

NAME: Ying Sheng
Andrew ID: yingshen

Report for HW2:
Design of logical architecture:
(highlight the design and feature of the system)

The architecture contains 5 parts which come in sequence in the pipeline:
QuestionAnnotator -> AnswerAnnotator-> TokenAnnotator->NGramAnnotator -
>AnswerScoreAnnotator->evaluatorAnnotator

As their name indicates, each annotator add certain types of annotations to the JCas.

Question Annotator: identified as the sequence between “Q ” and “?”, it includes the “?” in the end

Answer Annotator: identified as the sequence between “A 1/0 ” and “.”, it includes the ‘.’ in the end. The “1/0” is stored in an Boolean attribute called isCorrect.

Token Annotator: for region annotated by the previous two annotators, each word is annotated as a token.

NGram Annotator: within a question or answer annotation range, 1-, 2-, 3-Gram are annotated based on the token annotation.

AnswerScore Annotator: for each answer, the score is calculated by (NGram overlapped with question)/(total NGram number).

evaluatorAnnotator: rank all the AnswerScore Annotator by score, calculate precision, and print to stdout.