

Syntax natürlicher Sprachen

1: Syntaxanalyse und syntaktische Tests

A. Wisiorek

Centrum für Informations- und Sprachverarbeitung,
Ludwig-Maximilians-Universität München

14.10.2025

1. Parsing als Syntaxanalyse

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten
- 3 Konstituententests
 - Substitutionstest
 - Permutationstest
 - Eliminierungstest
 - Koordinationstest
- 4 Komplement/Adjunkt-Tests

Aufgaben eines Parsingalgorithmus

- 1 **Überprüfung** der grammatischen Struktur einer Eingabe auf Wohlgeformtheit
→ ***Suche einer Ableitung** gemäß der Regeln einer formalen Grammatik*
- 2 **Wiedergabe** der gefundenen grammatischen Struktur
→ ***Ableitungsbaum** (auch: Parsebaum, Syntaxbaum)*

Verarbeitung formaler Grammatiken

- Formale Grammatik = Syntaktisches Strukturmodell
→ *dessen Regeln sind aber nicht mehr als **Sammlung von Strings***
- **Verfahren** notwendig, um zu entscheiden, ob eine Eingabe gemäß einer gegebenen formalen Grammatik **wohlgeformt** ist

Parsing-Algorithmen

- Verfahren zur **Verarbeitung** von formalen Grammatiken zur **Strukturerkennung und -Analyse** der **Eingabe** (Satz als Tokensequenz)

Strukturerkennung

- **Überprüfung** der grammatischen Struktur einer Eingabe als Suche einer Ableitung aus den Regeln einer formalen Grammatik (ob Satz in formaler Sprache enthalten ist)

Strukturzuweisung

- gleichzeitig **Wiedergabe** der in der Suche aufgebauten grammatischen Struktur der Eingabe (Syntaxbaum)

Haupttypen

- 1 **CFGs** (kontextfreie Grammatiken)
 - Konstituentenstruktur (Phrasenstrukturgrammatik)
 - Strukturinformationen in den Knoten des Syntaxbaums
- 2 **Dependenzgrammatiken**
 - Abhängigkeitsrelationen zwischen Wörtern
 - Strukturinformationen in den Kanten des Syntaxbaums (Relationslabel)
- 3 **Chunk-Parser** (u.a. reguläre Grammatiken)
 - partielle Syntaxanalyse (nur wichtigste Konstituenten: NPs, VPs, PPs)
 - flache, nicht-hierarchische Strukturanalyse

Erweiterungen von CFGs

- 1 **PCFGs** = Probabilistische CFGs
- 2 **FCFGs** = Feature-basierte CFGs

Voraussetzungsschritte für automatische Syntaxanalyse

- Sentence Segmentation
- Tokenisierung
- Part-of-Speech-Tagging
- Stemming
- morphologisches Parsing (Kasus, Agreement)

Mögliche Folgeanwendungen

- **Entity Extraction**
- **Relation Extraction**
- **Semantic Parsing** (Natural Language Understanding)

2. Analyse syntaktischer Einheiten

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten**
- 3 Konstituententests
 - Substitutionstest
 - Permutationstest
 - Eliminierungstest
 - Koordinationstest
- 4 Komplement/Adjunkt-Tests

Segmentierung: Feststellung syntaktischer Einheiten

- **Zerlegung** Satz in **Konstituenten**
→ *Konstituenten = sprachliche Einheiten, die Teile einer größeren Einheit sind (insbes. Wörter, Phrasen oder Teilsätze)*
- Analyse **syntagmatischer** Beziehung zwischen Einheiten
- zusammengehörende Elemente feststellbar über **Konstituententests**

Klassifizierung: Feststellung syntaktischer Kategorien

- **syntaktische Kategorien = Mengen syntaktischer Einheiten** mit gleichen Eigenschaften (**Klassen**)
- morphologische, syntaktische und semantische **Klassifizierungskriterien**
 - syntaktisches Kriterium: **Austauschbarkeit im gleichen Kontext**
- Analyse **paradigmatischer** Beziehung zwischen Einheiten

Syntagmatische vs. paradigmatische Dimension

syntagmatisch

lineare Kombinierbarkeit

der { große
kleine } Hund jagt { die Katze.
Ferdinand. }

Ein großer { Hund
*Ferdinand } jagt die Katze.

{ Sie
die sich im Gehen nach dem Hund umschauende Frau } stürzt.

Er { sieht einen Hund
geht } auf dem Weg. Er { sieht
*geht } einen Hund.

paradigmatisch

vertikale Austauschbarkeit

3. Konstituententests

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten
- 3 Konstituententests**
 - Substitutionstest
 - Permutationstest
 - Eliminierungstest
 - Koordinationstest
- 4 Komplement/Adjunkt-Tests

Konstituententests

- zur Feststellung syntagmatischer Einheiten (Konstituenten)
- Konstituenten-Hierarchie: **unmittelbare vs. mittelbare Konstituenten**

Hauptkriterien

Ersetzbarkeit und Verschiebbarkeit unter Erhalt der **Grammatikalität**

3.1. Substitutionstest

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten
- 3 Konstituententests
 - **Substitutionstest**
 - Permutationstest
 - Eliminierungstest
 - Koordinationstest
- 4 Komplement/Adjunkt-Tests

auch: Ersatzprobe

- **eine unter Erhalt der Grammatikalität des Satzes durch eine andere Wortfolge ersetzbare Wortfolge ist Konstituente**
- **Feststellung von Einheiten** (Austauschbarkeit im gleichen Kontext; allgemeiner: auch für Wortarten, Flexionsparadigmen)
- **Pronominalisierung**
→ *Eine durch Pronomen ersetzbare Wortfolge ist Konstituente*
- **für Feststellung Verb mit Erweiterungen (komplexe VP) als Konstituente**
→ *Ersatz komplexer VP durch (intransitives) Verb*

Der Junge verkauft die Äpfel des Bauern.

Identifizierung der Wortfolgen *der Junge* und *die Äpfel des Bauern* als Konstituenten durch **Pronominalisierung**:

Er verkauft sie.

Identifizierung der Wortfolge *verkauft die Äpfel des Bauern* als **komplexe VP-Konstituente** durch Ersatz mit intransitivem Verb:

Der Junge arbeitet.

3.2. Permutationstest

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten
- 3 Konstituententests
 - Substitutionstest
 - **Permutationstest**
 - Eliminierungstest
 - Koordinationstest
- 4 Komplement/Adjunkt-Tests

auch: Verschiebeprobe

- eine im Satz ohne Zerstörung der Grammatikalität verschiebbare Wortfolge ist Konstituente
- im Deutschen: Verschiebung vor finites Verb (ins sog. **Vorfeld**)
→ **Topikalisierung**: rhetorisch-pragmatische Funktion
- Verwendung zum **Testen von Wortstellungsmöglichkeiten**

auch: Auflösung von Ambiguität

- *Der Junge beobachtete das Mädchen mit dem Fernglas.* (ambig)
- *Das Mädchen mit dem Fernglas beobachtete der Junge.* (NP-Attr.)
- *Mit dem Fernglas beobachtete der Junge das Mädchen.* (VP-Adj.)

Der Junge verkauft die Äpfel des Bauern.

Identifizierung der Wortfolgen *der Junge* und *die Äpfel des Bauern* als Konstituenten durch **Permutation**:

Die Äpfel des Bauern verkauft der Junge.

3.3. Eliminierungstest

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten
- 3 Konstituententests
 - Substitutionstest
 - Permutationstest
 - **Eliminierungstest**
 - Koordinationstest
- 4 Komplement/Adjunkt-Tests

auch: Weglassprobe

- **eine ohne Zerstörung der Grammatikalität eines Satzes weglassbare Wortfolge ist Konstituente**
- Feststellung syntaktisch weglassbarer bzw. optionaler Einheiten
 - (s.u. Eliminierungstest als Adjunkt-Test)

Der Junge verkauft die Äpfel des Bauern.

Identifizierung der Wortfolge *des Bauern* als **attributive Konstituente** (nominales Adjunkt) durch Eliminierung:

Der Junge verkauft die Äpfel.

3.4. Koordinationstest

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten
- 3 Konstituententests
 - Substitutionstest
 - Permutationstest
 - Eliminierungstest
 - **Koordinationstest**
- 4 Komplement/Adjunkt-Tests

Koordination: Verbindung mit *und* / *aber*

- eine mit einer anderen Wortfolge unter Erhalt der Grammatikalität des Satzes koordinierbare Wortfolge ist Konstituente
- geeignet für die **Ermittlung von Phrasenteilen** (Attributen usw.)
- Analyse der Struktur von **komplexen Konstituenten**

Konstituenten müssen vom gleichen Typ sein

- *Er schrieb einen Brief und eine Karte.*
- *Er schrieb an dich und an mich.*
- **aber nicht:** **Er schrieb einen Brief und an mich.*

Anwendung Koordinationstest

Der Junge verkauft die Äpfel des Bauern.

Identifizierung der Wortfolge *des Bauern* als **Konstituente** durch Koordination:

Der Junge verkauft die Äpfel des Bauern und der Bäuerin

Konstituentenstruktur als Ergebnis aller Konstituententests

Der Junge verkauft die Äpfel des Bauern

Der Junge verkauft die Äpfel des Bauern

verkauft die Äpfel des Bauern

die Äpfel des Bauern

4. Komplement/Adjunkt-Tests

- 1 Parsing als Syntaxanalyse
- 2 Analyse syntaktischer Einheiten
- 3 Konstituententests
 - Substitutionstest
 - Permutationstest
 - Eliminierungstest
 - Koordinationstest
- 4 Komplement/Adjunkt-Tests

- neben den Konstituententests gibt es weitere syntaktische Tests, die der Unterscheidung **notwendiger von nicht-notwendigen** Satzgliedern dienen (**Komplement vs. Adjunkt**)
 - als **Satzglieder** werden in der Germanistik die im Satz verschiebbaren (permutierbaren) Konstituenten mit zentraler syntaktischer Funktion (Subjekt, Objekt, Prädikat, Adverbial) bezeichnet, in die ein Satz unmittelbar zerlegt werden kann
- **Komplement:**
 - notwendiges Satzglied
 - Auftreten ist von Kopf gefordert (valenzgefordert)
 - z.B. Subjekt, Objekt als Verb-Komplemente: *die Katze sieht einen Hund*, nicht: **sieht einen Hund*
- **Adjunkt:**
 - nicht-notwendiges Satzglied
 - Auftreten ist nicht von Kopf gefordert
 - z.B. Adjektiv-Attribut als Adjunkt des Nomens: *das rote Auto*
 - z.B. PP-Adverbial als Adjunkt des Verbs: *er läuft im Park*

- **Auslagerung des Satzglieds in einen Satz mit dem Verb *geschehen*.**
 - wenn Grammatikalität erhalten \Rightarrow Adjunkt
 - wenn Grammatikalität *nicht* erhalten \Rightarrow Komplement

Er wartete im Park auf die Katze.

*Er wartete auf die Katze, und das geschah **im Park**.* (Adjunkt)

Er wartete im Park, und das geschah **auf die Katze.* (Komplement)

- **Auslagerung des Satzglieds in einen Adverbialsatz** (funktioniert nicht bei Zeitangaben).
 - wenn Grammatikalität erhalten \Rightarrow Adjunkt
 - wenn Grammatikalität *nicht* erhalten \Rightarrow Komplement (oder temporales Adjunkt! = Zeitangabe)

Er wartete im Park auf die Katze.

*Er wartete auf die Katze, als er **im Park** war.* (Adjunkt)

Er wartete im Park, als er **auf die Katze war.* (Komplement)

Er wartete auf die Katze, als er **bis Mitternacht war.* (temporales Adjunkt)

- **Eliminierung des Satzglieds** (Feststellung ihrer Weglassbarkeit).
 - wenn Grammatikalität erhalten \Rightarrow Adjunkt (oder fakultatives Komplement!, s.u.)
 - wenn Grammatikalität *nicht* erhalten \Rightarrow Komplement

Er beantwortet einen Brief im Arbeitszimmer.

Er beantwortet einen Brief ~~im Arbeitszimmer~~. (Adjunkt)

**Er beantwortet ~~einen Brief~~ im Arbeitszimmer.* (Komplement)

Er schreibt ~~einen Brief~~ im Arbeitszimmer. (fakultatives Komplement)

Obligatorische, fakultative und optionale Satzglieder

- **Obligatorisches Satzglied = Komplement** (valenzgefordert):
**Er beantwortet einen Brief*
- **Fakultatives Satzglied = Komplement, aber kontextabhängig eliminierbar:**
Er schreibt einen Brief
→ Ellipse eines an sich obligatorischen Satzglieds
→ oder Annahme von zwei Valenzrahmen
- **Optionales Satzglied = Adjunkt** (immer eliminierbar):

Er schreibt den ganzen Tag

- **Differenzierung fakultativer von optionalen Satzgliedern:**
beide: eliminierbar → Differenzierung über *geschehens*-Test:

Er schreibt, und es geschieht **einen Brief* (fakultatives Komplement)

*Er schreibt, und es geschieht **den ganzen Tag*** (Adjunkt, optionales Satzglied)