ID: Wyuan
Name: Wenqing Yuan

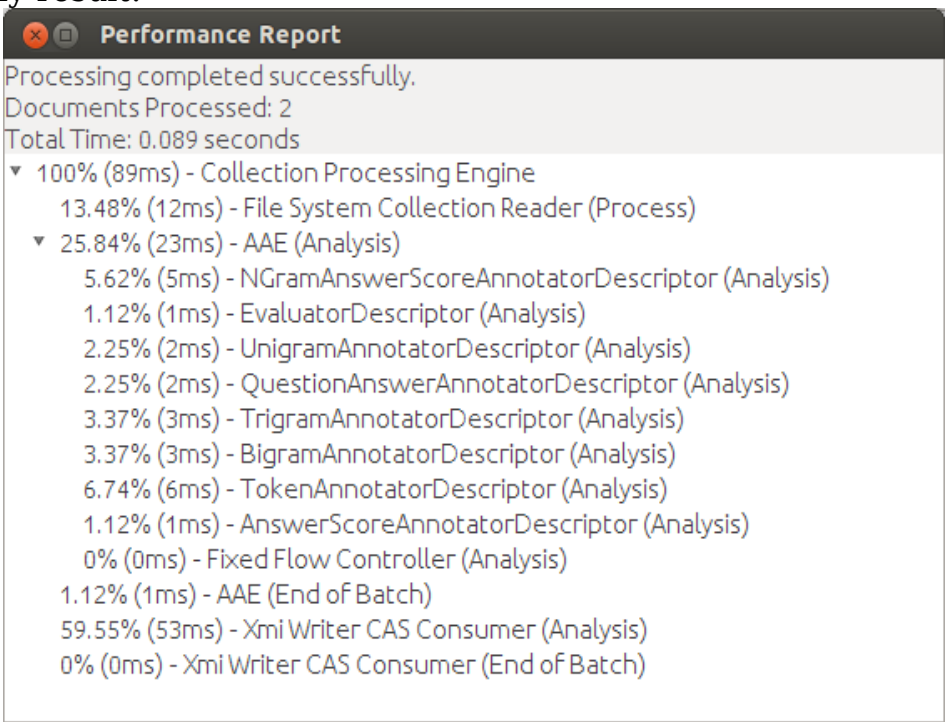
11-791 Design and Engineering of Intelligent Information System

homework 3

Task 1 Execution Architecture with CPE

I difined FileSystemCollectionReader.java and FileSystemCollectionReader.xml as collection reader to read raw input text data. And XmiWriterCASConsumer.java and XmiWriterCASConsumer.xml are used to consume what is produced by Analysis Engines. Then I ran the UIMA CPE GUI tool to agrregate collection reader, AAE and CAS consumer into the CPE pipline.

Below is my result:



Task 2 Deployment Architecture with UIMA-AS

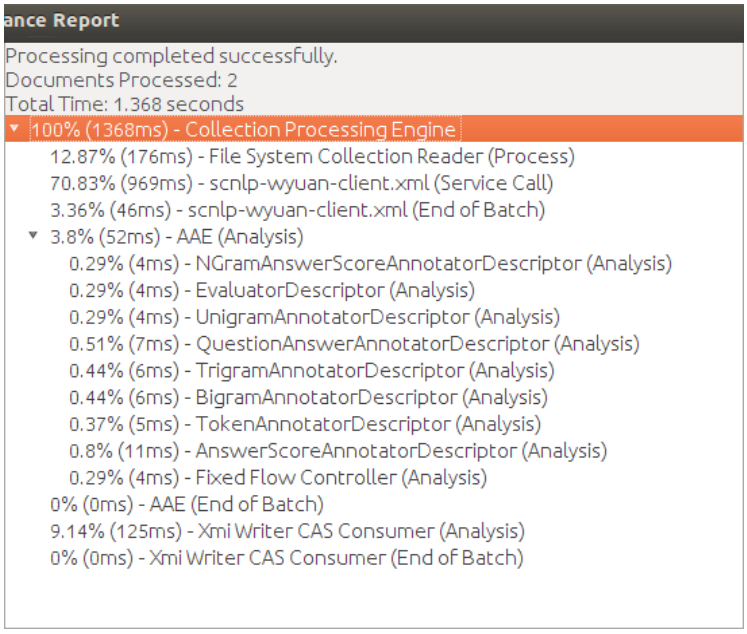
Task 2.2 Creating an UIMA-AS client

I added the cleartk-stanford-corenlp and uimaj-as-activemq dependency packages into Maven dependencies from pom.xml. These packages allows me to process input text contents with ready-to-use methods. Then I created Scnlp-wyuan-client.xml to tell UIMA to use remote AS service. To integrate the Name Entity annotations into my answer scoring component, I revised

answerScoreAnnotator.java.

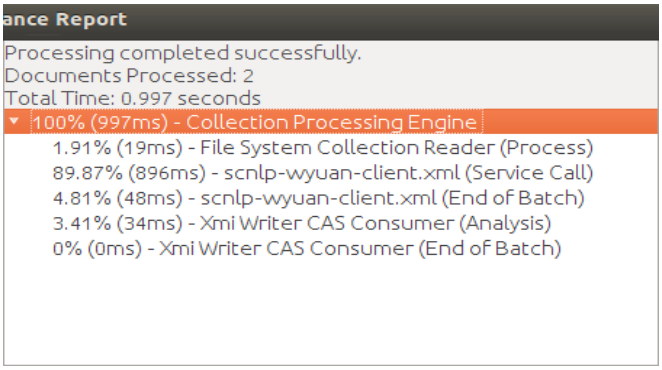
Below is the result of local UIMA-AS client and my annotators with Name Entity in them.

Task 2.2 takes much longer time than in task 1 because I used two analysis engines and one of them is remote. The accuracy is close to that of task 1.



Task 2.3 Deploying your own UIMA-AS service

In this task, I created deployment descriptor (hw2-wyuan-aae-deploy.xml), client descriptor (hw2-wyuan-aae-client.xml) and a CPE descriptor (hw3-ID-aae-as-CPE.xml) to integregate local UIMA-AS service into my annotator. The next step is to open broker service on my computer with startBroker.sh. Then, deploy UIMA-AS service with command line tool (deployAsyncService.sh hw2-wyuan-aae-deploy.xml -brokerURL tcp://localhost:61616). Now I can run CPE GUI tool eclipse to get result. The result is shown below.



The accuracy is similar with local UIMA-AS service in task 2.2 because they use the same service. The speed is a little faster than 2.2 because I did not used my own AAE this time. What is more, local service maybe faster than remote one because of latency.