

Parsing-Repetition



Übersicht

- ◆ Kurze Wiederholung von »Parsing-Einführung«

Linguistische Motivation

Warum sollen Grammatiken in einer formalen Notation festgehalten werden?

- ◆ Bei einer wissenschaftlichen Theorie soll klar sein, durch welche Beobachtungen sie widerlegt (*falsifiziert*) werden könnte.
- ◆ Beobachtbar bei einer Grammatik:
 - ◆ Ein korrekter Satz entspricht nicht der Grammatik
 - ◆ Ein falscher Satz entspricht der Grammatik→ klares, nachprüfbares Falsifikations-Kriterium

Parser/Akzeptor

- ◆ Akzeptor: Prüft maschinell, ob ein Satz der Grammatik entspricht
- ◆ Parser: Wie Akzeptor, gibt zusätzlich erkannte Struktur aus

Kontextfreie Grammatiken (CFG)

Kontextfreie Grammatiken (*Context-Free Grammars, CFG*) sind ein solcher Formalismus

- ◆ gut erforschte mathematische Eigenschaften
- ◆ Nachteile:
 - ◆ eine exotische Konstruktion des Zürichdeutschen sowie gewisse Ausdrücke von Bambara können mit CFG *nicht* beschrieben werden
 - ◆ äusserst umständlich (aber nicht unmöglich!), etwas komplexere sprachliche Phänomene mit CFG zu modellieren
- ◆ Es gibt andere, besser geeignete Formalismen als CFG
 - ◆ Lexikalisch-Funktionale Grammatik (LFG), Generalisierte Phrasenstruktur-Grammatik (GPSG), Kopf-getriebene (head-driven) Phrasenstruktur-Grammatik (HPSG), Kategorialgrammatik (CG), Baumadjunktionsgrammatik (TAG), etc. etc.

Verarbeitung vs. Formalismus



Mit einem Grammatikformalismus wird linguistisches Wissen festgehalten (»deklariert«)

- ◆ unabhängig davon, wie es verarbeitet wird
- ◆ ein gut entworfener Formalismus kann mit unterschiedlichen Verfahren verarbeitet werden
- ◆ »deklarativ«

Parser/Akzeptoren

- ◆ realisieren ein Verfahren, das mit einem solchen Formalismus arbeitet

Top-Down-Parsing

Ein Top-Down-Parser für eine kontextfreie Grammatik

- ◆ fängt mit dem Startsymbol an
- ◆ führt wiederholt Ableitungsschritte durch
- ◆ Ziel: Ableiten der zu analysierenden Kette

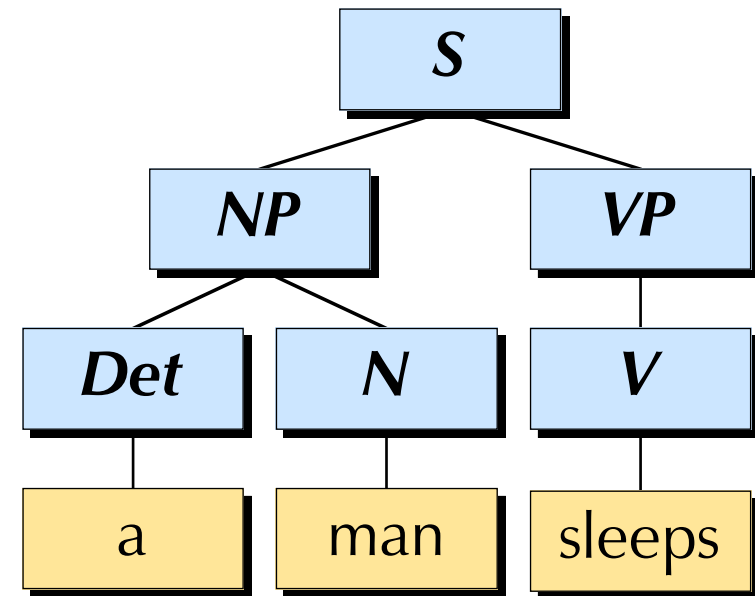
$S \Rightarrow NP VP \Rightarrow Det N VP \Rightarrow a N VP \Rightarrow a man VP \Rightarrow a man V \Rightarrow a man sleeps$

$S \rightarrow NP VP$	$Det \rightarrow the$	$V \rightarrow loves$
$NP \rightarrow Det N$	$Det \rightarrow a$	$V \rightarrow sleeps$
$VP \rightarrow V NP$	$N \rightarrow man$	$V \rightarrow sees$
$VP \rightarrow V$	$N \rightarrow woman$	$V \rightarrow thinks$

Top-Down-Parsing

Vorgehen beim Top-Down-Parsing:

- ◆ Ich suche ein S
- ◆ Um ein S zu erhalten, brauche ich eine NP und eine VP
- ◆ Um eine NP zu erhalten, brauche ich ein Det und ein N
- ◆ Um ein Det zu erhalten, kann ich das Wort »a« verwenden — gefunden
- ◆ Um ein N zu erhalten, kann ich das Wort »man« verwenden — gefunden
- ◆ Damit ist die NP vollständig
- ◆ Um eine VP zu erhalten, brauche ich ein V
- ◆ Um ein V zu erhalten, kann ich das Wort »sleeps« verwenden — gefunden
- ◆ Damit ist die VP vollständig
- ◆ Damit ist das S vollständig



$S \rightarrow NP VP$ $Det \rightarrow the$ $V \rightarrow loves$
 $NP \rightarrow Det N$ $Det \rightarrow a$ $V \rightarrow sleeps$
 $VP \rightarrow V NP$ $N \rightarrow man$ $V \rightarrow sees$
 $VP \rightarrow V$ $N \rightarrow woman$ $V \rightarrow thinks$

Bottom-Up-Parsing

Ein Bottom-Up-Parser für eine kontextfreie Grammatik

- ◆ fängt mit der zu analysierenden Kette an
- ◆ führt wiederholt Ableitungsschritte »rückwärts« durch
- ◆ Ziel: Erreichen des Startsymbols



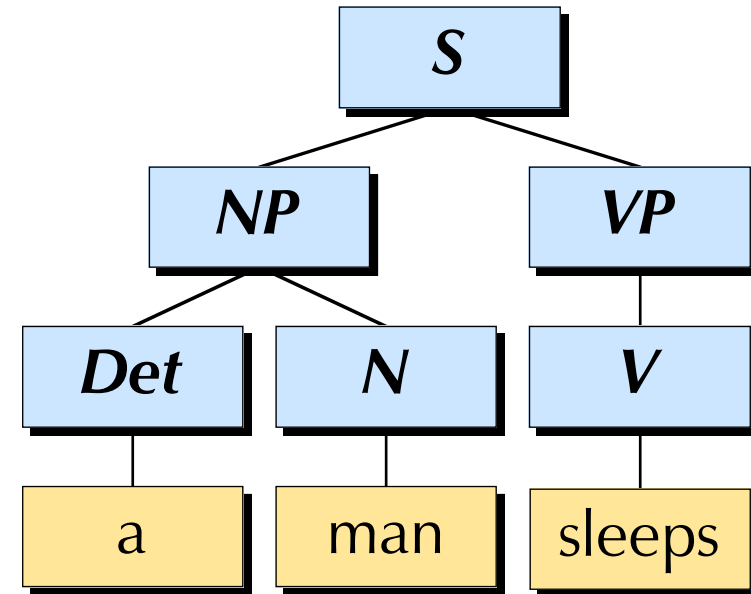
$S \Rightarrow NP VP \Rightarrow NP V \Rightarrow NP \text{ sleeps} \Rightarrow Det N \text{ sleeps} \Rightarrow Det \text{ man sleeps} \Rightarrow a \text{ man sleeps}$

$S \rightarrow NP VP$ $Det \rightarrow the$ $V \rightarrow loves$
 $NP \rightarrow Det N$ $Det \rightarrow a$ $V \rightarrow sleeps$
 $VP \rightarrow V NP$ $N \rightarrow man$ $V \rightarrow sees$
 $VP \rightarrow V$ $N \rightarrow woman$ $V \rightarrow thinks$

Bottom-Up-Parsing

Vorgehen beim Bottom-Up-Parsing:

- ◆ Nimm ein Wort — es ist »a«
- ◆ »a« ist ein Det
- ◆ Nimm ein weiteres Wort — es ist »man«
- ◆ »man« ist ein N
- ◆ Det und N bilden zusammen eine NP
- ◆ Nimm ein weiteres Wort — es ist »sleeps«
- ◆ »sleeps« ist ein V
- ◆ V bildet (für sich alleine) eine VP
- ◆ NP und VP bilden zusammen ein S



$S \rightarrow NP VP$ $Det \rightarrow the$ $V \rightarrow loves$
 $NP \rightarrow Det N$ $Det \rightarrow a$ $V \rightarrow sleeps$
 $VP \rightarrow V NP$ $N \rightarrow man$ $V \rightarrow sees$
 $VP \rightarrow V$ $N \rightarrow woman$ $V \rightarrow thinks$