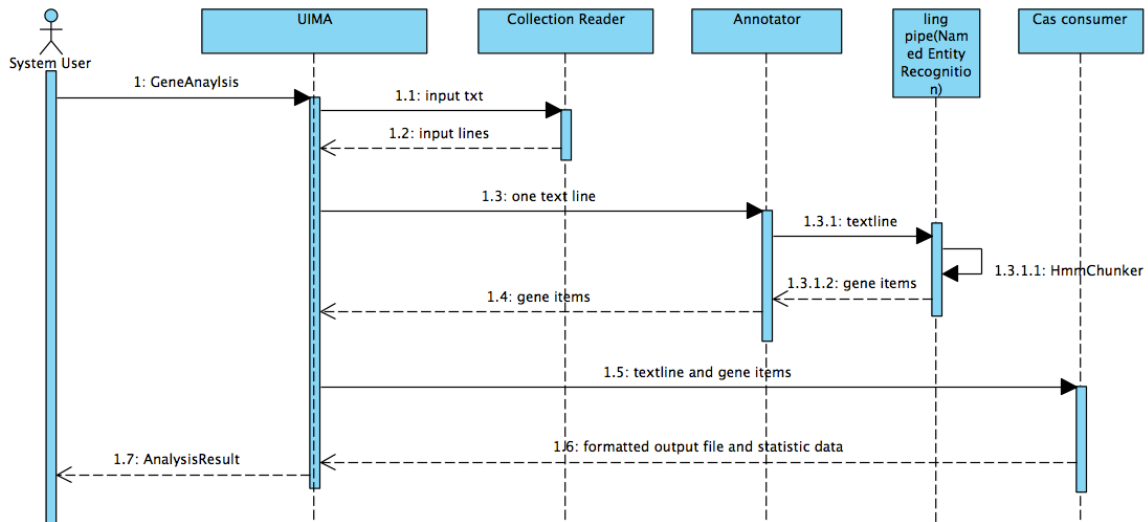


Gene Entity Recognizer Report

Design

As the gene entity is based on UIMA framework, the design is also related to UIMA framework. The basic workflow of entity recognizer is input, access and output. So UIMA framework is quite suitable for this kind of work. The similar process in UIMA is collection reader, annotator and consumer. The sequence diagram illustrated the idea as follow:



The applied recognition algorithm in Annotator is Lingpipe Entity Recognition.

Implementation

Define the type system according to the problem domain.

Get the annotator and annotator description file.

Implement collection reader, consumer and the according description file.

The CpeDescription file to make the different components together.