

Sujet de Projet

L'objectif est de réaliser en binôme ou trinôme un programme capable de conjuguer tous les verbes en français aux temps simples (sans auxiliaires).

Pour cela on s'appuiera sur le dictionnaire DELA du [LADL](#) (Laboratoire d'Automatique Documentaire et Linguistique). La première partie du travail consistera à construire un dictionnaire composé uniquement des verbes. La seconde partie consistera à réaliser le conjugueur, à partir des données du dictionnaire des verbes préalablement créé.

1/ Dictionnaire des verbes

Le dictionnaire utilisé pour le conjugueur sera extrait à partir du lexique *dela-fr-public.dic* qui vous est fourni. Il aura pour entrées les lemmes et non les formes fléchies. Il sera au format csv avec des tabulations comme séparateurs. Chaque entrée sera donc du type :

lemme → forme fléchiée → liste des codes des sous-catégories grammaticales

Un code sera ajouté pour les verbes commençant par un "h" afin de distinguer les "h" aspirés (pour lesquels il n'y a pas d'élision des pronoms) des "h" muets (avec élision). Vous déterminerez votre propre façon de coder cette information dans votre dictionnaire.

Le script devra également permettre :

- d'ajouter au début du dictionnaire la date de création du nouveau fichier ainsi que ses auteurs,
- d'indiquer l'ajout d'informations par rapport au données initiales du DELA ("h")

2/ Conjugueur

Pour un verbe donné (saisi par l'utilisateur), l'utilisateur pourra demander :

- toutes les conjugaisons (tous les modes, temps et personnes)
- toutes les conjugaisons d'un seul mode précisé par l'utilisateur (par ex. tous les temps de l'indicatif)
- la conjugaison à un mode et temps précis, donnés par l'utilisateur (par ex. présent du subjonctif)

Le programme affichera à chaque fois les 6 personnes, sauf lorsque l'utilisateur demande le participe présent ou passé, l'impératif, et pour les [verbes défectifs](#).

Le programme ne se terminera pas après l'affichage de la conjugaison. L'utilisateur pourra demander d'autres conjugaisons sans devoir relancer le script.

Les temps et modes disponibles seront :

- | | |
|-------------------------------------------------------|------------------------------|
| • Indicatif : présent, passé simple, imparfait, futur | • Impératif : présent |
| • Subjonctif : présent, imparfait | • Participe : présent, passé |
| • Conditionnel : présent | |

On tiendra compte des cas où "je" doit être élide.

On tâchera de ne pas trop contraindre les saisies de l'utilisateur et de ne pas lui demander trop de saisies si ce n'est pas nécessaire. Par exemple, il peut y avoir un fonctionnement par défaut où l'utilisateur n'aurait rien à taper, seulement valider en tapant sur "entrée". Soyez créatifs !

On veillera également à soigner les affichages.

Dictionnaire DELA

Chaque entrée est du type : `forme fléchie,lemme.codes`

La partie `codes` inclut la catégorie grammaticale et les traits morphosyntaxiques. Pour les verbes, le codage est du type : `catégorie + z1 : mode/temps personne`

Ex. : `V+z1:F2p` => *Verbe indicatif Futur 2^{ème} personne pluriel*

Une même forme fléchie peut correspondre à plusieurs conjugaisons. Dans ce cas elles sont indiquées les unes à la suite des autres, séparées par " : "

Par ex. la forme fléchie `mange` correspond à 5 conjugaisons :

`mange,manger.V+z1:P1s:P3s:S1s:S3s:Y2s`

Les codes sont les suivants :

Temps des verbes

- W (ou INF) = Infinitif
- K (ou PP) = Participe passé
- G (ou ANT) = Participe présent
- P = Indicatif Présent
- I = Indicatif Imparfait
- J = Indicatif Passé simple
- F = Indicatif Futur
- C = Conditionnel présent
- S = Subjonctif Présent
- Y = Impératif Présent

Genre, nombre, personne

- ms = masculin singulier
- fs = féminin singulier
- mp = masculin pluriel
- fp = féminin pluriel
- 1s = 1^{ère} personne du singulier
- 2s = 2^{ème} personne du singulier
- 3s = 3^{ème} personne du singulier
- 1p = 1^{ère} personne du pluriel
- 2p = 2^{ème} personne du pluriel
- 3p = 3^{ème} personne du pluriel

z1, z2 et z3 font référence au type de lexique (respectivement lexique de base, lexique secondaire, lexique résiduel), cette information ne nous intéressera pas.

Rendu des travaux

Une archive nommée `VOS_NOMS.zip`, déposée sur la plateforme Chamilo, rubrique travaux>Rendu projets, au plus tard le **17 décembre 2021** et contenant :

- ◆ Le script de création du dictionnaire de verbes
- ◆ Le fichier csv contenant les verbes extraits
- ◆ Le script du conjugueur
- ◆ Un court rapport (2 pages maxi) dans lequel vous indiquerez :
 - Le code utilisé pour les verbes en "h" avec une brève justification de votre choix de code ;
 - Les problèmes rencontrés et les solutions trouvées ;
 - Les éventuels problèmes de fonctionnement (bogues, certains cas non pris en charge, etc.) ;
 - Des propositions d'améliorations
- ◆ **NE PAS METTRE** le dictionnaire ABU dans l'archive, sinon ce sera trop lourd.

Votre code devra être clair et bien commenté. Les structures de données devront avoir des noms cohérents et explicites.

Vos commentaires incluront la date de création du script, ses auteurs, ainsi que ses fonctionnalités.