Simulace nedeterministických konečných automatů

Jan Holub
katedra počítačů
FEL ČVUT
Karlovo nám. 13
121 35 Praha 2
e-mail: holub@fel.cvut.cz
WWW: http://cs.felk.cvut.cz/~holub

Tři možnosti použití nedeterministických konečných automatů (NKA):

1. převod na ekvivalentní deterministický konečný automat (DKA)
(paměť $\mathcal{O}(2^m)$, čas $\mathcal{O}(2^m + n)$, kde
m – počet stavů NKA,
n – délka vstupního řetězce)

2. simulace NKA (menší paměťové nároky na úkor vyšší časové náročnosti)

3. NKA jako model pro algoritmy

Proč simulace NKA?

- převod NKA na DKA zabírá příliš času v porovnání s časem běhu DKA ($\mathcal{O}(2^m)$ vs. $\mathcal{O}(n)$)

- příliš vysoké paměťové nároky

- obojí

Metody simulace NKA:

1. Triviální metoda
   - udržování množiny aktivních stavů

2. Dynamické programování [Ukk85]
   - číselná proměnná pro určitou podmnožinu stavů (v podmožině může být nanejvýš jeden aktivní stav)

3. Bitový paralelismus [BG92]
   - využívá bitových operací k paralelní simulaci přechodů automatů (bitový vektor — posloupnost stavů)

BP — Hammingova metrika

$$R^j_{i+1} = (shl(R^j_i)\, or D[t_{i+1}])\, and\, (shl(R^{j-1}_i))$$

$$\begin{aligned} R^0_{i+1} &= (shl(R^0_i)\, or D[t_{i+1}]) \\ R^1_{i+1} &= (shl(R^1_i)\, or D[t_{i+1}])\, and\, (shl(R^0_i)) \end{aligned}$$

D	a	b	?
a	**0**	1	1
b	1	**0**	1
a	**0**	1	1

R^0_i	-	a	a	b	b	a	b	a	a	b	b	b
a	1	**0**	**0**	1	1	**0**	1	**0**	**0**	1	1	1
b	1	1	1	**0**	1	1	**0**	1	1	**0**	1	1
a	1	1	1	1	1	1	1	**0**	1	1	1	1

R^1_i	-	a	a	b	b	a	b	a	a	b	b	b
a	1	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
b	1	1	**0**	**0**	**0**	1	**0**	1	**0**	**0**	**0**	**0**
a	1	1	1	1	**0**	**0**	1	**0**	1	1	**0**	1

BP — Levenshteinova metrika

$$
\begin{aligned}
R^j_{i+1} &= (shl(R^j_i)\, or D[t_{i+1}])\, and \\
&\quad (shl(R^{j-1}_i\, and\, R^{j-1}_{i+1}))\, and\, (R^{j-1}_i) \\
R^0_{i+1} &= (shl(R^0_i)\, or D[t_{i+1}]) \\
R^1_{i+1} &= (shl(R^1_i)\, or D[t_{i+1}])\, and \\
&\quad (shl(R^0_i\, and\, R^0_{i+1}))\, and\, (R^0_i)
\end{aligned}
$$

D	a	b	?
a	**0**	1	1
b	1	**0**	1
a	**0**	1	1

R^0_i	-	a	a	b	b	a	b	a	a	b	b	b
a	1	**0**	**0**	1	1	**0**	1	**0**	**0**	1	1	1
b	1	1	1	**0**	1	1	**0**	1	1	**0**	1	1
a	1	1	1	1	1	1	1	**0**	1	1	1	1

R^1_i	-	a	a	b	b	a	b	a	a	b	b	b
a	1	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
b	1	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
a	1	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	1

Literatura:

[BG92] R. A. Baeza-Yates and G. H. Gonnet: A New Approach to Text Searching. *Communications of the ACM*, 35(10):74–82, 1992.

[Ukk85] E. Ukkonen: Finding Approximate Patterns in Strings. *Journal of Algorithms*, 6:132–137, 1985.