

11-791 Design and Engineering of Intelligent Information System

Execution Architecture with CPE and Deployment Architecture

with UIMA-AS

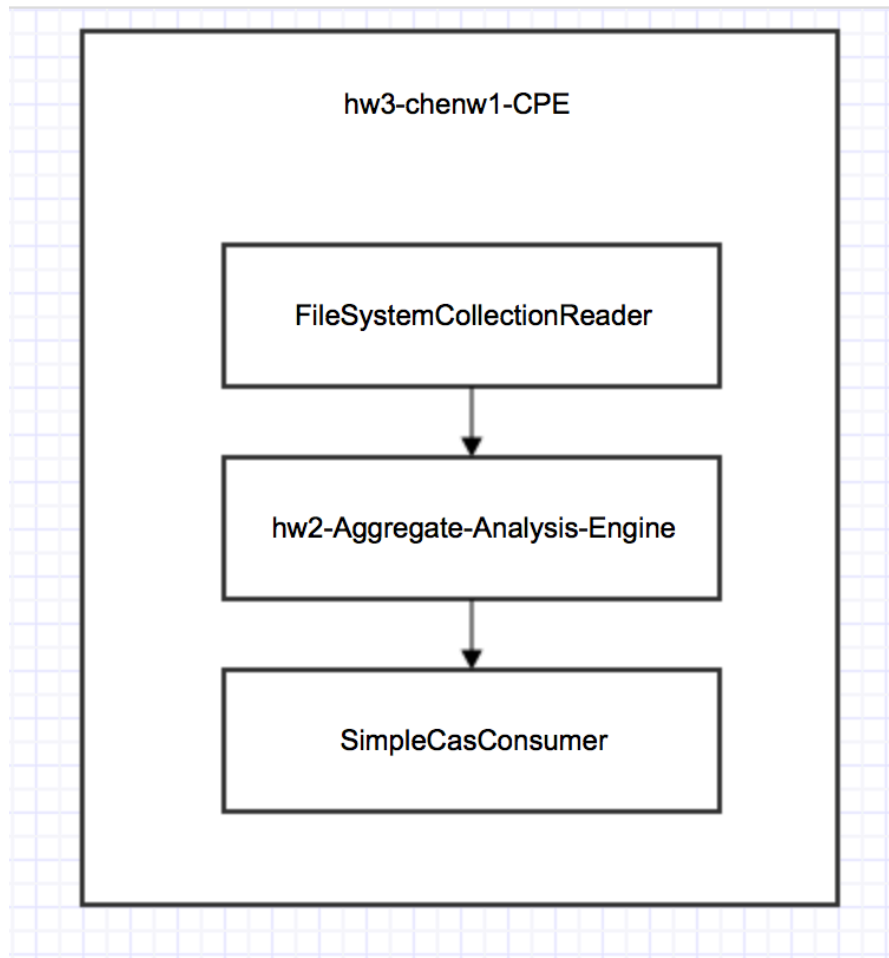
Name: Chen Wang

AndrewID: chenw1

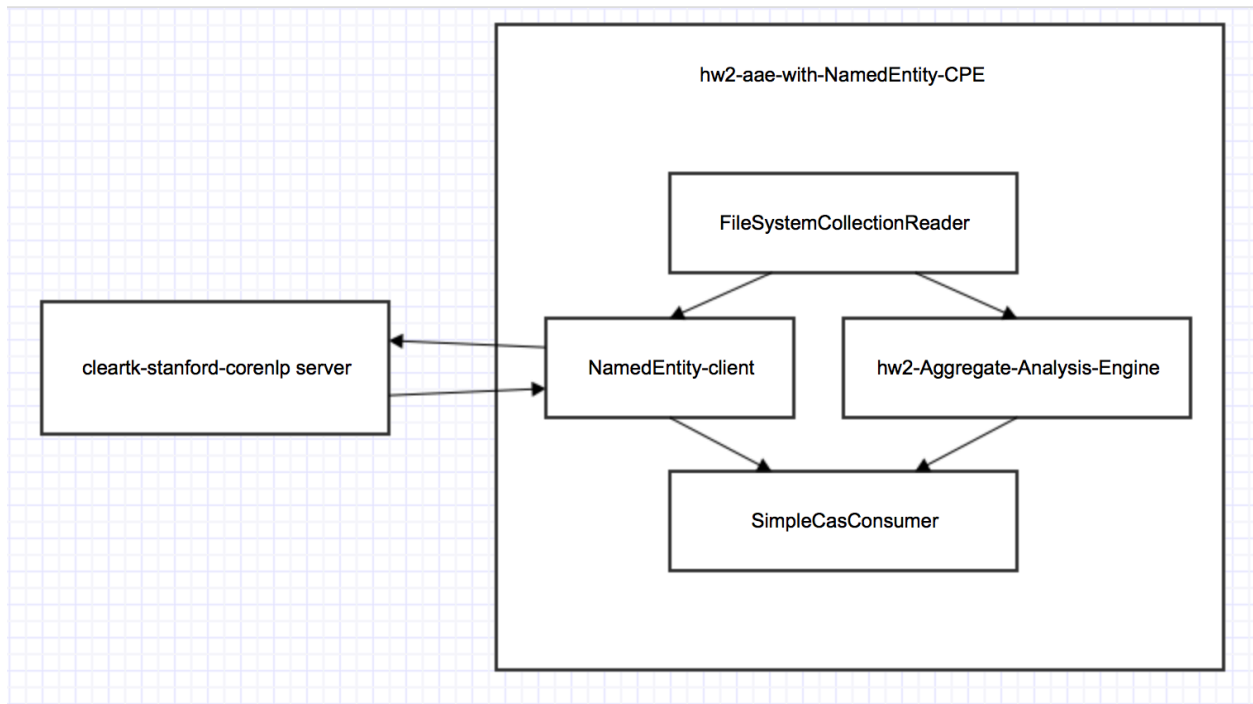
Requirement Analysis:

In homework 2, we built an Aggregate Analysis Engine (AAE) to rate different answers given a question, using N-Gram approach and running in UIMA document analyzer.

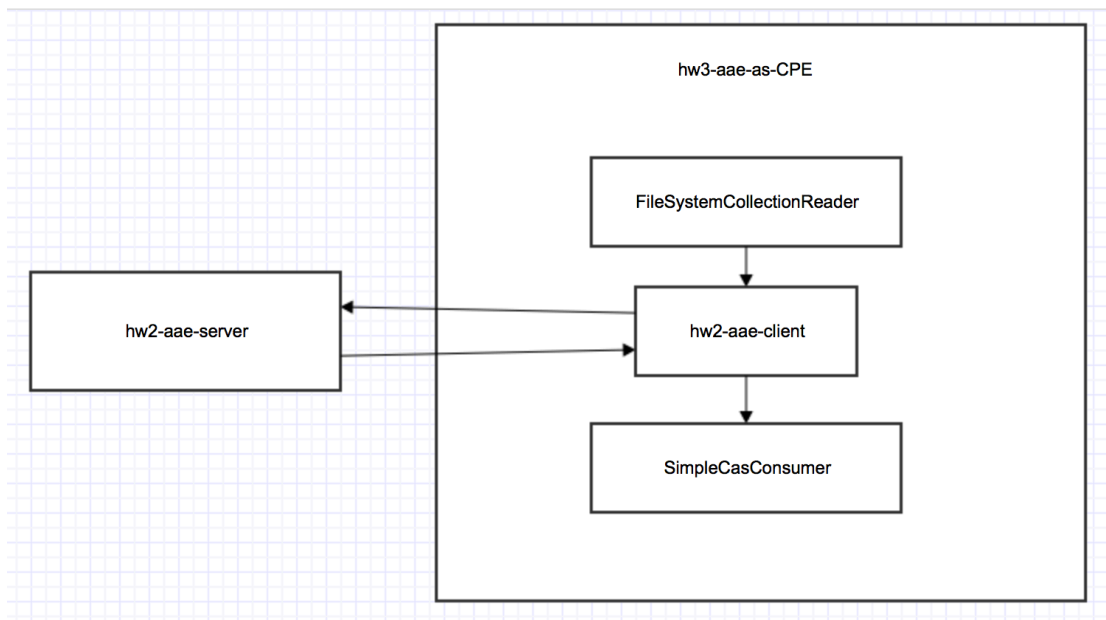
In the first part of homework 3, Task 1.2, we are asked to wrap the AAE from homework 2 into a Collection Processing Engine (CPE) by creating a Collection Reader and a Cas Consumer and putting them all together. The Collection Reader, FileSystemCollectionReader does the simple job of reading the content of all files under the input directory, put all text into a JCas object and pass it to AAE. The Cas Consumer does the job of, given the fully annotated document, print out the final result of our analysis, including the questions and answers analyzed, the accuracy of each question and answers, and the total accuracy. The structure of the CPE is demonstrated by the flow chart below.



The second part of homework 3, Task 2.2, we are asked to integrate Name Entity annotation into our existing answer scoring component. We do this by creating a UIMA-AS client. The client does the job of sending Cas to remote UIMA-AS server and collect the annotated Cas sent back from the server. There could be multiple ways of integrating Name Entity annotation. My approach is, as suggested by Di Wang on Piazza, taking those text covered by Name Entity annotations as N-Gram and score answers using the same N-Gram approach to score answers as before. The structure of Task 2.2 is shown below.



The last part of homework 3, Task 2.3, we are asked to deploy our AAE from homework2 as a service locally, and create a CPE with a UIMA-AS client as the AAE part. The Collection reader and Cas Consumer is the same as the CPE in Task 1.2 in this homework. The structure is shown as below.



Evaluation results and comparison:

	hw2-aae-CPE	aae-with-NameEntity	hw3-aae-as-CPE
Total time	1288 ms	2742 ms	1427 ms
Initialization time	1092 ms	1769 ms	1217 ms
Processing time	196 ms	973 ms	210 ms
Question 1 precision	0.5	0.5	0.5
Question 2 precision	0.67	0.33	0.67

The reason why AAE with NamedEntity performs worse is that, there are not many test data and a little variation of annotation could result in a big change of precision. Note that AAE with NamedEntity and AAE-as-CPE both take more time than hw2-aae-CPE. That is because the network overhead takes more time. And for AAE-with-NamedEntity specifically, the annotation at the server side also takes time.