



# NLP-Pipeline

Tobias Scheffer

Peter Haider

Uwe Dick

Paul Prasse

# NLP-Pipeline

- Folge von Verarbeitungsschritten für
  - ◆ Informationsextraktion,
  - ◆ Übersetzung,
  - ◆ Zusammenfassung,
  - ◆ Einige andere Sprachverarbeitungsoperationen.

# NLP- (Natural Language Processing-) Pipeline

1. Tokenisierung,
2. Wortarterkennung (Part-of-Speech-Tagging),
3. Eigennamenerkennung (Named Entity Recognition),
4. Parsing,
5. Chunking,
6. Informationsextraktion

# NLP- (Natural Language Processing-) Pipeline

- Sequentielle Verarbeitung von Sprache.
- Viele Varianten der Pipeline, je nach Problemstellung.
- Jede Stufe baut auf Ausgabe der vorangegangenen Stufe auf.
- Manche Verfahren werden für mehrere der Stufen eingesetzt.

# NLP- (Natural Language Processing-) Pipeline

- Problem: Fehler summieren sich über die Verarbeitungsschritte auf.
- Trend: Gesamte Verarbeitung als ein einziges Optimierungs-/Lernproblem formulieren.
- Einzelne Stufen können sich so besser beeinflussen.
  - ◆ Manchmal ergibt sich erst aus der semantischen Analyse, dass eine bestimmte Wortart für einen Term unplausibel ist.
- Problem: Als einziges Optimierungsproblem extrem aufwändig.
  - ◆ Stand der Forschung: Zwei Verarbeitungsschritte gleichzeitig.

# POS-Tagging

- Zuordnung der Wortart von Termen.
- Meist Eingabe für weitere Verarbeitungsschritte (Parser, Extraktionsregeln, Übersetzung, Sprachsynthese).
- „Tagset“: Definition der Wortarten, die unterschieden werden sollen.
- Im Prinzip: Nomen, Verben, Pronomen, Präposition, Adverb, Konjunktion, Partizip, Artikel, ...
  - ◆ John/NN saw/V her/PN duck/V oder
  - ◆ John/NN saw/V her/PN duck/N
- 95-97% korrekt gesetzte Tags üblich.

# Gebräuchliche Tags für das Englische

- AT Artikel
- BEZ „is“
- IN Preposition
- JJ Adjektiv
- NN Nomen
- NNP „Proper Noun“
- NNS Nomen, plural
- PERIOD „“ , „?“ , „!“
- PN Personalpronomen
- RB Adverb
- TO „to“
- VB Verb
- VBD Verb, Vergangenheit
- ...

# Wortarterkennung: Einflussfaktoren

- Manche POS-Folgen sind gebräuchlich
  - ◆ AT, JJ, NN.
- Andere hingegen weniger wahrscheinlich
  - ◆ AT, PN, VB.
- Die meisten Wörter treten nur in wenigen Wortarten auf, z.B.
  - ◆ „Lauf“/VB oder „Lauf“/NN, aber nicht „Lauf“/JJ.
- Vorhersage der Wortform auf Grundlage der Nachbartags: ca. 77% Genauigkeit.
- Vorhersage der wahrscheinlichsten Wortform jedes Wortes ohne Kontext: ca. 90% Genauigkeit.



# Wortartenerkennung: Methoden

- Historisch: Transformationsbasierte (regelbasierte) Tagger.
- HMMs
- Diskriminative Sequenzmodelle (Conditional Random Fields, Hidden Markov Support Vektor Machines).

# Wortarterkennung mit Markov-Modellen

- Ein Zustand pro POS-Tag.
- Trainingsdaten: „getaggttes“ Korpus. Zustände sind sichtbar.
- Anwendung des Taggers auf neuen Satz: Zustände sind noch nicht sichtbar.

# Training des POS-Taggers

- Startwahrscheinlichkeiten: Erste Tags eines Satzes abzählen.
- Übergangswahrscheinlichkeiten: Häufigkeiten der Tagübergänge abzählen.
- Beobachtungswahrscheinlichkeiten: Häufigkeiten aller Wörter in allen Zuständen abzählen.

# Anwenden des POS-Taggers

- Folge der Wahrscheinlichsten Einzelzustände  
gegeben Wortfolge.
  - ◆ Durch Forward-Backward-Algorithmus.
- Wahrscheinlichste Gesamtfolge von Zuständen  
gegeben Wortfolge.
  - ◆ Durch Viterbi-Algorithmus bestimmt.

# Trigramm-Tagger

- HMM = Bigramm-Tagger.
- Trigramm-Tagger: Jedes Paar von Tags bildet einen Zustand.
- Zustandsübergänge entsprechen Tag-Trigrammen.

# Transformationsbasierte Tagger

- Z.B. Brill Tagger
- Ausgangspunkt: jedes Wort mit dem dafür häufigsten Tag taggen.
- Anwendung von Transformationsregeln: „Wenn bestimmte Bedingungen im Kontext von  $X_t$  erfüllt sind, dann ersetze Tag von  $X_t$  durch  $X_t=x$ “.
- Konstruktion der Regeln durch Handarbeit, unterstützt durch Statistik (erhöht oder verringert Regel Anzahl der Fehler?)
- Weniger robust, historische Technik.

# Eigennamenerkennung – Named Entity Recognition

- Erkennung von:
  - ◆ Personen-, Firmen-, Ortsnamen
  - ◆ Daten, Uhrzeiten
  - ◆ Prozentangaben, Geldbeträgen
  - ◆ Genen, Proteinen, Loci, Viren, Krankheiten.
- Normierung von Eigennamen (Named Entity Resolution):
  - ◆ Identifizierung von Varianten:  
DTAG = Telekom = Deutsche Telekom AG
  - ◆ Abbildung in Taxonomie: Food/Fruit/Apple ≠  
Company/ComputerHardware/Apple.

# NER - Beispiel

The screenshot shows a Netscape browser window titled "Named Entity Recognition - Netscape". The address bar contains "Named Entity Recognition". The page displays the NCBI PubMed logo and navigation tabs for various databases. A search bar is visible with "PubMed" selected. Below the search bar, there are links for "Limits", "Preview/Index", "History", "Clipboard", and "Details". A list of search filters is shown, including "Virus", "DNA Domain or Region", "Protein", "DNA Family or Group", "Cell Line", and "Other". The "Protein" filter is selected. The search results show a single entry: "1: J Virol. 1993 Mar;67(3):1658-62." The title of the article is "Replication of type 1 human immunodeficiency viruses containing linker substitution mutations in the -201 to -130 region of the long terminal repeat." The authors are "Kim JY, Gonzalez-Scarano F, Zeichner SL, Alwine JC." The abstract text is annotated with colored boxes: "transfection analyses" (green), "chloramphenicol acetyltransferase" (red), "reporter gene" (grey), "linker substitution (LS) mutations" (green), "-201 and -130" (red), "long terminal repeat (LTR)" (green), "LTR transcriptional activity" (green), "T-cell line" (red), and "LS mutations" (green). The text also mentions "moderate decreases in" (grey), "S. L. Zeichner, J. Y. H. Kim, and J. C. Alwine, J. Virol. 65:2436-2444, 1991)." (grey), "In order to confirm the significance of this region in the context of viral replication, we" (grey), and "constructed several of these" (grey).



# Schwierigkeiten

- Anzahl möglicher Eigennamen im Prinzip unendlich groß.
- Viele Eigennamen sieht ein Erkenner zum ersten Mal (z.B. 3,3'-diaminobenzidine tetrahydrochlorine).
- Ständig werden neue Eigennamen erfunden (z.B. für neue chemische Verbindungen, neue Unternehmen).
- Nicht einheitliche Varianten des Namens einer Entität (Uni Potsdam, Universität Potsdam, Potsdamer Universität).

# Named Entity Recognition: Methoden

- Generierung von Merkmalen für Token:
  - ◆ Meist werden sehr viele Merkmale erzeugt.
  - ◆ Wort selbst, POS-Tag, orthographische Merkmale, ...
  - ◆ Gazetteer-Features.
- Erkennung von Eigennamen auf Grundlage der Merkmale:
  - ◆ „Sliding Window“,
  - ◆ Hidden-Markov-Modelle,
  - ◆ Conditional Random Fields.

# Named Entity Recognition

- Kernstück: „Gazetteers“
  - ◆ Listen bekannter Namen, auf Anwendungsbereich abgestimmt.
- Gazetteer aber keineswegs ausreichend.
  - ◆ Mit „Apple“ kann auch Obst gemeint sein.
  - ◆ Neu erfundene Namen tauchen nicht auf.
  - ◆ Mit „Ca“, „fat“ können Gene gemeint sein, müssen aber nicht.

# Generierung von Merkmalen

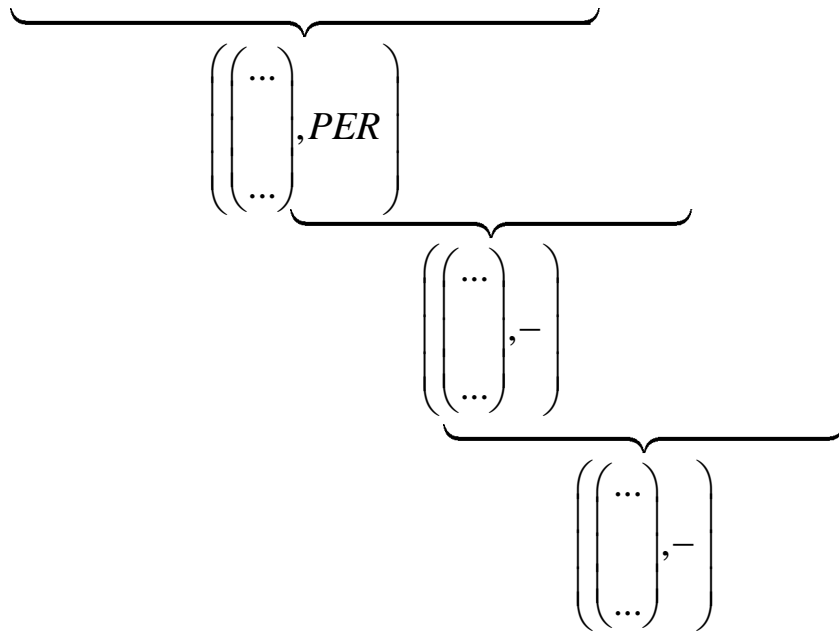
- Bag-of-Word-Features:
  - ◆ Wort ist „Aal“, ..., Wort ist „zutzeln“.
- Buchstaben-N-Gramm-Merkmale,
  - ◆ Wort enthält „aaa“, ..., Wort enthält „zzz“).
- Orthographische Merkmale:
  - ◆ „Beginnt mit Großbuchstaben“, ...,
  - ◆ „enthält Sonderzeichen“, „enthält Ziffer“, ...
- Gazetteer-Merkmale:
  - ◆ Wort taucht in Eigennamenliste 1 auf.

# Methoden: Sliding Window

- Merkmale werden in einem Fenster des Satzes generiert, der um ein zu klassifizierendes Token herumgelegt wird.
- Klassifikator entscheidet mit Hilfe dieser Merkmale, ob zentrales Token ein Eigenname ist.
- Klassifikator: Abbildung von Merkmalsvektor auf NER-Klassen.
- Fenster wird über Text geschoben.
- Trainieren: Klassifikator wird trainiert aus Paaren (Vektor, NER-Klasse). Beispiele werden aus getaggten Beispieltexten gewonnen.

# Methoden: Sliding Window

[PER **Wolff** ] , currently a journalist in [LOC **Argentina** ] , played with



- Ein Merkmalsvektor pro Fensterposition.
- Merkmale: Ein Merkmal pro Wort im Dictionary, POS, orthographische Merkmale, Element des Gazetteers?,...

# Methoden: HMMs und CRFs

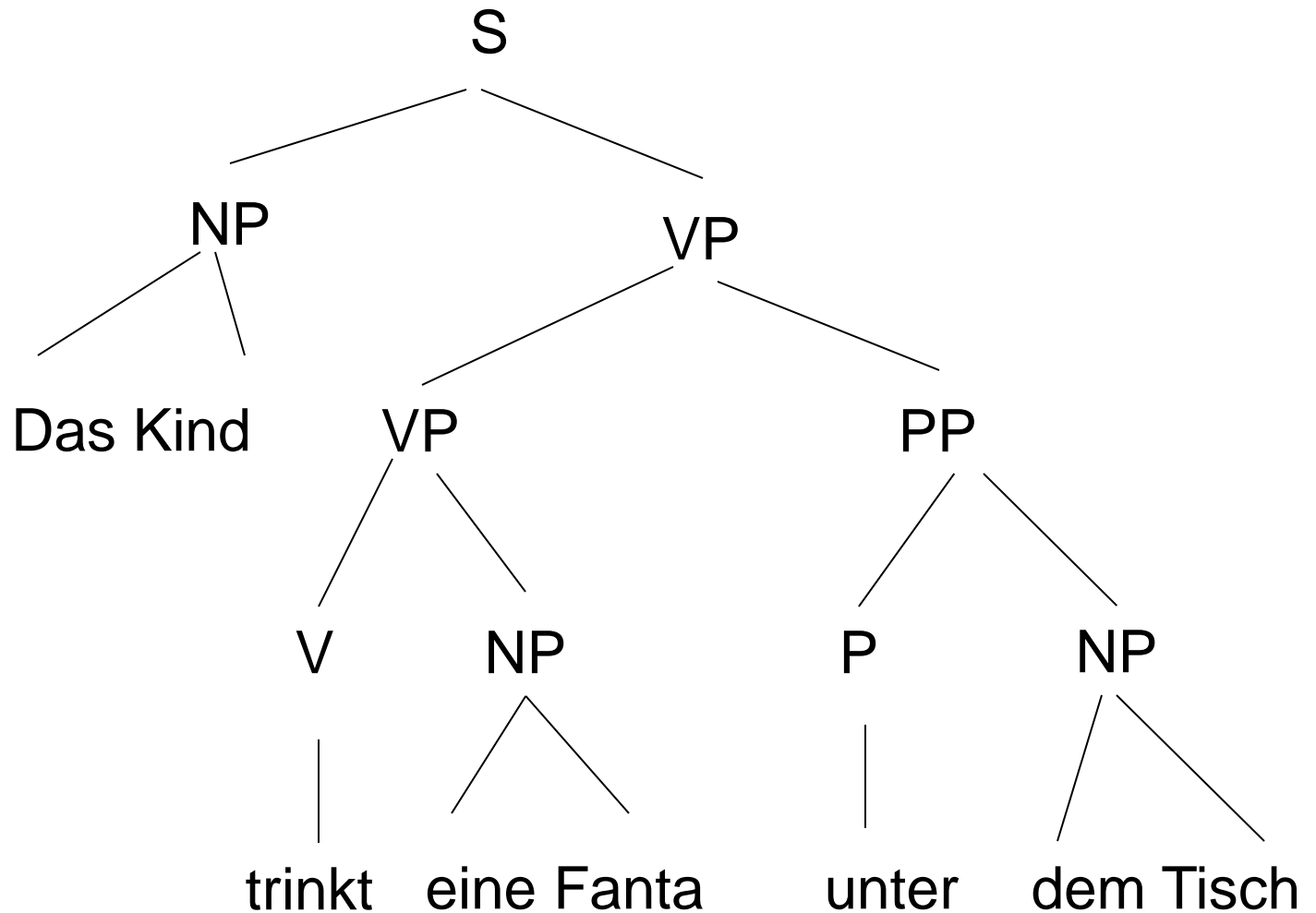
- $b_i(O_t)$ : Verteilung über Vektor der zuvor berechneten Merkmale.
- Trainieren durch abzählen aus getaggten Beispieltexten (kein Baum-Welch erforderlich).
- Siehe letzte Vorlesung.

# Parsing

- Wortarten werden in echter Sprache nicht aneinandergereiht sondern werden in Phrasen organisiert.
- Phrasen:
  - ◆ Nominalphrase: Das kleine Kind
  - ◆ Verbalphrase: trinkt eine Fanta unter dem Tisch.
    - ★ Präpositionalphrase: unter dem Tisch.
- In Phrasenstrukturgrammatiken können Phrasen baumartig dargestellt werden.



# Parsing



# Parsing: PCFGs

- Probabilistic context-free grammar: Statistisches Sprachmodell.
- Hidden-Markov-Modell:
  - ◆ Probabilistischer endlicher Automat  $\cong$  probabilistische reguläre Sprache
- Hidden-Markov-Modelle sind spezielle PCFG.
  - ◆ Jede reguläre Sprache ist gleichzeitig eine kontextfreie Sprache

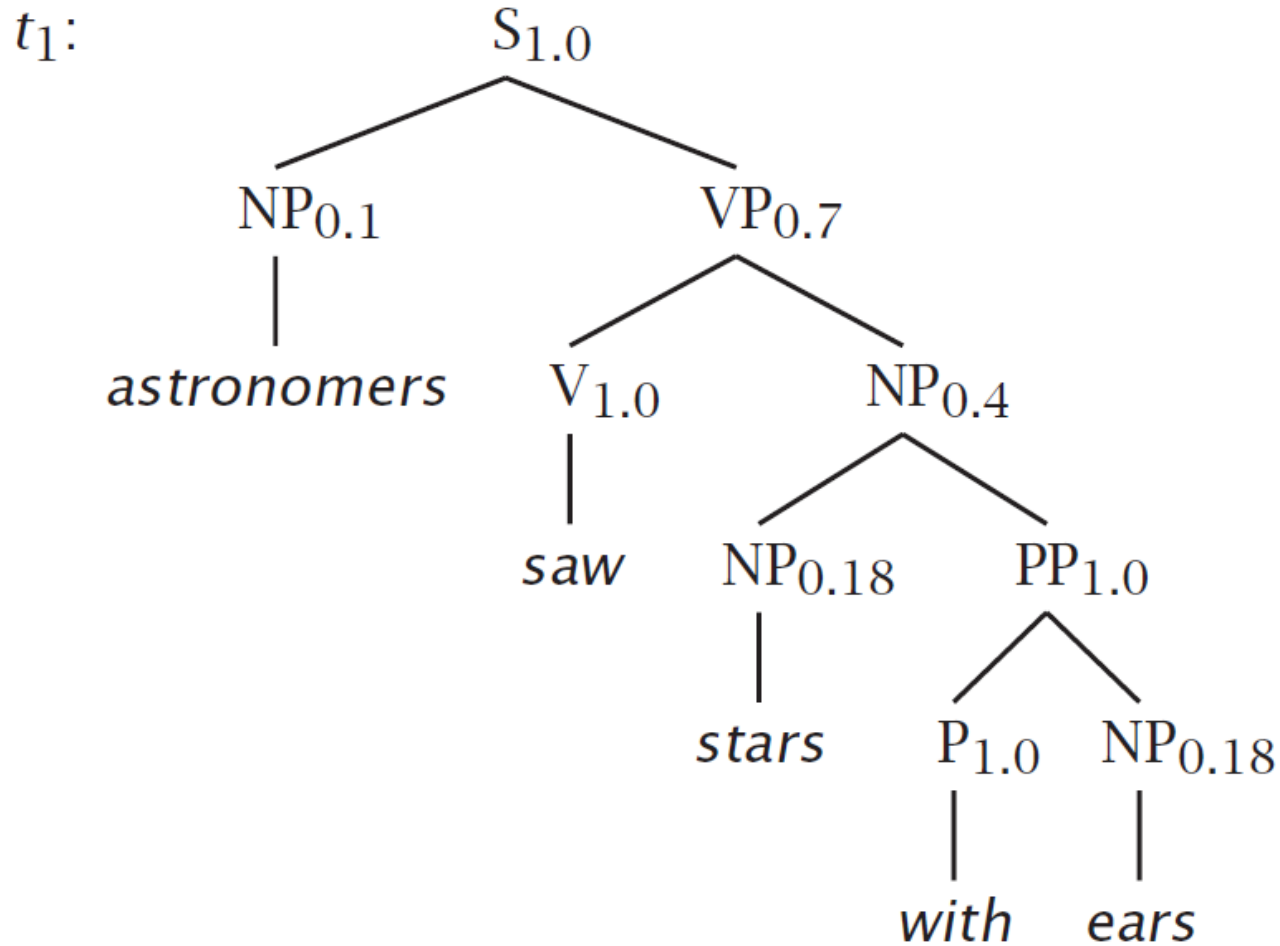
# PCFG: Definition

- Terminalsymbole  $\{w^k\}$ ,  $k=1 \dots V$ .
- Nichtterminale  $\{N^i\}$ ,  $i=1 \dots n$ .
- Startsymbol  $N^1$ .
- Regeln  $\{N^i \rightarrow \xi^j\}$ , wobei  $\xi^j$  eine Folge aus Terminalen und Nichtterminalen ist.
- Wahrscheinlichkeiten, so dass  $\forall i: \sum_j P(N^i \rightarrow \xi^j) = 1$
- $P(N^i \rightarrow \xi^j)$  steht für  $P(\xi^j \mid N^i)$ .
- Zu parsierender Satz:  $w_1, \dots, w_m$ .  $w_{ab}$  steht für  $w_a, \dots, w_b$ .
- $N^i \Rightarrow^* w_a \dots w_b$ , wenn Satz  $w_a \dots w_b$  aus Nichtterminal  $N^i$  abgeleitet werden kann.
- $N_{pq}^i$ :  $w_{pq}$  wird von Nichtterminal  $N^i$  abgeleitet.

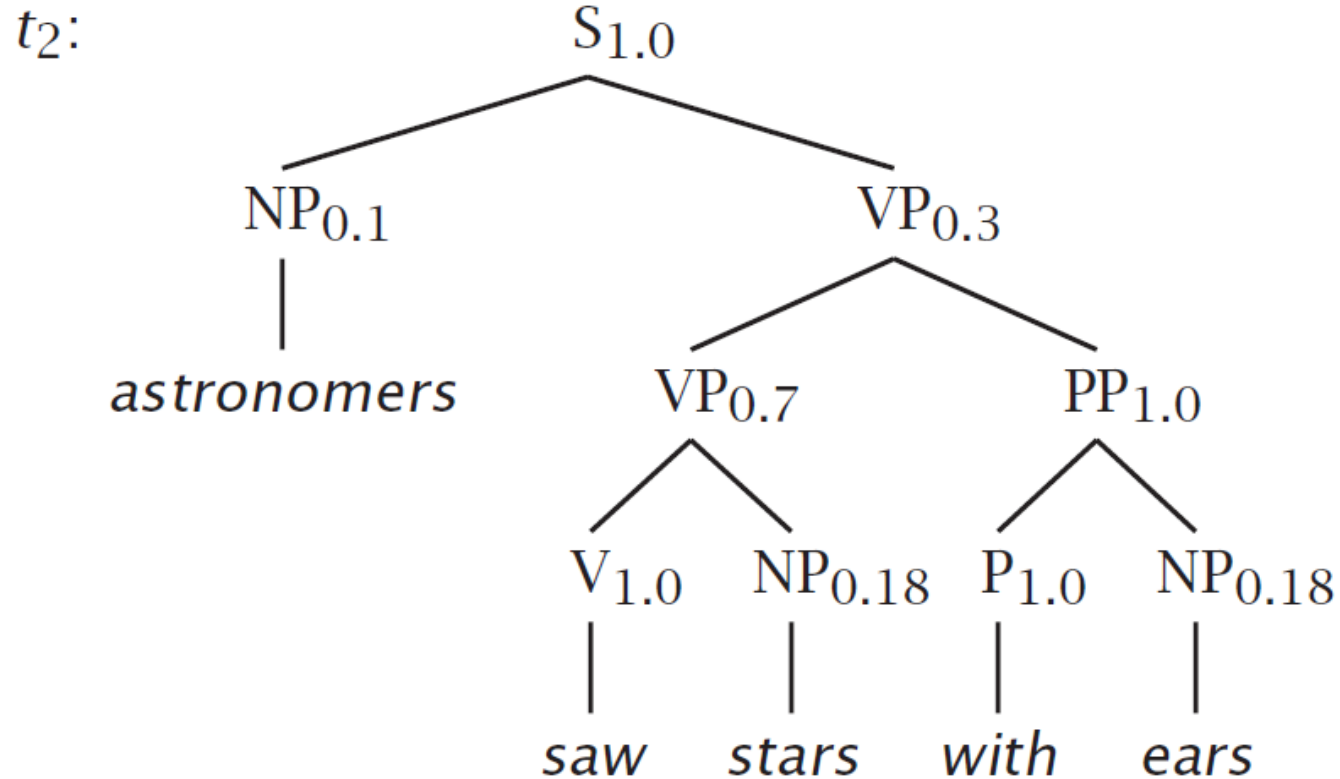
# Beispiel-PCFG

$S \rightarrow NP VP$	1.0	$NP \rightarrow NP PP$	0.4
$PP \rightarrow P NP$	1.0	$NP \rightarrow \textit{astronomers}$	0.1
$VP \rightarrow V NP$	0.7	$NP \rightarrow \textit{ears}$	0.18
$VP \rightarrow VP PP$	0.3	$NP \rightarrow \textit{saw}$	0.04
$P \rightarrow \textit{with}$	1.0	$NP \rightarrow \textit{stars}$	0.18
$V \rightarrow \textit{saw}$	1.0	$NP \rightarrow \textit{telescopes}$	0.1

# Beispiel-Parse



# Beispiel-Parse



# PCFG: Wahrscheinlichkeit eines Satzes

- Über alle Pars-Bäume  $t$ :  $P(w_{1m}) = \sum_t P(w_{1m}, t)$
- Beispiel:
  - ◆  $P(S \rightarrow \text{bla}) = 1/3$ .
  - ◆  $P(S \rightarrow S S) = 2/3$ .
  
  - ◆  $P(\text{„bla bla bla“}) = ?$

# PCFG: Elementare Fragen

- PCFGs beantworten (wie HMMs) drei Fragen:
  1. Was ist die Wahrscheinlichkeit eines Satzes  $w_{1m}$  gegeben eine Grammatik?
  2. Was ist der wahrscheinlichste Pars-Baum gegeben Satz und Grammatik?
  3. Welche Wahrscheinlichkeiten für die Grammatikregeln können wir aus einem gegebenen Trainingskorpus lernen?



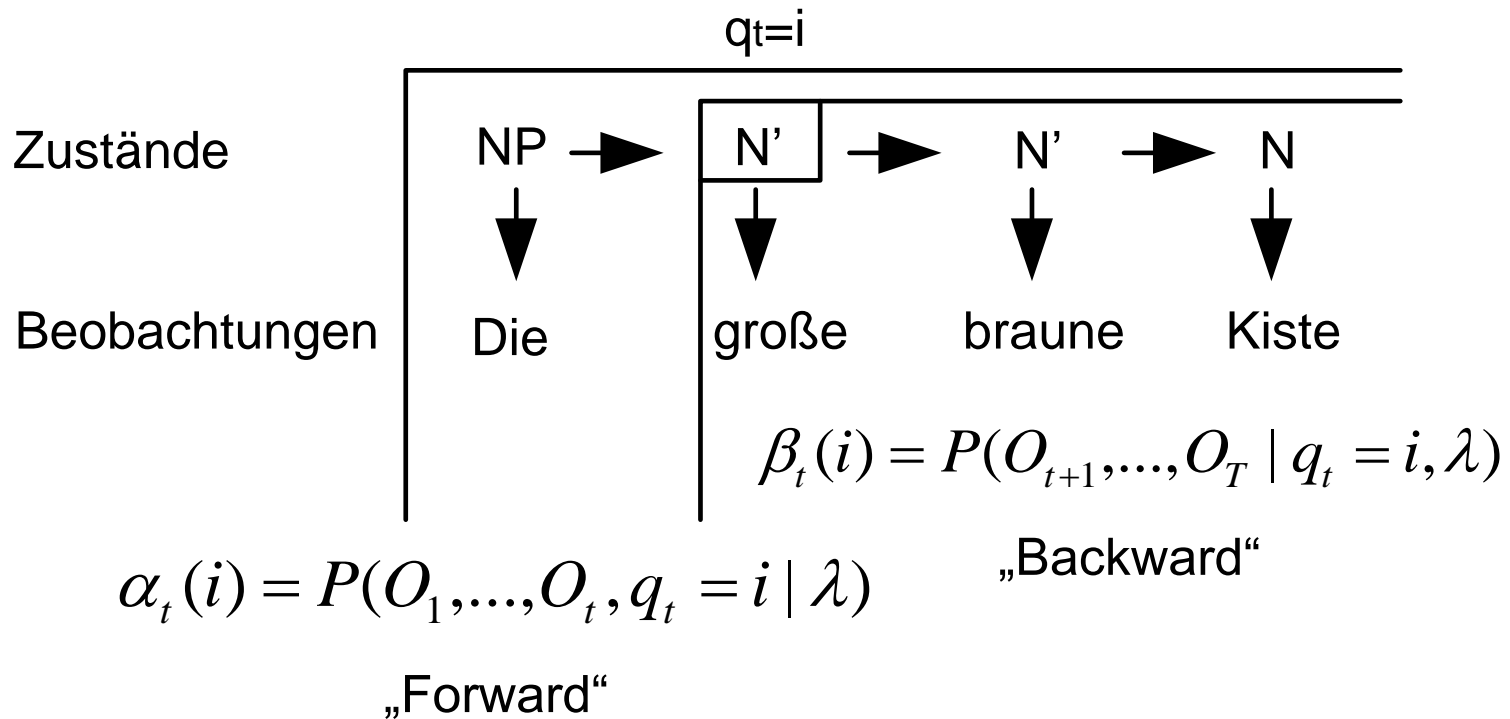
# Chomsky-Normalform

- Vereinfachung hier: Regeln haben eine der Formen:
  - ◆  $N^i \rightarrow N^j N^k$ ,
  - ◆  $N^i \rightarrow w^j$ .
- Für jede kontextfreie Grammatik lässt sich eine Grammatik in Chomsky-Normalform finden, die dieselbe Sprache erzeugt.
- (Allerdings entstehen nicht die gleichen Pars-Bäume).

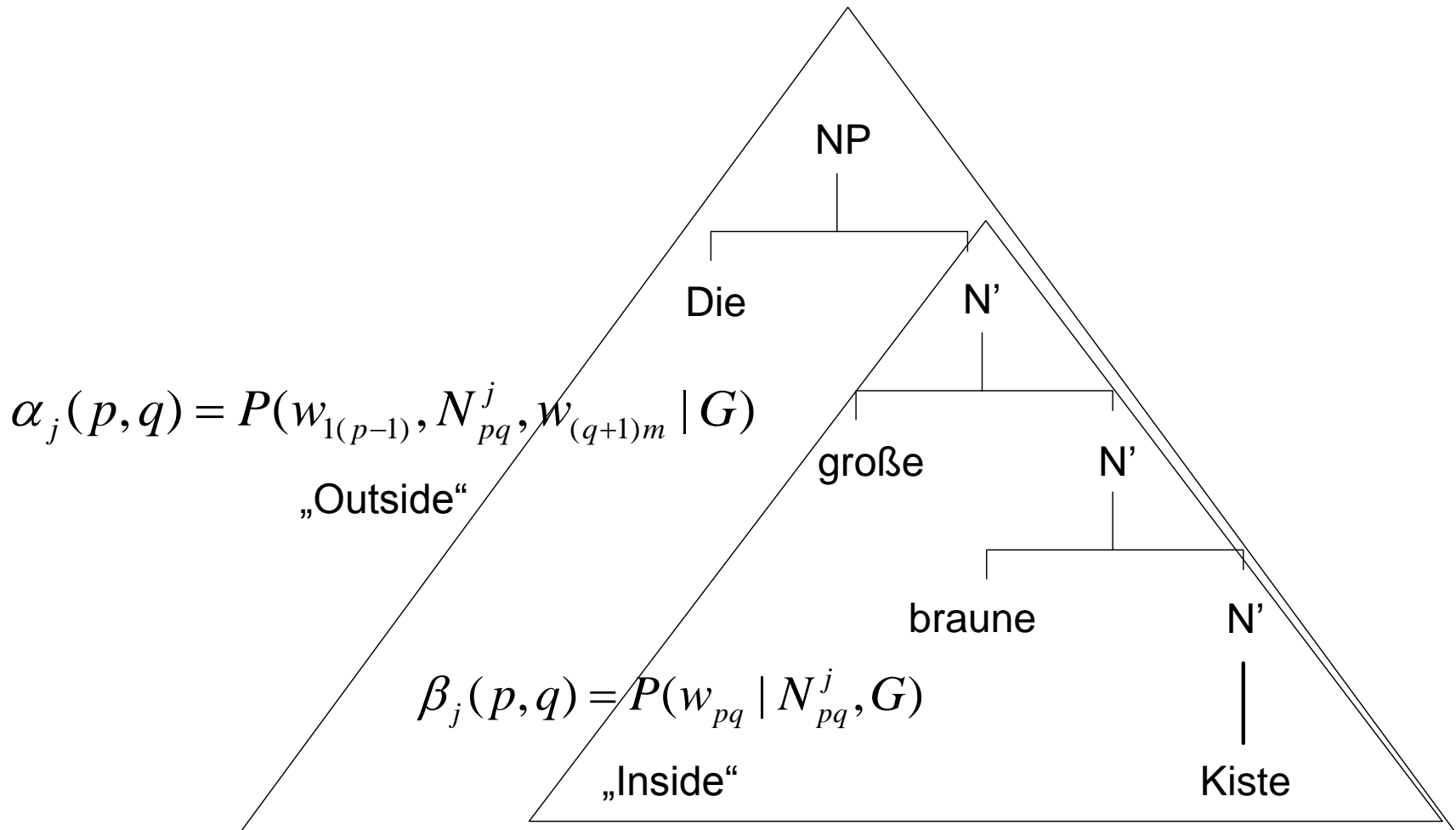
# Probabilistische reguläre Grammatiken

- Spezialfall: Jede Regel hat eine der Formen:
  - ◆  $N^i \rightarrow w^j N^k$ ,
  - ◆  $N^i \rightarrow w^j$ .
- Für jedes HMM existiert eine PCFG (Spezialfall PRG), die dieselbe Wahrscheinlichkeitsverteilung modelliert.

# Forward / Backward – Inside / Outside



# Forward / Backward – Inside / Outside



# Inside / Outside

- „Outside-Wahrscheinlichkeit“:

$$\alpha_j(p, q) = P(w_{1(p-1)}, N_{pq}^j, w_{(q+1)m} | G)$$

- „Inside-Wahrscheinlichkeit“

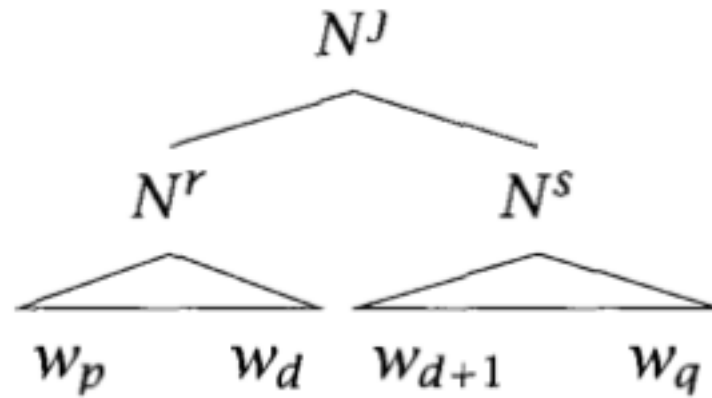
$$\beta_j(p, q) = P(w_{pq} | N_{pq}^j, G)$$

- Wahrscheinlichkeit einer Wortkette:

$$\begin{aligned} P(w_{1T} | G) &= P(N^1 \Rightarrow^* w_{1T}, G) \\ &= P(w_{1T} | N_{1T}^1, G) = \beta_1(1, T) \end{aligned}$$

# Inside-Berechnung anschaulich

$$\beta_j(p, q) = P(w_{pq} | N_{pq}^j, G)$$



# Berechnung der Inside-Wahrscheinlichkeit

- Verankerung:  $\beta_j(k, k) = P(w_k | N_{kk}^j, G)$   
 $= P(N_j \rightarrow w_k | G)$

- Induktionsschritt:

$$\begin{aligned}\beta_j(p, q) &= P(w_{pq} | N_{pq}^j, G) \\ &= \sum_{r,s} \sum_{d=p}^{q-1} P(w_{pd}, N_{pd}^r, w_{(d+1)q}, N_{(d+1)q}^s | N_{pq}^j, G) \\ &= \sum_{r,s} \sum_{d=p}^{q-1} P(N_{pd}^r, N_{(d+1)q}^s | N_{pq}^j, G) P(w_{pd} | N_{pq}^j, N_{pd}^r, N_{(d+1)q}^s, G) \\ &\quad \times P(w_{(d+1)q} | N_{pq}^j, N_{pd}^r, N_{(d+1)q}^s, w_{pd}, G) \\ &= \sum_{r,s} \sum_{d=p}^{q-1} P(N_{pd}^r, N_{(d+1)q}^s | N_{pq}^j, G) P(w_{pd} | N_{pd}^r, G) \\ &\quad \times P(w_{(d+1)q} | N_{(d+1)q}^s, G)\end{aligned}$$

# Beispiel: Inside-Berechnung

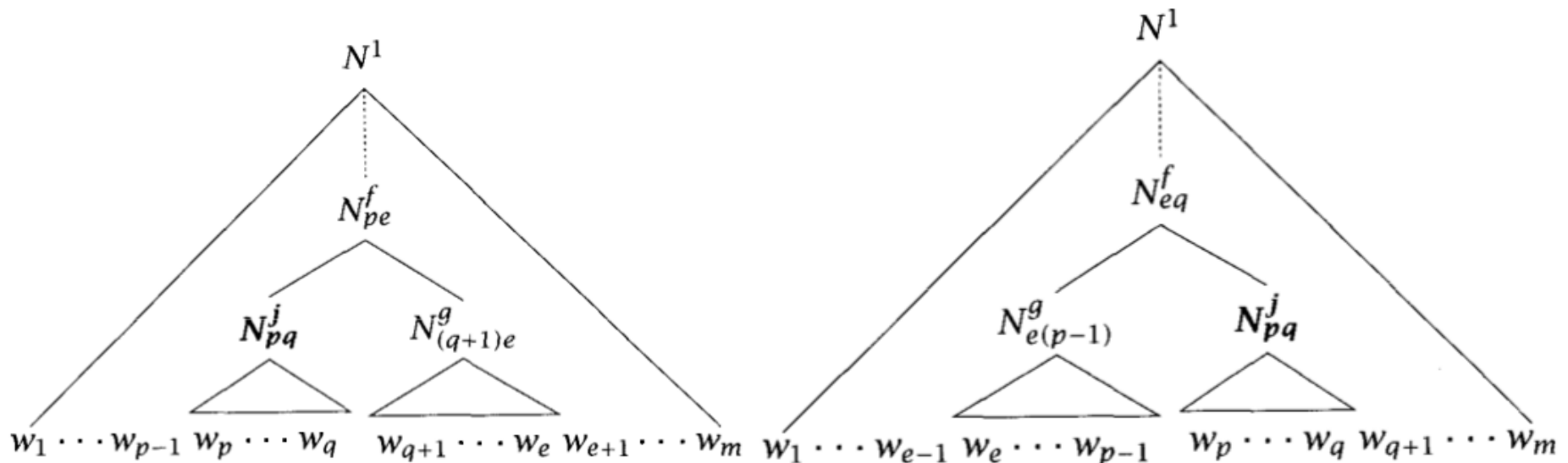
S → NP VP	1.0	NP → NP PP	0.4
PP → P NP	1.0	NP → <i>astronomers</i>	0.1
VP → V NP	0.7	NP → <i>ears</i>	0.18
VP → VP PP	0.3	NP → <i>saw</i>	0.04
P → <i>with</i>	1.0	NP → <i>stars</i>	0.18
V → <i>saw</i>	1.0	NP → <i>telescopes</i>	0.1

	1	2	3	4	5
1	$\beta_{NP} = 0.1$		$\beta_S = 0.0126$		$\beta_S = 0.0015876$
2		$\beta_{NP} = 0.04$ $\beta_V = 1.0$	$\beta_{VP} = 0.126$		$\beta_{VP} = 0.015876$
3			$\beta_{NP} = 0.18$		$\beta_{NP} = 0.01296$
4				$\beta_P = 1.0$	$\beta_{PP} = 0.18$
5					$\beta_{NP} = 0.18$
	<i>astronomers</i>	<i>saw</i>	<i>stars</i>	<i>with</i>	<i>ears</i>



# Outside-Berechnung anschaulich

$$\alpha_j(p, q) = P(w_{1(p-1)}, N_{pq}^j, w_{(q+1)m} | G)$$



# Berechnung der Outside-Wahrscheinlichkeit

- Verankerung:  $\alpha_1(1, T) = 1$ ;  $\alpha_j(1, T) = 0$  für  $j \neq 1$

- Induktionsschritt:

$$\alpha_j(p, q) = P(w_{1(p-1)}, N_{pq}^j, w_{(q+1)m} \mid G)$$

$$\begin{aligned}
 &= \left( \sum_{f, g} \sum_{e=q+1}^m P(w_{1(p-1)}, w_{(q+1)m}, N_{pe}^f, N_{pq}^j, N_{(q+1)e}^g) \right) \\
 &+ \left( \sum_{f, g} \sum_{e=1}^{p-1} P(w_{1(p-1)}, w_{(q+1)m}, N_{eq}^f, N_{e(p-1)}^g, N_{pq}^j) \right) \\
 &= \left( \sum_{f, g} \sum_{e=q+1}^m P(w_{1(p-1)}, w_{(e+1)m}, N_{pe}^f) P(N_{pq}^j N_{(q+1)e}^g \mid N_{pe}^f) P(w_{(q+1)e} \mid N_{(q+1)e}^g) \right) \\
 &+ \left( \sum_{f, g} \sum_{e=1}^{p-1} P(w_{1(e-1)}, w_{(q+1)m}, N_{eq}^f) P(N_{e(p-1)}^g, N_{pq}^j \mid N_{eq}^f) P(w_{e(p-q)} \mid N_{e(p-1)}^g) \right) \\
 &= \left( \sum_{f, g} \sum_{e=q+1}^m \alpha_f(p, e) P(N^f \rightarrow N^j N^g) \beta_g(q+1, e) \right) \\
 &+ \left( \sum_{f, g} \sum_{e=1}^{p-1} \alpha_f(e, q) P(N^f \rightarrow N^g N^j) \beta_g(e, p-1) \right)
 \end{aligned}$$

# Wahrscheinlichster Parse-Baum

- „Viterbi-Algorithmus für PCFG“
- $O(m^3n^3)$
- $\delta_i(p, q)$  = Höchste Inside-Wahrscheinlichkeit für einen Pars des Teilbaumes  $N_{pq}^i$ .
- Initialisierung:  $\delta_i(p, p) = P(N^i \rightarrow w_p)$
- Induktion:  $\delta_i(p, q) = \max_{j, k, p \leq r < q} P(N^i \rightarrow N^j N^k) \delta_j(p, r) \delta_k(r+1, q)$
- Besten Weg abspeichern:  
$$\psi_i(p, q) = \arg \max_{(j, k, r)} P(N^i \rightarrow N^j N^k) \delta_j(p, r) \delta_k(r+1, q)$$
- Parse-Baum rekonstruieren:  $\psi_i(p, q) = (j, k, r)$  dann sind Unterknoten  $N_{pr}^j$  (links) und  $N_{(r+1)q}^k$  (rechts) im Baum

# PCFG: Lernen aus Beispielkorpora

- Problem: Grammatik (CFG) muss schon gegeben sein, der Lernalgorithmus lernt nur die Wahrscheinlichkeiten  $P(N^i \rightarrow N^j N^k)$
  - Lernen von CFG: schwierig (Forschungsthema).
  - Lernen von CFG+Wahrscheinlichkeiten (komplettes Lernen einer PCFG aus Korpus: Forschungsthema.
1. Wenn Korpus mit Pars-Baum annotiert ist: Abzählen der Regelwahrscheinlichkeiten.
  2. Wenn Korpus nicht annotiert ist: EM-Algorithmus.

# Penn Treebank

- S simple clause
- SBARQ Wh- question
- SQ Yes/no question
- ADJP adjective phrase
- NP noun phrase
- PP prepositional phrase
- QP quantifier phrase
- VP verb phrase
- HWNP Wh- noun phrase
- -SBJ subject
- -LOC location
- CONJP conjunction phrase
- FRAG fragment
- INTJ interjection
- LST list marker
- NAC not a constituent grouping
- NX nominal constituent inside NP
- PRN parenthetical
- PRT particle
- RRC reduced relative clause
- ...

# Penn Treebank

( (S (NP-SBJ The move  
(VP followed  
(NP (NP a round)  
(PP of  
(NP (NP similar increases)  
(PP by  
(NP other lenders))  
(PP against  
(NP Arizona real estate loans))))))

- Insgesamt 4.5 Mio Wörter.
- Baum-Bänke aufzubauen und PCFGs zu lernen ist effektiver als Grammatiken von Hand zu entwickeln.



Show RDF Entry Page

Topics:

Politics	58%
War Conflict	57%

Social Tags:

- Islamic terrorism ★★☆☆
- 3rd millennium ★★☆☆
- Islam ★★☆☆
- War in Afghanistan ★★☆☆
- Salafis ★★☆☆
- Taliban ★★☆☆
- Afghan Civil War ★★☆☆
- September 11 attacks ★★☆☆
- Militant activity of Osama bin Laden ★★☆☆
- Responsibility for the September 11 attacks ★★☆☆
- Osama bin Laden ★★☆☆

Entities:

- City
- Company
- Continent
- Country
- Natural Feature
- Organization
- Person
- Position
- Province Or State

Events & Facts:

- Arrest
- Employment Relation
- Generic Relations
- Person Attributes
- Person Career

For a man on the run, Osama bin laden seemed to do very little running. Instead, he chose to spend long stretches — possibly years — in one place and often in the company of his family.

As details emerge of [bin Laden's](#) era as [America's](#) most-wanted man, it appears [he was often going in one direction](#) while the American-directed hunt was moving in another.

Pakistani authorities are pulling together a close-up view of [bin Laden's](#) final years from sources such as [his](#) three widows, including one who says she never left the upper floors of the walled compound in [Abbottabad](#) where [bin Laden was killed](#). But a far more sweeping narrative has taken shape from reports of [Guantanamo Bay](#) interrogations posted by WikiLeaks in late April just before the American raid on [bin Laden's](#) compound.

These documents — in addition to interviews by [The Associated Press](#) — indicate [bin Laden](#) relied on Afghan allies for years after the Sept. 11 attacks and possibly spent relatively limited time in [Pakistan's](#) rugged tribal areas, which had been the much-discussed focus of [U.S. intelligence](#) and military resources in the manhunt.

It also suggests that [Bin Laden](#) — either by design or chance — could have taken advantage of shortcomings in [America's](#) ability to gather timely leads on [his](#) movements or get credible sources within the patchwork of tribes and militia factions in [Afghanistan](#) and [Pakistan](#).

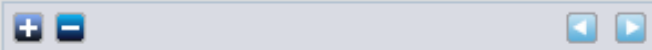
In perhaps the most striking dead-end chase, [U.S.](#) officials and others [strongly believed bin Laden slipped across the border](#) in [Pakistan](#) after dodging capture from an assault on Tora Bora in eastern [Afghanistan](#) in November 2001. But [he was still in Afghanistan](#) and galloping away on horseback in the opposite direction toward the northeastern [Kunar](#) province, according to [bin Laden's aide Awar Gul, who was arrested in December 2001](#) and eventually sent to Guantanamo.

But according to the documents released by WikiLeaks, [Gul gave the information](#) to interrogators from 2002 to 2006 — apparently too late to produce any active leads.

It also was assumed [bin Laden traveled light](#), accompanied by a few guards who were most likely Arabs. But it turns out [he kept close to his family](#) — or at least part of it — and [his most-trusted courier was a Kuwaiti-born Pakistani](#) who went by the nom de guerre [Abu Ahmed al-Kuwaiti](#), according to [U.S.](#) documents and investigators.

The discrepancies between the Western assumptions and the apparent details on [bin Laden's](#) movements go back the Sept. 11 attacks.

[Bin Laden](#) didn't go underground, as widely believed by intelligence agencies. [He stayed in Kandahar](#), mingled with [his](#) Arab fighters and met [Taliban leader Mullah Mohammed Omar](#), according to an AP interview with a former [Taliban intelligence chief](#).



### Topics:

Politics	58%
War Conflict	57%

### Social Tags:

Islamic terrorism	★★★★
3rd millennium	★★★★
Islam	★★★★
War in Afghanistan	★★★
Salafis	★★★
Taliban	★★★
Afghan Civil War	★★★
September 11 attacks	★★★
Militant activity of Osama bin Laden	★★★
Responsibility for the September 11 attacks	★★★
Osama bin Laden	★★★

### Entities:



<input checked="" type="checkbox"/> City
<input checked="" type="checkbox"/> Company
<input checked="" type="checkbox"/> Continent
<input checked="" type="checkbox"/> Country
<input checked="" type="checkbox"/> Natural Feature
<input checked="" type="checkbox"/> Organization
<input checked="" type="checkbox"/> Person
<input checked="" type="checkbox"/> Position

For a man on the run, Osama bin Laden seemed to do very little running. Instead, he chose to spend long stretches of time in one place.

As details emerge of [bin Laden's](#) era as [America's](#) most-wanted man, it appears [he](#) was often going in one direction.

Pakistani authorities are pulling together a close-

up view of [bin Laden's](#) final years from sources such as [his](#) three widows, including one who says she never met him.

But a far more sweeping narrative has taken shape from reports of [Guantanamo Bay](#) interrogations posted online.

These documents — in addition to interviews by [The Associated Press](#) — indicate [bin Laden](#) relied on Afghan

[Pakistan's](#) rugged tribal areas, which had been the much-discussed focus of [U.S. intelligence](#) and military researchers.

It also suggests that [Bin Laden](#) — either by design or chance — could have taken advantage of shortcomings in

the patchwork of tribes and militia factions in [Afghanistan](#) and [Pakistan](#).

In perhaps the most striking dead-

end chase, [U.S.](#) officials and others strongly believed [bin Laden](#) slipped across the border in [Pakistan](#) after

But [he](#) was still in [Afghanistan](#) and galloping away on horseback in the opposite direction toward the northeast

[in September 2001](#) and eventually sent to Guantanamo.

But according to the documents released by WikiLeaks, [Gul](#) gave the information to interrogators from 2002 to

It also was assumed [bin Laden](#) traveled light, accompanied by a few guards who were most likely Arabs. But

[trusted courier was a Kuwaiti-born Pakistani](#) who went by the nom de guerre [Abu Ahmed al-Kuwaiti](#),

The discrepancies between the Western assumptions and the apparent details on [bin Laden's](#) movements go

[Bin Laden](#) didn't go underground, as widely believed by intelligence agencies. [He](#) stayed in [Kandahar](#), mingling

to an AP interview with a former [Taliban intelligence chief](#).



# Fragen?