



# Computerlinguistische Grundlagen

Jürgen Hermes

Sommersemester 20

Sprachliche Informationsverarbeitung

Institut für Digital Humanities

Universität zu Köln



# Parser

- **Begriffe, Anwendungen, Strategien, Überblick**
- **Elementare Parsingalgorithmen**
- **Chart-Parser**
- **Probabilistische Parser**

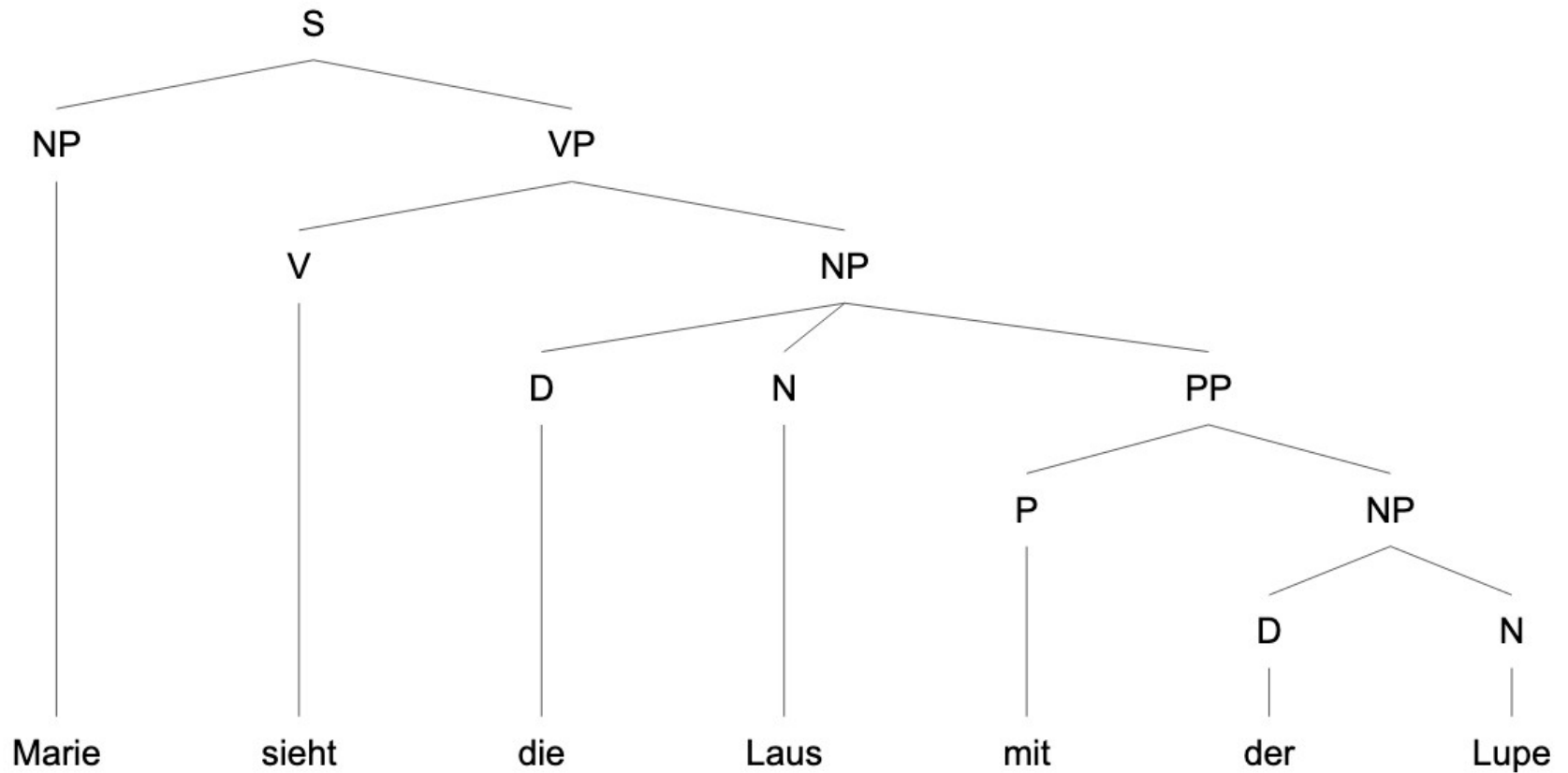


# Probabilistisches Parsen - Grundidee

- Ermittlung der wahrscheinlichsten Strukturbeschreibung für einen Eingabesatz durch Berechnung von Wahrscheinlichkeiten von Strukturen.
- Basiert auf probabilistischen Grammatiken (syntaktische und lexikalische Regeln sind mit Gewichten versehen; je wahrscheinlicher, desto gewichtiger)

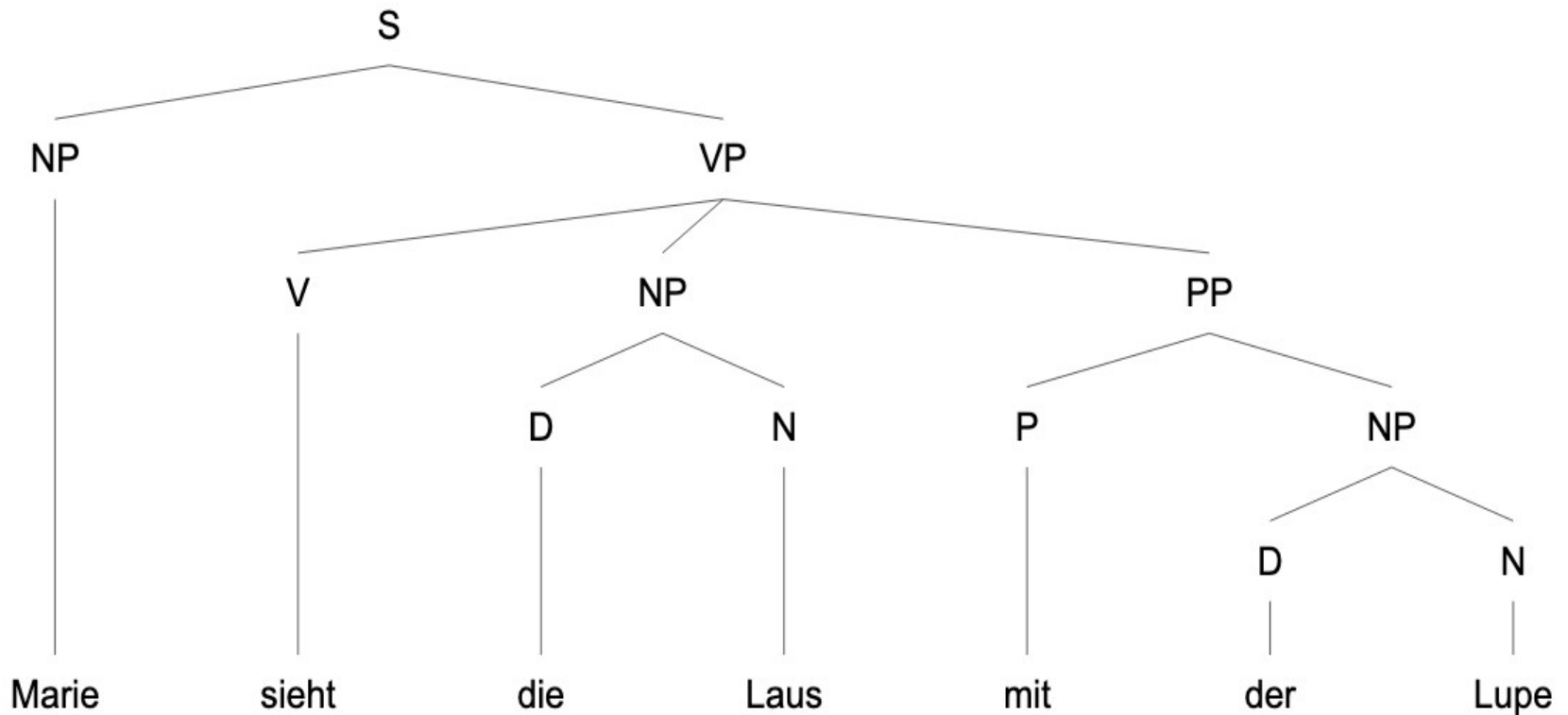


# Probabilistisches Parsen – ambige Strukturen



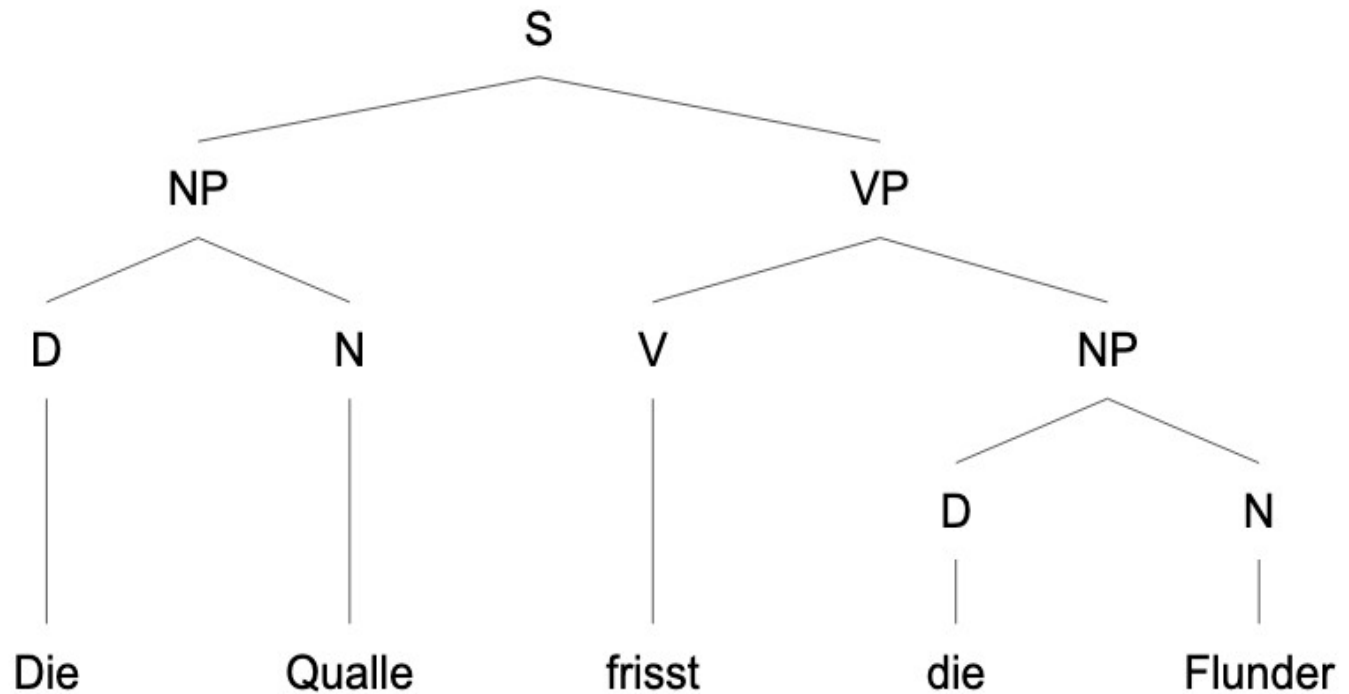


# Probabilistisches Parsen – ambige Strukturen



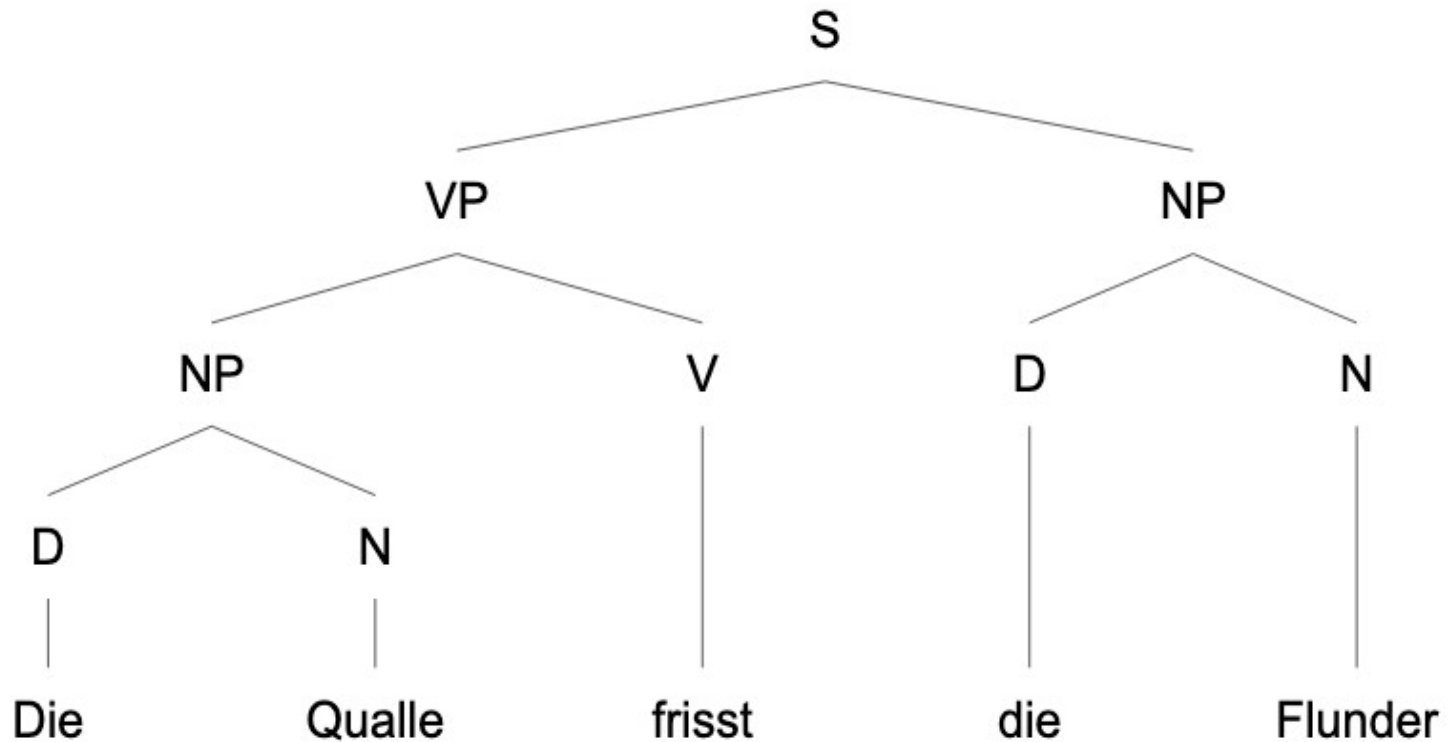


# Probabilistisches Parsen – ambige Strukturen





# Probabilistisches Parsen – ambige Strukturen





# Probabilistisches Parsing

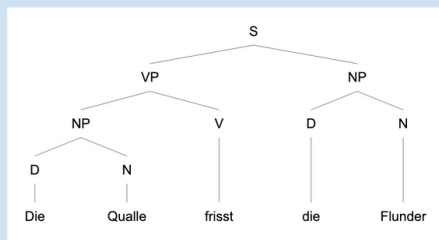
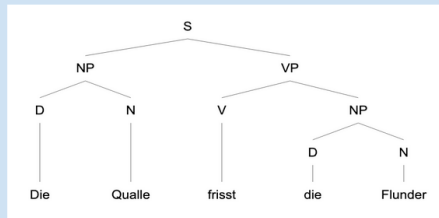
① Die ② Qualle ③ frisst ④ die ⑤ Flunder ⑥

## Syntax

$S \rightarrow NP VP$  0.8  
 $S \rightarrow VP NP$  0.2  
 $NP \rightarrow D N$  0.5  
 $VP \rightarrow V NP$  0.7  
 $VP \rightarrow NP V$  0.3

## Lexikon

$D \rightarrow die$  0.3  
 $N \rightarrow Qualle$  0.1  
 $N \rightarrow Flunder$  0.1  
 $V \rightarrow frisst$  0.1



	0	1	2	3	4	5	
0	<del><math>S \rightarrow NP VP</math> 0.8</del> <del><math>S \rightarrow VP NP</math> 0.2</del> <del><math>NP \rightarrow D N</math> 0.5</del> <del><math>VP \rightarrow V NP</math> 0.7</del> <del><math>VP \rightarrow NP V</math> 0.3</del>	<del><math>NP \rightarrow D N</math> 0.15</del>  $D \rightarrow die$ 0.3	<del><math>NP \rightarrow D N</math> 0.015</del> <del><math>VP \rightarrow NP V</math> 0.0045</del> <del><math>S \rightarrow NP VP</math> 0.012</del>	<del><math>VP \rightarrow NP V</math> 0.00045</del> <del><math>S \rightarrow NP VP</math> 0.0000054</del> <del><math>S \rightarrow VP NP</math> 0.00000</del>			<del><math>S \rightarrow NP VP</math> 0.0000126</del>  <del><math>S \rightarrow VP NP</math> 0.00000135</del>
1			$N \rightarrow Qualle$ 0.1				
2			<del><math>VP \rightarrow V NP</math> 0.7</del> <del><math>VP \rightarrow NP V</math> 0.3</del> <del><math>NP \rightarrow D N</math> 0.5</del>	<del><math>VP \rightarrow V NP</math> 0.07</del>  $V \rightarrow frisst$ 0.1		<del><math>VP \rightarrow V NP</math> 0.00105</del>	
3				<del><math>NP \rightarrow D N</math> 0.5</del>	<del><math>NP \rightarrow D N</math> 0.15</del>	<del><math>NP \rightarrow D N</math> 0.015</del>	
4					$D \rightarrow die$ 0.3	$N \rightarrow Flunder$ 0.1	





# Literatur / Hausaufgabe

## **Zur Nachbereitung:**

Jurafsky & Martin (2009): Statistical Parsing (S. 493-500)

Vervollständigen Sie außerdem die Chart auf S. 500 (Jurafsky/Martin).  
Näheres auf ILIAS.