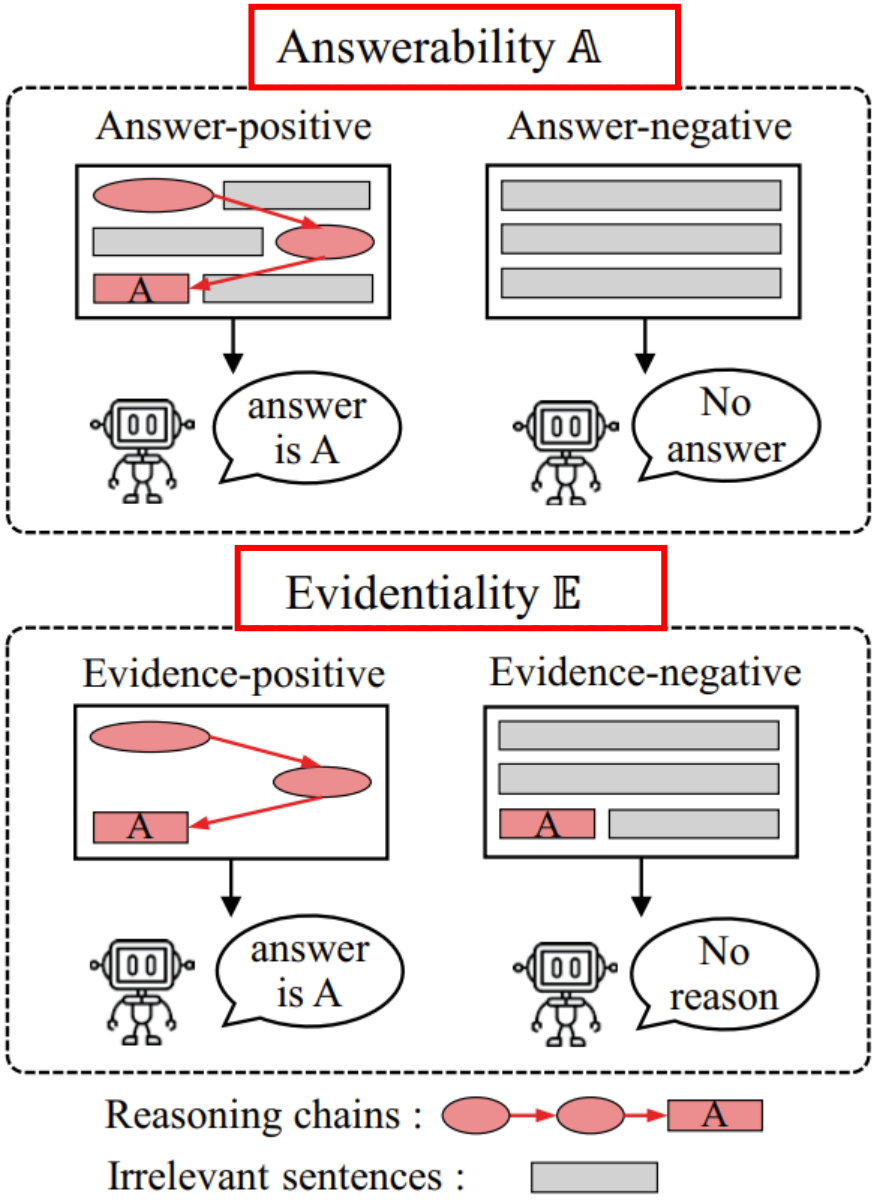
Dynamic Semantic Graph Construction and Reasoning for Explainable Multi-hop Science Question Answering. ACL2021Findings: 1044-1056

论文概括：有些多跳没有正确推理却选对了答案（走捷径）,模型会失去鲁棒性，作者想解决这个问题，通过一个正确的证据链来回答问题。

举例：

Q：Which country got independence when World War II ended?

Passage：…… Korea got independence in 1945 ……



Robustifying Multi-hop QA through Pseudo-Evidentiality Training. ACL/IJCNLP 2021: 6110-6119

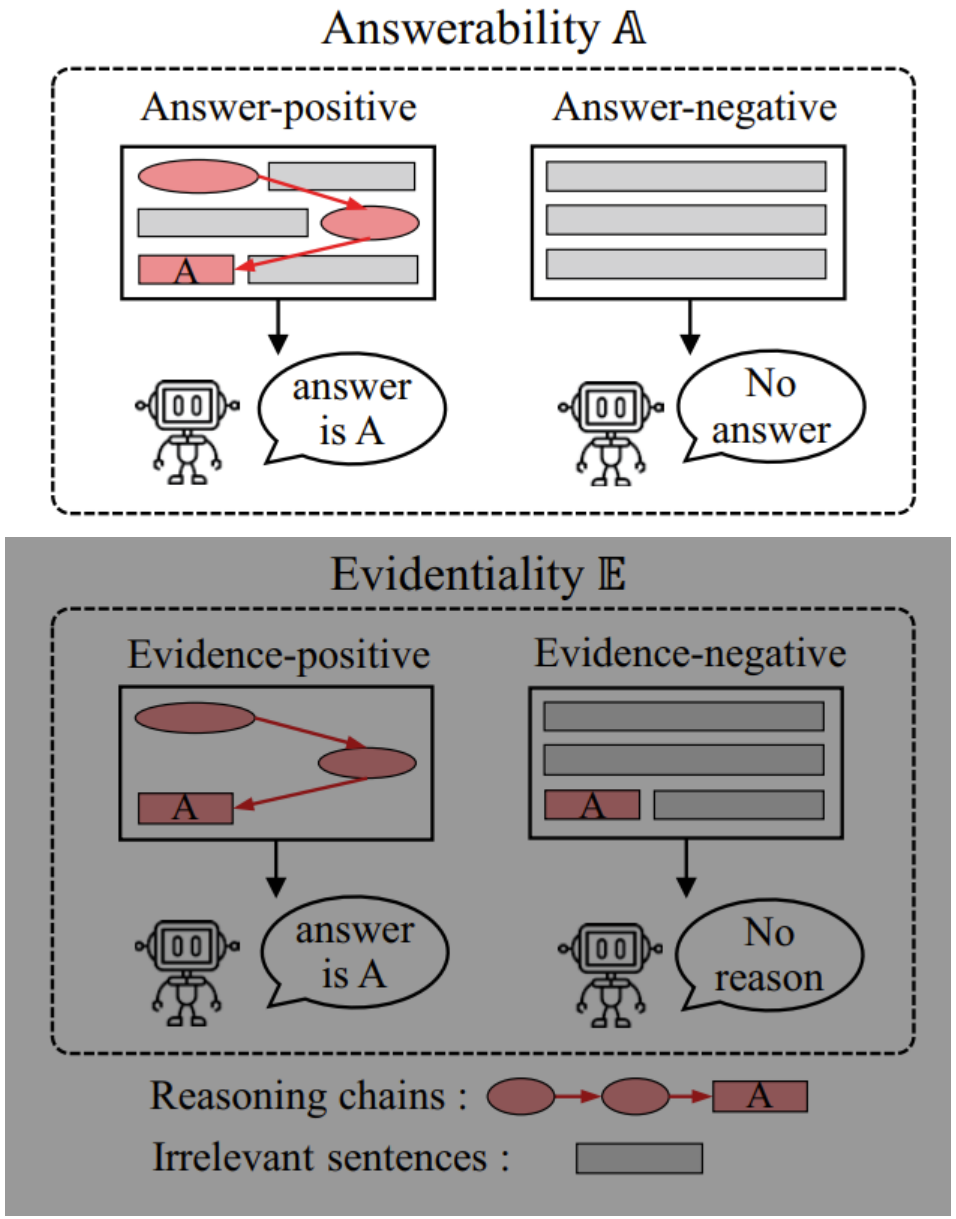
作者定义了两种类型的答案:

(1) Answer-positive

(2) Answer-negative

对于段落D，答案A，问题Q

如果(Q, A, D) 这个三元组可以找到正确答案，那就是A+，也就是Answer-positive。找不到就是A-，Answer-negative 。



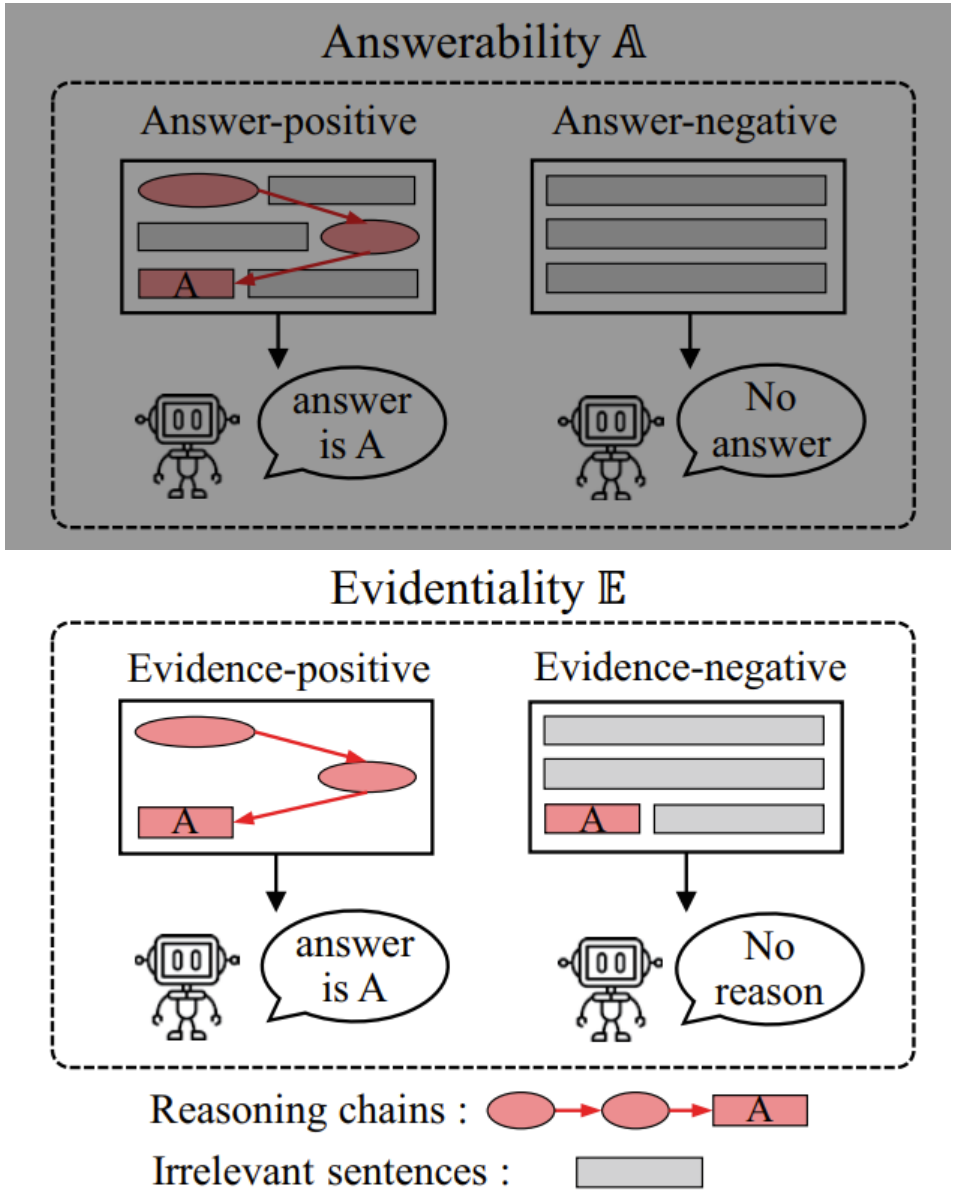
Robustifying Multi-hop QA through Pseudo-Evidentiality Training. ACL/IJCNLP 2021: 6110-6119

作者定义了两种类型的证据:

(1) Evidence-positive

(2) Evidence-negative

Evidence-positive 有答案也有证据，但是 Evidence-negative 的证据不Support答案（走捷径）

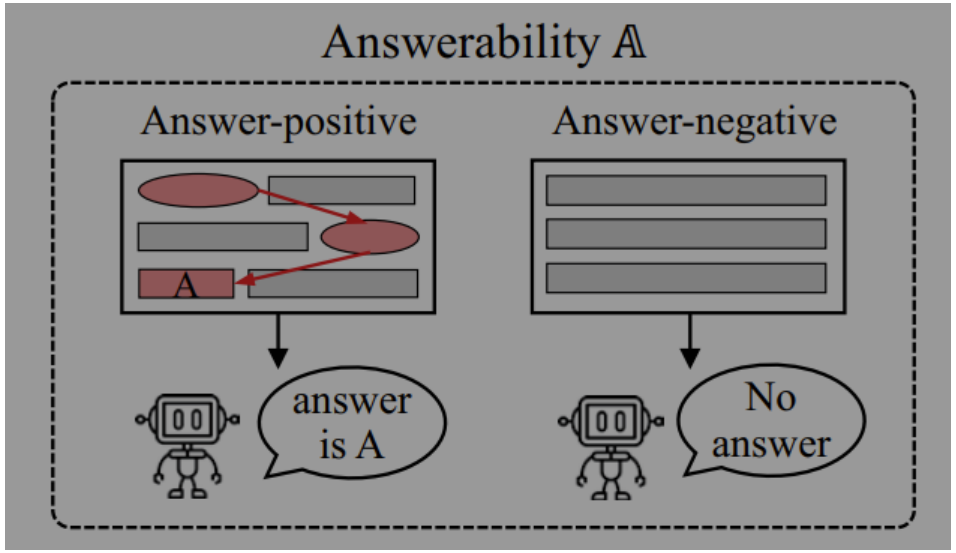


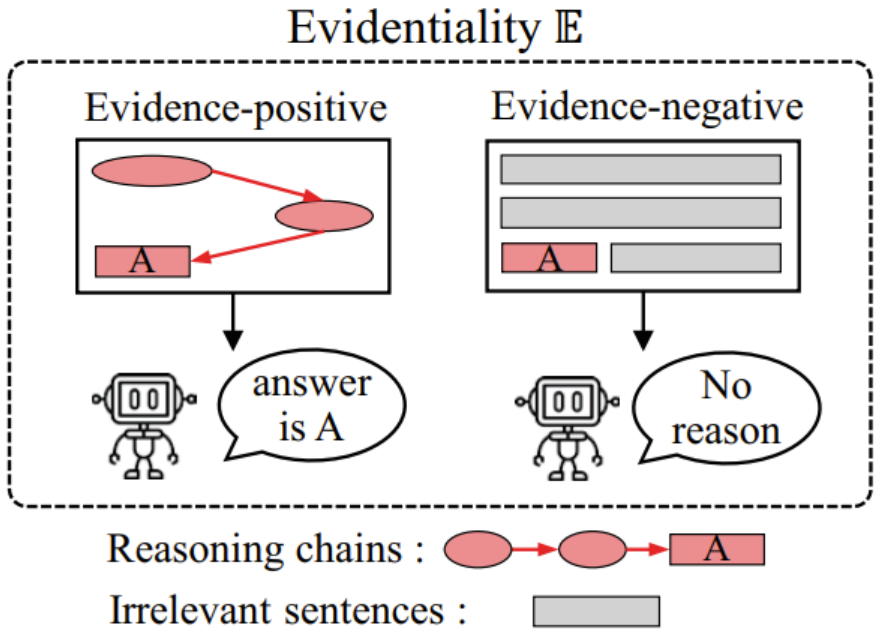
Robustifying Multi-hop QA through Pseudo-Evidentiality Training. ACL/IJCNLP 2021: 6110-6119

证据的构建:

E*是推断答案A的证据的基本事实，S*是包含A的句子，对应于Q。给定Q和A，预期的证据性的标签VE，表明回答的证据在段落中是否充分。

$$\mathcal{V}_E(\mathcal{Q},\mathcal{A},\mathcal{D}) \models True \quad \Leftrightarrow \quad E_* = \mathcal{D},\ \mathcal{A} \subset \mathcal{D}$$
$$\mathcal{V}_E(\mathcal{Q},\mathcal{A},\mathcal{D}) \models False \quad \Leftrightarrow \quad E_* \not\subset \mathcal{D},\ \mathcal{A} \subset \mathcal{D}$$





训练策略:

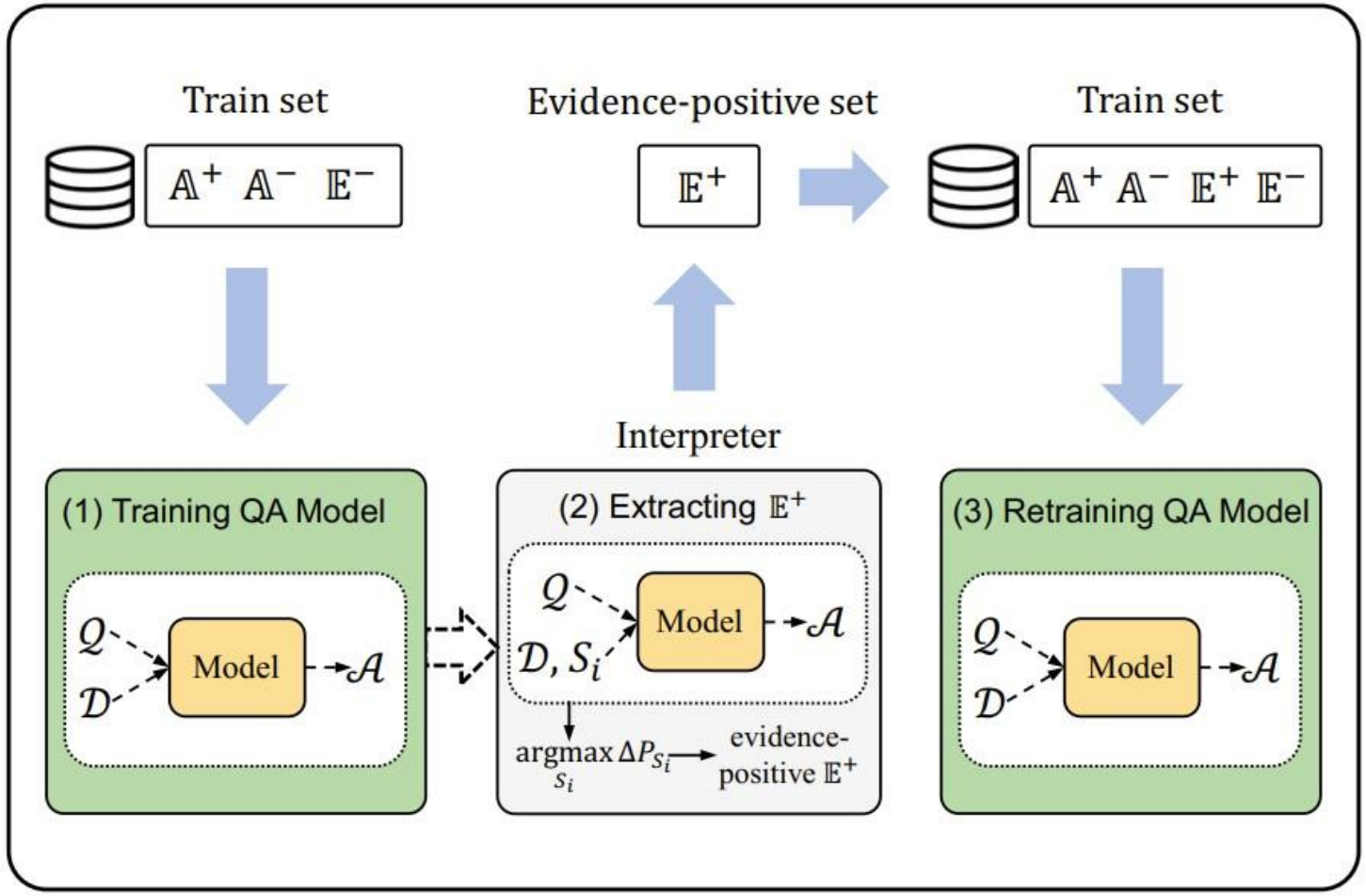
①：模型不应该对evidence-negative过度自信

②：对evidence-positive充满信心

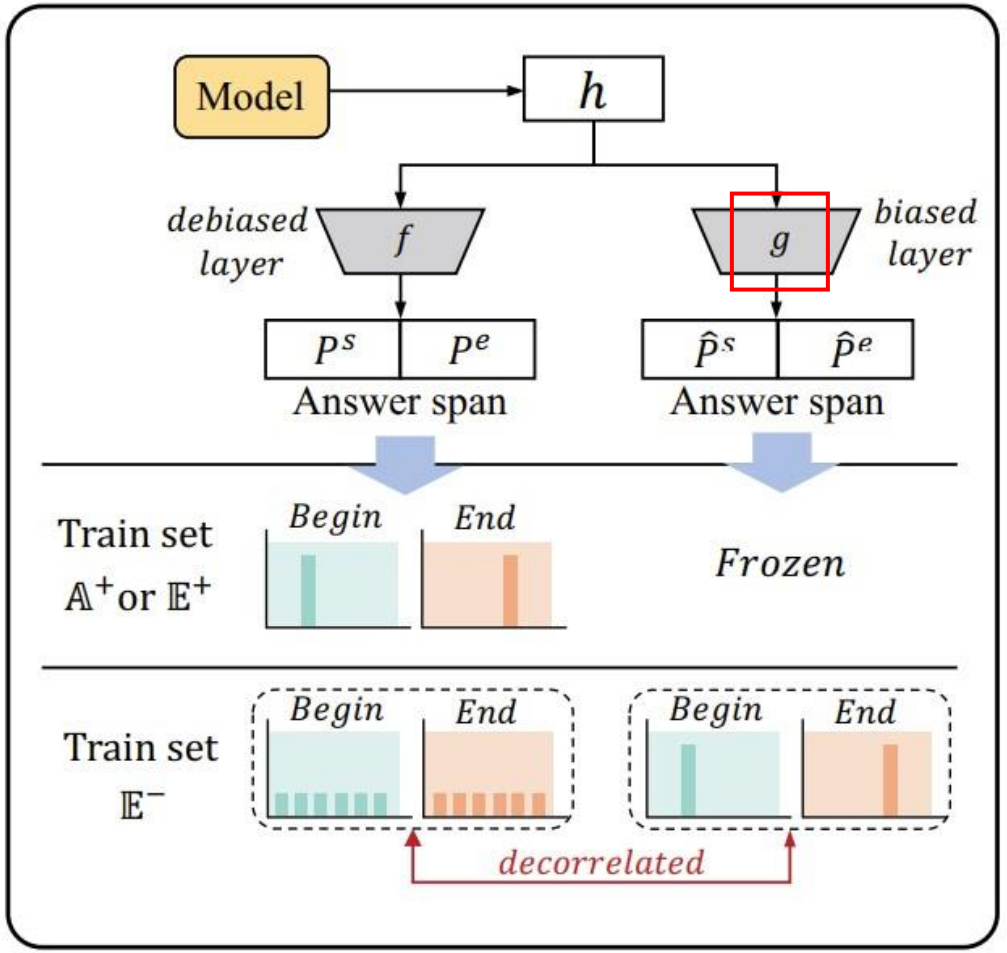
Robustifying Multi-hop QA through Pseudo-Evidentiality Training. ACL/IJCNLP 2021: 6110-6119

①：模型不应该对evidence-negative过度自信

②：对evidence-positive充满信心

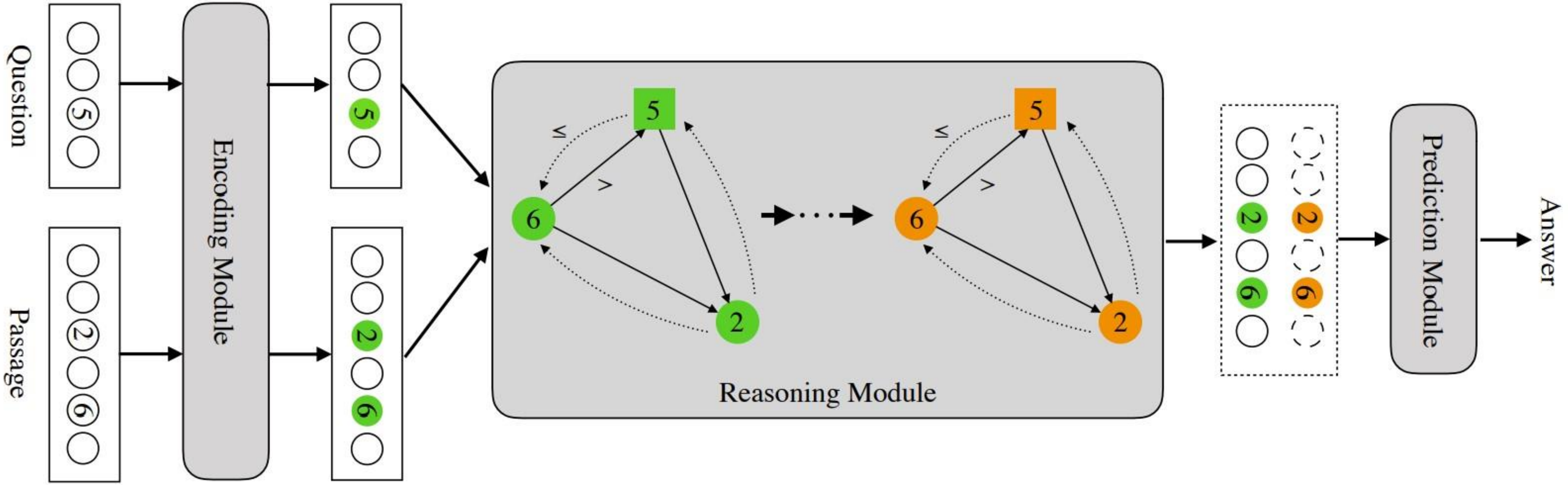


(a)



(b)

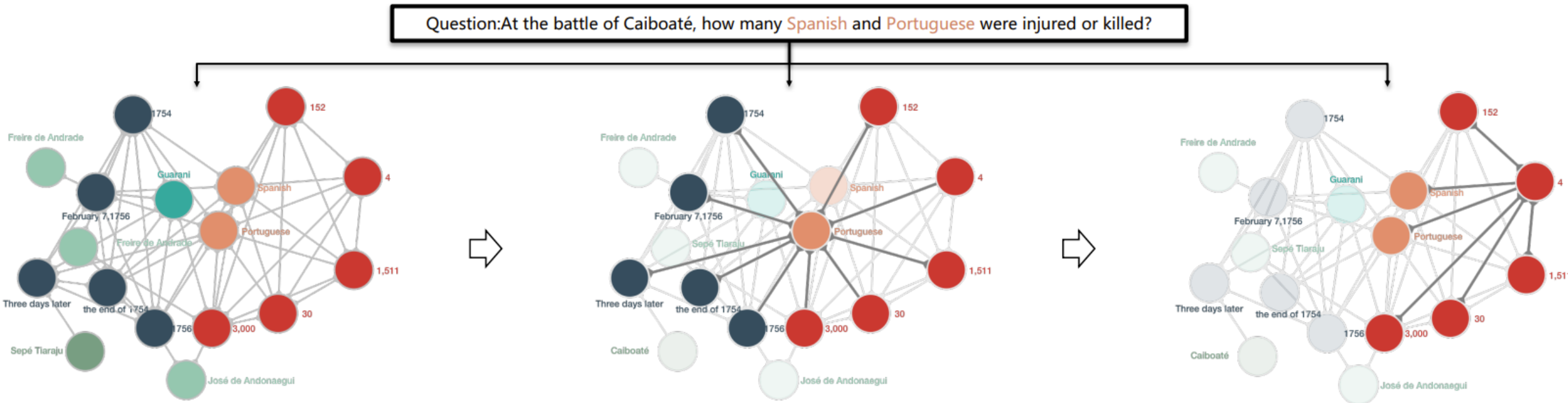
Robustifying Multi-hop QA through Pseudo-Evidentiality Training. ACL/IJCNLP 2021: 6110-6119



NumNet: Machine Reading Comprehension with Numerical Reasoning. EMNLP2019: 2474-2484

Question	Passage	Answer
At the battle of Caiboaté how many Spanish and Portuguese were injured or killed?	... In 1754 Spanish and Portuguese military forces were dispatched to force the Guarani to leave the area ... Hostilities resumed in 1756 when an army of 3,000 Spanish, Portuguese, and native auxiliary soldiers under José de Andonaegui and Freire de Andrade was sent to subdue the Guarani rebels. On February 7, 1756 the leader of the Guarani rebels, Sepé Tiaraju, was killed in a skirmish with Spanish and Portuguese troops. ... 1,511 Guarani were killed and 152 taken prisoner, while 4 Spanish and Portuguese were killed and about 30 were wounded...	34
In which quarter did Stephen Gostkowski kick his shortest field goal of the game?	The Cardinals' east coast struggles continued in the second quarter as quarterback Matt Cassel completed a 15-yard touchdown pass to running back Kevin Faulk and an 11-yard touchdown pass to wide receiver Wes Welker, followed by kicker Stephen Gostkowski's 38-yard field goal. In the third quarter, Arizona's deficit continued to climb as Cassel completed a 76-yard touchdown pass to wide receiver Randy Moss, followed by Gostkowski's 35- and 24-yard field goal. In the fourth quarter, New England concluded its domination with Gostkowski's 30-yard	third

Question Directed Graph Attention Network for Numerical Reasoning over Text. EMNLP 2020: 6759-6768



Question Directed Graph Attention Network for Numerical Reasoning over Text. EMNLP 2020: 6759-6768

论文概括：通过文字和表格进行计算金融问题

Page 91 from the annual reports of GRMN (Garmin Ltd.)

The fair value for these options was estimated at the date of grant using a Black-Scholes option pricing model with the following weighted-average assumptions for 2006, 2005 and 2004:

	2006	2005	2004
Weighted average fair value of options granted	$20.01	$9.48	$7.28
Expected volatility	0.3534	0.3224	0.3577
Distribution yield	1.00%	0.98%	1.30%
Expected life of options in years	6.3	6.3	6.3
Risk-free interest rate	5%	4%	4%

… The total fair value of shares vested during 2006, 2005, and 2004 was $9,413, $8,249, and $6,418 respectively. The aggregate intrinsic values of options outstanding and exercisable at December 30, 2006 were $204.1 million and $100.2 million, respectively. *(… abbreviate 10 sentences ...)*

Question: Considering the weighted average fair value of options , what was the change of shares vested from 2005 to 2006?

Answer: - 400

Calculations:

$$\left(\frac{9413}{20.01}\right) - \left(\frac{8249}{9.48}\right) = -400$$

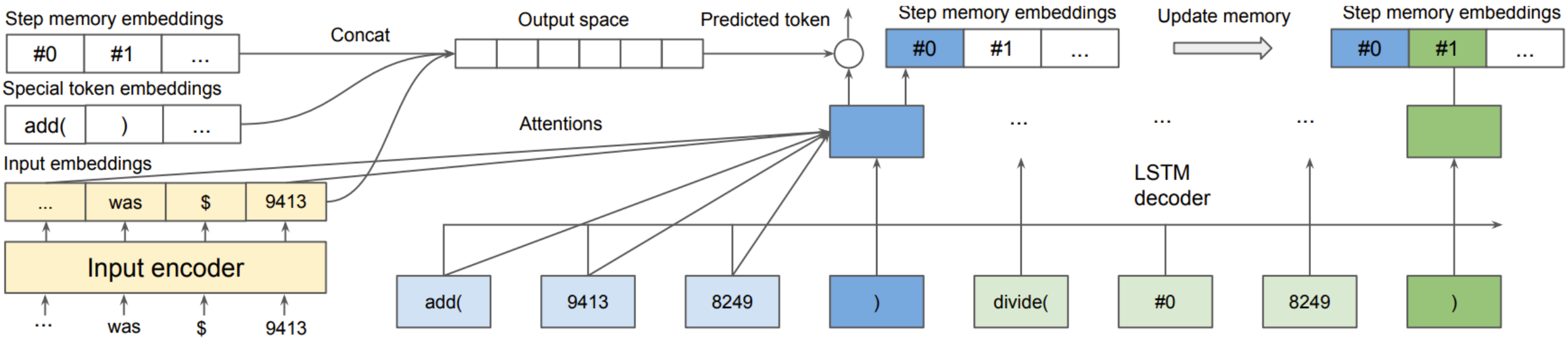
Program:

divide (9413, 20.01) | divide (8249, 9.48)

substract (#0, #1)

Figure 1: An example from FINQA: The system needs to learn how to calculate the number of shares, then select relevant numbers from both the table and the text to generate the reasoning program to get the answer.

FINQA: A Dataset of Numerical Reasoning over Financial Data EMNLP2021: 3697-3711



add, subtract, multiply, divide, greater, exp, table-max, table-min, table-sum, table-average

FINQA: A Dataset of Numerical Reasoning over Financial Data EMNLP2021: 3697-3711

感谢观看