

Travaux Pratiques de linguistique

Ressources linguistiques

—Master 1—

Recherche de motifs

Dans cette fiche, nous verrons comment effectuer des recherches de motifs complexes au moyen des expressions rationnelles et des graphes.

La recherche d'expressions rationnelles est détaillée dans le chapitre 4 du manuel.

► Exercice 1. Rappels sur les expressions rationnelles

Une expression rationnelle peut avoir une des 5 formes suivantes :

- AB : concaténation des expressions rationnelles A et B ;
- $A + B$: disjonction des expressions rationnelles A et B ;
- A^* : étoile de Kleene de l'expression rationnelle A ;
- (A) : expression rationnelle A ;
- xyz : séquence de lettres.

Rechercher dans le roman *Le tour du monde en 80 jours* les motifs suivants :

1. toutes les occurrences des pronoms personnels (*je, tu, il, ...*) ;
2. toutes les occurrences des pronoms personnels qui sont suivis par un verbe ;
3. toutes les suites d'au moins 3 adjectifs ;
4. toutes les suites de noms ; Pourquoi le motif $\langle N \rangle^*$ produit-il une erreur ? Que faire pour l'éviter ?

► Exercice 2. Motifs référant aux dictionnaires

Les tableaux 1, 2 et 3 décrivent les codes grammaticaux, sémantiques et flexionnels utilisés dans les dictionnaires. À l'aide de ces codes, et en vous aidant du chapitre 4 du manuel, rechercher les motifs suivants :

1. tous les adjectifs au féminin pluriel ;
2. tous les verbes, soit à l'imparfait, soit au présent ou à l'imparfait du subjonctif ;
3. tous les mots qui possèdent le trait sémantique "concret" ;
4. tous les noms possédant le trait sémantique "humain collectif" ;
5. tous les adjectifs qui ne sont pas très spécialisés ;
6. tous verbes un peu ou très spécialisés, soit au participe passé, soit à l'infinitif ;

7. tous les noms non humains ;

Les dictionnaires de mots simples et composés du texte se nomment `dlf` et `dlc`, et se trouvent dans le répertoire `/French/Corpus/80jours_snt/`. En utilisant la commande "Check Format" du menu "DELA", récupérer la liste des codes grammaticaux et sémantiques utilisés dans les dictionnaires du texte. À l'aide de cette liste, rechercher les motifs suivants :

8. tous les déterminants numériques ;

9. tous les noms de profession ;

10. tous les noms propres.

► Exercice 3. Méta-motifs

1. Consulter le chapitre 4.3.1 du manuel pour savoir ce que signifie les motifs `<NB>`, `<MOT>`, `<MIN>`, `<MAJ>`, `<PRE>` et `<DIC>`.
2. Rechercher tous les mots qui ne sont pas dans le dictionnaire.
3. Rechercher tous les mots qui ne sont pas écrits tout en minuscules.
4. Rechercher toutes les séquences commençant par le verbe *avoir*, se terminant par un verbe au participe passé, dans lesquelles peuvent s'insérer des séquences quelconques de mots entre virgules (exemple : *eût, au contraire, perdu*).

► Exercice 4. Grammaires locales

On souhaite construire une grammaire pouvant reconnaître les expressions de date et d'heure qui sont présentes dans le texte *Le tour du monde en 80 jours*.

1. Rechercher tous les nombres présents dans le texte.
2. En étudiant la concordance obtenue, construire un graphe reconnaissant les expressions de date.
3. Même question avec les expressions d'heure.
4. Factoriser les termes qui peuvent l'être dans des sous-graphes.
5. Généraliser vos grammaires à d'autres formes de date que celles présentes dans le texte.

Code	Signification	Exemples
A	adjectif	fabuleux
ADV	adverbe	réellement, à la longue
CONJC	conjonction de coordination	mais
CONJS	conjonction de subordination	puisque, à moins que
DET	déterminant	ses, trente-six
INTJ	interjection	adieu, mille millions de mille sabords
N	nom	prairie, vie sociale
PREP	préposition	sans, à la lumière de
PRO	pronom	tu, elle-même
V	verbe	continuer, copier-coller

TAB. 1 – Codes grammaticaux usuels

Code	Signification	Exemple
z1	langage courant	blague
z2	langage spécialisé	sépulcre
z3	langage très spécialisé	houer
Abst	abstrait	bon goût
An1	animal	cheval de race
An1Coll	animal collectif	troupeau
Conc	concret	abbaye
ConcColl	concret collectif	décombres
Hum	humain	diplomate
HumColl	humain collectif	vieille garde
t	verbe transitif	foudroyer
i	verbe intransitif	fraterniser
en	particule pré-verbale (PPV) obligatoire	en imposer
se	verbe pronominal	se marier
ne	verbe à négation obligatoire	ne pas cesser de

TAB. 2 – Quelques codes sémantiques

Code	Signification
m	masculin
f	féminin
n	neutre
s	singulier
p	pluriel
1, 2, 3	1 ^{re} , 2 ^e , 3 ^e personne
P	présent de l'indicatif
I	imparfait de l'indicatif
S	présent du subjonctif
T	imparfait du subjonctif
Y	présent de l'impératif
C	présent du conditionnel
J	passé simple
W	infinitif
G	participe présent
K	participe passé
F	futur

TAB. 3 – Codes flexionnels usuels