Intelligent Question-Answer System Format *

[Extended Abstract]

Shubhakar R Tipireddy
120101074
Department of CSE
IIT Guwahati
tipireddy@iitg.ernet.in

Abhishek Singh Gajuri
120101025
Department of CSE
IIT Guwahati
gajuri@iitg.ernet.in

Vinay Chandragiri
120101018
Department of CSE
IIT Guwahati
chandragiri@iitg.ernet.in

Vasu Ghatole
120101077
Department of CSE
IIT Guwahati
v.ghatole@iitg.ernet.in

1. PROBLEM STATEMENT

Question Answering Model(AI) which gives maximum probability for the most suitable answer available among the options

Description :

In recent times there has been an immense progress in data driven AI applications, but there is still a lack of machines that support reasoning and explanation. Developing a system that can make critical progress in fields of knowledge, reasoning and language is very important. Here we aim to develop a machine that has features like:

Question understanding, transforming questions, diagrams, and answer options from exams into computable input.

Integration with diagram processing systems to go beyond text to successfully understand exam questions.

Natural language parsing and processing of a variety of knowledge sources into a structured, computable knowledge base.

2. CHALLENGES

We want to tackle the following challenges in our work.

- Direct questions: These are questions testing direct simple knowledge like definitions and basic facts about the world for example, Sun rises in which direction? (A)East (B)West (C)South (D)North. These can be handled with existing methods of knowledge extraction from a corpus.

- Straightforward inference: These are questions where existing methods of learning from a corpus fail to an extent. Questions that need simple inferences like, Which of the habits is not good for human health? (A)Going for a jog (B)Eating excessive red meat (C)Sleeping everyday for 8 hours (D) Laughing. To answer this question the system has to understand that "eating" is a way through which living things take in food and that "red meat" contains high content of saturated fats which is bad for health.

- Complex knowledge: These questions need more complex analysis and reasoning. For example, A student conducts an experiment with two identical plants in the same type and amount of soil and water. One of these plants is near a sunny window and the other in a dark room. This experiment tests how the plants respond to (A)light (B)air (C)water (D) soil Again, information retrieval methods and word correlations do poorly. Rather, to answer this system is required to recognize a model of experimentation(perform two tasks, differing in only one condition), inferring that plant near window has access to sunlight while the one in dark room doesn't.

- Answering questions with figures: Questions can contain figures and these need to be handled. Diagrams may be tables, charts, figures, etc. This contains two issues, first one is part identification in figures and second is recognizing and interpretation of spatial information.

3. CURRENT METHODS

Many tasks in Natural Language Processing can me modelled into Question answering tasks by a language input.

Dynamic Memory Networks are a neural network based architecture which works based on episodic memories and generates answers accordingly. Understanding meaning of a text and making sense out it is one of the most complex task that a machine can do and requires reasoning over relevant facts. This neural network architecture is designed to address the sequence tagging tasks which essentially redirect to factual reasoning. This network first computes a representation of the input. This is done by the input module. There are other modules such as question module, Episodic memory module and Answer module.

4. REFERENCES

Discourse Complements Lexical Semantics for Non-factoid Answer Reranking

Modeling Biological Processes for Reading Comprehension Automatic Construction of Inference-Supporting Knowledge Bases Elementary School Science and Math Tests as a Driver for AI: A Study of the Knowledge Base Requirements for Passing an Elementary Science Test

5. ADDITIONAL AUTHORS

6. REFERENCES

[1] P. Jansen, M. Surdeanu, P. Clark Discourse Complements Lexical Semantics for Non-factoid Answer Reranking. *ACL*, 2014.

[2] Jonathan Berant, Vivek Srikumar, Pei-Chun Chen, Brad Huang, Christopher D. Manning, Abby Vander Linden, Brittany Harding, Peter Clark Modeling Biological Processes for Reading Comprehension. *EMNLP*, 2014.

[3] P. Clark. Elementary School Science and Math Tests as a Driver for AI: Take the Aristo Challenge! In *IAAI*, 2015.

APPENDIX