NGUYEN Thi Minh Tuyen

Théorie des automates et langages formels

Décidabilité

Exemples de problèmes indécidables

- Est-ce que deux grammaires non contextuelles sont équivalentes (définissent le même langage)?
- Est-ce qu'une grammaire non contextuelle est ambigu?
- Si une grammaire non contextuelle est ambigu, est-ce qu'il existe une autre grammaire, qui défini un même langage, n'est pas ambigu?
- Est-ce que le complément d'un langage non contextuel est un langage non contextuel?
- Est-ce que l'intersection de deux langages non contextuels est un langage non contextuel?
- Est-ce que l'intersection de deux langages non contextuels est vide?
- Est-ce qu'il existe un mot qui n'est pas dans le langage engendré par une grammaire non contextuelle?

Problèmes décidables

1. Est-ce que le langage engendré par une grammaire non contextuelle est vide?
2. Est-ce que le langage engendré par une grammaire non contextuelle est fini ou infini?
3. Est-ce qu'un mot **w** appartient au langage engendré par une grammaire non contextuelle?

Problèmes décidables

1. Est-ce que le langage engendré par une grammaire non contextuelle est vide?
2. Est-ce que le langage engendré par une grammaire non contextuelle est fini ou infini?
3. Est-ce qu'un mot **w** appartient au langage engendré par une grammaire non contextuelle?

Théorème 1

Déterminer si le langage engendré par une grammaire non contextuelle est vide est un problème décidable.

Preuve:

- $S \Rightarrow^* \Lambda$?

Sinon, on transforme la grammaire en une grammaire sous la forme normal de Chomsky.

- une production de la forme: S → terminal?
- Sinon, on répète:

1. On choisit une variable N tel que: N → t (suite de terminaux). On remplace N à droite dans une production par t. On enlève les autres productions pour N.
2. On arrête s'il y a une production S → t ou s'il n'y a plus de variables à choisir.

Exemple

$S \rightarrow XY$
$X \rightarrow AX$
$X \rightarrow AA$
$A \rightarrow a$
$Y \rightarrow BY$
$Y \rightarrow BB$
$B \rightarrow b$

Étape 1: Remplace A par a et B par b:

$S \rightarrow XY$
$X \rightarrow aX$
$X \rightarrow aa$
$Y \rightarrow bY$
$Y \rightarrow bb$

Étape 1: Remplace X par aa et Y par bb:

$S \rightarrow aabb$

Étape 1: Remplace S par aabb.

Problèmes décidables

1. Est-ce que le langage engendré par une grammaire non contextuelle est vide?
2. Est-ce que le langage engendré par une grammaire non contextuelle est fini ou infini?
3. Est-ce qu'un mot **w** appartient au langage engendré par une grammaire non contextuelle?

Algorithme 2.1

Cet algorithme sert à déterminer si une variable X peut générer une suite de terminaux.

- On échange S et X dans toutes les productions de la grammaire et on utilise l'algorithme de Théorème 1.

Si X ne peut pas générer une suite de terminaux, on dira que X est non productif.

Théorème 2.1

Il existe un algorithme qui peut déterminer si une variable **X** peut être utilisé dans une dérivation.

Démonstration du Théorème 2.1

- On détermine si X est non productif.
- Sinon,
  1. On trouve toutes les variables non productives.
  2. On enlève toutes les productions qui contiennent des variables non productives.
  3. On peint X en bleu (à droite et à gauche dans toute production).
  4. On peint en bleu toute variable à gauche tel que la production contient une variable bleue à droite. On peint en bleu toutes ces variables à droite aussi.
  5. On répète l'étape 4 tant que cette étape crée une nouvelle variable bleue.
  6. Si S est bleu, il existe une dérivation qui contient X.

Si X ne peut pas être utilisé dans une dérivation, X est inutile.

Exemple

$S \rightarrow ABa \mid bAZ \mid b$
$A \rightarrow Xb \mid bZa$
$B \rightarrow bAA$
$X \rightarrow aZa \mid aaa$
$Z \rightarrow ZAbA$

$S \rightarrow ABa \mid \cancel{bAZ} \mid b$
$A \rightarrow Xb \mid \cancel{bZa}$
$B \rightarrow bAA$
$X \rightarrow \cancel{aZa} \mid aaa$
$\cancel{Z \rightarrow ZAbA}$

X peut générer une suite de terminaux
Z est non productif

$S \rightarrow ABa \mid \cancel{bAZ} \mid b$
$A \rightarrow Xb \mid \cancel{bZa}$
$B \rightarrow bAA$
$X \rightarrow \cancel{aZa} \mid aaa$
$\cancel{Z \rightarrow ZAbA}$

$S \rightarrow ABa \mid \cancel{bAZ} \mid b$
$A \rightarrow Xb \mid \cancel{bZa}$
$B \rightarrow bAA$
$X \rightarrow \cancel{aZa} \mid aaa$
$\cancel{Z \rightarrow ZAbA}$

$S \rightarrow ABa \mid \cancel{bAZ} \mid b$
$A \rightarrow Xb \mid \cancel{bZa}$
$B \rightarrow bAA$
$X \rightarrow \cancel{aZa} \mid aaa$
$\cancel{Z \rightarrow ZAbA}$

$S \rightarrow ABa \mid \cancel{bAZ} \mid b$
$A \rightarrow Xb \mid \cancel{bZa}$
$B \rightarrow bAA$
$X \rightarrow \cancel{aZa} \mid aaa$
$\cancel{Z \rightarrow ZAbA}$

Il existe une dérivation qui contient X

Algorithme 2.2

Cet algorithme sert à déterminer si une variable X peut apparaître deux fois sur la même branche dans une dérivation.

- On remplace les X à gauche par Ж.
- On peint X en bleu et on fait l'algorithme de la peinture bleue (étapes 3 à 5 de l'algorithme dans slide 10).
- Si Ж est bleu, X peut apparaître 2 fois sur la même branche dans une dérivation.

Théorème 2

Il existe un algorithme qui peut déterminer si le langage engendré par une grammaire non contextuelle est fini ou infini.

1. On trouve toutes les variables inutiles et on enlève les productions qui les contiennent.
2. On cherche une variable qui n'est pas inutile et qui peut apparaître deux fois sur la même branche (s'il existe une telle variable).
3. Si à l'étape 2 on trouve une telle variable, alors le langage est infini, sinon il est fini.

Exemple

S → ABa I bAZ I b
A → Xb I bZa
B → bAA
X → aZa I bA I aaa
Z → ZAbA

la variable Z est inutile

Le langage engendré par une grammaire non contextuelle est infini.

S → ABa I b
A → Xb
B → bAA
X → bA I aaa

S → ABa I b
A → Xb
B → bAA
Ж → bA I aaa

S → ABa I b
A → Xb
B → bAA
Ж → bA I aaa

S → ABa I b
A → Xb
B → bAA
Ж→ bA I aaa

S → ABa I b
A → Xb
B → bAA
Ж → bA I aaa

S → ABa I b
A → Xb
B → bAA
Ж →bA I aaa

Problèmes décidables

1. Est-ce que le langage engendré par une grammaire non contextuelle est vide?
2. Est-ce que le langage engendré par une grammaire non contextuelle est fini ou infini?
3. Est-ce qu'un mot **w** appartient au langage engendré par une grammaire non contextuelle?

Problème 3

Étant données une grammaire non contextuelle G et un mot $w = w_1w_2 \ldots w_n$. Est-ce que G génère le mot w?

Algorithme CYK (Cocke, Kasami(1965),Younger(1967))

- On transforme **G** en Forme Canonique de Chomski.
- A chaque étape i de cet algorithme on cherche tous les sous mots de $w=w_1w_2 \ldots w_n$ de longueur i qu'on peut générer avec cette grammaire.

Exemple

S → XY
X → XA
Y → AY
A → a
X → a l b
Y → a

w = baaaa

S	X	Y	A
	a b	a	a
aa ba	aa ba	aa	
aaa baa	aaa baa	aaa	
aaaa baaaa	baaa		
aaaaa baaaa			

Question?