Fundamentele Informatica 3

voorjaar 2012

http://www.liacs.nl/home/rvvliet/fi3/

Rudy van Vliet
kamer 124 Snellius, tel. 071-527 5777
rvvliet(at)liacs.nl

college 11, maandag 16 april 2012

8. Recursively Enumerable Languages
8.3. More General Grammars
8.4. Context-Sensitive Languages and the Chomsky Hierarchy

Huiswerkopgave 3,
inleverdatum 16 april 2012, 13:45 uur

Definition 8.1. Accepting a Language and Deciding a Language

A Turing machine T with input alphabet Σ accepts a language $L \subseteq \Sigma^*$,
if $L(T) = L$.

T *decides* L,
if T computes the characteristic function $\chi_L : \Sigma^* \to \{0, 1\}$

A language L is *recursively enumerable*,
if there is a TM that accepts L,

and L is *recursive*,
if there is a TM that decides L.

8.3. More General Grammars

reg. languages	reg. grammar	FA	reg. expression
cf. languages	cf. grammar	PDA	
re. languages	unrestr. grammar	TM	

Definition 8.10. Unrestricted grammars

An unrestricted grammar is a 4-tuple $G = (V, \Sigma, S, P)$, where V and Σ are disjoint sets of variables and terminals, respectively, S is an element of V called the start symbol, and P is a set of productions of the form

$$\alpha \to \beta$$

where $\alpha, \beta \in (V \cup \Sigma)^*$ and α contains at least one variable.

Notation as for CFGs:

$$\alpha \Rightarrow_G^* \beta$$

$$L(G) = \{x \in \Sigma^* \mid S \Rightarrow_G^* x\}$$

but...

Example 8.11. A Grammar Generating $\{a^{2^k} \mid k \in \mathbb{N}\}$

$\{a, a^2, a^4, a^8, a^{16}, \ldots\} = \{a, aa, aaaa, aaaaaaaa, aaaaaaaaaaaaaaaa, \ldots\}$

Example 8.11. A Grammar Generating $\{a^{2^k} \mid k \in \mathbb{N}\}$

$\{a, a^2, a^4, a^8, a^{16}, \ldots\} = \{a, aa, aaaa, aaaaaaaa, aaaaaaaaaaaaaaaa, \ldots\}$

$S \to LaR \quad L \to LD \quad Da \to aaD \quad DR \to R \quad L \to \Lambda \quad R \to \Lambda$

Example 8.12. A Grammar Generating $\{a^nb^nc^n \mid n \geq 1\}$

Example 8.12. A Grammar Generating $\{a^nb^nc^n \mid n \geq 1\}$

$$S \to SABC \mid LABC$$

$$BA \to AB \quad CB \to BC \quad CA \to AC$$

$$LA \to a \quad aA \to aa \quad aB \to ab \quad bB \to bb \quad bC \to bc \quad cC \to cc$$

Theorem 8.13.
For every unrestricted grammar G, there is a Turing machine T with $L(T) = L(G)$.

Exercise.
Use (the second part of) the construction from Theorem 8.13 to obtain a TM simulating a derivation in the unrestricted grammar with productions

$$S \to aBS \mid \Lambda \quad aB \to Ba \quad Ba \to aB \quad B \to b$$

Theorem 8.14.
For every Turing machine T with input alphabet Σ, there is an unrestricted grammar G generating the language $L(T) \subseteq \Sigma^*$.

Proof.

1. Generate (every possible) input string for T.
2. Simulate computation of T for this input string as derivation in grammar.
3. If T reaches accept state, reconstruct original input string.

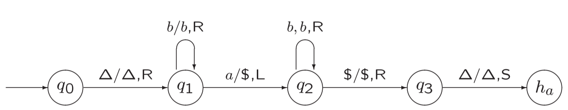
Notation:

description of tape contents: $x\underline{\sigma}y$ or $x\underline{y}$

configuration $xqy = xqy\Delta = xqy\Delta\Delta$

initial configuration corresponding to input x: $q_0\Delta x$

In the third edition of the book, a configuration is denoted as $(q, x\underline{y})$ or $(q, x\underline{\sigma}y)$ instead of xqy or $xq\sigma y$.
This old notation is also allowed for Fundamentele Informatica 3.



8.4. Context-Sensitive Languages and the Chomsky Hierarchy

reg. languages	reg. grammar	FA	reg. expression
cf. languages	cf. grammar	PDA	
cs. languages	cs. grammar	LBA	
re. languages	unrestr. grammar	TM	

Definition 8.16. Context-Sensitive Grammars
A *context-sensitive grammar* (CSG) is an unrestricted grammar in which no production is length-decreasing.
In other words, every production is of the form $\alpha \to \beta$, where $|\beta| \geq |\alpha|$.

A language is a context-sensitive language (CSL) if it can be generated by a context-sensitive grammar.

Example 8.12. A Grammar Generating $\{a^n b^n c^n \mid n \geq 1\}$

$$S \to SABC \mid LABC$$

$$BA \to AB \quad CB \to BC \quad CA \to AC$$

$$LA \to a \quad aA \to aa \quad aB \to ab \quad bB \to bb \quad bC \to bc \quad cC \to cc$$

Not context-sensitive.

Example 8.17. A CSG Generating $L = \{a^n b^n c^n \mid n \geq 1\}$

$$S \to SABC \mid \mathcal{A}BC$$

$$BA \to AB \quad CB \to BC \quad CA \to AC$$

$$\mathcal{A} \to a \quad aA \to aa \quad aB \to ab \quad bB \to bb \quad bC \to bc \quad cC \to cc$$

Definition 8.10. Linear-Bounded Automata
A *linear-bounded automaton* (LBA) is a 5-tuple $M = (Q, \Sigma, \Gamma, q_0, \delta)$ that is identical to a nondeterministic Turing machine, with the following exception.

There are two extra tape symbols [and], assumed not to be elements of the tape alphabet Γ.
The initial configuration of M corresponding to input x is $q_0[x]$, with the symbol [in the leftmost square and the symbol] in the first square to the right of x.
During its computation, M is not permitted to replace either of these brackets or to move its tape head to the left of the [or to the right of the].

Theorem 8.19.
If $L \subseteq \Sigma^*$ is a context-sensitive language, then there is a linear-bounded automaton that accepts L.

Proof. . .