**Problem**

Question Answering: In reading comprehension, a paragraph and a question about the paragraph is provided to a model as input and an answer to the question is provided as output.

Question answering is important for these reasons:
- Almost all NLP tasks can be posed as question answering problems
- Question answering is the first step to reading comprehension
- NLP tasks and therefore QA tasks form the basis of most human computer interactions

**Background**

SQUAD Dataset:
- Paragraphs from Wikipedia and 150k questions
- Questions and answers crowdsourced using Turk
- Half of the questions cannot be answered using paragraph

Often, successful SQUAD models condition the end token on the start token. Answer-Pointer generates probability distribution based on attention across context tokens. Answer-Pointer derivatives include:
- R-Rec: Machine Reading Comprehension with Self-Matching Networks
- Dynamic Coattention Networks for Question Answering
  While models should apply an Answer-Pointer layer in order to improve performance, Answer-Pointer output might be too deep or sophisticated for the size of the dataset.

Goal: Save time for future modelers and have a benchmark for which architectures benefit from Answer-Pointer by examining the effect of an Answer-Pointer output in the performance of the model.

**Methods**

**Character-Level Embeddings**

Resulting improvements added to the baseline BIDAF model. Improvements were mixed and matched according to the ablation testing schedule described in methods. Answer-pointer output layer was substituted in for BIDAF output layer and the models were compared.

**Analysis**

Qualitative Observations about Answer-Pointer Models vs BIDAF Output Models:
- Better at encapsulating articles like "the" or "and"
- Seeks more verbose answers
- More biased against giving "NA" as an answer
- Sensitive to punctuation schemas
- Question: What type of group is "The Islamic State"?
- Context: "Islamic State.....is a Wahhabi/Salafi jihadi extremist group...
- Answer: Wahhabi/Salafi jihadi extremist millitant
- C: Prediction: extremist millitant
- C: Prediction: Wahhabi/Salafi jihadi extremist millitant

Fig 5. Example of QA pair from C/C.

Answer-Pointer Specific Qualitative Explanations:
- Attention based categorical distributions over start and end tokens with argmax of these as start and end tokens output
- Very concentrated attention at the start token suggests that the passage has an answer
- Uses attention knowledge to predict less likely tokens from the context causing end token attention distributions to become more concentrated
- OOV token has less attention weight by comparison in the end token layer making model less likely to predict "NA"
- Understands test well when the question text is directly in the context paragraph or the wording is close to the input question
- Does not understand well when question test incorporates a word or abbreviation that the model has not seen

**Conclusions**

Model Observations:
- Chocolate answer pointer provided inconsistent benefit to models
- Conject that B, C, and D already capture information that answer pointer could provide, leading to answer pointer effects being dampened in B, C, D, Nest steps

Propose aggressively regularizing answer pointer RNN (perhaps applying dropout on start token inputs)
- Explore whether C underforms C with different modeling seeds

**References**