

Syntax natürlicher Sprachen

Tutorium

Partielles Parsen

Sarah Anna Uffelmann

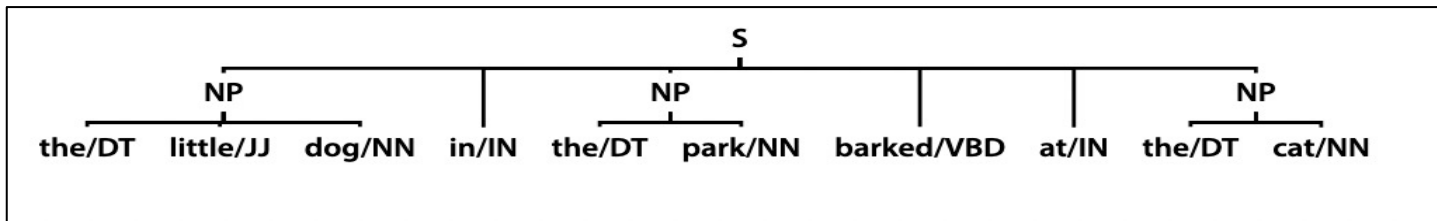
26.01.2024

Chunking

Für viele Anwendungen (z.B. im Bereich Information Extraction oder Information Retrieval) ist keine vollständige Syntaxanalyse notwendig, so dass **partielles Parsing** **ausreichend** ist:

Statt einen Satz komplett zu analysieren, wird dabei **nur nach bestimmten Konstituenten gesucht** --- **NPs, VPs, PPs, (AdJPs)**

Beim Parsing werden dann nur diejenigen Wörter berücksichtigt, die Teil der gesuchten Konstituenten sind.



Suchen wir beispielsweise nach NPs, erhalten wir nur NP-Chunks

Chunking

Beim Chunking erhalten wir **flache, nicht-hierarchische Analysen**:
Chunk-Bäume haben die Tiefe 1

Bsp: „**Das Huhn mit dem weißen Gefieder** ist im Stall.“

Wenn wir nach NPs suchen, erhalten wir:

- „das Huhn“
- „dem weißen Gefieder“
- „Stall“

aber nicht: „das Huhn mit dem weißen Gefieder“ als komplexe NP

-> **Komplexe Phrasen werden nicht erkannt!**

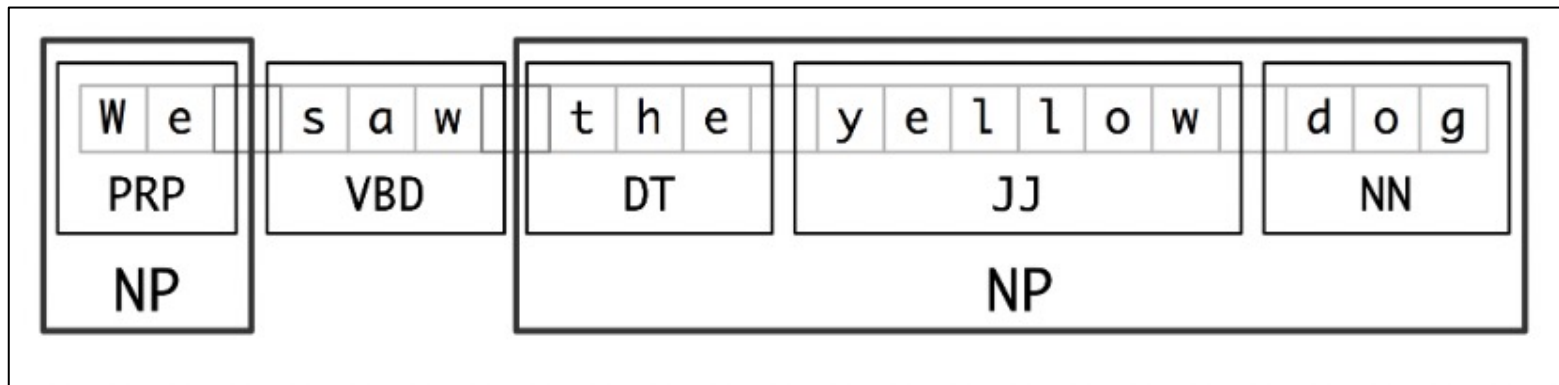
Aber: Ein **kaskadierender Chunk-Parser** kann durch Kombination einfacher Chunk-Parser auch **komplexe Einheiten** finden.

Chunking

Was leistet ein Chunk-Parser?

Segmentierung: Er findet die Grenzen von Chunks in einer Folge von Tokens

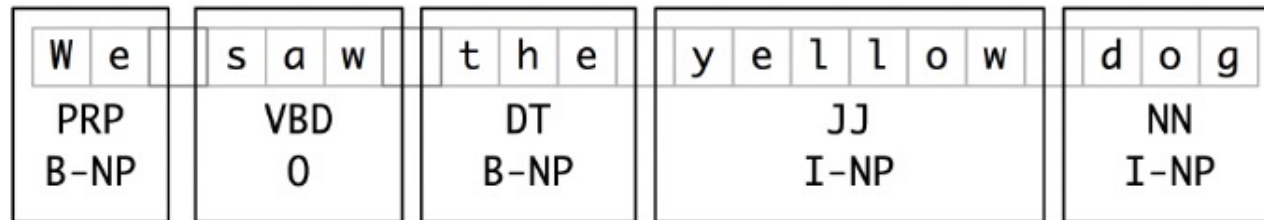
Tagging: Er labelt die gefundenen Chunks



IOB-Format

IOB = Inside – Outside – Beginning

Ein IOB-Tag repräsentiert gleichzeitig die Segmentierung und das Label.



Wir benötigen **kein Ende-Tag**, denn das Ende einer Einheit ist implizit kodiert:

- wenn ein O auf ein B oder ein I folgt; eine Einheit endet:

we ----- saw

B-NP O

- wenn ein B auf ein B oder ein I folgt; eine Einheit endet und gleichzeitig beginnt eine neue:

we ----- gave ----- him ----- the ----- letter

B-NP O B-NP B-NP I-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. **Markiere dabei alle NPs.**

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .

B-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .

B-NP

I-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .

B-NP

I-NP

O

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .
B-NP I-NP O B-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .
B-NP I-NP O B-NP I-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .
B-NP I-NP O B-NP I-NP O

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .
B-NP I-NP O B-NP I-NP O B-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .
B-NP I-NP O B-NP I-NP O B-NP I-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema. Markiere dabei alle NPs.

Das ----- Mädchen ----- sieht ----- das ----- Huhn ----- mit ----- dem ----- Fernglas ----- .

B-NP

I-NP

O

B-NP

I-NP

O

B-NP

I-NP

O

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.

Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch

B-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch

B-NP B-VP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- .

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- .

B-PP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- .
B-PP B-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- .
B-PP B-NP O

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

Alternative:

im ----- Park ----- , ----- denn ----- er ----- hatte ----- ihn ----- gesucht ----- .
B-PP B-NP O

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- , ----- denn ----- er ----- hatte ----- ihn ----- gesucht ----- .
B-PP B-NP O O

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- , ----- denn ----- er ----- hatte ----- ihn ----- gesucht ----- .
B-PP B-NP O O B-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- , ----- denn ----- er ----- hatte ----- ihn ----- gesucht ----- .
B-PP B-NP O O B-N B-VP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- , ----- denn ----- er ----- hatte ----- ihn ----- gesucht ----- .
B-PP B-NP O O B-NP B-VP B-NP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- , ----- denn ----- er ----- hatte ----- ihn ----- gesucht ----- .
B-PP B-NP O O B-NP B-VP B-NP B-VP

IOB-Format: Übung

Annotiere den folgenden Satz nach dem IOB-Tagging-Schema.
Markiere dabei alle NPs, VPs und PPs.

Er ----- gibt ----- dem ----- Mann ----- mit ----- der ----- Brille ----- das ----- rote ----- Buch
B-NP B-VP B-NP I-NP B-PP B-NP I-NP B-NP I-NP I-NP

im ----- Park ----- , ----- denn ----- er ----- hatte ----- ihn ----- gesucht ----- .
B-PP B-NP O O B-NP B-VP. B-NP B-VP O