

# OPTIMALIZACE SYNTAKTICKÉ ANALÝZY STAVOVÝCH GRAMATIK ZALOŽENÁ NA TERMINÁLOVÉ PREDIKCI

**Paulík Miroslav**

*Fakulta informačních technologií VUT v Brně, Božetěchova 1/2, 612 66 Brno,  
tel.: +420 732 648 112, xpauli00@stud.fit.vutbr.cz*

## Abstrakt

Tento abstrakt obsahuje nástin nové optimalizační metody syntaktické analýzy založené na stavových gramatikách. Pro plné porozumění následujícího textu je nutná znalost základních pojmů z oblasti formálních jazyků, překladačů a mechanismů syntaktické analýzy. Definice těchto termínů spolu s příslušnými matematickými operacemi jsou vysvětleny např. v [1] a [2].

Terminálová predikce je rozšířením syntaktické analýzy pomocí modifikovaného hlubokého zásobníkového automatu [3], jejímž účelem je co nejvčasnější detekce tzv. „slepých větví“ syntaktické analýzy. Taková větev je typická právě pro stavové gramatiky, jenž umožňují aplikaci pravidel gramatiky na zanořené neterminální symboly. U takových symbolů ale principiálně nelze jednoznačně určit, jaké pravidlo má být uplatněno. Proto dochází ke zmíněnému větvení, kde v každé větvi je použito právě jedno z aplikovatelných pravidel s tím, že maximálně jedna větev povede k přijetí vstupního řetězce. Tu zajišťuje právě výše zmíněná modifikace, která využívá transformace pravidel stavové gramatiky na taková, jejichž derivační řetězec nezačíná neterminálovým symbolem (derivují prázdný řetězec nebo řetězec začínající terminálem).

Standardně se během analýzy aplikují pravidla tak dlouho, dokud není přijat celý vstupní řetězec nebo nedojde ke konfliktu vstupního a očekávaného terminálu. K tomu může dojít až po aplikaci velkého množství pravidel, které mohou znamenat další potencionální větvení. Terminálová predikce se naproti tomu snaží o co nejvčasnější detekci situace, kdy není možné vstupní řetězec v dané větvi přijmout. Princip této optimalizace spočívá ve využívání informací z aktuálního očekávaného vstupu ukládaného na zásobník a speciálních údajů z lexikální analýzy dokončené ještě před začátkem samotné syntaktické analýzy.

Pro fungování terminálové predikce je potřeba nejprve získat údaje z lexikální analýzy o relativních pozicích všech výskytů jednotlivých terminálů ze vstupního řetězce (uložené např. do tabulky). Po aplikaci každého pravidla je pak nutné porovnat derivační řetězec se statistickými údaji z výše zmíněné tabulky (dále jen tabulky). Podle úrovně predikce se tyto kontroly zaměřují buď na pouhé porovnání počtu terminálů z derivovaného řetězce s jejich skutečným počtem z tabulky nebo na pořadí a zároveň i vzájemné pozice očekávaných symbolů. V nejextrémnější verzi lze kontrolovat i neterminální symboly a jejich terminálové uzávěry (těmi se rozumí existence alespoň jedné z posloupností všech terminálů daných derivovačnými řetězci jednotlivých aplikovatelných pravidel ve stupním řetězci).

Pravděpodobně největší použitelnost bude mít predikce využívající jak pořadí, tak i vzájemné pozice terminálů (přesnou efektivitu je nutné dokázat podrobným výzkumem). U této modifikace je vyžadován kromě tabulky a zásobníku s očekávanými symboly ještě druhý, který bude obsahovat místo terminálů jejich konkrétní pozice z tabulky. Při dohledávání pozice dalšího derivovaného terminálu se vezme z tabulky taková pozice požadovaného terminálu, která je větší, než hodnota jeho předchůdce v tomto zásobníku. Pro všechny následující terminály v tomto zásobníku je nutné provést test, zda jejich hodnoty pozic tvoří rostoucí posloupnost (tedy zda jsou konzistentní vzhledem ke skutečnému vstupnímu řetězci). V případě porušení této posloupnosti je nutné dohledat rekurzivně nové pozice těchto symbolů, aby byla tato podmínka splněna. Dojde-li k situaci, že se v tabulce nenachází žádný terminál s požadovanou hodnotou, je aktuální větev syntaktické analýzy zamítnuta.

Terminálová predikce je tedy optimalizační metodou syntaktické analýzy založené na stavových gramatikách, která vyžaduje předzpracování lexikální analýzy. Tato modifikace má několik úrovní lišících se složitostí kontroly a nároky na zdroje. Predikce je obecně použitelná, nicméně největší efektivitu lze předpokládat právě u stavových gramatik. Doposud však nebyl proveden žádný podrobný výzkum, který by jednoznačně určil skutečnou míru použitelnosti této optimalizace.

***Klíčová slova:*** *terminálová predikce; hluboký zásobníkový automa, stavová gramatika*

## **Literatura**

- [1.] Meduna, A.: Elements of Compiler Design, New York, US, T & F, 2008, s. 304, ISBN 978-1-4200-6323-3
- [2.] Meduna, A.: Automata and Languages: Theory and Applications [Springer, 2000], London, GB, Springer, 2005, s. 892, ISBN 1-85233-074-0
- [3.] Meduna, A.: Deep Pushdown Automata, In: Acta Informatica, roč. 2006, č. 98, DE, s. 114-124, ISSN 0001-5903 2005, s. 892, ISBN 1-85233-074-0