

Tourism Domain QA System

Krishnkant Swarnkar, Harsh Rangwani, A.K. Singh, Rajeev Sangal

Computer Science And Engineering Indian Institute of Technology (BHU) Varanasi



Introduction/Background

Question Answering is an Al hard problem which involves understanding of user's natural language queries and answering them in natural language. Various attempts have been made to solve different QA problems like SQUAD, IBM Watson, YODA etc.

Objectives



- Reading Unstructured
 Wikipedia Text On Internet.
- Annotating the text.
- **Converting** information to Knowledge Graph.
- **Exhaustive Search** in Knowledge Graph.
- Provide Concise Answer.

Approach/Methods

Pipeline Architecture

Large Amount of Data

Annotation (NER, POS, Coreference)

Data Structuring

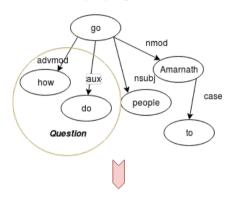
Augmented Structured Query

Candidate Retrieval

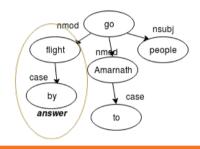
Factoid Answer Retrieval

Result

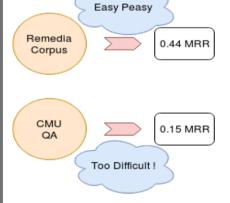
1. How do people go to Amarnath?



People go to Amarnath by flight.



Results + Discussions



* MRR : Mean Reciprocal Rank

The complexity and noisiness of documents in CMUQA makes it difficult for our system to find the best answers for the gueries.

From the observations on the results on above, it was evident that our QA model was able to perform some **basic level of inferencing**, with the help of coreferencing module.

- Good Performance On Grammatical Texts (Remedia).
- Not Scales for Wikipedia.
- Need For Development of Answer Pruning.
- No training required.
- **Scalable** to low resource languages.

Novelty/Innovation

Annotation - A fast annotation scheme which combines Stanford NER and DbPedia NER.

Data Structuring - A novel approach for indexing text in a structured form using augmented dependency graphs.

Query Retrieval - A novel dependency based sugraph matching technique based on

Social Impact

(i) The stakeholders are tourists who have many queries about the new place they are going to. (ii) Using the techniques of HCI and NLP, this system sims to answer the queries of tourists. (iii) Deployability on smart devices will make it more accessible to the society. (iv) It will promote Tourism in the country.

Acknowledgement

We are thankful to **Prof. AK Singh Sir** and **Prof. Rajeev Sangal Sir** for mentoring the project.