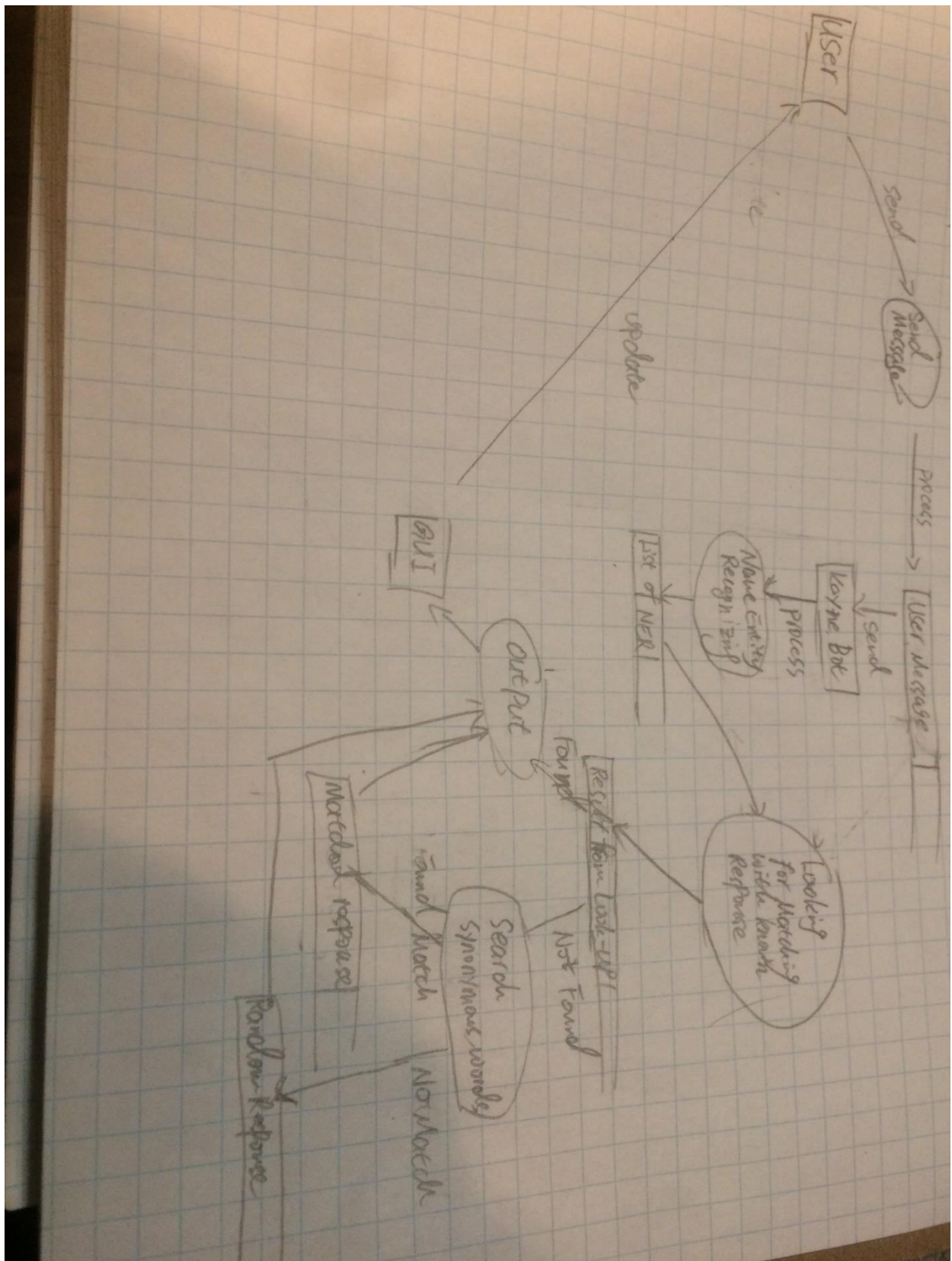
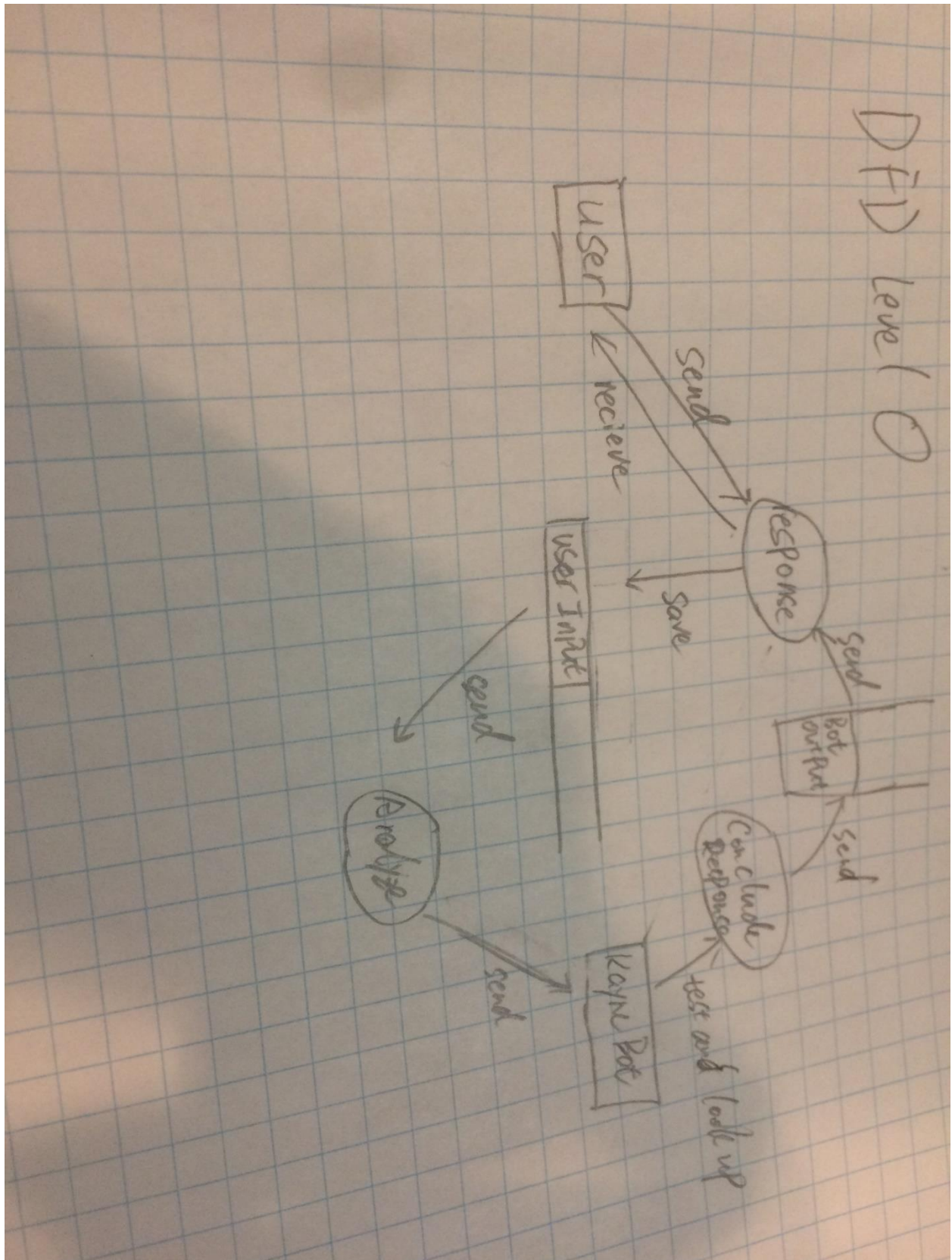


DFD LEVEL 0:

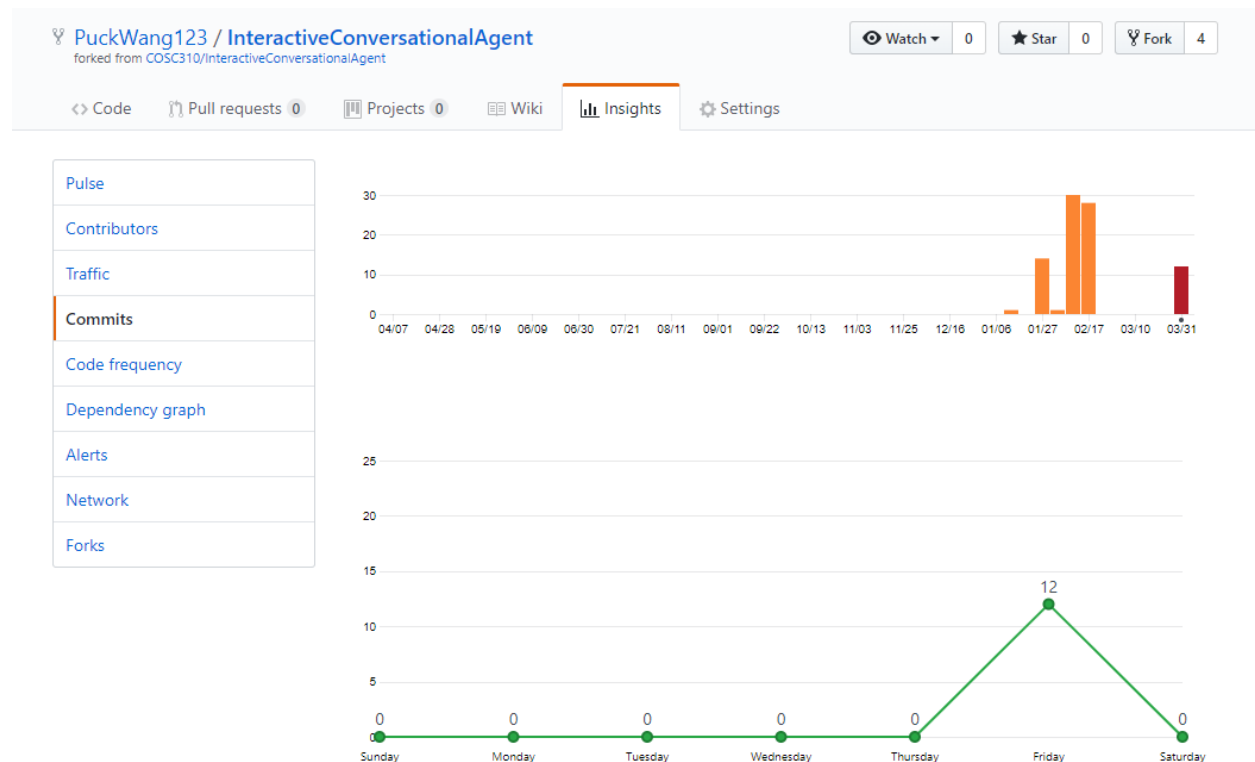


DFD LEVEL 1:



GitHub repository:

<https://github.com/PuckWang123/InteractiveConversationalAgent>



Features:

Name Entity Recognition:

```
how are y
[main] INFO edu.stanford.nlp.tagger.maxent.MaxentTagger - Loading POS tagger from edu/stanford/nlp/models/pos-tagger/english-left3words/english-left3words-distsim.tagger ... done [0.6 s]
[main] INFO edu.stanford.nlp.ie.AbstractSequenceClassifier - Loading classifier from edu/stanford/nlp/models/ner/english.all.3class.distsim.crf.ser.gz ... done [1.3 sec]
[main] INFO edu.stanford.nlp.ie.AbstractSequenceClassifier - Loading classifier from edu/stanford/nlp/models/ner/english.muc.7class.distsim.crf.ser.gz ... done [2.1 sec]
[main] INFO edu.stanford.nlp.ie.AbstractSequenceClassifier - Loading classifier from edu/stanford/nlp/models/ner/english.conll.4class.distsim.crf.ser.gz ... done [0.7 sec]
[main] INFO edu.stanford.nlp.time.JollyDayHolidays - Initializing JollyDayHoliday for SUTime from classpath edu/stanford/nlp/models/sutime/jollyday/Holidays_sutime.xml as sutime.binder.
[main] INFO edu.stanford.nlp.time.TimeExpressionExtractorImpl - Using following SUTime rules: edu/stanford/nlp/models/sutime/defs.sutime.txt,edu/stanford/nlp/models/sutime/english.sutime
[main] INFO edu.stanford.nlp.pipeline.TokensRegexNERAnnotator - ner.fine.regexner: Read 588794 unique entries out of 581863 from edu/stanford/nlp/models/kbp/english/gazetteers/regexner_
[main] INFO edu.stanford.nlp.pipeline.TokensRegexNERAnnotator - ner.fine.regexner: Read 4869 unique entries out of 4869 from edu/stanford/nlp/models/kbp/english/gazetteers/regexner_
[main] INFO edu.stanford.nlp.pipeline.TokensRegexNERAnnotator - ner.fine.regexner: Read 585573 unique entries from 2 files
VRB
VBP
NN
```

The original output was “How are you”. By using StanfordCoreNLP, it allows the bot to recognize each word’s entity as being Partial-of-speech. Later, NER can be used to collaborate with Synonym recognition based on the type of the POS. Noticeably, the structure of StanfordCoreNLP is so large that it consumes almost 4GB of RAM to load.

Stemmer:



When user greets to the bot the first time, it is appropriate. The second greeting however contains a typo as being “hi” becomes “his”. The chatbot can still recognize it and stem it to “hi”.

Limitation of the program:

- Unable to use the WordNet dictionary due to Windows default decompression format, which conflicts with the format of JWI.
- The chatbot does not have neural network which limits its learning based on human inputs.
- Limited topic available
- Refactor is required to improve the execution time