
Formale Sprachen und Automaten

Motivation und Hintergrund

Formale Sprachen: der Gegenstand

- Kann man die Begriffe der **Sprache** und der **Grammatik** auf eine mathematische Grundlagen stellen?
- “Die Theorie formaler Sprachen, [...], betrachtet Sprachen als Mengen von Zeichenreihen oder Wörtern und befasst sich mit formalen Methoden zu ihrer Erkennung und Erzeugung.”
(Loos 1993, 1)
- “The basic assumption made in generative grammar is that languages can be regarded as collections whose membership is definitely and precisely specifyable.”
(Gazdar, Klein, Pullum & Sag 1985, 1)

Formale Sprachen und Informatik

- “Die Anfänge der theoretischen Informatik lagen in verschiedensten Bereichen: Biologen untersuchten Modelle für Neutronen-Netze, Elektroingenieure entwickelten die Schaltkreistheorie [...], Mathematiker arbeiteten an den Grundlagen der Logik und Linguisten untersuchten die Grammatiken für natürliche Sprachen.”

(Hopcroft & Ullman 1990, 9)

- “Eines [sic] der Haupteinsichten der Theorie [Berechenbarkeitstheorie] ist es, dass es nicht-berechenbare Funktionen (bzw. algorithmisch unlösbare Probleme) gibt.”

(Schöning 1992, 9)

Formale Sprachen und Linguistik

- “In brief, mathematical study of formal properties of grammars is, very likely, an area of linguistics of great potential. It has already provided insight into questions of empirical interest and will perhaps some day provide much deeper insights.”

(Chomsky 1964, 62)

- “Much work has been devoted [...] to the question of constraining [...] grammars for natural languages. But [...] this work has been free of serious mathematical content. In view of this fact, claims in the current literature implying that one variant of TG [transformational grammar] [...] is more restrictive than another [...] are about as sensible as claims to the effect that Turing machines which employ narrow grey tape are less powerful than ones employing wide orange tape.”

(Gazdar 1982, 131)

Formale Sprachen und Linguistik 2

- “Ebenso verfehlt ist es, natürliche Grammatiken von vornherein auf ein bestimmtes Format zu trimmen, etwa das kontextfreie, mit der Begründung, daß jede solche Grammatik als Entscheidungsverfahren zu benutzen ist. [...] das muss sich aus empirischen, nicht apriorischen Gründen ergeben.”

(Stechow & Sternefeld 1988, 56)

- “Diese Fragen dürfen deshalb nicht aufgrund irgendwelcher mathematischer Vorüberlegungen beantwortet werden, etwa von dem Gesichtspunkt aus, daß kontextfreie Grammatiken den Vorzug haben, daß man leicht Analysealgorithmen für sie konstruieren kann. [...] so argumentiert Gazdar in mehreren Schriften – eine fehlgeleitete Methodologie, wie wir heute meinen.”

(Stechow & Sternefeld 1988, 102)

Formale und natürliche Sprachen

- Gazdar (1982) motiviert, wieso Linguisten etwas über formale Sprachen wissen sollten.
- Postal (1964) versucht anhand von Mohawk zu zeigen, dass natürliche Sprachen nicht **kontextfrei** sein können.
- Pullum & Gazdar (1982) behaupten, dass alle bisherigen Argumente (u.a. das von Postal), welche Nicht-Kontextfreiheit zeigen sollten, fehlerhaft sind.
- Higginbotham (1984) und Postal & Langedoen (1984) versuchen anhand des Englischen neue Argumente zu konstruieren.
- Pullum (1984) attackiert diese Versuche.
- Shieber (1985) und Culy (1985) schließlich bringen wieder neue Argumente (aus dem Schweizerdeutschen und aus Bambara).

Quellenangaben

- Der Stoff der in dieser Veranstaltung benutzten Folien speist sich aus folgenden Quellen:
 1. Bhatt & Kuhlmann (2002)
 2. Hopcroft & Ullman (1990)
 3. Partee, ter Meulen & Wall (1990)
 4. Schöning (1992)

Literatur

- Bhatt, Rajesh & Gregory Kuhlmann (2002): Automata Theory. Vorlesungsmanuskript, University of Texas, Austin.
- Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Massachusetts.
- Culy, Christopher (1985): 'The Complexity of the Vocabulary in Bambara', *Linguistics and Philosophy* **8**, 345–351.
- Gazdar, Gerald (1982): Phrase Structure Grammar. *In: P. Jacobson & G. Pullum, eds, The Nature of Syntactic Representation*. Reidel, Dordrecht, pp. 131–186.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum & Ivan Sag (1985): *Generalized Phrase Structure Grammar*. Blackwell, Oxford.
- Higginbotham, James (1984): 'English is not a Context-Free Language', *Linguistic Inquiry* **15**, 225–234.

Hopcroft, John E. & Jeffrey D. Ullman (1990): *Einführung in die Automatentheorie, formale Sprachen und Komplexitätstheorie*. Addison-Wesley.

Loos, Rüdiger (1993): *Formale Sprachen – Theorie und Anwendung*. Vorlesungsmanuskript, Universität Tübingen.

Partee, Barbara H., Alice ter Meulen & Robert E. Wall (1990): *Mathematical Methods in Linguistics*. Vol. 30 of *Studies in Linguistics and Philosophy*, Kluwer, Dordrecht.

Postal, Paul (1964): Limitations of Phrase Structure Grammars. In: J. A. Fodor & J. J. Katz, eds, *The Structure of Language: Readings in the Philosophy of Language*. Prentice Hall, Englewood Cliffs, New Jersey, pp. 137–151.

Postal, Paul & D. Terrence Langedoen (1984): 'English and the Class of Context-Free Languages', *Computational Linguistics* **10**, 177–181.

Pullum, Geoffrey (1984): 'On Two Recent Attempts to Show that English is not a CFL', *Computational Linguistics* **10**, 182–186.

Pullum, Geoffrey & Gerald Gazdar (1982): 'Natural Languages and Context-Free Languages', *Linguistics and Philosophy* **4**, 471–504.

Schöning, Uwe (1992): *Theoretische Informatik kurz gefasst*. Wissenschaftsverlag, Mannheim.

Shieber, Stuart (1985): 'Evidence Against the Context-freeness of Natural Language', *Linguistics and Philosophy* **8**, 333–343.

Stechow, Arnim von & Wolfgang Sternefeld (1988): *Bausteine syntaktischen Wissens*. Westdeutscher Verlag, Opladen.