

M1 WIC – M2 DCISS – PMO-TALN

TD – Analyse syntaxique

Hervé Blanchon

1 Grammaire ATIS très simplifiée

On considère la grammaire hors contexte suivante dont S est l'axiome (le symbole de départ) :

S → NP VP	Det → that this a the
S → Aux NP VP	Noun → book flight meal man
S → VP	Verb → book include read
NP → Det NOM	Aux → does
NOM → Noun	
NOM → Noun NOM	
VP → Verb	
VP → Verb NP	

1.1 Analyse top-down

Écrire le(s) arbre(s) d'analyse top-down des énoncés suivants :

- the man read this book
- book that flight

1.2 Analyse bottom-up

Écrire le(s) arbre(s) d'analyse bottom-up des énoncés suivants :

- the man read this book
- book that flight

2 Grammaire ATIS simplifiée

On considère la grammaire hors contexte suivante dont S est l'axiome (le symbole de départ) :

S → NP VP	Nominal → Nominal PP
S → Aux NP VP	PP → Prep NP
S → VP	
VP → Verb	Det → that this a the
VP → Verb NP	Noun → book flight meal man
NP → Det Nominal	Verb → book include includes read
NP → ProperNoun	Prep → from to on
Nominal → Noun	ProperNoun → Houston TWA
Nominal → Noun Nominal	Aux → does

2.1 Analyse top-down

Écrire le(s) arbre(s) d'analyse top-down des énoncés suivants :

- the flight from Houston includes a meal
- does the flight from Houston includes a meal

2.2 Analyse bottom-up

Écrire le(s) arbre(s) d'analyse bottom-up des énoncés suivants :

- the flight from Houston includes a meal
- does the flight from Houston includes a meal

3 Autour du verbe « voir »

On considère la grammaire hors contexte suivante dont S est l'axiome (le symbole de départ) :

S → NP VP	DT → the
NP → DT NN	NN → man dog cat park
NP → NP PP	VB → saw
PP → IN NP	IN → in with under
VP → VB NP	

3.1 Nombre d'analyses pour un énoncé

3.1.1 Avec cette grammaire, la phrase « the man saw the dog in the park » a combien d'analyses ?

3.1.2 Avec cette grammaire, la phrase « the man saw the dog in the park with the cat » a combien d'analyses ?

3.2 Prédiction du nombre d'analyses

Soit C_n définit comme le n -ième nombre de Catalan définit comme suit¹ :

$$C_n = \frac{(2n)!}{(n+1)!n!}$$

On pourrait montrer que C_n est le nombre d'arbres binaires qui possèdent $n + 1$ feuilles.

Considérons maintenant une phrase qui est grammaticale sous la grammaire hors contexte ci-dessus, et qui a exactement k prépositions suivant le verbe, et 0 préposition avant le verbe (une préposition est tout mot avec l'étiquette IN). Combien d'arbres d'analyse cette phrase aura-t-elle ?

4 Le singulier et le pluriel

On considère la grammaire hors contexte suivante dont S est l'axiome (le symbole de départ) :

S → NP VP	DT → the
NP → DT NN	NN → man dog cat park
NP → DT NNS	NNS → dogs cats parks
NP → NP PP	VB → see sees
PP → IN NP	IN → in with
VP → VB NP	
VP → VP PP	

Cette grammaire permet de produire des énoncés non grammaticaux, tels que :

- the dog see the cat
- the dog in the park see the cat
- the dog in the park see the cat in the park the dogs sees the cat
- the dogs in the park sees the cat
- the dogs in the park sees the cat in the park

4.1 Modification de la grammaire pour qu'elle ne produise que des énoncés corrects

Comment modifieriez-vous la grammaire pour que toutes les phrases qu'elle génère soient grammaticales ?

Indice : il convient de mieux gérer le singulier et le pluriel.

¹ Voir Wikipedia pour plus d'information

5 Autour du verbe « voir » en cascade

On considère la grammaire hors contexte suivante dont S est l'axiome (le symbole de départ) :

S → NP VP	VP-GAP → Vdt NP
NP → DT NN	WH → that
VP → Vt DT	DT → the
VP → Vdt NP NP	NN → man
NP → NP RELC	NN → dog
RELC → WH S-GAP	NN → cat
S-GAP → VP	NN → park
S-GAP → NP VP-GAP	Vt → saw
VP-GAP → Vt	Vdt → gave

5.1 Arbres d'analyse sous la grammaire proposée

- 5.1.1 Donner les arbres d'analyse pour les énoncés « the man that saw the dog saw the cat » et « the man that the cat saw saw the dog »
- 5.1.2 Donner une phrase qui soit grammaticale sous cette grammaire et qui contienne le trigramme « saw saw saw »

5.2 Modification de la grammaire

Supposons que nous ajoutons les règles suivantes à la grammaire, de sorte que l'énoncé « the man said the cat saw the dog » puisse être analysée correctement :

VP → V3 S
V3 → said

Quelle(s) règle(s) ajouteriez-vous à la grammaire pour que la phrase suivante puisse être analysée ? :
« the dog that the man said the cat saw saw the park »

6 Importation en Suisse

On considère la grammaire hors contexte partielle suivante dont S est l'axiome (le symbole de départ) :

S → NP VP	N → 2012 Switzerland USA exports
NP → Det N	V → exports
NP → N	VBP → are
NP → NP PP	VBG → increasing
VP → V	Det → the
VP → VP PP	P → to from in
VP → VBP VDG PP	
PP → P NP	

6.1 Vers la forme normale de Chomsky

Transformer la grammaire pour la mettre sous forme normale de Chomsky.

6.2 Analyse avec l'algorithme CYK

En utilisant l'algorithme CYK, analysez l'énoncé suivant en utilisant la grammaire sous forme normale de Chomsky que vous aurez proposée :

the exports from the USA to Switzerland are increasing in 2012

Fournir à la fois la structure de données complète, entièrement remplie, utilisée par l'algorithme, ainsi que le résultat de l'analyse sous la forme d'un ou de plusieurs arbres.

7 Time flies like an arrow

On considère la grammaire hors contexte suivante dont S est l'axiome (le symbole de départ) :

$S \rightarrow NP VP$	$NP \rightarrow \text{time}$
$S \rightarrow Vst NP$	$Vst \rightarrow \text{time}$
$S \rightarrow S PP$	$NP \rightarrow \text{flies}$
$VP \rightarrow V NP$	$VP \rightarrow \text{flies}$
$VP \rightarrow VP PP$	$P \rightarrow \text{like}$
$NP \rightarrow \text{Det N}$	$V \rightarrow \text{like}$
$NP \rightarrow NP PP$	$\text{Det} \rightarrow \text{an}$
$NP \rightarrow NP NP$	$N \rightarrow \text{arrow}$
$PP \rightarrow P NP$	

7.1 Analyse avec l'algorithme CYK

En utilisant l'algorithme CYK, analysez l'énoncé suivant en utilisant la grammaire proposée qui est déjà sous forme normale de Chomsky :

time flies like an arrow

Fournir à la fois la structure de données complète, entièrement remplie, utilisée par l'algorithme, ainsi que le résultat de l'analyse sous la forme d'un ou de plusieurs arbres.

8 Sources

Les exercices 1 et 7 sont inspirés du cours CS 585 de Andrew McCallum, University of Massachusetts Amherst

Les exercices 3, 4 et 5 sont inspirés du cours COMS 4705 de Michael Collins, Columbia University

L'exercice 6 est inspiré du cours Introduction to NLP de J.-C. Chappelier et Martin Rajman, École Polytechnique Fédérale de Lausanne