

Site : Luminy St-Charles St-Jérôme Cht-Gombert Aix-Montperrin Aubagne-SATIS
 Sujet de : 1^{er} semestre 2^{ème} semestre Session 2 Durée de l'épreuve : 2h
 Examen de : M2 Nom du diplôme : Master Intelligence Artificielle
 Code du module : Libellé du module : TAL et linguistique
 Calculatrices autorisées : OUI Documents autorisés : OUI

1 Graph-based Parsing par programmation dynamique

Soit S_1 la phrase *Il pleuvait sur la route* et son analyse A_1 au format conll¹ :

```
1 Il      il      PRON  2 nsubj
2 pleuvait pleuv    VERB  0 root
3 sur     sur     ADP   5 case
4 la     le     DET   5 det
5 route  route  NOUN  2 obl
```

On souhaite faire l'analyse de cette phrase avec l'algorithme de Eisner.

Pour cela, on définit les matrices `rightClose`, `leftClose`, `rightOpen` et `leftOpen`. La case `rightClose[g][d]` correspond à la case `E[g][d][1][0]` dans la page 53 du livre *Dependency Parsing*, `rightOpen[g][d]` correspond à `E[g][d][1][1]`, `leftClose[g][d]` correspond à `E[g][d][0][0]` et `leftOpen[g][d]` correspond à `E[g][d][0][1]`.

On considère que les poids des dépendances (ROOT, pleuvait)², (pleuvait, il), (pleuvait, route), (route, la) et (route, sur) ont pour poids 1 et que les autres dépendances ont pour poids 0.

1. Quel est le poids de A_1 ?
2. Donner le contenu des quatre matrices à l'issue de leur remplissage par l'algorithme³.

2 Graph-based Parsing par arbre couvrant

On souhaite analyser le groupe prépositionnel *sur la route* avec l'algorithme de Chu-Liu-Edmonds. Les poids des dépendances sont représentés dans la matrice ci-dessous, où les lignes correspondent aux gouverneurs et les colonnes aux dépendants (la dépendance (ROOT, sur), par exemple, vaut -2.).

	ROOT	sur	la	route
ROOT	0	-2	0	1
sur	0	0	0	2
la	0	0	0	0
route	0	2	1	0

1. Dessinez le graphe complet sur lequel se fera la recherche de l'arbre couvrant.
2. Dessinez le graphe aux différentes étapes de l'algorithme.

3 Transition-based Parsing

Soit S_2 la phrase *Il pleuvait fort sur la grand route* et son analyse A_2 au format conll.

1. On a indiqué dans l'analyse les étiquettes des dépendances, mais celles-ci ne seront pas utilisées pour l'exercice.
2. Dans la dépendance (ROOT, pleuvait), *ROOT* est le gouverneur et *pleuvait* est le dépendant.
3. On adoptera la convention suivante : en ligne le gouverneur et en colonne le dépendant.

1	Il	il	PRON	2	nsubj
2	pleuvait	pleuvoir	VERB	0	root
3	fort	fort	ADV	2	advmod
4	sur	sur	ADP	7	case
5	la	le	DET	7	det
6	grand	grand	ADJ	7	amod
7	route	route	NOUN	2	obl

Donner la séquence de configurations correspondant à A_2 ainsi que les transitions permettant de passer d'une configuration à la configuration suivante avec :

1. Un système de transitions arc eager
2. Un système de transitions arc standard

4 Cocke Younger Kasami

Soit la grammaire hors-contexte probabiliste suivante : $G = \langle \{S, GN, GV, GP, D, A, V, N, NP, P\}, \{Jean, commande, une, glace, à, la, fraise\}, P, S \rangle$ où P est l'ensemble de productions suivant :

1	S	→	GN GV	1	NP	→	Jean
0.4	GN	→	NP	1	V	→	commande
0.2	GN	→	D N	0.5	D	→	une
0.3	GN	→	D A N	0.5	N	→	glace
0.1	GN	→	GN GP	1	P	→	à
0.3	GV	→	V GN	0.5	D	→	la
0.2	GV	→	V	0.5	N	→	fraise
0.5	GV	→	GV GP				
1	GP	→	P GN				

1. Transformez G sous forme normale de Chomsky.
2. Expliquez comment vous avez traité les probabilités lors de la transformation.
3. Effectuez l'analyse de la phrase $S_3 = Jean commande une glace à la fraise$ à l'aide de l'algorithme CYK.
4. Utilisez la matrice CYK pour calculer la probabilité de la phrase S_3 .

5 Universal Dependencies

Donner au format conll deux analyses syntaxiques pour chacune de deux phrases suivantes :

- *Il a commandé une glace à la fraise.*
- *J'ai vu le chien et le chat de mon voisin.*