

Report for Homework 1

Jingwei Shen

js1

1. Overview

For this homework, I design a typesystem to meet the requirements of the homework task. In my design, I consider suggestions in homework guide. My types are designed in an inheriting way. And I use a namespace for my types.

I designed 8 types in total: Model.Base, Model.Document, Model.Answer, Model.Question, Model.Token, Model.AnswerScore, Model.Evaluation, Model.NGram. The picture below shows all of them, their superTpye and features.

Type Name or Feature Name	SuperType or Range		Element Type
[-] Model.Base	uima.tcas.Annotation		
Sentence	uima.cas.String		
Confidence	uima.cas.Double		
Model.Document	uima.tcas.DocumentAnnotation		
[-] Model.Answer	Model.Base		
IsCorrect	uima.cas.Boolean		
Model.Question	Model.Base		
Model.Token	Model.Base		
[-] Model.AnswerScore	Model.Answer		
Score	uima.cas.Double		
[-] Model.NGram	Model.Base		
Elements	uima.cas.FSArray	→	Model.Token
[-] Model.Evaluation	uima.tcas.Annotation		
Precision	uima.cas.Double		
MethodName	uima.cas.String		

Next part will be a detailed introduction of all the types and their relations.

2. Design Details

2.1 Model.Document

This is a type which contains the whole input file. So it can be viewed as an input type.

2.2 Model.Base

This is a basic type of annotation since many other types are inherited from it. It has 2 features: *sentence* and *confidence*. *Sentence* is a string telling us where is the annotation from. Of course it is a whole sentence by the name. And *confidence* is numeric value indicating how confident the origin is.

2.3 Model.Question

This is an annotation inherited from *Model.Base*. It simply contains

the question in the input file.

2.4 Model.Answer

It is also an annotation inherited from the *Model.Base*. Besides the inherited features, it also has a new feature named *IsCorrect*, a Boolean feature which tells us if the answer is correct.

2.5 Model.Token

It is an annotation inherited from *Model.Base*. It shows the results of token analysis.

2.6 Model.NGram

It is It is an annotation inherited from *Model.Base*. It shows the results of NGram analysis. It has a new feature called *Elements* which show us the tokens inside it. *Elements* is an array of *Model.Tokens*.

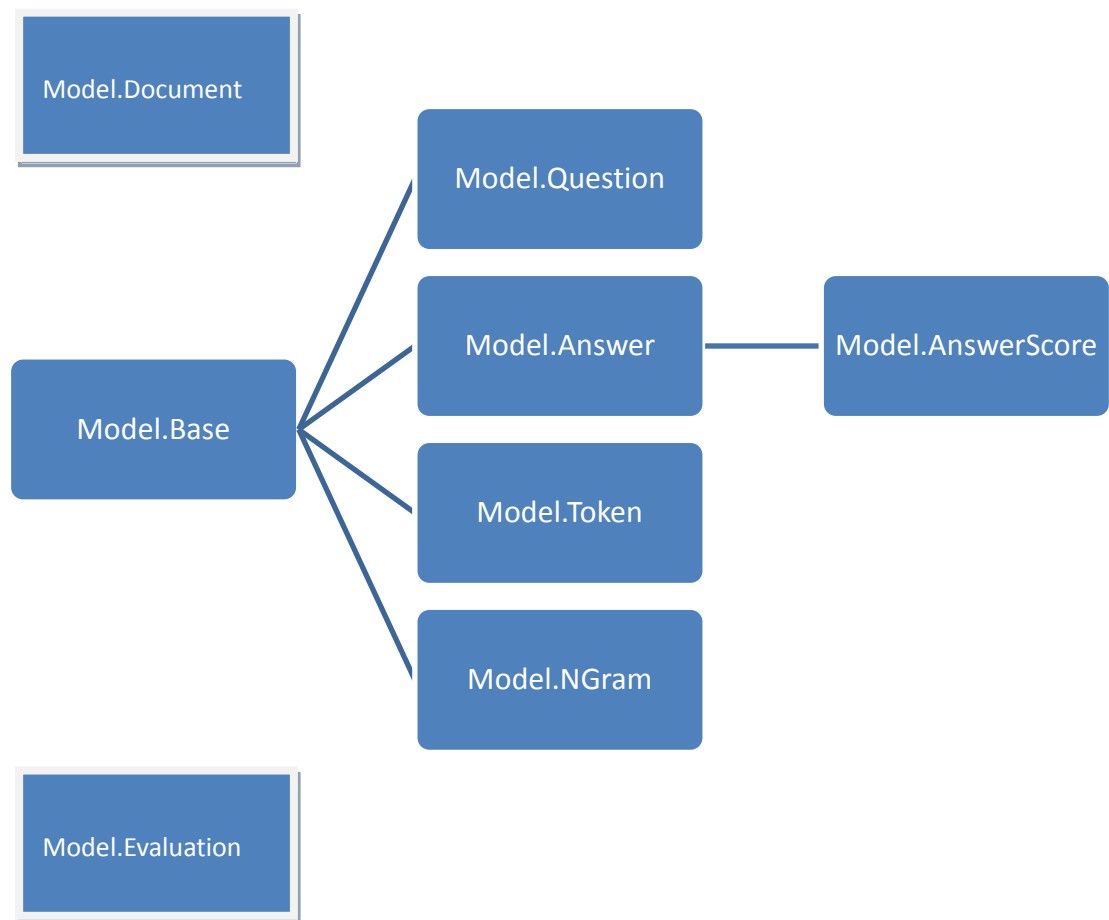
2.7 Model.AnswerScore

It is an annotation inherited from *Model.Answer*. It has a numeric feature named *score* indicating the score of one particular answer.

2.8 Model.Evaluation

This annotation is in fact an output type. It tells us the precision of one method. Its two features are *precision* and *MethodName*. As the name suggests, *precision* is a numeric value of how precise the method is and *MethodName* is a String of the name of applied method.

Here is graph to show the relations of these types.



3. Summary

This homework helps me better understand the concept of typeSystem, which is like a data structure. Thanks for the help from TAs.