Solving MultiHop Reading Comprehension Task using RL algorithms

Krishna Manglani, Md Imbesat Hassan Rizvi

Agenda

- Introduction
- Multihop Reasoning over Knowledge Bases (KBs)
- Challenges in MultiHop Reading Comprehension (RC) setting
- Environment Creation
- Our Model Architecture
- Training Results
- Inference Details

Introduction

- Reading Comprehension Task:
  - Given Supports (paragraphs) and question from paragraphs, select the best answer from the candidates
  - It can be modelled as graph reasoning problem
  - Methods for graph reasoning:
    - Embedding based (Attention over Graphs)
    - Graph Walk based (Random walk models, RL)

MultiHop Reasoning over Knowledge Graphs using RL (MINERVA)

- Agent starts with a hub node
- All the outward edges from the node are the actions that an agent can take
- All the previous nodes visited by the agent and the current node forms the state for the agent
- Agent walks for fixed number of steps and at the last step, agent receives terminal reward of 1 if last node == answer_node for the question
- Agent Learns this policy using REINFORCE algorithm

Training Details (MINERVA)

Policy:

$$\mathbf{h_t} = \mathrm{LSTM}(\mathbf{h_{t-1}}, [\mathbf{a_{t-1}}; \mathbf{o_t}])$$

$$\mathbf{d_t} = \mathrm{softmax}(\mathbf{A_t}(\mathbf{W_2}\mathrm{ReLU}(\mathbf{W_1}[\mathbf{h_t}; \mathbf{o_t}; \mathbf{r_q}]))),$$
$$A_t \sim \mathrm{Categorical}(\mathbf{d_t}).$$

Reward function:

$$J(\theta) = \mathbb{E}_{(e_1, r, e_2)\sim D}\mathbb{E}_{A_1,\ldots,A_{T-1}\sim\pi_\theta}[R(S_T)|S_1 = (e_1, e_1, r, e_2)],$$

Problems in MultiHop RC setting

- In MINERVA, there is only single Knowledge Graph so there is only single environment, agent has to learn a policy to navigate in this environment only
- In MultiHop RC, **Every document forms a new environment**, agent needs to learn environment agnostic policy to navigate in test time

Environment Details (GraphGen)

- WikiHop dataset:
  - Several support documents to search for a given query
  - Question, Answer and Answer Candidates
- Intra-document graph between sentences with coreference resolution and NER entity match
- Inter-document graph between sentences to root node of matching NER across documents
- Single Hub Node functioning as search start node or search root node
- Answer matched to sentence containing all answer tokens

Environment Details (GraphGen)


```
{'id': 'WH_train_0',
 'question': 'participant_of juan rossell',
 'answer': '1996 summer olympics',
 'candidates': ['1996 summer olympics', 'olympic games', 'sport'],
 'supports': ['The 2004 Summer Olympic Games, officially known as the Games of the
 XXVIII Olympiad and commonly known as Athens 2004, was a premier international
 multi-sport event held in Athens, Greece, from 13 to 29 August 2004 with the motto
 "Welcome Home." 10,625 athletes competed, some 600 more than expected, accompanied
 by 5,501 team officials from 201 countries. There were 301 medal events in 28
 different sports. Athens 2004 marked the first time since the 1996 Summer Olympics
 that all countries with a National Olympic Committee were in attendance. 2004 marked
 the return of the games to the city where they began.',
  'The Pan-American or Pan American Games (also known colloquially as the Pan Am
  Games) is a major sporting event in the Americas featuring summer sports, in which
  thousands of athletes participate in a variety of competitions. The competition is
  held among athletes from nations of the Americas, every four years in the year
  before the Summer Olympic Games. The only Winter Pan American Games were held in
```


Sample Data

Our Model Architecture

Policy:

$$d_t = softmax(A_t(W_2 ReLU(W_1[e_t, Q])))$$
$$A_t \sim Categorical(d_t)$$

Reward function:

$$J(\theta) = E_{e_1,\ldots,e_{T-1}\sim\pi_\theta}(R(e_T)|e_0 = Hub)$$

Note: Sentence and Query Embeddings are preprocessed via BERT

Training Results

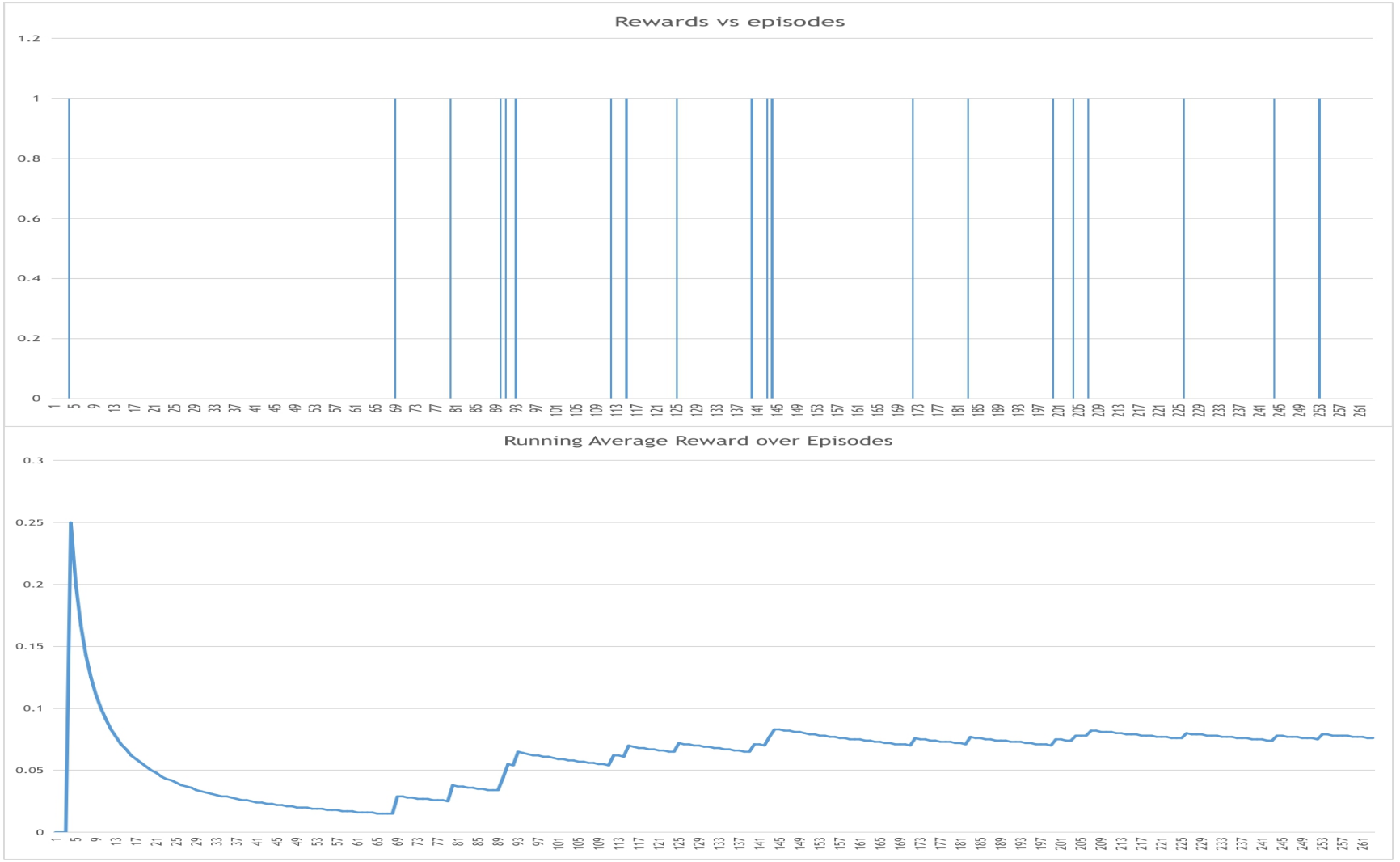

Figure (Top): Reward vs Episode
(Bottom): Running Average Reward vs Episode

Inference Details

- Given a Question and supports at the test time, Environment is made using the GraphGen Procedure.
- Beam search $(\text{beam\_width} = k)$ for T steps is done to find out the k-most probable graph paths starting from hub node.
- If last_node of most probable path contains answer then error$=0$ else error$=1$ for that example