

Computerlinguistik I

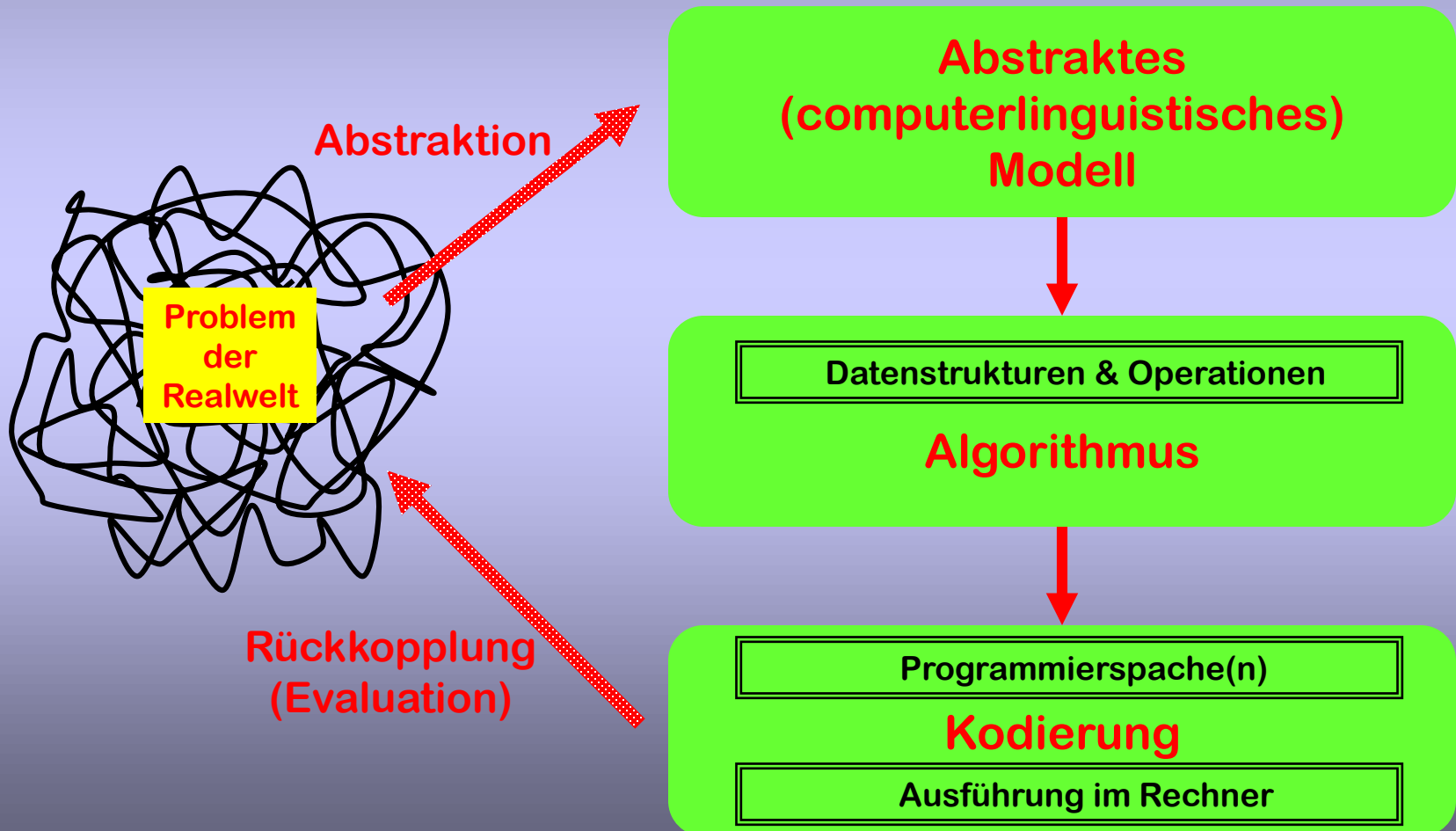
Vorlesung im WiSe 2018/19
(M-GSW-09)

Prof. Dr. Udo Hahn

Lehrstuhl für Computerlinguistik
Institut für Germanistische Sprachwissenschaft
Friedrich-Schiller-Universität Jena

<http://www.julielab.de>

Informatischer Problemlösungszyklus



Informatischer Problemlösungszyklus

- **Modellbildung**
 - **Abstraktion** von allen unwesentlichen Details der Problemstellung im Hinblick auf die algorithmische Lösung
 - Spezifikation der **logischen** Abhängigkeiten zwischen problemlösungsrelevanten Objekten
 - **(computer)linguistisches Wissen**

Informatischer Problemlösungszyklus

- **Algorithmisierung**
 - Übersetzung der modellbezogenen Spezifikation in
 - eine Menge von **Objekten** (Datenstrukturen) mit bestimmten Eigenschaften und Beziehungen zueinander
 - die erlaubten **Operationen** auf diesen Objekten
 - **Algorithmus**: (möglichst präzise) Beschreibung einer Folge zulässiger Operationen auf den Objekten, um das Problem zu lösen
 - **Computerlinguistische Kernexpertise**

Informatischer Problemlösungszyklus

- **Kodierung (Programmierung)**
 - Übersetzung der algorithmischen Spezifikation in Konstrukte einer (geeigneten) Programmiersprache
- **Ausführung des Programms**
 - Hier erst Bezug auf konkrete Maschinen (Datenstrukturen und Algorithmen sind abstrakte Konstruktionen)
 - Test-Modifikationszyklus ... Dokumentation !
 - **Informatisches Know-How**

Morphologische Prozesse: Flexion - Deflexion

- Kombination von **Grundformen** mit **Flexionsaffixen** (Kasus, Numerus, Tempus)
 - Deklination
 - **Land**: Land, Land**es**, Land**e**, L**ä**nder, L**ä**nder**n**
 - Konjugation
 - **landen**: land**e**, land**est**, land**et**, land**eten**, **gelandet**
- primär syntaktische, nur minimale semantische Information, keine grundlegenden Wortartwechsel

Morphologische Prozesse: Derivation - Dederivation

- Kombination von **Grundformen** mit **Derivationsaffixen**
 - **Land**: landen, verlanden, anlanden,
 - **Land**: Landung, Verlandung , Anlandung
 - **Land**: ländlich, verländlichen, Verländlichung
- modifizierende semantische Information, häufig mit Wortartwechsel verbunden

Morphologische Prozesse: Komposition - Dekomposition

- Kombination von Grundformen mit Grundformen (mittels Fugeninfixen)
 - Land: Landnahme, Landflucht, Landgang
 - Land: Heimatland, Ausland, Bauland
 - Land: Landesrekord, Landesverrat, Landsmann
 - Land: Inlandsflug, Landesratspräsidentengattin
- starke semantische Modifikation, fast keine Wortartwechsel
 - ... aber: Rotkehlchen, Weichteile

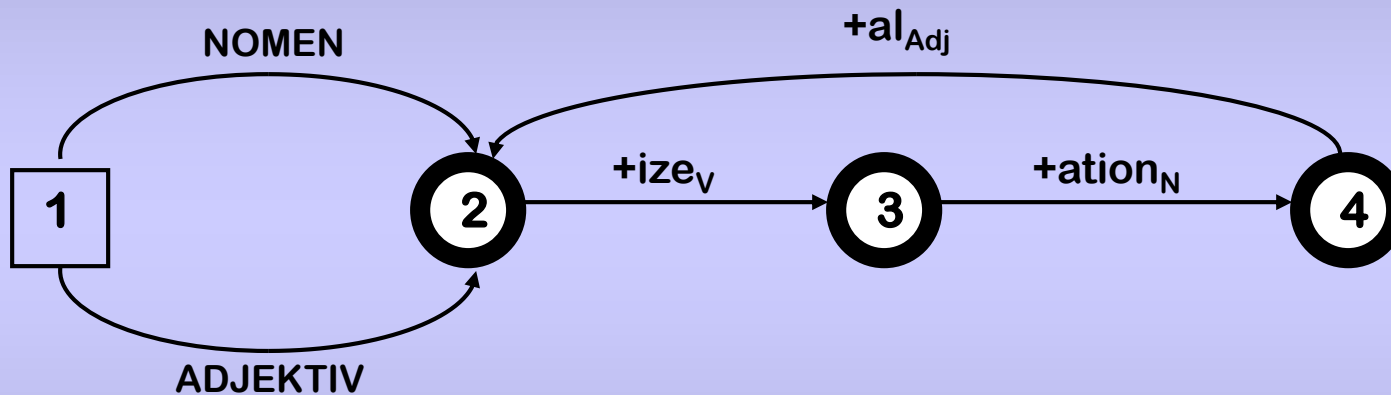
Lemmatisierung vs. Wort-Parsing

Eingabe	Lemma	Wort-Parse
Töchtern	Tochter	
Houses	Haus	
sagte	sagen	
Spiegelungen	Spiegelung	
leichter	leicht	
verlängerte	verlängert	
	verlängern	

Lemmatisierung vs. Wort-Parsing

Eingabe	Lemma	Wort-Parse
Töchtern	Tochter	Tochter [+N, +FEM, +PL, +DAT]
Hauses	Haus	Haus [+N, +NEU, +SG, +GEN]
sagte	sagen	sagen [+V, +SG, {1P,3P}, +PAST]
Spiegelungen	Spiegelung	[Spiegel] _N [ung] _{ds} [+N, +FEM, +PL, {NOM,GEN,DAT,AKK}]
leichter	leicht	leicht [+Adj, +POS, +MAS, +SG, +NOM] [+Adj, +KOM]
verlängerte	verlängert	[ver] _{dp} [[lang] _{Adj} [er] _{ds} Adj[t] _{ds} [+Part, {MAS,FEM,NEU}, +SG, + DEF, +NOM] [+Part, {FEM,NEU}, +SG, + DEF, +AKK]
	verlängern	[ver] _{dp} [[lang] _{Adj} [er] _{ds} Adj[n] _{ds} [+V, +SG, {1P,3P}, +PAST]

Automat für Dederivation



NOMEN: hospital, motor, category, ...

ADJEKTIV: moral, concrete, tender, ...



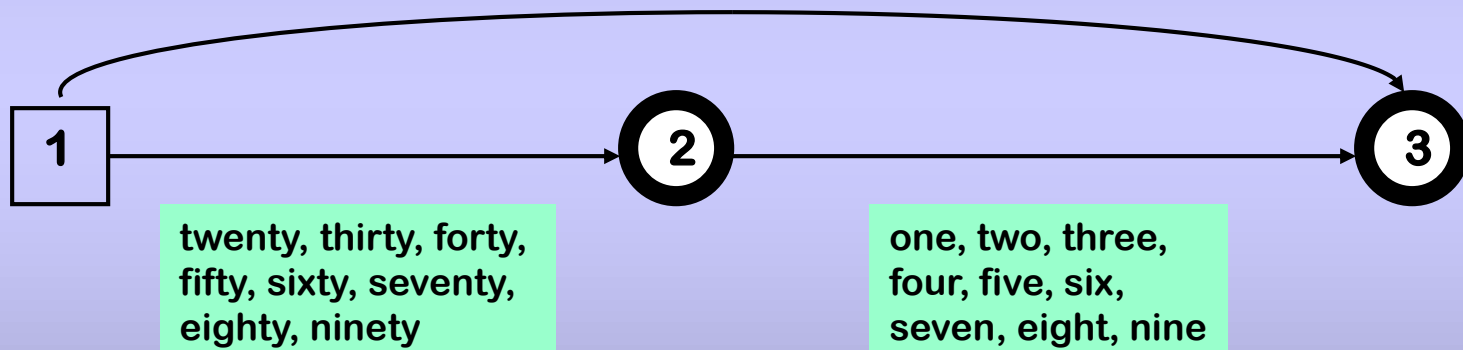
Anfangszustand



möglicher Endzustand

Automat für englische Zahlen von 1 bis 99

one, two, three, four, five, six, seven, eight, nine, ten,
eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen



Mengentheoretische Grundbegriffe

- Die Zusammenfassung aller Elemente x , die eine Eigenschaft \mathcal{E} haben, wird als **Menge** M bezeichnet:

$$M := \{x \mid x \text{ hat die Eigenschaft } \mathcal{E} \}$$

Beispiele:

LAUF := $\{x \mid x \text{ ist deutsches Lexem, das mit „LAUF“ beginnt} \}$

EoR := $\{x \mid x \text{ ist deutsches Lexem, das auf „E“ oder „R“ endet} \}$

Mengentheoretische Grundbegriffe

- Seien M_1 und M_2 Mengen. M_1 ist **Teilmenge** von M_2 , falls aus $x \in M_1$ stets $x \in M_2$ folgt; symbolisch: $M_1 \subseteq M_2$.
- Gilt für zwei Mengen, M_1 und M_2 , einerseits $M_1 \subseteq M_2$ und andererseits $M_1 \neq M_2$, dann ist M_1 **echte Teilmenge** von M_2 ; symbolisch: $M_1 \subset M_2$

Beispiele:

$\text{LAUF}^* := \{\text{Laufbahn, laufen, Lauffeuer, Laufmaschine, Laufsteg}\} \subseteq \text{LAUF}$

$\text{LAUF} \subset \text{LA} := \{x \mid x \text{ ist deutsches Lexem, das mit „LA“ beginnt}\}$

$\text{R} := \{x \mid x \text{ ist deutsches Lexem, das auf „R“ endet}\} \subseteq \text{EoR}$

Mengentheoretische Grundbegriffe

- Gilt für zwei Mengen, M_1 und M_2 , sowohl $M_1 \subseteq M_2$ als auch $M_2 \subseteq M_1$, so folgt: $M_1 = M_2$ (**Mengengleichheit**).
- Die **leere Menge** ist die Menge, die kein Element enthält; symbolisch: $\{\}$ oder \emptyset .
 - Bemerkung: \emptyset ist Teilmenge jeder Menge.
- Die **Kardinalität** einer endlichen Menge M ist die Anzahl ihrer Elemente; symbolisch: $|M|$

Mengentheoretische Grundbegriffe

- Wenn M und N Mengen sind, dann charakterisiert die Menge

$$M \cap N := \{x \mid x \in M \text{ und } x \in N\}$$

den **Durchschnitt**

$$M \cup N := \{x \mid x \in M \text{ oder } x \in N\}$$

die **Vereinigung**

von M und N

Mengentheoretische Grundbegriffe

- **Beispiele:**

LAUF* := {Laufbahn, laufen, Lauffeuer, Laufmaschine, Laufsteg}

LAUF* \cap EoR

= { Lauffeuer, Laufmaschine }

{ Lauffeuer, Laufmaschine } \cup { Lauffeuer,
Laufpass }

= { Lauffeuer, Laufmaschine, Laufpass }

Mengentheoretische Grundbegriffe

- Wenn $I = \{1, \dots, n\}$ eine nichtleere Indexmenge ist und jedes $i \in I$ für M_i eine Menge charakterisiert, dann gilt als
 - Verallgemeinerung des **Durchschnitts**

$$\bigcap_{i \in I} M_i := \{x \mid x \in M_i \text{ für alle } i \in I\} = \bigcap_{i=1}^n M_i$$

- Verallgemeinerung der **Vereinigung**

$$\bigcup_{i \in I} M_i := \{x \mid x \in M_i \text{ f. mind. ein } i \in I\} = \bigcup_{i=1}^n M_i$$

Mengentheoretische Grundbegriffe

- Die Menge aller Teilmengen einer Menge M heißt **Potenzmenge**:

$$\wp(M) := \{ N \mid N \subseteq M \} = 2^M$$

Beispiel:

$\text{LAUFS} := \{ \text{Laufschritt, Laufstall, Laufsteg} \}$

$2^{\text{LAUFS}} = \{ \emptyset, \{ \text{Laufschritt} \}, \{ \text{Laufstall} \}, \{ \text{Laufsteg} \},$
 $\{ \text{Laufschritt, Laufstall} \}, \{ \text{Laufschritt, Laufsteg} \},$
 $\{ \text{Laufstall, Laufsteg} \}, \text{LAUFS} \}$

$| 2^{\text{LAUFS}} | = 2^3 = 8$

Mengentheoretische Grundbegriffe

- Das **Kartesische Produkt** von endlich vielen Mengen M_1, \dots, M_n , $n \geq 2$, ist die Menge aller **n-tupel**:

$$M_1 \times M_2 \times \dots \times M_n := \{ (m_1, \dots, m_n) \mid m_i \in M_i, 1 \leq i \leq n \}$$

Beispiel:

LAUFB := { Laufbahn, Laufbursche }

LAUFS := { Laufschrift, Laufstall, Laufsteg }

LAUFB \times LAUFS = { (Laufbahn, Laufschrift), (Laufbahn, Laufstall),
(Laufbahn, Laufsteg), (Laufbursche, Laufschrift),
(Laufbursche, Laufstall), (Laufbursche, Laufsteg) } ²⁰

Grundbegriffe zu Relationen

- Eine (zweistellige) **Relation** ρ zwischen zwei Mengen M_1 und M_2 ist eine Teilmenge von $M_1 \times M_2$, d.h. $\rho \subseteq M_1 \times M_2$. Man schreibt auch $m \rho n$ für $(m,n) \in \rho$.

Beispiel:

GleicheLänge \subseteq DLexeme x DLexeme

GleicheLänge = { (du, da), (da, Ei), (er, es), (Dom, Bor),
(Aal, Tor), (Bild, Tier), (Tiger, Sekte),... }

Grundbegriffe zu Relationen

- Eine Relation ρ auf einer nichtleeren Menge M heißt **Äquivalenzrelation**, wenn
 - $m \rho m$ für jedes $m \in M$ (reflexiv)
 - aus $m \rho n$ folgt $n \rho m$ (symmetrisch)
 - aus $k \rho m$ und $m \rho n$ folgt $k \rho n$ (transitiv)

Beispiel:

Gleiche Länge ist Äquivalenzrelation:

(1) reflexiv: (du, du), (da, da), (Ei, Ei), (Aal, Aal), (Tiger, Tiger), ...

(2) symmetrisch: (Aal, Tor) \Rightarrow (Tor, Aal), (Tiger, Sekte) \Rightarrow (Sekte, Tiger), ...

(3) transitiv: (du, da), (da, Ei) \Rightarrow (du, Ei),

(Bild, Tier), (Tier, Rand) \Rightarrow (Bild, Rand), ...

Grundbegriffe zu Relationen

- Ist ρ eine Äquivalenzrelation auf einer Menge M , dann heißt jede Menge $[m] := \{ n \mid m \rho n \}$ für ein $m \in M$ die von m repräsentierte **Äquivalenzklasse**.
- Jede Äquivalenzrelation auf M bewirkt eine Einteilung von M in paarweise disjunkte (d.h. elementfreie) Äquivalenzklassen.

Beispiele:

$[da] = \{ Ei, er, es, du, \dots \}$

$[Aal] = \{ Tor, Bor, elf, vom, \dots \}$

$[Bild] = \{ Tier, Rand, grün, hell, \dots \}$

Grundbegriffe zu Relationen

- Eine **Halbordnung** (**partielle Ordnung**) auf einer Menge M ist eine Relation „ $<$ “ auf M mit den Eigenschaften
 - aus $k < m$ und $m < n$ folgt: $k < n$ (**transitiv**)
 - für kein $m \in M$ gilt: $m < m$ (**irreflexiv**)

Beispiel:

Ist_Unterbegriff \subseteq DLexeme x DLexeme

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt),
(Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Grundbegriffe zu Relationen

- Eine **lineare Ordnung** (**totale Ordnung**) auf einer Menge M ist eine Halbordnung „ $<$ “ auf M , bei der für beliebige $m, n \in M$ entweder $m < n$ oder $n < m$ oder $m = n$ gilt.

Beispiel:

Lexikographisch_Vor \subseteq DLexeme x DLexeme

Lexikographisch_Vor

= { (Aal, Bild), (Bild, Rand), (Rand, Tiger), (Stelle, Stiel),
(Stiel, Stil), (Stil, Stunde), (Klee, Zone), ... }

Grundbegriffe zu Relationen

- Das **Produkt** zweier Relationen, ρ und σ auf M , ist festgelegt durch

$$\rho \sigma := \{ (x, z) \mid (x, y) \in \rho \text{ und } (y, z) \in \sigma \text{ f.e. } y \in M \}$$

Grundbegriffe zu Relationen

- Für eine beliebige Relation ρ auf M definiert
 - $\rho^0 := \{ (m,m) \mid m \in M \}$ die **Diagonale**,
 - $\rho^1 := \rho$ und $\rho^i := \rho^{i-1} \rho$ für $i > 1$
 - $\rho^+ := \bigcup_{i \geq 1} \rho^i = \rho^1 \cup \rho^2 \cup \dots \cup \rho^n$
die **transitive Hülle** von ρ ,
 - $\rho^* := \bigcup_{i \geq 0} \rho^i = \rho^0 \cup \rho^1 \cup \rho^2 \cup \dots \cup \rho^n$
die **reflexive und transitive Hülle** von ρ

Transitive Hülle von „*Ist_Unterbegriff*“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Transitive Hülle von „*Ist_Unterbegriff*“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

*Ist_Unterbegriff*¹ = *Ist_Unterbegriff*

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff¹ = Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff¹ = Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ) }

Transitive Hülle von „*Ist_Unterbegriff*“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff¹ = Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff¹ = Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff¹ = Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt) }

Transitive Hülle von „*Ist_Unterbegriff*“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff¹ = Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Transitive Hülle von „*Ist_Unterbegriff*“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt),
(KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Ist_Unterbegriff³ = Ist_Unterbegriff² Ist_Unterbegriff

= { (VW-Golf, KFZ) }

Transitive Hülle von „*Ist_Unterbegriff*“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt),
(KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Ist_Unterbegriff³ = Ist_Unterbegriff² Ist_Unterbegriff

= { (VW-Golf, KFZ), (VW-PKW, Artefakt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Ist_Unterbegriff³ = Ist_Unterbegriff² Ist_Unterbegriff

= { (VW-Golf, KFZ), (VW-PKW, Artefakt), (PKW, Objekt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Ist_Unterbegriff³ = Ist_Unterbegriff² Ist_Unterbegriff

= { (VW-Golf, KFZ), (VW-PKW, Artefakt), (PKW, Objekt), (Schreibtisch, Objekt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ),
(KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch,
Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt),
(KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Ist_Unterbegriff³ = Ist_Unterbegriff² Ist_Unterbegriff

= { (VW-Golf, KFZ), (VW-PKW, Artefakt), (PKW, Objekt),
(Schreibtisch, Objekt) }

Ist_Unterbegriff⁴ = Ist_Unterbegriff³ Ist_Unterbegriff

= { (VW-Golf, Artefakt) }

Transitive Hülle von „*Ist_Unterbegriff*“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

*Ist_Unterbegriff*² = *Ist_Unterbegriff*¹ *Ist_Unterbegriff*

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

*Ist_Unterbegriff*³ = *Ist_Unterbegriff*² *Ist_Unterbegriff*

= { (VW-Golf, KFZ), (VW-PKW, Artefakt), (PKW, Objekt), (Schreibtisch, Objekt) }

*Ist_Unterbegriff*⁴ = *Ist_Unterbegriff*³ *Ist_Unterbegriff*

= { (VW-Golf, Artefakt), (VW-PKW, Objekt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Ist_Unterbegriff³ = Ist_Unterbegriff² Ist_Unterbegriff

= { (VW-Golf, KFZ), (VW-PKW, Artefakt), (PKW, Objekt), (Schreibtisch, Objekt) }

Ist_Unterbegriff⁴ = Ist_Unterbegriff³ Ist_Unterbegriff

= { (VW-Golf, Artefakt), (VW-PKW, Objekt) }

Ist_Unterbegriff⁵ = Ist_Unterbegriff⁴ Ist_Unterbegriff

= { (VW-Golf, Objekt) }

Transitive Hülle von „Ist_Unterbegriff“

Ist_Unterbegriff

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt) }

Ist_Unterbegriff² = Ist_Unterbegriff¹ Ist_Unterbegriff

= { (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt) }

Ist_Unterbegriff³ = Ist_Unterbegriff² Ist_Unterbegriff

= { (VW-Golf, KFZ), (VW-PKW, Artefakt), (PKW, Objekt), (Schreibtisch, Objekt) }

Ist_Unterbegriff⁴ = Ist_Unterbegriff³ Ist_Unterbegriff

= { (VW-Golf, Artefakt), (VW-PKW, Objekt) }

Ist_Unterbegriff⁵ = Ist_Unterbegriff⁴ Ist_Unterbegriff

= { (VW-Golf, Objekt) }



Ist_Unterbegriff⁶ = Ist_Unterbegriff⁵ Ist_Unterbegriff = ⁴³{ }

Transitive Hülle von „Ist_Unterbegriff“

$$\bigcup_{i \geq 1} \text{Ist_Unterbegriff}^i = \text{Ist_Unterbegriff}^1 \cup \text{Ist_Unterbegriff}^2 \cup \dots \cup \text{Ist_Unterbegriff}^n$$

= { (VW-Golf, VW-PKW), (VW-PKW, PKW), (PKW, KFZ), (KFZ, Artefakt), (Artefakt, Objekt), (Schreibtisch, Möbel), (Möbel, Artefakt), (VW-Golf, PKW), (VW-PKW, KFZ), (PKW, Artefakt), (KFZ, Objekt), (Schreibtisch, Artefakt), (Möbel, Objekt), (VW-Golf, KFZ), (VW-PKW, Artefakt), (PKW, Objekt), (Schreibtisch, Objekt), (VW-Golf, Artefakt), (VW-PKW, Objekt), (VW-Golf, Objekt) }

Grundlagen formaler Sprachen: Alphabet

- Sei Σ ein beliebiges **Alphabet**, d.i. eine Menge von Symbolen oder Zeichen
 - **Beispiele für verbreitete Alphabete:**
 - {A,B,C, ..., X,Y,Z} lateinisches Alphabet
 - {1,2,3, ..., 7,8,9, 0} indisch-arabisches Zahlensystem
 - {0,1} Binärzahlen
 - {  ,  ,  } internat. Ampelalphabet
 - {A[denin], G[uanin], T[hymin], C[ytosin]} Basen-Alphabet der DNA

Grundlagen formaler Sprachen: Wörter

- Seien **Wörter** (**Sätze**, **Strings**, **Ketten**) über einem Alphabet Σ in der folgenden Weise definiert:
 1. ε ist ein Wort über Σ (ε ist das Leerwort, das keine Symbole hat)
 2. falls χ ein Wort über Σ und $\alpha \in \Sigma$ ist, dann ist $\chi \alpha$ ein Wort über Σ
 3. γ ist ein Wort über Σ genau dann, wenn sein Bildung aus (1) oder (2) folgt

Grundlagen formaler Sprachen: Konkatenation von Wörtern

- Das Wort $\omega \circ \tau := \omega \tau := \omega_1 \dots \omega_m \tau_1 \dots \tau_n$ heißt **Konkatenation** von ω und τ , falls $\omega = \omega_1 \dots \omega_m$ und $\tau = \tau_1 \dots \tau_n$ ($\omega_i, \tau_j \in \Sigma$) Wörter über Σ sind; „o“ (sprich: „Kringel“) ist der Konkatenationsoperator.
 - Für alle Wörter ω gilt: $\omega \circ \varepsilon = \varepsilon \circ \omega = \omega$

Beispiele für Konkatenationen:

- **ABC** o **X** = **ABCX**
- **24** o **24** = **2424**
- **wort** o **stamm** = **wortstamm**

Formale Grundlagen von Automaten: formale Sprache

- Eine (**formale**) **Sprache** \mathcal{L} über einem Alphabet Σ ist eine Menge von Ketten über Σ .
- Sei ferner Σ^* (bzw. Σ^+) die Menge *aller* Ketten über Σ unter Einschluss (bzw. Ausschluss) von ε .
- Dann gilt für jede Sprache \mathcal{L} über Σ :

$$\mathcal{L} \subseteq \Sigma^*$$

Beispiele für formale Sprachen

- $\Sigma = \{A, B, C, \dots, X, Y, Z, _ \}$
 - $\mathcal{L}_1 = \{GUTEN_TAG, GUTEN, TAG\}$
= $\{GUTEN, GUTEN_TAG, TAG\}$
 - $\mathcal{L}_2 = \{GTNTG, GTN, TG\}$
 - $\mathcal{L}_3 = \{TNT, TN, T\}$
 - $\mathcal{L}_4 = \{GG, GTTG, GGGG, GGTTGG, \dots\}$
 - $\mathcal{L}_5 = \{A, C, G, T, ACCGTG, \dots\}$

Beispiele für formale Sprachen

- $\Sigma = \{0, 1, 2, 3, \dots, 7, 8, 9, +, -, =\}$
 - $\mathcal{L}_1 = \{5+7=20-8, 5+7=5+8, 5759=57-59\}$
 - $\mathcal{L}_2 = \{5+=7-2=, +++, 1-11782---3\}$
 - $\mathcal{L}_3 = \{3, 33, 333, 3333, 33333, \dots\}$