

Universités Bordeaux 1 — Master Informatique
UE INF 316 – Traitement Automatique des Langues
Examen du vendredi 13 mai, 14h–17h
documents non autorisés / calculatrice bienvenue

Exercice A

On considère les mots suivants *tu, le, la, les, charcutier, charcutiers, sale, sales, salent, tranche, tranches, tranchent, petite, petites* et les catégories suivantes: *det, adj, nom, pronom, verbe*.

(A.i) Quelles sont les catégories possibles pour chaque mot?

(A.ii) Construire un Modèle de Markov Caché (HMM) en supposant qu'il y a équi-probabilité des transitions et des mots dans chaque catégorie.

(A.iii) Pour les phrases suivantes:

- (1) Les charcutiers salent les petites tranches.
- (2) Le charcutier sale les petites tranches.
- (3) Le charcutier sale les petites.
- (4) Le charcutier sale les tranches.
- (5) Le charcutier sale les tranche.
- (6) Le charcutier sale la tranche.
- (7) Tu sales les tranches.
- (8) Tu les tranches.

donner leur probabilités et leur étiquetage grammatical.

(A.iv) En utilisant les phrases ci-dessus comme corpus d'apprentissage, réestimer les paramètres du HMM par la méthode de Viterbi.

(A.v) En utilisant les phrases ci-dessus comme corpus d'apprentissage, réestimer les paramètres du HMM par la méthode de Baum-Welsh.

(A.vi) Comparer les deux méthodes.

(A.vii) Comment pourrait-on raffiner ce modèle afin que les prédictions soient meilleures?

Exercice B

(B.i) Ecrire une grammaire hors-contexte engendrant les phrases correctes sur le vocabulaire de l'exercice précédent.

(B.ii) Quel est le type du langage engendré? Si c'est possible, donner une expression décrivant ce langage.

(B.iii) Utiliser l'algorithme d'Earley pour analyser 1 et 6.

Exercice C

On souhaite modéliser par des grammaires de clauses définies écrites en Prolog la grammaire de phrases simples avec des relatives qui sont aussi bien introduites par le pronom « qui » (dont l'antécédent est le sujet de la relative) que par le pronom « que » (dont l'antécédent est l'objet de la relative). Les verbes sont

présenter, connaître, les noms propres *Adeline, Benoît, Caroline, Didier*, les noms communs *personne, ami*. Pour chaque grammaire produite on donnera deux analyses de phrases correctes et un exemple de phrase non analysée en expliquant pourquoi elle ne l'est pas.

On souhaite pouvoir analyser des phrases comme:

(9) Adeline, qui connaît une personne que Pierre connaît, présente un ami à Didier que Benoît apprécie.

(C.i) Ecrire une DCG qui rende compte de ces phrases, en tenant compte de l'accord entre un verbe et son sujet.

(C.ii) Modifier la DCG pour que l'analyse produise les arbres d'analyse.

(C.iii) Etendre cette grammaire pour permettre que les verbes soient au passé composé, et en respectant l'accord entre le participe passé et le complément d'objet antéposé (*que*).

(10) Caroline présente des amis que Benoît a connus à Adeline.

(C.iv) Etendre cette grammaire pour autoriser l'inversion du sujet dans les relatives introduites par *que* ainsi que la permutation des compléments de *présenter*

(11) Adeline, qui connaît une personne que connaît Pierre, présente à Didier qu'apprécie Benoît un ami.

Exercice D

Donner un transducteur qui analyse la conjugaison des verbes réguliers du premier groupe (*chanter*) aux temps suivant de l'indicatif: présent, futur, imparfait.

(D.i) Etendre le transducteur pour prendre en compte les verbes comme *appeler* (*j'appelle, nous appelons*)?

(D.ii) Comment procéder pour gérer les temps composés?

Exercice E

Une grammaire catégorielle est la donnée pour chaque mot d'un ensemble de formules définies ainsi: n, sn, S sont des formules, et si U et V sont deux formules, U/V est une formule et $V \setminus U$ sont des formules. Une suite de mots $m_1 \cdots m_n$ est une phrase si on peut trouver pour chaque mot m_i une formule f_i parmi celles qui lui sont associées de sorte que la suite de formules $f_1 \cdots f_n$ se réduisent en S au moyen des règles suivantes: $(U/V)V \rightarrow U$ et $V(V \setminus U) \rightarrow U$.

Pierre	sn
Marie	sn
un	sn/n
chat	n
dort	$sn \setminus S$
regarde	$(sn \setminus S)/sn$

(E.i) Analyser les phrases

(12) Pierre dort.

(13) Un chat dort.

(14) Marie regarde un chat.

(E.ii) Adapter l'algorithme de Cocke-Younger-Kasami aux grammaires catégorielles.

(E.iii) Si on a le temps on réfléchira aux formules à associer aux adjectifs (antéposés ou postposés), et aux pronoms relatifs.