

Sous le cadre du module de théorie des langages et compilation ,et en utilisant les outils flex/bison et le language de programmation C,

Projet Traducteur: Dialecte Marocain => Francais

EL-HARSAL Mohamed Ali

GUOURCH Otmane

DAHMOU Hicham

Composants du projet

Scripts python	122 lignes
Bibliotheque des listes	151 lignes
flex	408 lignes
Bison	479 lignes
Differents fichiers des ressources	-
Total	1160 lignes

C'est pas le nombre des lignes qui rend un projet bon mais c'est un indicateur parmi d'autre ,donc je vous invite a le lire.



People think this is a bit odd, but if they knew why I was doing it. They would understand

Que ce qu'on a appris le long de la réalisation de ce projet:

- Renforcement de notre niveaux dans le language C.
- L'importance des outils de gestions de versions (meme pour un group de 3 personnes).
- Organiser le code et simuler un comportement “orienté-objet” dans un language non orienté objet.
- Certains outils sont plus adaptés à certaines taches .

...

Vous voulez dire
que vous ne savez
pas déjà ceci ???



Quelques problèmes que nous avons rencontré ?

- Plusieurs choix dans un temp très limité.
- Le code source du flex et bison “n'est pas jolie”.
- Le code est développé en parallèle sur deux plateformes différentes.
- Le dialecte marocain n'est même pas standardisée.

“Pas jolie” c'est subjectif, soyez reconnaissant au moins vous avez un outil fonctionnel



Rappels de cours:

Langage formel

Le **langage formel (LF)** est un langage obéissant à une syntaxe formelle stricte, permettant d'exposer des énoncés de manière précise.

Dans un langage formel, on fait abstraction du sens et de la signification des mots (la sémantique) ☐ **Ce qui facilite la transformation et la manipulation.**

☐ les machines peuvent les comprendre et les interpréter facilement.



Langage naturel

Un **langage naturel (LN)**, est une *langue* parlée par *un être humain*.

Il est défini par un ensemble de règles et sur un vocabulaire fini, très large et évolutif.

Avec un LN on peut formuler une même idée avec différentes manières et d'avoir pour une même formulation plusieurs sens ☐ **Cela peut donner lieu à des problèmes de communication dus à des interprétations multiples.**

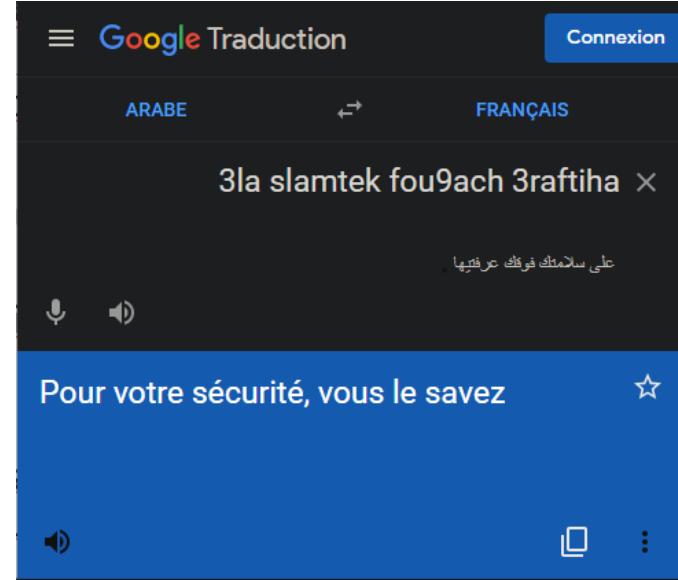


urban
DICTIONARY



Generation des données/ Traitement primitif:

- Scrapping des commentaires facebook ou des postes déterminés.
- Extraction de la liste des mots (avec un filtrage primaire).
- Transformation, prenant en considération nos règles d'écritures.
- Implémentation d'une “norme” qui sera utile au cas des mots auxquels il y a une ambiguïté dans la façon d'écriture (distance de Levenshtein)
- Programme en ligne des commandes qui aide à la classification des mots.



Quels outils pour faire ces prétraitements sur le text brut ???



Analyse lexical

- On fetch les données à partir de notre “base de données” (disque dur) vers notre “structure de données” (RAM).
- On essaye d'extraire le maximum d'informations qu'on donne à l'entrée d'analyseur syntaxique sous forme des structures, qui seront empilées implicitement au cours du traitement d'une phrase quelconque.

Quelle structure de données ? Et pourquoi ?



Analyse syntaxique/Analyse Semantique

- listPhrase : phrase
 - | listPhrase phrase ;

- phrase : sujet description EOL
 - | verb EOL
 - | verb complement EOL
 - | sujet verb complement EOL
 - | sujet verb EOL
 - | QUESTION SPACE verb '?'
 - | QUESTION SPACE verb complement '?"

- Sujet : PRONOUN SPACE
 - | PRONOUN SPACE 'w' SPACE NOUNP SPACE
 - | NOUNP SPACE
 - | NOUNP SPACE 'w' SPACE NOUNP SPACE
 - | NOUN SPACE
 - | NOUN SPACE 'w' SPACE NOUN SPACE ;

- Verb : VERBE SPACE VERBE SPACE
 - | NEGATION VERBE SPACE VERBE SPACE
 - | VERBE SPACE
 - | NEGATION VERBE SPACE ;

- Description : ADJ SPACE
 - | NEGATION ADJ SPACE ;

- complement: sujet
 - | sujet ADJ SPACE
 - | ADV SPACE
 - | ADV SPACE ADV SPACE
 - | description ADV SPACE
 - | 'I' SPACE NOUN SPACE
 - | 'f' SPACE NOUN SPACE
 - | 'm' '3' 'a' SPACE NOUNP SPACE ;

Noter bien que notre programme a deux modes d'executions :

- un mode normal ou on affiche la traduction et en cas d'erreur un message l'erreur (vous devais ajouter les mots utilisé au fichiers de resources).
- un mode “verbose” auquelle on affiche tout les details sur la phrase que vous avez entree ,pour l'activer il suffit d'ajouter le flag ‘-d’ a la commande sans oublier de passer en parametre le nom du fichier a traduire.

```
[T1945@silver listeV4]$ ./darijaF ./text.txt -d

*****
question sent fu9ax
verbe sent sala
pronoun sent ana
    Verb: AR>>>sala FR>>>terminer Genre(0:M|1:F|2:U)>2 personne(1|2|3)>>>1
          Nombre(S:singulier|P:pluriel)>>>P Temps(Present:0|Past:1|Future:2)>>>1 Negatif>>>Non
nounP sent mouna
    Complement: AR>>>ana w mouna FR>>>mouna et moi Genre(0:M|1:F|2:U)>2
          Nombre(S:singulier|P:pluriel)>>>P personne(1|2|3)>>>1
quand-est-ce-que terminer(Present|1Personne|Pluriel) mouna et moi ?
```

```
≡ text.txt ×  
≡ text.txt  
1 fu9ax salina ana w mouna ?  
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL 1: bash  
  
noun sent drhm  
noun sent doular  
verbe sent sala  
    Verb: AR>>>sala FR>>>terminer Genre(0:M|1:F|2:U)>2 personne(1|2|3)>>>3  
        Nombre(S:singulier|P:pluriel)>>>P Temps(Present:0|Past:1|Future:2)>>>1 Negatif>>>Non  
  
noun sent lbanka  
    Subject: AR>>>drhm w doular FR>>>le dirham et le dollard Genre(0:M|1:F|2:U)>2  
        Nombre(S:singulier|P:pluriel)>>>P personne(1|2|3)>>>3  
  
    Complement: AR>>>avec lbanka FR>>>dans la banque Genre(0:M|1:F|2:U)>M  
        Nombre(S:singulier|P:pluriel)>>>S personne(1|2|3)>>>N  
  
le dirham et le dollard terminer(Present|3Personne|Pluriel) dans la banque  
*****
```

```
≡ text.txt ×  
≡ text.txt  
1 fu9ax salina ana w mouna ?  
2 drhm w doular salaw f lbanka .  
3 taha w mouna ma5rjux klaw m3a hasna .  
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL 1: bash  
  
*****  
  
nounP sent taha  
nounP sent mouna  
verbe sent 5rj  
verbe sent kla  
    Verb: AR>>>5rj FR>>>sortir Genre(0:M|1:F|2:U)>2 personne(1|2|3)>>>3  
        Nombre(S:singulier|P:pluriel)>>>P Temps(Present:0|Past:1|Future:2)>>>1 Negatif>>>Oui  
  
    Verb: AR>>>kla FR>>>manger Genre(0:M|1:F|2:U)>2 personne(1|2|3)>>>3  
        Nombre(S:singulier|P:pluriel)>>>P Temps(Present:0|Past:1|Future:2)>>>1 Negatif>>>Non  
  
nounP sent hasna  
    Subject: AR>>>taha w mouna FR>>>taha et mouna Genre(0:M|1:F|2:U)>2  
        Nombre(S:singulier|P:pluriel)>>>P personne(1|2|3)>>>3  
  
    Complement: AR>>>avec hasna FR>>>avec hasna Genre(0:M|1:F|2:U)>M  
        Nombre(S:singulier|P:pluriel)>>>S personne(1|2|3)>>>N  
  
taha et mouna n'est pas sortir manger(Present|3Personne|Pluriel) avec hasna
```

≡ text.txt X ≡ origine.txt

≡ text.txt

```
1 wax mxaw taha ?
2 mxaw klina .
3 fu9ax salina ana w mouna ?
4 taha w mouna ma5rjux klaw m3a hasna .
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL

1: bash

```
[T1945@silver listeV4]$ ./darijaF ./text.txt
```

```
*****
```

Error | Line: 1
le sujet et le verbe doivent de refaire au meme nombre

```
*****
```

Error | Line: 2
les deux verbes n'ont pas la meme structure des relations de personne

```
*****
```

```
quand-est-ce-que terminer(Present|1Personne|Pluriel) mouna et moi ?
```

```
*****
```

```
taha et mouna n'est pas sortir manger(Present|3Personne|Pluriel) avec hasna
```

```
*****
```

Vos question ?

