



Computerlinguistische Grundlagen

Jürgen Hermes

Wintersemester 19/20

Sprachliche Informationsverarbeitung

Institut für Linguistik

Universität zu Köln



Morphologie

- Wort – Morphem – Allomorph
- Sprachtypologie anhand der Wortstruktur
- Flexion und Wortbildung
- Morphologische Analyse
- **Referate: Morphologische Analysetools**
- **Grundmodelle der generativen Morphologie**
- **Morphologie mit endlichen Automaten**
- **Morphologie mit Default-Vererbungsnetzen**



Grundmodelle der generativen Morphologie

- **Morphembasierter Ansatz:** Kombination von Morphemen zu vollständigen Formen
- **Wortbasierter Ansatz:** Bildung von Wörtern aus Stammformen (z.B. dem Infinitiv)
- **Realisierungsbasierter Ansatz:** Realisierung einer Wortform aus vorgegebener Form



Morphologie mit endlichen Automaten

- Beispiel für morphembasierten Ansatz
- einfache endliche Automaten
- finite-state-Transducer
- Probleme:
 - Nichtkonkatenative Phänomene
 - Nichtlokale Abhängigkeiten



Morphologie mit endlichen Automaten

- Beispiel für morphembasierten Ansatz
- einfache endliche Automaten
- finite-state-Transducer
- Probleme:
 - Nichtkonkatenative Phänomene
 - Nichtlokale Abhängigkeiten



Morphologie mit Default-Vererbungsnetzen

- Beispiel für realisierungsbasierten Ansatz
- Reichere Syntax (Typ-0-Sprache)
- Gebrauch von Defaultmechanismen
- Repräsentation des lexikalischen Wissens in Knoten, die Pfaden Werte zuweisen



Wissensrepräsentation in DATR: Grundlegende Syntax

WARTEN:

<form präs sg eins> == warte

<form präs sg zwei> == wartest

(...)

.



Wissensrepräsentation in DATR: Resolutionspfade

WARTEN:

<form> == <wurzel> <endung>

<wurzel> == wart

<endung präs sg eins> == e

<endung präs sg zwei> == est

<endung präs sg drei> == et

<endung präs pl eins> == en

<endung präs pl zwei> == et

<endung präs pl drei> == en

.



Wissensrepräsentation in DATR: Redundanzen

WARTEN:

<form> == <wurzel> <endung>

<wurzel> == wart

<endung präs sg eins> == e

<endung präs sg zwei> == est

<endung präs sg drei> == et

<endung präs pl eins> == en

<endung präs pl zwei> == et

<endung präs pl drei> == en

.



Wissensrepräsentation in DATR: Defaults und Neutralisierung

WARTEN:

| | | | |
|--|-----------------|--------------------------------|-----------------------------|
| <code><form></code> | <code>==</code> | <code><wurzel></code> | <code><endung></code> |
| <code><wurzel></code> | <code>==</code> | <code>wart</code> | |
| <code><endung sg></code> | <code>==</code> | <code>e</code> | |
| <code><endung sg zwei></code> | <code>==</code> | <code>est</code> | |
| <code><endung präs sg drei></code> | <code>==</code> | <code>et</code> | |
| <code><endung pl></code> | <code>==</code> | <code>en</code> | |
| <code><endung pl zwei></code> | <code>==</code> | <code>et</code> | |
| <code><endung präs></code> | <code>==</code> | <code><endung></code> | |
| <code><endung prät></code> | <code>==</code> | <code>et <endung></code> | <code>.</code> |



Wissensrepräsentation in DATR: Verallgemeinerung

VERB_SCHWACH_T:

| | | | |
|------------------------------------|-----------|-----------------------|-----------------------|
| <form> | == | <wurzel> | <endung> |
| <endung sg> | == | e | |
| <endung sg zwei> | == | est | |
| <endung präs sg drei> | == | et | |
| <endung pl> | == | en (...). | |

WARTEN:

| | | |
|-----------------------|-----------|-----------------------|
| <> | == | VERB_SCHWACH_T |
| <wurzel> | == | wart. |



Wissensrepräsentation in DATR: Behandlung nichtlokaler Abhängigkeiten

VERB_SCHWACH_T:

<form> == <wurzel> <endung>

<form part> == ge <wurzel> <endung part>

<endung part> == et

.



Wissensrepräsentation in DATR: Behandl. nichtkonkatenativer Phänomene

VERB_STARK:

`<wurzel> == <w_anf> <w_vok> <w_end>`

`<endung sing> == ∅`

`<endung präs sing eins> == e`

`<endung sing zwei> == st`

`<endung plur> == en`

`<endung plur zwei> == et`

.



Wissensrepräsentation in DATR: Behandl. nichtkonkatenativer Phänomene

HALTEN:

<> == VERB_STARK

<w_anf> == h

<w_end> == lt

<w_vok> == a

<w_vok präs sg zwei> == ä

<w_vok präs sg drei> == ä

<w_vok prät> == ie

.



Literatur / Hausaufgabe

Zur Nachbereitung:

Lesen Sie: Carstensen et al. (2004):

Kapitel 3.2 (S. 190-217)

Entwerfen Sie einen jeweils einen Finite-State-Transducer und eine DATR-Repräsentation für die Präsens- und Präteritumformen der Verben „bleiben“ und „holen“. Wie weit Sie jeweils ins Detail gehen, bleibt Ihnen überlassen.

Zur Vorbereitung:

Lesen Sie: Ramers (2000): Kapitel 1 (S. 11-34)