

## Der PATR II Formalismus

## PATR II: Funktionale Beschreibungen

Für PATR II gilt:

- Sprachliche Ausdrücke werden durch **funktionale Beschreibungen** oder **Deskriptionen** (engl. *functional descriptions*, FD) beschrieben
- Eine einfache funktionale Deskription besteht aus einer Menge von **Deskriptoren** (engl. *descriptor*)
- Es gibt nur eine Form von Deskriptoren:
  - ▶ Merkmale, d.h. Attribut-Wert-Paare



## PATR II: Pfade

- Eine Folge von Attributen (dargestellt von Atomen) in Spitzklammern bildet einen **Pfad**:  
⟨Pred Obj Qual Head cat⟩
- Für jeden Wert in einer FD gibt es wenigsten einen Pfad durch den er identifiziert wird, z.B. für den Wert **dog**: ⟨Pred Obj Head lex⟩
- Der Pfad  $\langle a_1, a_2, \dots, a_k \rangle$  identifiziert den Wert des Attributs  $a_k$  in der FD, die den Wert des Pfades  $\langle a_1, a_2, \dots, a_{k-1} \rangle$  darstellt. Es kann gelesen werden als das  $a_k$  des  $a_{k-1}, \dots$  des  $a_1$ .

## Merkmale

- Ein Paar, das aus einem Pfad in einer FD und einem Wert besteht, zu dem der Pfad führt, ist ein **Merkmal** (engl. *feature*) des beschriebenen Objektes.
- Ist der Wert ein Symbol, ist das Paar ein **Basismerkmal** (engl. *basic feature*) der FD.
- Jede FD kann als eine Liste von Basismerkmalen dargestellt werden.

## Merkmallisten

$\langle \text{cat} \rangle = S$	$\langle \text{Pred Obj Head lex} \rangle = \text{dog}$
$\langle \text{Subj cat} \rangle = NP$	$\langle \text{Pred Obj Qual cat} \rangle = PP$
$\langle \text{Subj case} \rangle = \text{Subjective}$	$\langle \text{Pred Obj Qual Head cat} \rangle = P$
$\langle \text{Pred cat} \rangle = VP$	$\langle \text{Pred Obj Qual Head lex} \rangle = \text{in}$
$\langle \text{Pred Head cat} \rangle = Vt$	$\langle \text{Pred Obj Qual Obj cat} \rangle = NP$
$\langle \text{Pred Head lex} \rangle = \text{kicked}$	$\langle \text{Pred Obj Qual Obj Spec cat} \rangle = Det$
$\langle \text{Pred Obj cat} \rangle = NP$	$\langle \text{Pred Obj Qual Obj Spec lex} \rangle = \text{the}$
$\langle \text{Pred Obj Spec cat} \rangle = Det$	$\langle \text{Pred Obj Qual Obj Head cat} \rangle = N$
$\langle \text{Pred Obj Spec lex} \rangle = \text{the}$	$\langle \text{Pred Obj Qual Obj Head lex} \rangle = \text{garden}$
$\langle \text{Pred Obj Head cat} \rangle = N$	$\langle \text{Pred Obj Qual Obj case} \rangle = \text{Objective}$
	$\langle \text{Pred Obj case} \rangle = \text{Objective}$

## Unifikation

- Wenn zwei oder mehr einfache FDD kompatibel sind, können sie zu einer einfachen FD kombiniert werden, welche die Dinge beschreibt, die welche die Ausgangs-FDD beschreiben. Dieser Prozess heißt **Unifikation**.
- In der FUG ist "=" das Zeichen für die Unifikation. Der Ausdruck  $\alpha = \beta$  bezeichnet also das Ergebnis der Unifikation von  $\alpha$  und  $\beta$

## PATR II: Verkettung und Unifikation

Der PATR-II-Formalismus, der am internationalen Forschungsinstitut in Stanford (SRI International) als linguistisches Werkzeug entwickelt worden ist, basiert im wesentlichen auf den Operationen der Verkettung und der Unifikation:

- ◆ *Verkettung*: Die einzige zulässige Operation zur Kombination von Zeichenketten ist die Verkettung. Dadurch erhält der Formalismus die Form einer kontextfreien Grammatik, obwohl die formale Mächtigkeit darüber hinausgeht.
- ◆ *Unifikation*: Die einzige zulässige Operation zur Kombination von Informationsstrukturen ist die Unifikation. Dadurch wird der Formalismus rein deklarativ und unabhängig von der Reihenfolge der Auswertung.

## Merkmalestrukturen - Merkmallisten

<i>Kat :</i>	<i>NP</i>						
<i>Kongruenz :</i>	<table><tr><td><i>Numerus :</i></td><td><i>Singular</i></td></tr><tr><td><i>Person :</i></td><td><i>3</i></td></tr><tr><td><i>Genus :</i></td><td><i>Maskulin</i></td></tr></table>	<i>Numerus :</i>	<i>Singular</i>	<i>Person :</i>	<i>3</i>	<i>Genus :</i>	<i>Maskulin</i>
<i>Numerus :</i>	<i>Singular</i>						
<i>Person :</i>	<i>3</i>						
<i>Genus :</i>	<i>Maskulin</i>						

$\langle Kat \rangle = NP$

$\langle Kongruenz Numerus \rangle = Singular$

$\langle Kongruenz Person \rangle = 3$

$\langle Kongruenz Genus \rangle = Maskulin$

## Kombinationsregeln

Mit der Definition der Merkmalstrukturen ist der Bereich der Formen charakterisiert, mit deren Elementen die zu analysierenden Objekte, nämlich Zeichenketten, beschrieben werden. Die Mengen der Objekte und Beschreibungen sind normalerweise nicht endlich. Wir brauchen daher eine Möglichkeit nicht-endliche Mengen von Objekten und Beschreibungen endlich zu charakterisieren. Dies geschieht durch eine Grammatik. Die Regeln der Grammatik müssen zwei Dinge klären:

1. Wie werden Zeichenketten zu größeren Zeichenketten verkettet.
2. In welcher Beziehung stehen die assoziierten Merkmalstrukturen zueinander.



## Annotierte kontextfreie Regeln

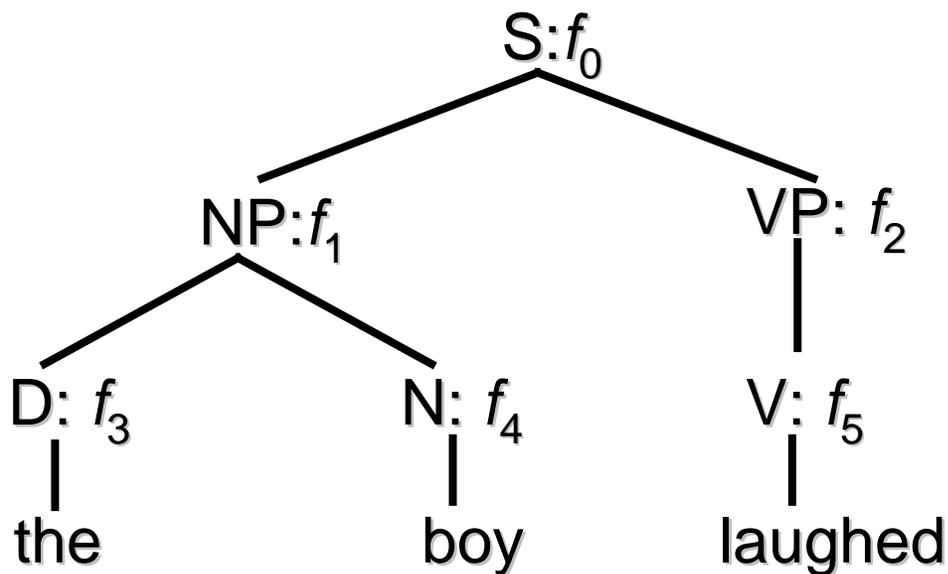
Alle genannten unifikationsbasierten Grammatikformalismen benutzen zur Lösung der beiden genannten Aufgaben im Prinzip annotierte kontextfreie Regeln, wobei allerdings große Unterschiede in der Form der Darstellung dieser Regeln bestehen.

Die im folgenden gewählte Form zeichnet sich durch ihre besondere Einfachheit aus; sie ist jedoch ausdrucksstark genug, so daß alle gebräuchlichen Notationen auf sie zurückgeführt werden können.



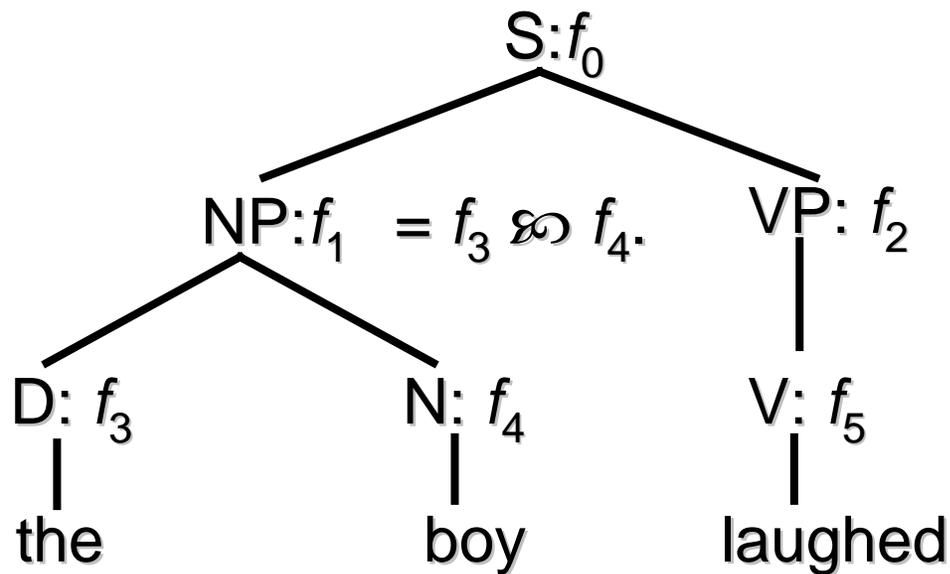
## Konstituentenstruktur - Merkmalstruktur

Vereinfachend nehmen wir zunächst an, daß jede Konstituente eines Satzes wie *The boy laughed* durch eine Merkmalstruktur charakterisiert ist, und die Merkmalstruktur einer Konstruktion das Ergebnis der Unifikation der Merkmalstrukturen seiner Konstituenten ist.



## Beispiel

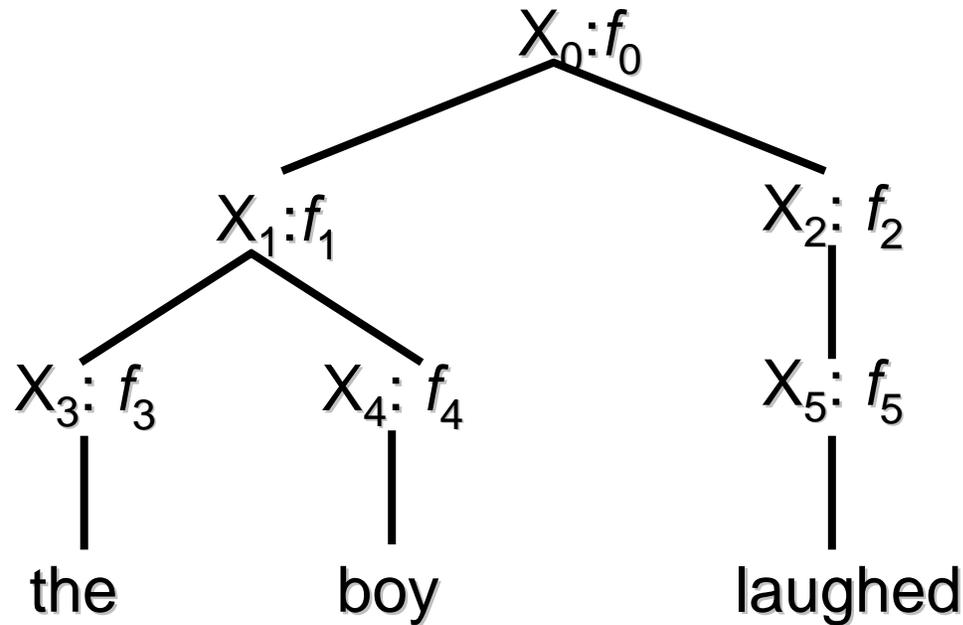
Die mit der NP *the boy* assoziierte Merkmalstruktur  $f_1$  setzt sich also aus denen vom Determinator *the* ( $f_3$ ) und vom Nomen *boy* ( $f_4$ ) zusammen, wobei vereinfachend gelten soll:  $f_1 = f_3 \otimes f_4$ .



## Beispiel

In der im folgenden gewählten Form des Formalismus sind die syntaktischen Kategorien selbst Merkmale, d.h. werden durch Attribut-Wert-Paare wie [*Kat:* *S*], [*Kat:* *NP*] etc. charakterisiert. Wir erhalten somit die genereller Darstellung:

## Kategorien als Merkmale



Dabei soll vereinfacht gelten  $f_0 = f_1 \bowtie f_2$ ,  $f_1 = f_3 \bowtie f_4$ , etc.

## Unifikationsbasierte Grammatik

Eine unifikationsbasierte Grammatik, wie wir sie im folgenden verwenden werden, besteht also aus zwei Komponenten:

- ◆ Einer kontextfreien Grammatik, und
- ◆ einer Menge von Gleichungsmengen, deren Elemente den Regeln der kontextfreien Grammatik zugeordnet sind.

## Grammatikregel

Kontextfreie Regel

Beispiel:

$X_0 \rightarrow X_1 X_2$

$(R_1)$

$\textcircled{1} X_0 \text{ Kat} \textcircled{2} = S$

$\textcircled{1} X_1 \text{ Kat} \textcircled{2} = NP$

$\textcircled{1} X_2 \text{ Kat} \textcircled{2} = VP$

$\textcircled{1} X_0 \text{ Kopf} \textcircled{2} = \textcircled{1} X_2 \text{ Kopf} \textcircled{2}$

$\textcircled{1} X_0 \text{ Kopf Subjekt} \textcircled{2} = \textcircled{1} X_2 \text{ Kopf} \textcircled{2}$

Merkmalsbeschränkungen

## Beispiele

John sleeps a  $\left[ \begin{array}{l} \textit{Kat} : S \\ \textit{Kopf} : \boxed{1} \left[ \begin{array}{l} \textit{Form} : \textit{finit} \\ \textit{Subj} : \boxed{2} \left[ \textit{Kongr} : \left[ \begin{array}{l} \textit{Num} : \textit{Sg} \\ \textit{Pers} : 3 \end{array} \right] \right] \end{array} \right] \end{array} \right]$

John a  $\left[ \begin{array}{l} \textit{Kat} : NP \\ \textit{Kopf} : \boxed{2} \end{array} \right]$

sleeps a  $\left[ \begin{array}{l} \textit{Kat} : VP \\ \textit{Kopf} : \boxed{1} \end{array} \right]$

## Destruktive Unifikation

$$\left[ \begin{array}{l} \textit{Kat} : \textit{NP} \\ \textit{Kopf} : \left[ \textit{Kongr} : \left[ \begin{array}{l} \textit{Num} : \textit{Sg} \\ \textit{Pers} : 3 \end{array} \right] \right] \end{array} \right]$$

$$\left[ \begin{array}{l} \textit{Kat} : \textit{VP} \\ \textit{Kopf} : \left[ \begin{array}{l} \textit{Form} : \textit{finit} \\ \textit{Subj} : \left[ \textit{Kongr} : \left[ \begin{array}{l} \textit{Num} : \textit{Sg} \\ \textit{Pers} : 3 \end{array} \right] \right] \end{array} \right] \end{array} \right]$$