



INF580: Mathematical Programming

Projet : Générateur de poésie

Matthieu LEQUESNE & Quentin LISACK

24 février 2016

Table des matières

1	Objectif du projet	2
2	Implementations	2
2.1	Gérer les contraintes	2
2.2	Génération des données	3
2.3	Autres approches	3
3	Conclusion	3

Résumé

Natural Language Processing ... blah

1 Objectif du projet

Notre objectif a été d'obtenir un programme capable d'écrire des poèmes. Pour cela, on a réduit le problème à un programme qui écrit deux vers qui riment. En répétant cette opération, on peut créer des poèmes de taille arbitraire. Nous nous sommes placé dans le cadre classique de vers de 12 syllabes, mais cette donnée importe assez peu. Le programme doit donc avoir les spécifications suivantes :

- Entrée : un corpus de textes poétiques.
- Sortie : deux vers, tels que :
 - chaque vers comporte 12 syllabes ;
 - les deux vers riment (mais sont différents) ;
 - le texte a un sens.

On pourra ajouter qu'on désire un programme qui comporte une part d'aléatoire puisqu'on souhaite générer des poèmes différents à chaque appel.

2 Implementations

2.1 Gérer les contraintes

2.1.1 Stockage des mots

Pour chaque vers on stocke les mots qui apparaissent. Or, on ne connaît pas a priori le nombre de mots utilisés mais on sait uniquement que le nombre de syllabes est fixé à 12. Or en français il ne peut pas y avoir plus d'un mot sur deux qui ne contient pas de voyelle (donc qui compte pour 0 syllabes). Ceci donne une borne supérieure de 24 mots par phrase.

Pour chaque vers, on stocke donc les 24 mots du vers, avec la possibilité que certains mots soient le *mot vide*. Pour plus de simplicité, tous les mots vides sont regroupés au début du vers.

En réalité, on ne travaille pas avec les mots (les chaînes de caractères) mais chaque mot qui apparaît dans le corpus initial est identifié par un numéro. Soit N le nombre de mots différents. On a donc des variables :

$$x_{i,j}^{(k)} = \begin{cases} 1 & \text{si le } i^{\text{ème}} \text{ mot du vers } k \text{ est } j \\ 0 & \text{sinon.} \end{cases}$$

avec $1 \leq i \leq 24, 1 \leq j \leq N$ et $1 \leq k \leq 2$.

On impose qu'il y a un et un unique mot par position :

$$\forall k \in \{1, 2\}, \forall i \leq N, \quad \sum_{j=1}^N x_{i,j} = 1.$$

2.1.2 Le nombre de syllabes

On stocke pour chaque mot i son nombre de syllabes $syl(i)$. On impose donc la contrainte suivante :

$$\forall k \in \{1, 2\}, \quad \sum_{i=1}^{24} \sum_{j=1}^N x_{i,j}^{(k)} syl(j) = 12.$$

2.1.3 Les rimes

A chaque mot i du corpus, on lui assigne sa classe de rime $rim(i)$, qui correspond aux 4 dernières lettres du mot (ou moins s'il est de taille inférieur). On décide donc que deux mots a et b riment si $rim(a) = rim(b)$.

La condition de rime des deux vers devient alors :

$$\sum_{j=1}^N x_{24,j}^{(1)} rim(j) = \sum_{j=1}^N x_{24,j}^{(2)} rim(j).$$

2.1.4 Le sens du texte

2.1.5 Caractère aléatoire

2.2 Génération des données

.dat files

2.3 Autres approches

2.3.1 Approche globale

2.3.2 Syllabe par syllabe

2.3.3 Par groupes de mots

3 Conclusion

*SI BIEN RACE D UN SOMBRE ET DU DIFFORME
CE CHAT I RACE D OU SUR LE VIN INFORME
VIE ET SOMBRE ET DE MA VIE ET D UN ORDRE
LE COEUR ET BRULE JUSQU AU FOND DE PERDRE*

*PLEIN DE TOI BIZARRE DEITE DEITE BRUNE
DE LA VIE ET SOMBRE ET DE LA VIE ET BRUNS
COEUR PLEIN DE MA VIE ET QUE JE TE COMPARER
NATURE AINSI QU A LA VIE ET PAREE*

*YEUX DE LA VIE ET DE L ON MEPRISE SALUE
ET D AUSSI PETITS RACE D UN DECOR SALUT
ET VOTRE PURE LUMIERE ET QUI S EFFRAYA*

ET BRULE PREND PITIE DE LA VIE ET LES FRAYE

INF580 : GÉNÉRATEUR DE POÉSIE

LA VIE ET DE TA VIE ET DANS LES NATIVE
LA VIE ET COMME UNE NUIT FUGITIVE