Teoria da Computação

Aula Teórica 21
Linguagens independentes de contexto:
Autómatos de Pilha e Lema da Bombagem

António Ravara

Departamento de Informática

17 de Maio de 2016

Modelo Operacional de Linguagens Indep. de Contexto

- Autómatos Finitos Deterministas reconhecem exactamente Linguagens Regulares.
- Existe também um modelo computacional executável que reconhece as Linguagens Independentes de Contexto.
- Autómatos de Pilha (Deterministas) reconhecem Linguagens Independentes de Contexto (Deterministas).
- É necessário ter memória para guardar (partes da) palavra que está a executar.

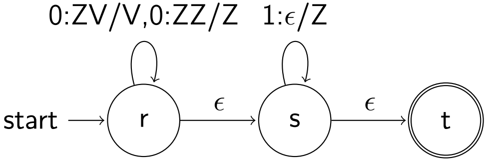
Considere-se a linguagem $\{0^n 1^n \mid n \in \mathbb{N}_0\}$

- Antes de começar a executar os 1s é preciso saber quantos 0s foram executados.
- Guarda-se cada 0 numa estrutura auxiliar (uma pilha) e vão-se consumindo à medida que se executam os 1s.

Um autómato

- Começa-se no estado inicial com a pilha vazia.
- Enquanto se leem 0s não se sai do estado inicial mas vão-se guardando na pilha.
- Para ler 1s passa-se (por ϵ) para outro estado.
- Leem-se 1s enquanto houver 0s na pilha (ler 1 retira um 0).
- Com a pilha estiver vazia, passa-se (por ϵ) para o estado final.

Representação gráfica do autómato

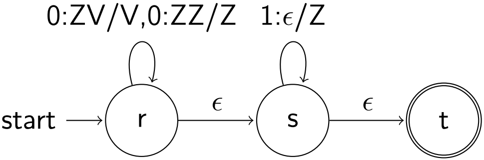


Como "processar" (aceitar) 0011:

$$
\begin{array}{rcl}
\langle r, V, |0011\rangle & \xrightarrow{0{:}ZV/V} & \langle r, ZV, 0|011\rangle \\
& \xrightarrow{0{:}ZZ/Z} & \langle r, ZZV, 00|11\rangle \\
& \xrightarrow{\epsilon} & \langle s, ZZV, 00|11\rangle \\
& \xrightarrow{1{:}\epsilon/Z} & \langle s, ZV, 001|1\rangle \\
& \xrightarrow{1{:}\epsilon/Z} & \langle s, V, 0011|\rangle \\
& \xrightarrow{\epsilon} & \langle t, V, 0011|\rangle
\end{array}
$$

Aceita a palavra porque a pilha está vazia e está no estado final.

Palavras não aceites pelo autómato

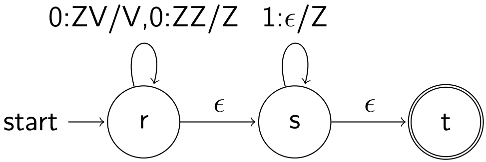


011:

$$
\begin{array}{ccl}
\langle r, V, |011\rangle & \xrightarrow{0:ZV/V} & \langle r, ZV, 0|11\rangle \\
& \xrightarrow{\epsilon} & \langle s, ZV, 0|11\rangle \\
& \xrightarrow{1:\epsilon/Z} & \langle s, V, 01|1\rangle \\
& \not\xrightarrow{1:\epsilon/Z} &
\end{array}
$$

Não aceita a palavra porque a pilha está vazia, mas não consegue chegar ao estado final.

Palavras não aceites pelo autómato



001:

$$
\begin{aligned}
\langle r, V, |001\rangle &\xrightarrow{0:ZV/V} \langle r, ZV, 0|01\rangle \\
&\xrightarrow{0:ZZ/Z} \langle r, ZZV, 00|1\rangle \\
&\xrightarrow{\epsilon} \langle s, ZZV, 00|1\rangle \\
&\xrightarrow{1:\epsilon/Z} \langle s, ZV, 001|\rangle \\
&\xrightarrow{\epsilon} \langle t, ZV, 001|\rangle
\end{aligned}
$$

Não aceita a palavra porque a pilha não está vazia, apesar de estar no estado final.

Autómato de Pilha

Considera-se o séptuplo $A = \langle S, \Sigma, \Gamma, s, Z, \delta, F \rangle$, sendo:

Definição

- S um conjunto finito, dos *estados* de A (memória interna).
- Σ um conjunto finito, das *acções* de A (o seu *alfabeto*).
- Γ um conjunto finito (o *alfabeto* da pilha, memória externa do autómato).
- $s \in S$ é o *estado inicial* de A.
- $Z \in \Gamma$ é o *símbolo inicial* da pilha de A.
- $\delta \subseteq S \times (\Sigma \cup \{\epsilon\}) \times \Gamma \times S \times \Gamma^*$ é a *relação* de *transição* de A.
- $F \subseteq S$ é o conjunto dos *estados finais* (ou de aceitação) de A.

- O conjunto δ deve ser finito.
- Se δ é uma função, o autómato é *determinista*.

Linguagem de um Autómato de Pilha

Define-se uma generalização da relação de transição para "processar" palavras – permite "dar" vários passos no autómato.

Relação de transição extendida

$$\delta^* \subseteq S \times \Sigma^* \times \Gamma^* \times S \times \Gamma^*$$

δ^* é definida indutivamente:

- $(r, \epsilon, W, r, W) \in \delta^*$;
- $(r, aw, VW, t, W') \in \delta^*$, se $(r, a, V, s, W'') \in \delta$ e $(s, w, W''W, t, W') \in \delta^*$.

Linguagem

$L = \{w \in \Sigma^* \mid (s, w, Z, f, Z) \in \delta^* \wedge f \in F\}$

Resultados de Expressividade

- A linguagem de um Autómato de Pilha (Determinista) é Independente de Contexto (Determinista).
- Qualquer Autómato de Pilha pode ser convertido numa Gramática Independente de Contexto (e vice-versa).
- Há Autómato de Pilha Não-Deterministas que não podem ser convertidos em Deterministas equivalentes.
  As Linguagens Independente de Contexto Deterministas estão estritamente contidas nas Não-Deterministas.

Lema da Bombagem para Linguagens Independentes de Contexto

Uma linguagem L é Independente de Contexto, se existe um natural $n \geq 1$ tal que qualquer palavra $w \in L$ com $|w| \geq n$ é tal que $w = uvxyz$, sendo:

1. $|vxy| \leq n$
2. $|vy| \geq 1$
3. $\forall k \in \mathbb{N}_0 . uv^k xy^k z \in L$

- Pode-se usar o Lema para provar por absurdo que dada linguagem não é Independente de Contexto.
- No entanto, o Lema não caracteriza completamente as Linguagens Independentes de Contexto:
  Há linguagens que satisfazem o Lema e não o são.

A linguagem $L = \{a^k b^k c^k \mid \ k \in \mathbb{N}_0\}$ não é Independente de Contexto

1. Fixa-se um natural $n \geq 1$ e considera-se a palavra $a^n b^n c^n \in L$.
2. Pelas condições do Lema, vxy não pode conter mais de 2 símbolos distintos (pois $|vxy| \leq n$) . Logo:
   2.1 $vxy = a^i$ com $i \leq n$
   2.2 $vxy = a^i b^j$ com $i + j \leq n$
   2.3 $vxy = b^i$ com $i \leq n$
   2.4 $vxy = b^i c^j$ com $i + j \leq n$
   2.5 $vxy = c^i$ com $i \leq n$

   Em cada caso é fácil verificar que $uv^2xy^2z \notin L$