Ekvivalence konečných automatů a regulárních gramatik

Thursday, May 30, 2013 8:38 AM

Regulární gramatiky popisují všechny *regulární jazyky* a v tomto smyslu (ve schopnosti popisu jazyka) jsou ekvivalentní s konečnými automaty a regulárními výrazy. Regulární gramatika ja buď pravá regulární (neterminály jsou vpravo) nebo levá regulární (neterminály jsou vlevo).

Regulární jazyk je formální jazyk (množina (i nekonečná) slov složených z omezené abecedy), který:

- může být akceptován deterministickým/nedeterministickým konečným stavovým automatem
- lze popsat regulárním výrazem
- lze ho generovat regulární gramatikou

Příkladem **ne**regulárního jazyka je $a^n b^n$, kde n > 1 (alespoň jedno *a* následované stejným počtem *b*), gramatika pro palindromy apod. - *lze určit na základě Nerodovy věty, která se užívá v důkazech, že nějaký jazyk není rozpoznatelný konečným automatem*

Každý regulární jazyk je rozpoznatelný konečným automatem; každý jazyk rozpoznatelný konečným automatem je regulární.

Kleenova věta: Libovolný jazyk je regulární, právě když je rozpoznatelný konečným automatem. Přechodový graf je T nad **S** je konečný orientovaný graf, jehož každá hrana je pojmenována jistým slovem wÎ**S***. Alespoň jeden z uzlů grafu je počáteční a některé uzly jsou koncové. Ke každému přechodovému grafu T nad abecedou S existuje regulární výraz R nad S takový, že

$$\tilde{R} = \tilde{T}$$

a ke každému regulárnímu výrazu R nad S existuje konečný automat A takový, že

$$\tilde{A} = \tilde{R}$$

.

Postup převodu gramatiky na konečný automat

Potřebujeme získat gramatiku typu 3 ve standardní formě.

Regulární gramatika je ve **standardní formě**, jestliže obsahuje pouze pravidla tvaru $X \rightarrow aY$ a $X \rightarrow a$, $X \rightarrow e$ kde *X,Y* jsou neterminály, *a* je právě jeden terminál, *e* je prázdný symbol. Toho dosáhneme takto:

- Původní gramatika typu 3 (lineární): G = (N, T, S, P)
- Požadovaná regulární gramatika: G' = (N', T, S, P')
- Požadovaná gramatika G' bude mít stejné terminální symboly a stejný počáteční stav.
- Konstrukce přechodů *P'*:
  - do *P'* zařadíme všechna pravidla z *P* ve tvaru $X \rightarrow aY$ a $X \rightarrow e$
  - za každé pravidlo $X \rightarrow x1\ x2\ x3\ Y$ zařadíme do *P'* soustavu pravidel:
    - $X \rightarrow x1\ X1$
    - $X1 \rightarrow x2\ X2$
    - $X2 \rightarrow x3\ Y$
      - za každé pravidlo $X \rightarrow z1\ z2$, zařadíme do *P'* soustavu:
    - $X \rightarrow z1\ Z1$
    - $Z1 \rightarrow z2\ Z2$
    - $Z2 \rightarrow e$

- Místo pravidel tvaru $X \rightarrow Y$ musíme zajistit to, aby z každého stavu X pro který máme $X \rightarrow Y$, bylo možné odvodit všechny řetězce, které lze odvodit z Y.
- N' vznikne obohacením N o všechny nově vytvořené neterminální symboly

Zkonstruujeme automat z nově vytvořené gramatiky

- stavy budou odpovídat neterminálním symbolům
- vstupy budou odpovídat terminálním symbolům
- přechodovou funkci zkonstruujeme na základě analogií
    - $X \rightarrow aY \leftrightarrow$ přechod ze stavu X do stavu Y při vstupu symbolu a
- počáteční stav bude odpovídat počátečnímu symbolu
- množinu koncových stavů určíme z pravidel $X \rightarrow e$

Tímto jsme získali **nedeterministický konečný automat**, který lze převést na **deterministický konečný automat.**

From <https://d.docs.live.net/e3534876709763a3/Dokumenty/ZCU/Statnice/Statnice.docx>