CSI3120: Programming Language Concepts

Fall 2014

Assignment 3

Due Date: November 17, 2014 Due Time: 11:59pm

As you know, the University of Ottawa is a bilingual institution which means that most courses are offered in both English and French and must be evaluated in the same way. This means that there are a lot of assignments, exams, websites and so on to be translated from English to French. The purpose of this assignment is to see how much of this translation you are able to automate using Perl in order to help professors manage this challenge.

Problem 1 [50 marks]:

In the first problem, you will be asked to translate the structure of assignment sets based on some basic knowledge of the differences between English and French specifically at the University of Ottawa.

Here is the knowledge that I will ask you to rely upon:

Course labeling: the difference in labeling between an English and a French course at the University of Ottawa is only shown in the second digit of the 4-digit number which follows the three letter discipline code, where the "1" is replaced by a "5". Example, the French equivalent of csi3120 is csi3520. The French equivalent of SEG4105 is SEG4505. (the three letter discipline code can appear in upper or in lower cases).

Time display: The time display system can be different in English and in French at the University of Ottawa. In English, we often rely on the AM/PM 12-hour system, whereas in French, we always rely on the 24-hour system. For example 3:15pm in English should be translated into 15h15 in French (the colon, in English become and "h" in French as well).

Date display: Similarly, the date display is different in English and in French. While the month comes before the day in English, the day comes first in French. As well, the month needs to be translated (please see the lexicon of common words below). For example, "January 3, 2014" becomes "3 Janvier 2014" (note that the comma also disappears in French).

Lexicon of common terms: Here is a lexicon (word/expression dictionary) for some terms and expressions used in Assignments:

English	French	English	French	English	French
Winter	Hiver	Due Date	Date de Remise	Problem	Problème
Summer	Été	Due Time	Heure de Remise	Mark	Point
Fall	Automne	Assignment	Devoir	November	Novembre

Your job in this problem is to take a text version of this assignment sheet and translate the course labeling, the time and date display as well as words and terms shown in the above lexicon from English to French in the text. To indicate where the translations have occurred, please surround every translated term but the labels <French> … </French>. For example the English text

"Due Date: November 17, 2014 Due Time: 11:59pm" should become

"<French> Date de Remise </French> : <French> 17 Novembre 2014 </French> <French> Heure de Remise </French> : <French> 23h59 </French>"

Note that this blind translation should create problems in the text. If your program works correctly, the sentence:

"Example, the French equivalent of csi3120 is csi3520." should become

"Example, the French equivalent of <French> csi3520 </French> is csi3520."

Do not worry about this.

Problem 2 [50 Marks]:

In the second question, you have to build a lexicon yourself. To do so, please integrate the advice from the following article

http://www.techrepublic.com/blog/how-do-i/how-do-i-perform-translations-with-perl-and-babel-fish/

on how to integrate babel-fish translations in your program.

Although babel-fish translates sentences and paragraphs directly, here, I will ask you to use it on individual words. In particular, your program will have to take as input some text and build a hash which takes as a key each word of that text and as its value the babel-fish translation of that word. You will then be asked to use your lexicon to translate a new text.

The input I would like to test your system on is:

"As you know, the University of Ottawa is a bilingual institution which means that most courses are offered in both English and French and must be evaluated in the same way. This means that there are a lot of assignments, exams, websites and so on to be translated from English to French. The purpose of this assignment is to see how much of this translation you are able to automate using Perl in order to help professors manage this challenge."

The new text to be translated into French (try other languages that you are familiar with if you want: it could be fun! But please show your output in French):

"The professors are able to manage websites. You know that the purpose of the assignment is the translation of exams in French. The translation of "websites" from English to French is "websites".

Translate the new text directly using Babel-fish. Did your program obtain the same result? The truth is that translation of sentences is much more involved than word-for-word translation. If that's a problem that interests you, there is a lot of research to be done in this domain!!!!