Kapitel 6

Baumsprachen und Logik

6.1 Motivation

Bei der Beschreibung von formalen Sprachen haben wir (wenigstens) drei zunächst grundverschiedene Begriffsapparate zur Verfügung.

1. Automaten: wir beschreiben Algorithmen, die Wortmengen erzeugen oder akzeptieren
2. Algebra: wir beschreiben Eigenschaften von Wortmengen, z. B. durch (reguläre) Ausdrücke oder Gleichungen (Constraints) oder Abbildungen (Homomorphismen) in andere Strukturen (endliche Algebren, Monoide)
3. Logik: wir betrachten Wortmengen als Modellmengen von bestimmten Formeln

Für *reguläre* (Baum- und Wort-)Sprachen benutzen wir insbesondere

1. endliche Automaten
2. reguläre Ausdrücke
3. monadische Logik zweiter Stufe

Die Beschreibungskraft dieser Modelle stimmt überein.

Wir begeben uns in diesem Kapitel in die Logik. Den logischen Formeln (an deren Modellmengen wir interessiert sind) sieht man nun kaum noch an, in welchem Universum sie zu interpretieren sind. Das erleichtert das Übertragen von Resultaten für Wortsprachen auf Baumsprachen, oder gibt doch wenigstens die dabei einzuschlagende Richtung an.

Wir verfolgen diesen Weg hin zu Teilklassen regulärer Sprachen (definierbar erster Ordnung, aperiodisch, sternfrei).

Der Text dieses Kapitels besteht derzeit nur aus Stichpunkten, die als Anhaltspunkte beim Studium der Literatur dienen mögen.

6.2 Inhalt

6.3 Logische Klassifikation regulärer Sprachen

- D. Perrin: Finite Automata,
  in: Handbook of Theoretical Computer Science, Elsevier, 1990
- J.-E. Pin: Syntactic Semigroups,
  in: Handbook of Formal Languages, Springer, 1996
- A. Potthoff: Logische Klassifizierung regulärer Baumsprachen,
  Diss., Univ. Kiel, 1994
- W. Thomas: Languages, Automata, and Logic,
  in: Handbook of Formal Languages, Springer, 1996

6.4 Sprachen, Automaten, *Logik*

- Universum: Positionen in Wort
  (Baum, Graph, ...)

- Relationen $<, <_1$ für Positionen
- (für Bäume: $<_1, <_2, \ldots, <_d$)
- für alle $x \in \Sigma$: Prädikat $Q_x(p)$

- Variablen für Positionen
  (FO – first order)
- und Mengen von Positionen
  (MSO – monadic second order)
- Relationen $x \in X$

6.5 Beispiele

- FO: alle Wörter, die rechts von jedem a ein b haben
- $<$ ist MSO-definierbar aus $<_1$
- MSO: $(aa)^*$

6.6 MSO_0

- Elemente $\rightarrow$ Einermengen

6.7 Büchi 1960

L erkennbar durch endlichen Automaten
$\iff$ L definierbar in MSO($<_1$).

Alphabet erweitern: $\Sigma \times \{0,1\}^k$
(k: Anzahl der Variablen/Zustände)

Korollar (Quantor-Elimination):
Jede MSO($<_1$)-Formel hat EMSO($<_1$)-Form.

Das gleiche für Bäume
(Thatcher, Wright, Doner, 1968-70)

6.8 allgemeine ESO, Fagin 1974

L definierbar in ESO $\iff$ $L \in \mathbf{NP}$

6.9 FO($<$)

$<_1$ ist durch $<$ definierbar

- Boolesche Operationen
- Verkettung
- kein $*$

6.10 Sternfreie Ausdrücke

- $\emptyset$, ϵ, x (für all $x \in \Sigma$)
- Komplement, $\cap$, $\cup$
- Verkettung

- McNaughton, Papert 1972:
  FO($<$) $\iff$ sternfrei.
- ist Sternfreiheit entscheidbar?

6.11 Syntaktisches Monoid einer Sprache L

vergleiche: Rechtskongruenz
$x \sim_L y \iff \forall u : ux \in L \leftrightarrow uy \in L$

endlich viele Äq.-Klassen (= Zustände
des vollst. min. det. Automaten für L)

jetzt: syntaktische Kongruenz rechts und links
$x \approx_L y \iff \forall u, w : uxw \in L \leftrightarrow uyw \in L$

endlich viele Äq.-Klassen
aus v.m.d. Aut. von L zu berechnen

diese bilden Monoid (unter $\cdot$)

"genauso" für Bäume: $\approx_L$ auf Kontexten!

6.12 Aperiodizität

H eine Halbgruppe.

- $x \in H$ *idempotent*: $x^2 = x$
- H endlich $\Rightarrow \forall x : \exists k : x^k$ ist idemp.
- (bezeichnen dieses x^k mit x^ω)
- H *aperiodisch*: $\forall x : x^\omega = x^{\omega+1}$

Schützenberger 1965:
aperiodisch $\iff$ sternfrei

6.13 Erweiterte Sternhöhe

- (A)periodizität entscheidbar
- es gibt periodische Sprachen ($\geq$ ein $*$)
- *offen*: ex. Sprachen mit Sternhöhe > 1?

6.14 Das gleiche für Bäume?

Kleene-Stern für Bäume

Thomas 1984:
FO($<$) $\Rightarrow$ aperiodisch $\Rightarrow$ sternfrei

Potthoff 1994: jede reguläre Baumsprache
über Σ ohne unäre Symbole ist sternfrei.

offen: Entscheidbarkeit von:
"ist L FO-definierbar?"

6.15 Straubing-Hierarchie f. aperiod. Spr.

Halbring (Sprachen, $\cup, \cdot$)
vgl. Ring (Polynome, $+, \cdot$)

M eine Menge von Sprachen:

- polynomiale Hülle $\mathbf{PC}(M)$:
  $\{\bigcup L_0 a_1 L_1 \dots a_n L_n \mid a_i \in \Sigma, L_i \in M\}$
- Boolesche Hülle $\mathbf{BC}(M)$:
  Komplement, $\cup$, $\cap$

Straubing-Hierarchie $M_0 = \{\emptyset, \Sigma^*\}$

$M_{k+1/2} = \mathbf{PC}(M_k)$; $M_{k+1} = \mathbf{BC}(M_{k+1/2})$

- Hierarchie ist echt
- schöpft die aperiodischen Sprachen aus

6.16 Beginn der Straubing-Hierarchie

Stufe 0: trivial: $\emptyset, \Sigma^*$

Stufe 1/2: endl. Vereinigung von
$\Sigma^* a_1 \Sigma^* \dots a_n \Sigma^*$.
($\Rightarrow$ kleine Schleifen in Automaten!)

Stufe 1: Boolesche Kombinationen von 1/2.

endliche Sprachen $\subseteq$ Stufe 1.

Stufe 3/2: endl. Vereinigung von
$A_0^* a_1 A_1^* \dots a_n A_n^*$ mit $A_i \subseteq \Sigma$.
($\Rightarrow$ kleine Schleifen in Automaten!)

bis hierher auch entscheidbare Charakterisierungen
durch geordnetes syntakisches Monoid.

offen: wie weiter? (z. b. Stufe 2)

offen: und für Bäume?
(kleine Schleifen in Baum-Automaten)