Interim Progress Report

Haya Al-Kuwari, Akhyar Kamili, Mohammed Nurul Hoque, Abubaker Omer

November 20, 2018

This report presents our group's progress in the project and our implementation plan so far.

1   Initial Experiments

We experimented with Stanford CoreNLP library using simple sentences. We found the idea `tokregex &` `semgrex` particularly appealing because it simplifies a lot of processing as regex matching-like operations. Unfortunately, the CoreNLP server takes up too much space and time when these features are used, while it is quite snappy when used just for regular operations like sentence-splitting, tagging and parsing.

2   Implementation Plan

We decided to implement a subset of the `semgrex` module that is sufficient for our purpose. ...

3   Results

We implemented a prototype for a splitting sentences and matching a template agains each sentence. We tested it with an actual template and document. Figure <insert ref> shows some of the results.