Komplexität und natürliche Sprache

Kompetenz versus Performanz

Timm Lichte & Christian Wurm

Heinrich-Heine-Universität Düsseldorf

Sommersemester 2017, 24.05.2017

SFB 991

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Outline

Outline

Kompetenz versus Performanz

Kompetenz	**Performanz**
Was wird charakterisiert	Wie wird charakterisiert
Wissensinhalt	Wissensanwendung
Sprachklassen	Regelmengen von Grammatiken (Kompetenz?)
	Regelanwendungen für ein konkretes Wort

Sprachklassen-Komplexität

$\hat{L}_1$ ist komplexer als $\hat{L}_2$, falls $\hat{L}_1 \supset \hat{L}_2$.

Grammatik-Komplexität (Regelanzahl)

Sei $L(G_1) = L(G_2)$, G_1 ist komplexer als G_2, falls $|\mathcal{P}_1| > |\mathcal{P}_2|$.

Grammatik-Komplexität (Ambiguität)

G_1 ist komplexer als G_2 bezüglich $\overline{w}$, falls $|PTR_1(\overline{w})| > |PTR_2(\overline{w})|$.

Kompetenz versus Performanz (cont.)

Grammatik/Wort-Komplexität (Ableitungslänge)

$\overline{w_1}$ ist komplexer als $\overline{w_2}$ bezüglich G, falls

$$|S \vdash \ldots \vdash \overline{w_1}| > |S \vdash \ldots \vdash \overline{w_2}|$$

Outline

Schwache und starke Äquivalenz von Grammatiken

Gegeben zwei kontextfreie Grammatiken $G_1 = \langle \mathcal{N}_1, \Sigma_1, \mathcal{P}_1, S_1 \rangle$ und $G_2 = \langle \mathcal{N}_2, \Sigma_2, \mathcal{P}_2, S_2 \rangle$.

Schwache Äquivalenz

G_1 und G_2 sind schwach äquivalent gdw. G_1 und G_2 erzeugen dieselbe Sprache.

$$G_1 \equiv_w G_2 \Leftrightarrow L(G_1) = L(G_2)$$

Starke Äquivalenz

G_1 und G_2 sind stark äquivalent gdw. G_1 und G_2 erzeugen dieselbe Sprache mit **denselben** Ableitungen.

$$G_1 \equiv_s G_2 \Leftrightarrow \mathcal{P}_1 = \mathcal{P}_2 \wedge S_1 = S_2$$

Schwache und starke Äquivalenz von Grammatiken (cont.)

Strukturelle Äquivalenz

G_1 und G_2 sind strukturell äquivalent gdw. G_1 und G_2 erzeugen dieselbe Sprache mit den **gleichen** Ableitungen.

$G_1 \equiv_P G_2 \Leftrightarrow$ Es gibt eine Bijektion $f : \mathcal{N}_1 \longrightarrow \mathcal{N}_2$, wobei $f(S_1) = S_2$, und eine Bijektion $g : \mathcal{P}_1 \longrightarrow \mathcal{P}_2$, so dass gilt: Wenn $\alpha \longrightarrow \beta_1 ... \beta_n \in \mathcal{P}_1$ mit $\alpha \in \mathcal{N}_1$ und $\beta_i \in \Sigma_1 \cup \mathcal{N}_1$, dann gibt es $g(\alpha \longrightarrow \beta_1 ... \beta_n) = f(\alpha) \longrightarrow \gamma_1 ... \gamma_n$ mit $\gamma_i = \beta_i$, falls $\gamma_i \in \Sigma_2$, oder $\gamma_i = f(\beta_i)$, falls $\gamma_i \in \mathcal{N}_2$.

Schwache und starke Äquivalenz von Grammatikklassen

Gegeben zwei Grammatikklassen $\hat{G}_1$ und $\hat{G}_2$.

Schwache Äquivalenz

$\hat{G}_1$ und $\hat{G}_2$ sind schwach äquivalent gdw. es gibt eine Bijektion $f : \hat{G}_1 \to \hat{G}_2$, so dass für jedes $G \in \hat{G}_1$ gilt:

$$G \equiv_w f(G)$$

Starke Äquivalenz

$\hat{G}_1$ und $\hat{G}_2$ sind stark äquivalent gdw. es gibt eine Bijektion $f : \hat{G}_1 \to \hat{G}_2$, so dass für jedes $G \in \hat{G}_1$ gilt:

$$G \equiv_s f(G)$$

Äquivalenz von Grammatiken aus verschiedenen Grammatikformalismen?

Normalformen für kontextfreie Grammatiken

Normalform heißt: Die **Regeln** haben eine bestimmte Form.

Chomsky-Normalform (CNF)

$A \to B\ C \mid a$ (mit $A, B, C \in \mathcal{N}$, $a \in \Sigma$)

Normalformen für kontextfreie Grammatiken

Normalform heißt: Die **Regeln** haben eine bestimmte Form.

Chomsky-Normalform (CNF)

$A \longrightarrow B\ C \mid a$ (mit $A, B, C \in \mathcal{N}$, $a \in \Sigma$)

Greibach-Normalform (GNF)

$A \longrightarrow a\ \beta$ (mit $A \in \mathcal{N}$, $a \in \Sigma$, $\beta \in \mathcal{N}^*$)

Normalformen für kontextfreie Grammatiken

Normalform heißt: Die **Regeln** haben eine bestimmte Form.

Chomsky-Normalform (CNF)

$A \longrightarrow B\ C \mid a$ (mit $A, B, C \in \mathcal{N}$, $a \in \Sigma$)

Greibach-Normalform (GNF)

$A \longrightarrow a\ \beta$ (mit $A \in \mathcal{N}$, $a \in \Sigma$, $\beta \in \mathcal{N}^*$)

Lexikalisierung

$A \longrightarrow \alpha\ a\ \beta$ (mit $A \in \mathcal{N}$, $a \in \Sigma$, $\alpha, \beta \in \mathcal{N}^*$)

Kontextfreie Grammatiken in CNF und GNF und lexikalisierte kontextfreie Grammmatiken sind **schwach** äquivalent zu kontextfreien Grammatiken.

Normalformen für kontextfreie Grammatiken (cont.)

Nicht schwach äquivalent zu kontextfreien Grammatiken:

lineare CFG

$A \longrightarrow a\ B \mid B\ a \mid a$ (mit $A, B \in \mathcal{N}$, $a \in \Sigma$)

oder

$A \longrightarrow a\ B\ b \mid a$ (mit $A, B \in \mathcal{N}$, $a, b \in \Sigma$)

Normalformen für kontextfreie Grammatiken (cont.)

Nicht schwach äquivalent zu kontextfreien Grammatiken:

lineare CFG

$A \longrightarrow a\ B\ |B\ a\ |a$ (mit $A, B \in \mathcal{N}$, $a \in \Sigma$)

oder

$A \longrightarrow a\ B\ b\ |a$ (mit $A, B \in \mathcal{N}$, $a, b \in \Sigma$)

Lineare CFGs können nicht die Dyck-Sprachen erzeugen.

$S \longrightarrow \epsilon|SS|(S)$

Ordentliche ("proper") CFGs[8]

Eine kontextfreie Grammatik $G = \langle \mathcal{N}, \Sigma, \mathcal{P}, S \rangle$ ist ordentlich gdw.:

- keine unerreichbare Nichtterminale

$$\forall A \in \mathcal{N} \colon \exists \alpha, \beta \in (\Sigma \cup \mathcal{N})^* \colon S \vdash_G^* \alpha A \beta$$

- keine unproduktive Nichtterminale

$$\forall A \in \mathcal{N} \colon \exists \overline{w} \in \Sigma^* \colon A \vdash_G^* \overline{w}$$

- keine ϵ-Produktionen

$$\nexists A \in \mathcal{N} \colon A \vdash_G \epsilon \quad \text{oder}$$

$$\mathcal{P} \subseteq \mathcal{N} \to (\Sigma \cup \mathcal{N} \setminus \{S, \epsilon\})^+ \cup \{S \to \epsilon\}$$

- keine Zyklen

$$\nexists A \in \mathcal{N} \colon A \vdash_G^+ A$$

Ordentliche (“proper”) CFGs

Die Klasse der ordentlichen kontextfreien Grammatiken ist schwach äquivalent zur Klasse der kontextfreien Grammatiken.

- Unerreichbare und unproduktive Nichtterminale können eliminiert werden. [5: §7.1.1]
- ϵ-Produktionen können eliminiert werden. [5: §7.1.3]
- Zyklen (allgemeiner “unit pairs”) können eliminiert werden. [5: §7.1.4]

Ordentliche ("proper") CFGs

Die Klasse der ordentlichen kontextfreien Grammatiken ist schwach äquivalent zur Klasse der kontextfreien Grammatiken.

- Unerreichbare und unproduktive Nichtterminale können eliminiert werden. [5: §7.1.1]
- ϵ-Produktionen können eliminiert werden. [5: §7.1.3]
- Zyklen (allgemeiner "unit pairs") können eliminiert werden. [5: §7.1.4]

Aber

Durch die Eliminierung der ϵ-Produktionen und der Zyklen (bzw. "unit pairs") kann $\mathcal{P}$ erheblich größer werden.

Outline

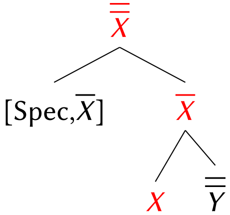
X-bar-Schema

Chomsky (1970)

“To introduce a more uniform notation, let us use the symbol $\overline{X}$ for a phrase containing X as its head. Then the base rules introducing N, A, and V will be replaced by a schema (48), where in place of … there appears the full range of structures that serve as complements and X can be any one of N, A, or V:

(48) $\overline{X} \rightarrow X \ldots$

Continuing with the same notation, the phrases immediately dominating $\overline{N}$, $\overline{A}$ and $\overline{V}$ will be designated $\overline{\overline{N}}$, $\overline{\overline{A}}$, $\overline{\overline{V}}$ respectively.“



Anwendungsbeispiel

ϵ-freie Regeln für NPs

$NP \rightarrow$ *Alpakas* | *die Alpakas* | *freche Alpakas* | *die frechen Alpakas* | *die hungrigen frechen Alpakas* | ...

Anwendungsbeispiel

ϵ-freie Regeln für NPs

$NP \longrightarrow$ *Alpakas* | *die Alpakas* | *freche Alpakas* | *die frechen Alpakas* | *die hungrigen frechen Alpakas* | ...

x-bar-Regeln für NPs

?

Weiterentwicklung in Jackendoff (1977) und Stowell (1981).

Kritische und formal adequate Besprechung in Kornai & Pullum (1990).

X-bar-Schema nach Kornai & Pullum (1990)

Lexicality

A CFG observes Lexicality iff every nonterminal is X^i, where $X \in V_T$ and $i > 0$.

Succession

A Lexicality observing CFG observes Succession iff every rule rewriting some nonterminal X^n has a daughter labeled X^{n-1}.

Weak Succession

A Lexicality observing CFG observes Weak Succession iff every rule rewriting some nonterminal X^n has a daughter labeled X^{n-1} or X^n.

Uniformity

A Lexicality-observing CFG observes Uniformity iff $\exists m \in N[V_N = \{X^i | 1 \leq i \leq m, X \in V_T\}]$

X-bar-Schema nach Kornai & Pullum (1990) (cont.)

Maximality

A CFG observing Lexicality and Succession observes Maximality iff for every rule $X^n \rightarrow YX^{n-1}Z$, the strings Y and Z are in V_M^*, where $V_M = \{X^m | X \in V_T\}$.

Centrality

A Lexicality-observing CFG observes Centrality iff the start symbol is the maximal projection of a distinguished preterminal.

Peripherality

A Lexicality-observing CFG observes Peripherality iff in any rule rewriting X^1 as YX^0Z, either $Y = \epsilon$ or $Z = \epsilon$.

X-bar-Schema nach Kornai & Pullum (1990) (cont.)

Optionality

A CFG $G = \langle V_N, V_T, P, S \rangle$ observes Optionality iff for every rule in P of the form $\alpha \longrightarrow W$ there exist β, W_1, W_2 such that

1. $\beta \in (V_N \cup V_T)$;
2. $W_1, W_2 \in (V_N \cup V_T)^*$;
3. $W = W_1 \beta W_2$; and
4. the rule $\alpha \longrightarrow W_1' \beta W_2'$ in P for all strings W_1' and W_2' constructible by deleting some elements from W_1 and W_2, respectively. In such rules, β will be called the CORE.

Weniger Komplexität? (Kornai & Pullum 1990)

Standard X-bar Grammar (SXBG)

A Standard X-bar Grammar (SXBG) is a Lexicality-observing CFG in which the rules have the form $X^n \longrightarrow YX^{n-1}Z$, where $Y, Z \in V_M^*$.

Falls ϵ-frei: Ja! (wenige endliche Sprachen)
Beispiel: SXBG mit genau einem Präterminal σ:

- Sei $G = \langle\{\sigma^1\}, \{\sigma\}, \{\sigma^1 \longrightarrow \sigma\}, \sigma^1\rangle$, dann $L(G) = \{\sigma\}$.
- Sei $G = \langle\{\sigma^1\}, \{\sigma\}, \{\sigma^1 \longrightarrow \sigma^1\sigma, \sigma^1 \longrightarrow \sigma\}, \sigma^1\rangle$, dann $L(G) = \{\sigma^n | n > 0\}$.

Standard X-bar Grammar mit ϵ-Produktionen (SXBG$^\epsilon$)

A Standard X-bar Grammar (SXBG) is a Lexicality-observing CFG in which the rules have the form $X^n \longrightarrow YX^{n-1}Z$, where $Y, Z \in V_M^*$, or $X^0 \longrightarrow \epsilon$.

Falls **nicht** ϵ-frei: Nein! (schwach äquivalent zu CFG)

Weniger Komplexität? (Kornai & Pullum 1990) (cont.)

CAVEAT

- X-bar-Schema wird sehr unterschiedlich eingesetzt.
- Optionalität senkt die Ausdrucksstärke, wird aber kaum verwendet.
- Implizite Grundannahme scheint zu sein: X^n mit $n > 0$ sind derivationell nicht ambig,
- Nur Kompetenz, d.h. Sprachklassen-Komplexität, wird betrachtet.

Was die Regelanzahl und die Ableitungslänge betrifft, sind SXBGs sicherlich **nicht** optimal.

Anliegen ist, die **Ableitungsambiguität** (auch grammatikübergreifend) zu minimieren.

Outline

Derivational Theory of Complexity (DTC)

Vermutung bezogen auf die Transformationsgrammatik

Verarbeitungskomplexität korreliert mit der Anzahl der nötigen Transformationen.

Grammatik/Wort-Komplexität (Ableitungslänge)

$\overline{w_1}$ ist komplexer als $\overline{w_2}$ bezüglich G, falls

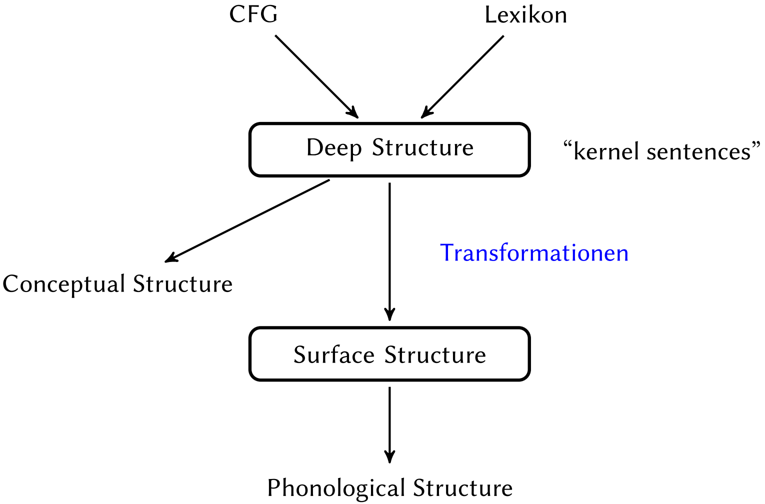
$$|S \vdash \ldots \vdash \overline{w_1}| > |S \vdash \ldots \vdash \overline{w_2}|$$

Hoffnung auf

- explizite Performanzmodelle
- Adäquatheitstests für linguistische Theorien

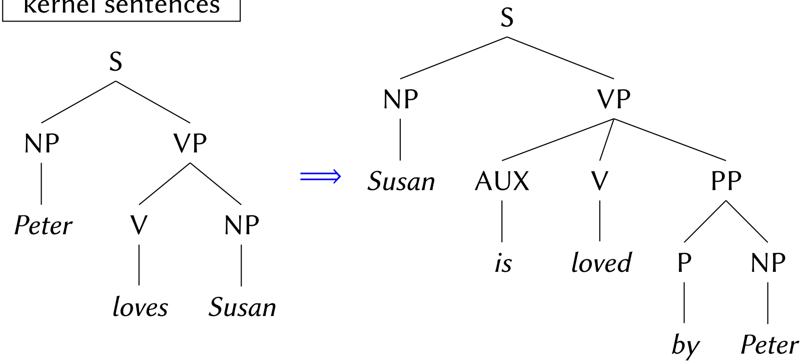
Transformationsgrammatik

Chomsky (1957; 1965)



Chomsky (1957; 1965)

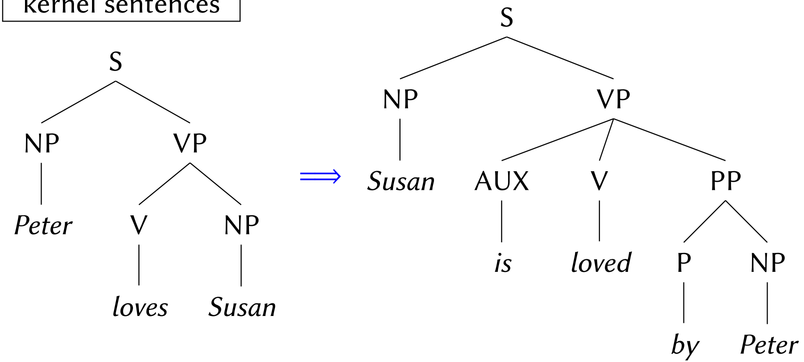
“kernel sentences”



$$\boxed{NP_1 \ V \ \ NP_2 \longrightarrow NP_2 \ \ AUX \ \ V \ \ by \ \ NP_1}$$

Chomsky (1957; 1965)

“kernel sentences”



$$\boxed{NP_1\ \ V\ \ NP_2 \longrightarrow NP_2\ \ AUX\ \ V\ \ by\ \ NP_1}$$

- Komplexitätstheoretisch sehr problematisch: Das Wortproblem für TG ist **unentscheidbar** (Peters & Ritchie 1973).

Kontra DTC

Gegenevidenz (Fodor, Bever & Garrett 1974):

Passiv

(1) a. The first shot the tired soldier the mosquito bit fired missed.
    b. The first shot fired by the tired soldier bitten by the mosquito missed.

Kontra DTC

Gegenevidenz (Fodor, Bever & Garrett 1974):

Passiv

(1) a. The first shot the tired soldier the mosquito bit fired missed.
    b. The first shot fired by the tired soldier bitten by the mosquito missed.

Ellipse

(2) a. John swims faster than Bob swims.
    b. John swims faster than Bob.
    c. John swims faster than Bob does.

Kontra DTC

Gegenevidenz (Fodor, Bever & Garrett 1974):

Passiv

(1) a. The first shot the tired soldier the mosquito bit fired missed.
    b. The first shot fired by the tired soldier bitten by the mosquito missed.

Ellipse

(2) a. John swims faster than Bob swims.
    b. John swims faster than Bob.
    c. John swims faster than Bob does.

Hängt stark von der Analyse ab![10]

Zusammenfassung

- Unterscheidung von “Kompetenz” und “Performanz”
- “Performanz” bezieht sich auf Regeln.
- Äquivalenz und Normalformen
- regelbezogenen Komplexitätsbegriffe (Regelanzahl, Ambiguität, Ableitungslänge)
- Beispiele aus der formalen Linguistik: X-bar-Syntax und DTC

Nächste Sitzung: andere wichtige performanzbezogene Komplexitätsbegriffe (Automatentheorie, Raum- und Zeitkomplexität)

Literaturverzeichnis

[1] Chomsky, Noam. 1957. *Syntactic structures.* Den Haag: Mouton.

[2] Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, MA: The MIT Press.

[3] Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.

[4] Fodor, Jerry A., Thomas G. Bever & Merrill F. Garrett. 1974. *The psychology of language: an introduction to psycholinguistics and generative grammar.* McGraw-Hill.

[5] Hopcroft, John E., Rajeev Motwani & Jeffrey D. Ullman. 2001. *Introduction to automata theory, languages and computation.* Addison-Wesley.

[6] Jackendoff, Ray. 1977. *X′ syntax: A case study of phrase structure.* Cambridge, MA: MIT Press.

[7] Kornai, András & Geoffrey K. Pullum. 1990. The x-bar theory of phrase structure. English. *Language* 66(1). 24–50. `http://www.jstor.org/stable/415278`.

[8] Nijholt, Anton. 1980. *Context-free grammars: Covers, normal forms, and parsing.* (Lecture Notes in Computer Science 93). Berlin: Springer.

[9] Peters, Stanley P. & Robert W. Ritchie. 1973. On the generative power of transformational grammar. *Information Sciences* 6. 49–83.

[10] Phillips, Colin. 1996. *Order and structure.* Cambridge, MA: MIT Doctoral dissertation. `http://dspace.mit.edu/handle/1721.1/10666`.

[11] Stowell, Timothy A. 1981. *Origins of phrase structure.* Cambridge, MA: MIT MIT dissertation.