重慶理工大學

CHONGQING UNIVERSITY OF TECHNOLOGY

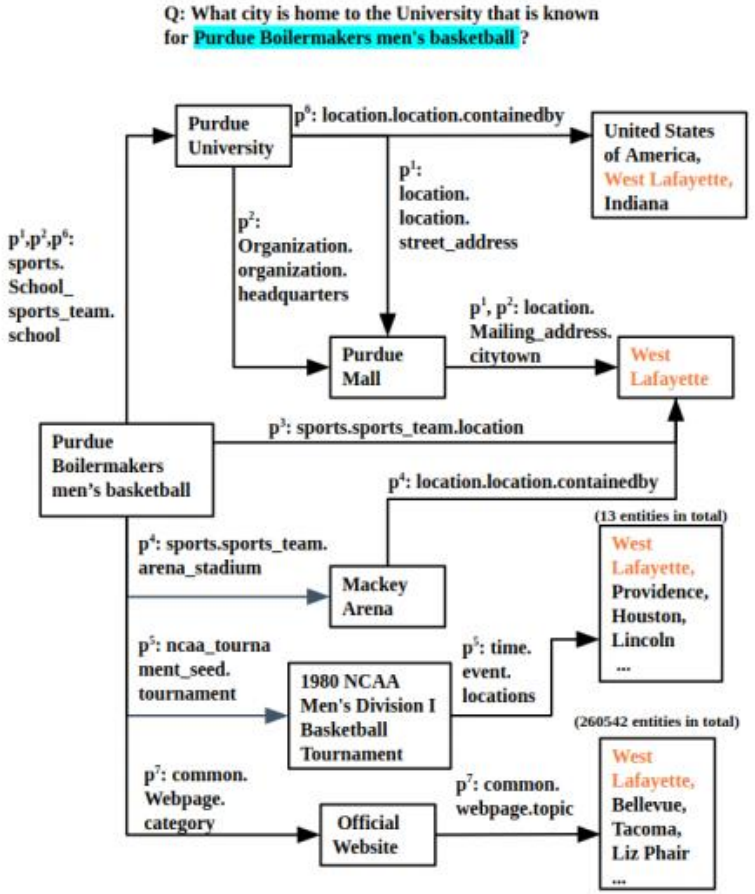
A New Concept of Knowledge based Question Answering (KBQA) System for Multi-hop Reasoning

日期：2022.12.02

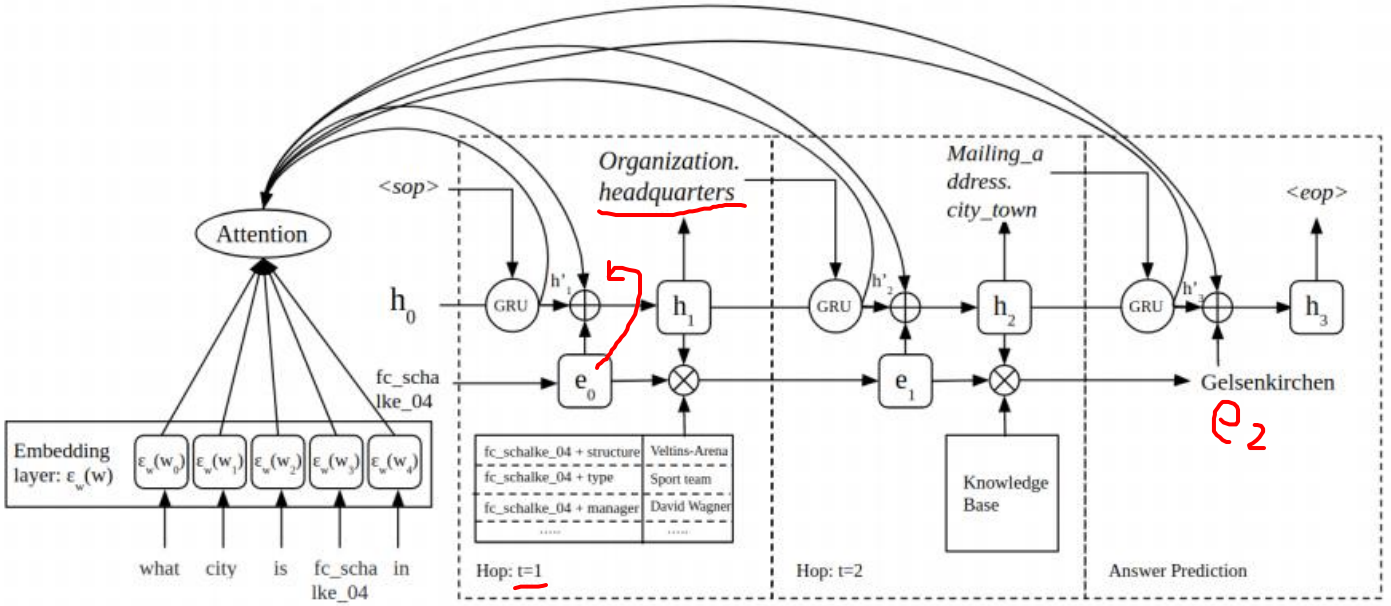
胡洋

Introduction

Figure 1: One QA example with Multiple Reason_x0002_ing Paths (MRPs) from COMPLEXWEBQUESTION-1.1.The blue color highlighted is the extracted topic entity.Each square represents an entity, and the arrows represent the relations. Reasoning path p1to p4are the correct ones containing meaningful reasoning paths to the final answer. p5and p6 are the "second choice" paths that generate a larger final answer set containing some wrong entities. p7is the wrong one as its reasoning path is totally not interpretable and the answer set is huge.



method

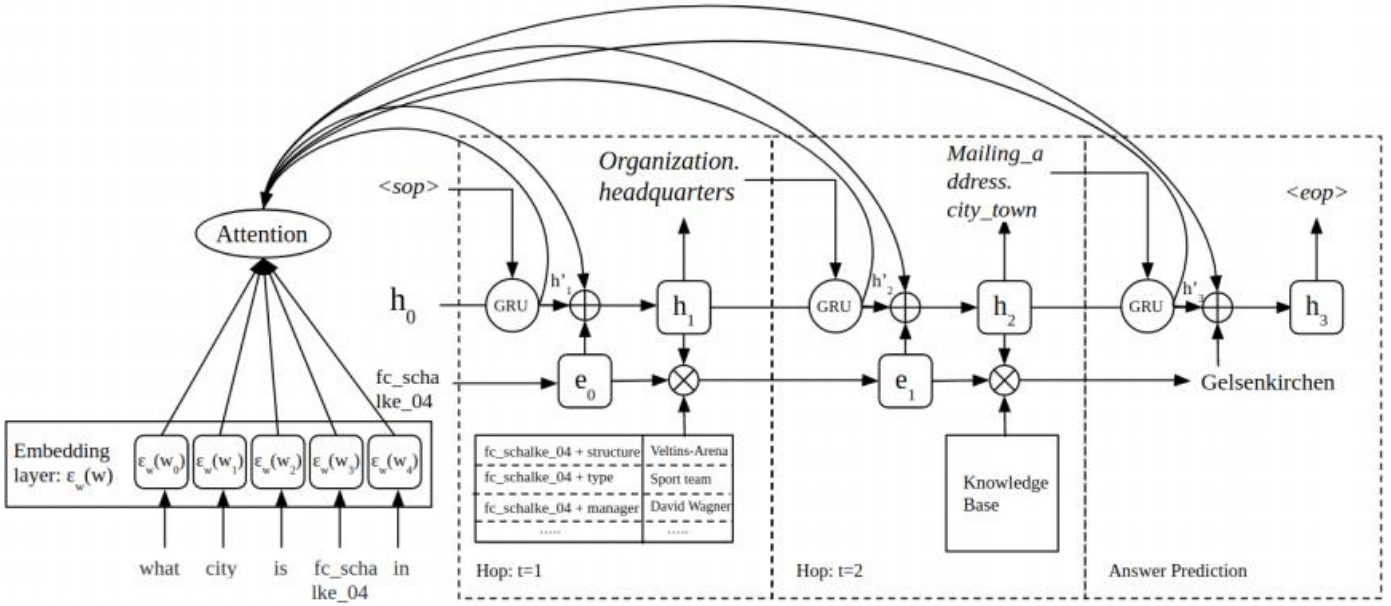


$$u_{tk} = a(h'_t, \varepsilon_w(w_k))$$

$$\alpha_{tk} = \frac{\exp(u_{tk})}{\sum_{0 \le j \le |q|-1} \exp(u_{tj})}$$

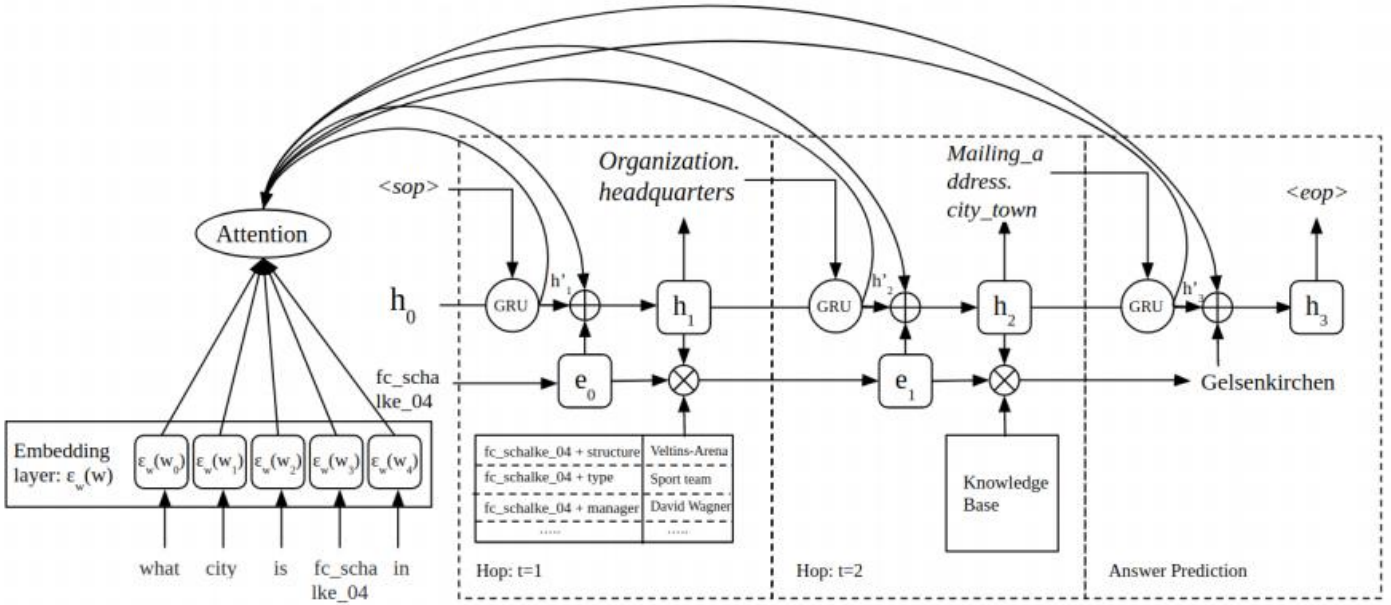
$$c_t = \sum_{0 \le j \le |q|-1} \alpha_{tj} \varepsilon_w(w_j)$$

method



$$p(r_t = \gamma_k | q, e_0, r_1, \cdots, e_{t-1})$$
$$= \frac{\exp < h_t, \varepsilon_r(\gamma_k) >}{\sum_j \exp < h_t, \varepsilon_r(\gamma_j) >}$$

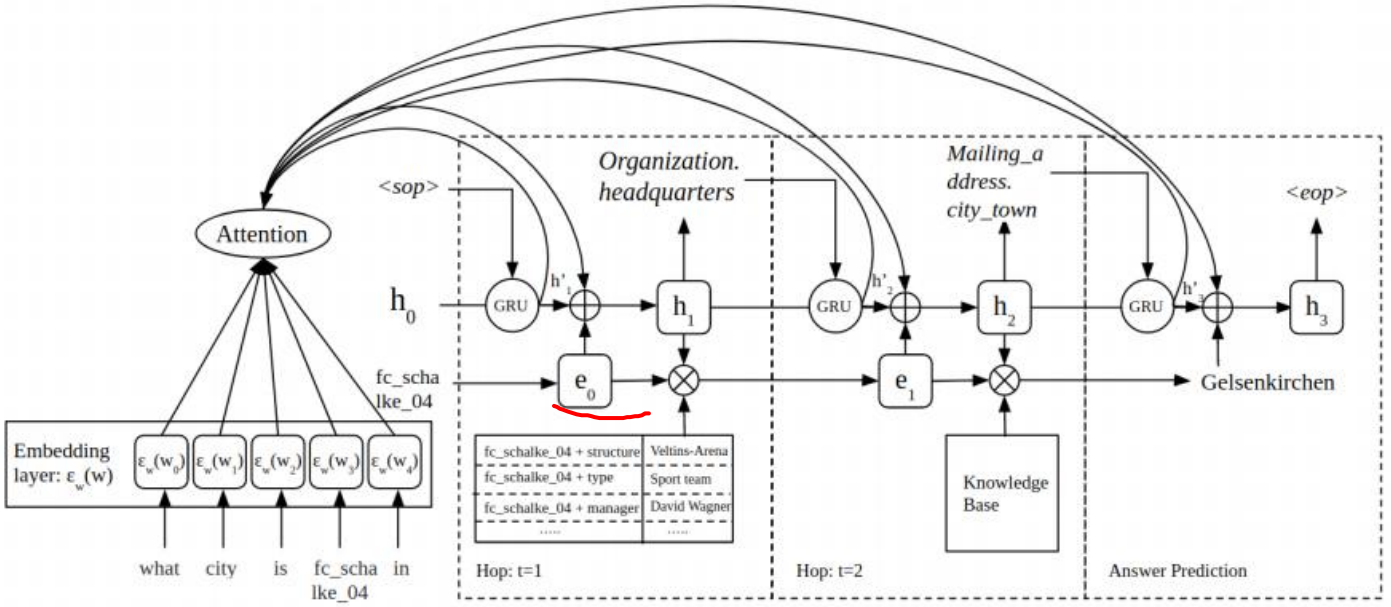
method



$$p(e_t|e_{t-1}, r_t) \qquad (1)$$

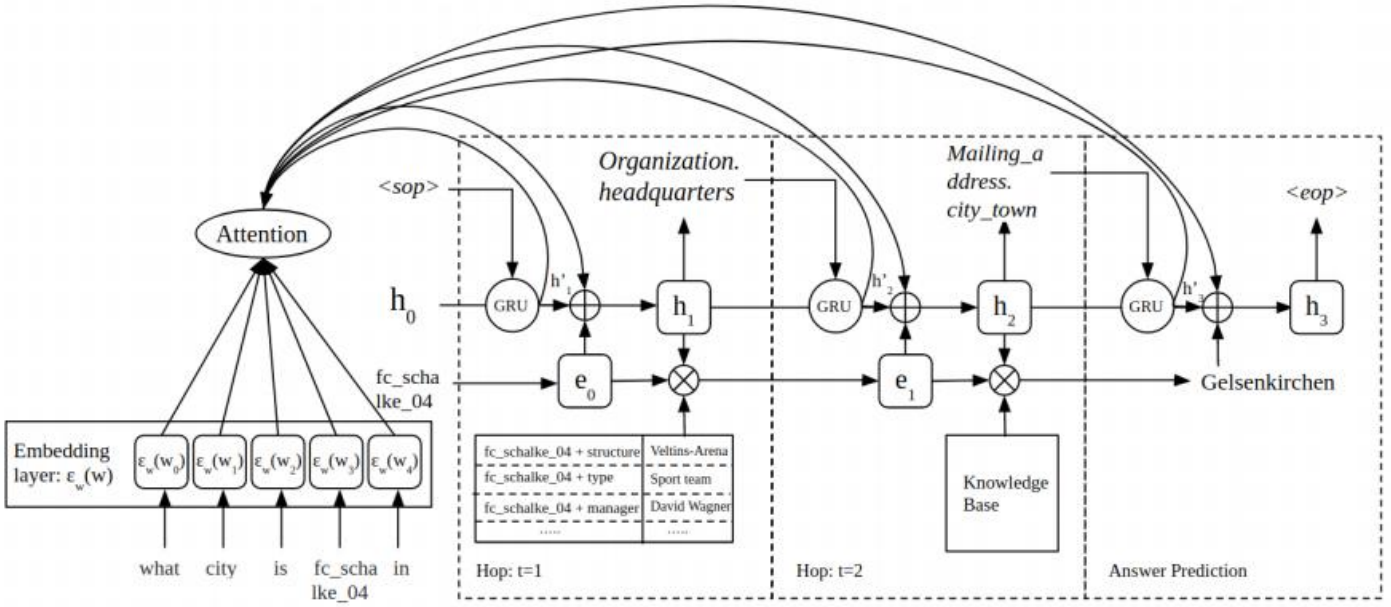
$$= \begin{cases} 1/M & \text{if } e_t \text{ is one of the } M \text{ matched entities} \\ 0 & \text{if } e_t \text{ is not a matched entity} \end{cases}$$

method



$$p(\mathbf{p}|q) = \prod_{t=1}^{T-1} p(e_t|e_{t-1}, r_t) \prod_{t=1}^{T} p(r_t|q, e_0, r_1, \cdots, e_{t-1}) \quad (2)$$

method



$$
\begin{aligned}
& p(y|q) \\
&= \sum_{\mathbf{p}\in\mathcal{P}} [p(e_{T(\mathbf{p})} = y|\mathbf{p}, q)p(\mathbf{p}|q)] \\
&= \sum_{\mathbf{p}\in\mathcal{P}} \prod_{t=1}^{T(\mathbf{p})} [p(e_t|e_{t-1}, r_t)p(r_t|q, e_0, r_1, \cdots, e_{t-1})]
\end{aligned}
\tag{3}
$$

method

Algorithm 1: Training method for a MRP-QA System

Input : KBQA dataset $(q^n, y^n, e_0^n), n = 1, 2, \cdots, N$, Knowledge Base $\mathcal{KB}$, Threshold k_1 and k_2.

Output : Trained model parameters

foreach *instance* (q^n, y^n, e_0^n) **do**
  Use DFS algorithm to get a set of paths $\mathcal{P}^n$ from e_0^n to y^n.
  Remove from $\mathcal{P}^n$ paths that point to more than k_1 entities.
end
foreach *batch* **do**
  foreach (q^n, y^n, e_0^n) *in the batch* **do**
    Get top k_2 paths in $\mathcal{P}$ sorted by $p(\mathbf{p}|q)$ based on current model: $\tilde{\mathcal{P}}^n = \{\mathbf{p}_1^n, \cdots, \mathbf{p}_{k_2}^n\}$
  end
  Update model parameters by maximizing

$$\sum_{(q^n, y^n, e_0^n)} \log \sum_{\mathbf{p} \in \tilde{\mathcal{P}}^n} p(y^n|\mathbf{p}, q^n) P(\mathbf{p}|q^n)$$

end

Experiment

	#train	#valid	#test	max_hops	>1 path
WQSP	2677	297	1639	2	79.4%
CWQ	27639	3519	3531	6	83.4%
PQL2H	1275	159	160	2	12.5%
PQL3H	1649	206	207	3	45.2%
PQL+	2924	365	367	3	30.6%

Table 1: Statistics of datasets. To count the data percentage with more than one path, *i.e.* >1 path, we use graph search algorithm to calculate what percentage of QA pairs can be solved with MRPs.

Experiment

	WQSP	CWQ
STAGG_SP (Yih et al., 2016)	71.7	-
HR-BiLSTM (Yu et al., 2017)	62.3	31.2
KBQA-GST (Lan et al., 2019)	67.9	36.5
KV-MemNN* (Miller et al., 2016)	38.6	-
STAGG_answer* (Yih et al., 2016)	66.8	-
NSM* (Liang et al., 2017)	69.0	-
GRAFT-Net* (Sun et al., 2018)	62.8	26.0
PullNet*(Sun et al., 2019)	68.1	47.2
TransferNet*(Shi et al., 2021)	71.4	48.6
NSM_{+h}*(He et al., 2021)	74.3	48.8
MRP-QA-marginal_prob*	**74.9**	**49.9**

Table 2: We report F1 (%) on WQSP and CWQ test sets. Methods labeled with $*$ only require the final answer as the supervision, and they are directly comparable to our MRP-QA model. As references, We also report the performance of methods that requires extra supervisions in the first block.

Experiment

Setting	Δ F1 (std)
MRP-QA – $entity\_in\_RNN$	−2.1 (0.21)
MRP-QA – $marginal\_prediction$	−1.8 (0.32)
MRP-QA – $inference\_in\_training$	−3.4 (0.15)
MRP-QA – $mutual\_information$	−1.8 (0.16)

Table 3: Feature ablation study on the dev set with a mean of 5 runs.

Experiment

Method	Objective	Path $\mathbf{p}$
single ground truth	$p(y\|\mathbf{p}, q)p(\mathbf{p}\|q)$	single ground truth path leading to y
single random	$p(y\|\mathbf{p}, q)p(\mathbf{p}\|q)$	single random path leading to y
multiple product	$\prod_{\mathbf{p}\in\mathcal{P}} p(y\|\mathbf{p}, q)p(\mathbf{p}\|q)$	all valid paths leading to y
multiple marginal (MRP-QA)	$\sum_{\mathbf{p}\in\mathcal{P}} p(y\|\mathbf{p}, q)p(\mathbf{p}\|q)$	all valid paths leading to y

Table 4: Different choices of paths and objectives.

Experiment

	WQSP			CWQ		
	1 path	>1 path	all	1 path	>1 path	all
single ground truth	66.8	69.3	68.1	40.8	49.2	46.4
single random	65.7	64.1	64.8	40.8	46.9	44.9
multiple product	69.1	70.2	69.7	40.9	50.7	47.5
multiple marginal (MRP-QA)	73.0	76.3	74.9	43.7	53.0	49.9

Table 5: We break test set into two groups based on number of paths associated with them and report F1(%).

Experiment

	PQL2H	PQL3H	PQL+
HR-BiLSTM (Yu et al., 2017)	97.5	87.9	92.9
IRN (Zhou et al., 2018)	72.5	71.0	52.9
ABWIM (Zhang et al., 2018)	94.3	89.3	92.6
UHop (Chen et al., 2019)	97.5	89.3	92.3
KV-MemNN* (Miller et al., 2016)	72.2	67.4	-
Our MRP-QA Method-marginal_prob*	**98.4**	**97.8**	**98.0**

Table 6: We report set accuracy (%) on PQL. Similar to Table 2, we use $*$ to highlight the methods which only requires the answer as supervision.

Experiment

Question: what state does romney live in? Answer: Massachusetts Topic entity: romney		
SINGLE GROUND TRUTH	MULTIPLE PRODUCT	MULTIPLE MARGINAL (OUR)
.89:children	.29:education_institution/ state_province_region	.83:**places_lived/ location**
.06:**government_positions/ jurisdiction_of_office**	.25:**places_lived/ location**	.12:**government_positions/ jurisdiction_of_office**
.04:**government_positions/ office_position_or_title**	.25:**government_positions/ district_represented**	.04:**government_positions/ district_represented**
.00:**government_positions/ district_represented**	.01:**government_positions/ jurisdiction_of_office**	.01:place_of_birth/ state
.00:place_of_birth	.01:place_of_birth/ state	.00:education/ degree
.00:jurisdiction_of_office	.01:sibling/ place_of_birth	.00:election_campaigns

Table 7: A running examples from WEBQUESTIONSP dataset. We show the probability $P(r_0, \cdots, r_T|q)$ before the inferred relations. Paths that lead to the correct answers are highlighted in bold. We use / to split two relations. The three columns are corresponding to the results by using different training settings as it is in Table 4. Due to space limit, we only show the partial name of a relation in the example and the probability less than .01 is shown as .00. We do not show $P(e_0, \cdots, e_{T-1}|q)$ because they are not determined by our MRP-QA model.

重慶理工大學

CHONGQING UNIVERSITY OF TECHNOLOGY

Thanks!