Callie Vaughn (ANDREWID cvaughn)
Homework 1, September 11, 2013
Design and Engineering of Intelligent Information Systems

This type system was created for the task of identifying the correct answer to a given question out of a set of provided potential answers. For this system I have created the following five types, each with their own features:

1. The BaseAnnotation type has only two features: *source*, which is a String describing the document from which an object is extracted; and *confidence*, which is a Double that indicates how much confidence the system places in the object's designation. All of the remaining four types extend BaseAnnotation, and therefore they all have their own *source* and *confidence* features.

2. The Token type represents a single token from the CAS file, delimited by whitespace or punctuation. It has three features: *beginning* is an Integer that represents the character-offset index of the first character in the token; *ending* is an Integer representing the character-offset index of the last character in the token; and *content* is a String that contains the string of characters in the token itself.

3. The Ngram type represents a single instance of an N-gram from the CAS file. It has four features: *beginning* is an Integer that represents the character-offset index of the first character in the N-gram; *ending* is an Integer representing the character-offset index of the last character in the N-gram; *content* is a StringArray of length N that contains one string of content characters for each token in the N-gram; and *n* is an integer denoting the number of tokens in the Ngram.

4. The Question type represents a single question from the CAS file. It has two features: *tokenContent* is the words of the Question stored as tokens; and *ngramContent* is the words of the Question stored as Ngrams.

5. The Answer type represents a single potential answer from the CAS file. It has four features: *tokenContent* is the words of the Answer stored as an array of tokens; *ngramContent* is the words of the Answer stored as an array of Ngrams; *score* is the answer score assigned by the system, and *isRight* is a Boolean that is set to True if the Answer is right and False if the Answer is wrong.

For the evaluation step, all of these annotations will be utilized by two methods:
- Answer[] sortAnswers(Answer[]), which will sort the array of Answers according to their scores
- Double calculatePrecision(Answer[]), which will calculate the precision of the system based on the sorted array of Answers.