

***** Projet : un wiki-traducteur latin <-> français. *****

Attention : bonne nouvelle, NUL besoin de connaître le latin pour ce projet !
(tout comme il n'est pas besoin d'être médecin pour écrire dans une
SSII un logiciel pour des médecins).

Autre bonne nouvelle, le client a déjà beaucoup réfléchi sur ce projet
et sera toujours réactif pour aider les élèves (réponse aux mails le jour-même).

*** L'historique du projet**

A l'heure du web 2.0 (c'est-à-dire le web coopératif qui offre des services dont le succès repose sur le vieil adage "l'union fait la force"),
il a semblé intéressant à Cyril Banderier, le concepteur de ce projet,
de développer un logiciel de traduction qui tirerait bénéfice de cette technologie web 2.0
d'une part, et qui reposerait d'autre part sur une approche purement algorithmique pour la
traduction.

La première partie de cette démarche existe déjà en partie, et des bases de données de
vocabulaire ont ainsi été compilées collectivement pour diverses langues (autres que le
latin).

La deuxième partie existe déjà aussi : cette approche algorithmique est notamment celle
utilisée avec succès pour les langages informatiques (théorie de la compilation), c'est-à-
dire donc pour des langues qui ne sont nullement des "langues naturelles". Les langues
naturelles (donc le latin, le français...) présentent l'inconvénient d'offrir parfois une certaine
ambiguïté, et donc, pour comprendre (parser) une phrase ("Le jeune pilote (sujet) la
(COD) porte (verbe) " ou "Le jeune (sujet) pilote (verbe) la porte (COD) " ?), il faut
parfois utiliser le sens des mots, on ne peut se contenter de la nature et de la fonction
possible de chaque mot pour trouver le "vrai" sens de la phrase.

Bien sûr, demander à un logiciel de faire cette approche sémantique serait un projet bien
trop ambitieux, et nous ne sommes pas prêt de voir des traducteurs (il faudrait faire
d'immenses progrès en linguistique et en intelligence artificielle) capables de lever toutes
ces ambiguïtés.

Notre approche algorithmique consiste donc à se donner un modèle simplifié de la
grammaire du français :

nous nous restreignons aux phrases du type "sujet verbe complément" et toutes les
variantes classiques. Dans le cas où il y aurait deux façons de lire (grammaticalement
parlant) la phrase, notre logiciel traitera les deux cas à égalité. C'est ensuite à l'humain
d'utiliser son cerveau pour tirer partie du sens, afin de sélectionner quelle traduction
proposée lui semble la plus adéquate.

*** Les objectifs de ce projet**

L'originalité de ce projet réside donc le fait de pousser à fond ces deux approches (bases
de données et analyse grammaticale indépendante du contexte sémantique), le tout
couplé à une grande modularité, flexibilité et extensivité.

En résumé, notre logiciel a le mérite :

- d'offrir une base de données pour le vocabulaire et la grammaire latin
- de permettre aux utilisateurs d'élargir la base du vocabulaire
- de permettre aux utilisateurs d'élargir la base de la grammaire
- de traduire des phrases (latin <-> français)
- de présenter une interface logicielle multilingue
- de présenter une interface logicielle conviviale et sécurisée (identification des contributeurs, etc)
- de pouvoir étendre par la suite sa fonctionnalité à la traduction d'autres langues
- d'être un logiciel libre

Il vaut bien voir que le traducteur ne donnera pas *la* traduction, mais plutôt une liste de traductions grammaticalement possibles, dont la plupart ne feront pas sens si on replace la phrase traduite dans son contexte. En clair, le traducteur ne s'attaque pas à la cohérence sémantique.

*** Le traducteur en ligne : ennemi ou ami des professeurs de latin ?**

Le public sera naturellement l'ensemble des personnes étudiant ou enseignant le latin, ou encore des personnes souhaitant une aide pour traduire un texte latin rencontré au cours d'une lecture.

Le décortiquage grammatical fait par le logiciel, avec son jeu de couleur pour les fonctions grammaticales, sera une aide pédagogique appréciable pour toute personne faisant l'apprentissage du latin.

En un sens, ce logiciel peut révolutionner la façon d'enseigner le latin, tout comme la calculatrice ou les logiciels de calcul formel sur ordinateur permettent de révolutionner la façon d'enseigner, de faire, des mathématiques. Une partie des connaissances (extraction à la main d'une racine carrée ?) est perdue, mais on gagne en temps et en cerveau disponible pour d'autres apprentissages.

Bref comme tout outil, ce logiciel sera ce que les gens en feront. Une aide à la maison pour "tricher" sur les "versions" et les thèmes ? Un cyber-professeur corrigeant sa version et son thème ? Une raison pour l'éducation nationale de ne plus recruter de professeur de latin ?

Arrêtons la plaisanterie, chacun sait que l'apprentissage d'une langue ne s'arrête pas à quelques règles de grammaires et des listes apprises par coeur. La culture d'une langue et d'une civilisation, l'histoire, la pratique au quotidien, l'émulation entre les élèves sont autant d'éléments qui feront que l'outil que nous proposons ici sera un plus pour les enseignants, et donc un ami plutôt qu'un ennemi.

*** Le traducteur en ligne : le latin et... les autres langues.**

De part sa modularité, notre logiciel permet de s'attaquer assez aisément à toutes les langues indo-européennes :

grec, russe, allemand, anglais, néerlandais, suédois, danois, norvégien, islandais, italien, espagnol, roumain...

D'autres langues (arabe, hébreux, chinois, japonais...) sont également attaquables à l'avenir, mais nous obligent à écrire plus de code, et il est préférable d'attendre que les standards unicode

soient vraiment répandus pour éviter les nombreux bugs dûs aux pages de caractères non compatibles. Par ailleurs, ces langues ont parfois un degré d'ambiguïté plus important, et augmente le poids (dans le sens de traduction langue non contrainte -> langue contrainte) de l'importance de l'approche sémantique, point sur lequel justement ce projet fait l'impasse.

EXISTANT : une petite partie du logiciel est déjà partiellement rédigée/fonctionnelle (essentiellement : une portion de l'interface).

CONNAISSANCES UTILES : PHP , MySQL, Lex/YACC (ou tout autre analyseur lexical/syntaxique).

NB DE PARTICIPANTS : de 4 à 6 étudiants
(suivant le niveau des élèves, et leur nombre, on pourra ajouter ou enlever certaines fonctionnalités au cahier des charges).

ENCADRANT : Cyril.Banderier at lipn.fr <http://lipn.fr/~banderier/>

"Le champion du monde de Clabbers"

Ce projet a pour but de créer le premier programme informatique qui sera champion du monde du jeu de "clabbers".

Nota bene : il n'est nullement besoin d'être joueur de clabbers ou de Scrabble pour réaliser ce projet !

Clabbers est une variante relativement populaire (surtout sur Internet Scrabble Club) du jeu de Scrabble.

Tous les mots formés doivent être des anagrammes de mots admis dans le dictionnaire choisi, mais les lettres peuvent ne pas être dans le bon ordre. Par exemple, le mot JOUANT peut être écrit AOUJNT, NAJTUO ou JUNOTA, pour maximiser le score du coup. Par la suite, un joueur peut par exemple rajouter "ED" à la fin de JUNATO, car cela devient une anagramme du mot "DEJOUANT", etc.

Il faut jouer une partie de clabbers pour comprendre à quel point cette variante révolutionne le jeu de scrabble, le plateau devient bien plus dynamique ! Il devient ainsi bien difficile à un humain d'imaginer tous les coups licites.

C'est pourquoi il ne vous sera pas difficile d'exploiter la brute force de l'ordinateur pour battre les humains, (même les champions actuels !).

Il existe une liste officielle des mots que l'on a le droit de jouer au scrabble, celle-ci est librement disponible: <http://www.isc.ro/lists/ods.zip>

La partie algorithmique du projet consistera donc en un parcours "astucieux" de cette liste (par exemple sous la forme d'un arbre) afin d'avoir une vitesse de jeu quasi-instantanée.

Le programme devra être capable de jouer sur Internet Scrabble Club (dont l'interface est facile : juste une dizaine de commandes en mode texte pour interagir avec le serveur), et sera à déposer sous sourceforge en tant que logiciel libre.

Division possible des tâches :

- 1 personne qui travaille sur l'interface (qui devra être en français et anglais).
- 1 personne qui travaille sur l'algorithmique
- 1 personne qui travaille sur l'interaction avec le serveur ISC

Pour plus de renseignements : <http://fr.wikipedia.org/wiki/Clabbers>

EXISTANT : la liste des mots jouables (pas de saisie fastidieuse à faire !!) + Internet Scrabble Club (bien documenté).

CONNAISSANCES UTILES : le langage de programmation de votre choix !

NB DE PARTICIPANTS : de 4 à 6 étudiants (suivant le niveau des élèves, et leur nombre, on pourra ajouter ou enlever certaines fonctionnalités au cahier des charges).

ENCADRANT : Cyril.Banderier at lipn.fr <http://www-lipn.univ-paris13.fr/~banderier/>