

УДК 004.4'413

## ПРОГРАМНИЙ ЗАСІБ ДЛЯ НАЛАГОДЖЕННЯ ФОРМАЛЬНИХ ГРАМАТИК ПАРАЛЕЛЬНОГО ПЕРЕДУВАННЯ ТА ПОДАЛЬШОГО СИНТАКСИЧНОГО АНАЛІЗУ

Сичков Віталій Сергійович

к.т.н., доц. і зав. кафедри ІІЗ Комлева Наталія Олегівна  
Національний університет «Одеська політехніка», УКРАЇНА

**АНОТАЦІЯ.** У цій роботі розглянуто деякі методи синтаксичного аналізу ланцюжків та виконано огляд конструктора *Formal Grammar*, використовуваного під час лабораторних занять із дисципліни «Теорія формальних граматики і трансляції». На основі цього створено програмний засіб *FGPP*, який дозволяє скоротити час перевірки робіт студентів нашої кафедри за певними темами цієї дисципліни.

**Вступ.** Нині існує чимало досліджень із підвищення та управління якістю освітнього процесу, зокрема [1-3]. До актуальних проблем також можна віднести покращення продуктивності праці викладачів закладів вищої освіти. Під час виконання певних професійних завдань вони можуть стикатися з певними незручностями, які заважають робити це ефективніше. Одним із шляхів подолання останніх є автоматизація відповідних процесів шляхом упровадження та використання прикладних засобів інформаційно-комунікаційних технологій. Так, у межах дисципліни «Теорія формальних граматики і трансляції» (ТФГТ), яку викладають студентам 2-го курсу кафедри інженерії програмного забезпечення (ІІЗ) Одеської політехніки, цілком можна пришвидшити процес перевірки лабораторних робіт. Далі в роботі надано пояснення важливості та запропоновано метод розв'язання цього питання.

**Мета роботи.** Скорочення часу перевірки лабораторних робіт студентів із ТФГТ, які передбачають налагодження формальної граматики та висхідний синтаксичний аналіз (СА) вхідного ланцюжка методом паралельного передування, за рахунок розроблення і застосування програмного засобу, який автоматизує ці процеси за описаними в протоколах правилами граматики та ланцюжками мов.

**Основна частина роботи.** Формальна, або формалізована, мова – це штучна (на відміну від природної, наприклад, української чи англійської) мова, яка характеризується точними правилами побудови виразів та їх розуміння. Неформально кажучи, це множина кінцевих ланцюжків (слів, рядків тощо) кінцевої довжини над кінцевим алфавітом.

Формальна граматика – це метод опису формальної мови, який передбачає виділення певної підмножини ланцюжків із деякого алфавіту. Існують два типи граматики: породжувальні та розпізнавальні (аналітичні). Породжувальні визначають правила, які дозволяють створити будь-який ланцюжок у мові, тоді як розпізнавальні – входження цього ланцюжка до мови.

Формальну граматику  $G$  можна представити як сукупність із чотирьох об'єктів

$$\{V_T, V_N, R, S\}, \quad (1)$$

де  $V_T$  – це набір (алфавіт, або словник) термінальних символів – елементів-терміналів, які безпосередньо присутні в ланцюжках мови, яка відповідає граматиці, та мають певне, незмінне значення (подібні до поняття «літера»);  $V_N$  – набір нетермінальних символів – елементів-нетерміналів, призначених для позначення певних сутностей у мові та пояснення процесу утворення конструкцій;  $R$  – множина породжувальних правил виду  $a \rightarrow b$ , де  $a$  і  $b$  – ланцюжки, побудовані з літер алфавіту  $V = V_T \cup V_N$ , який називають повним алфавітом, або словником, граматики;  $S$  – початковий нетермінальний символ граматики (аксіома),  $S \in V_N$ .

Надалі матимемо справу з контекстно-вільними (КВ-) граматиками – граматиками другого типу за ієрархією (класифікацією граматики) Хомського. Ліва частина кожного з їх породжувальних правил складається з одного нетермінала, а права є ланцюжком, який складається з символів  $V$ , зокрема порожнього ланцюжка  $\lambda$  (позначають так:  $V^* = V \cup \{\lambda\}$ ).

Фразою є послідовність символів, яка виводиться безпосередньо з одного нетермінального символу. Проста фраза містить лише термінальні символи. Найлівіша проста фраза є основою ланцюжка мови (сентенціальної форми).

Один із найпростіших та наочних способів висхідного аналізу КВ-мов заснований на використанні відношень передування (ВП), які визначаються між символами граматики на етапі розроблення аналізатора, з огляду на правила граматики. У процесі синтаксичного аналізу ці відношення дозволяють послідовно виділяти прості фрази і згортати їх, забезпечуючи висхідне розпізнавання вхідного ланцюжка.

Процес налагодження початкової граматики передбачає побудову відповідної матриці відношень символів її повного словника, позбавленої конфліктів. Це означає, що будь-які два символи можуть бути пов'язані не більше, ніж одним ВП. Саме відсутність конфліктів у цій матриці дозволяє виконати подальший синтаксичний аналіз вхідного ланцюжка відповідним методом передування. Він необхідний для визначення приналежності цього ланцюжка до множини можливих речень мови, яку описують правила заданої граматики [4].

Існує низка методів аналізу на основі ВП. На лабораторних заняттях із ТФГТ студенти поглиблено вивчають два методи: простого та паралельного передування. Граматика простого передування – це КВ-граматика без порожніх або однакових правих частин та конфліктів у матриці передування.

Для обчислення таких матриць можна використовувати синтаксичні дерева (СД) або визначати їх на основі спеціальних бінарних відношень (БВ) *FIRST*, *LAST* та *EQUAL*. Нехай є три символи *A*, *B* і *C*, тоді:

1. *A FIRST B*, якщо граматика має правило виду  $A \rightarrow B\alpha$ ,  $A \in V_N$ ,  $B \in V$ ,  $\alpha \in V^*$ .
2. *A LAST B*, якщо граматика має правило виду  $A \rightarrow \alpha B$ ,  $A \in V_N$ ,  $B \in V$ ,  $\alpha \in V^*$ .
3. *A EQUAL B*, якщо граматика має правило виду  $C \rightarrow \alpha A B \beta$ ,  $C \in V_N$ ,  $A, B \in V$ ,  $\alpha, \beta \in V^*$ .

Граматика паралельного передування містить певні зміни у відношеннях та виконанні синтаксичного аналізу ланцюжків порівняно з граматиною простого передування. У таблиці 1 наведено порівняльну характеристику розглянутих методів, де символи *A*, *B*  $\in V$ .

Таблиця 1– Порівняльна характеристика методів синтаксичного аналізу

Характеристика	Метод простого передування	Метод паралельного передування
Кількість відношень	3	4
ВП «дорівнює» ( $\doteq$ ) та зв'язок із БВ	Символи <i>A</i> і <i>B</i> належать основі та розташовані поруч ( <i>B</i> слідує за <i>A</i> ); <i>EQUAL</i>	Символи <i>A</i> і <i>B</i> належать простій фразі та розташовані поруч ( <i>B</i> слідує за <i>A</i> ); <i>EQUAL</i>
ВП «менше» ( $\leq$ ) та зв'язок із БВ	Якщо <i>B</i> – перший символ основи, а <i>A</i> передує йому та не входить в основу; <i>EQUAL</i> $\times$ <i>FIRST</i> <sup>+</sup>	Якщо <i>B</i> – перший символ простої фрази, а <i>A</i> передує йому та не входить в цю фразу; <i>EQUAL</i> $\times$ <i>FIRST</i> <sup>+</sup>
ВП «більше» ( $\geq$ ) та зв'язок із БВ	Якщо <i>A</i> – останній символ основи, а <i>B</i> слідує за ним та не входить в основу; ( <i>LAST</i> <sup>+</sup> ) <sup>T</sup> $\times$ <i>EQUAL</i> $\times$ <i>FIRST</i> <sup>*</sup>	Якщо <i>A</i> – останній символ простої фрази, а <i>B</i> слідує за ним та не входить в цю фразу; ( <i>LAST</i> <sup>+</sup> ) <sup>T</sup> $\times$ <i>EQUAL</i> <sup>*</sup> $\times$ <i>FIRST</i> <sup>*</sup>
ВП «хрестик» ( $\otimes$ ) та зв'язок із БВ	Відсутнє	Якщо <i>A</i> – останній символ однієї простої фрази, а <i>B</i> – перший символ наступної простої фрази; <i>LAST</i> <sup>+</sup> $\times$ <i>EQUAL</i> <sup>*</sup> $\times$ <i>FIRST</i> <sup>*</sup>
Переміщення СД	Зліва направо – знизу догори	Знизу догори

Тут верхні індекси позначають таке: “*T*” – транспонування матриці, “+” – транзитивне замикання, “\*” – рефлексивно-транзитивне замикання. Як бачимо, синтаксичний аналіз другим методом передбачає визначення не лише основ, як у першому, а всіх простих фраз, які на кожному кроці одразу замінюють лівими частинами правил, що зумовлює швидшу перевірку приналежності вхідного ланцюжка до описаної мови.

Під час виконання лабораторних робіт студентам пропонується використовувати *Formal Grammar* – програмний конструктор КС-грамматик, що містить текстовий редактор із підсвічуванням синтаксису, модуль виявлення нескінченної рекурсивності правил і формувач матриць простого передування. За його допомогою вони можуть перевірити розроблену граматику на наявність помилок і конфліктів та налагодити її як граматику простого передування, за умови дотримання синтаксису метамови цього конструктора [5].

Для налагодження початкової граматики паралельного передування для подальшого виконання синтаксичного аналізу студентам необхідно визначити матрицю передування, проте вже самостійно шляхом побудови відповідних синтаксичних дерев або на основі описаних БВ.

Під час перевірки лабораторних робіт студентів із дисципліни ТФГТ, що передбачають роботу з методом паралельного передування, було помічено певні труднощі. Викладачам немає потреби пересвідчуватися в правильності побудови матриць простого передування, оскільки його виконує вже випробуваний програмний засіб. Проте для паралельного методу необхідно дізнатися, чи всі відношення та конфлікти між символами були виявлені, для уникнення неможливості виконання СА, який є окремим завданням. Також потрібно перевірити кожен крок останнього та його виконання за мінімальну їх кількість. Це займає досить багато часу, що особливо відчутно зі зростанням обсягу граматики, довжини вхідних ланцюжків та кількості робіт, які перевіряються.

Для розв'язання цієї проблеми було розроблено програмний засіб *FGPP*, який може відображати помилки граматики, матриці бінарних відношень, конфлікти відношень із матрицею паралельного передування, хід та результат синтаксичного аналізу (для налагодженої граматики). Він орієнтований на використання не студентами, а викладачами кафедри ІПЗ, та покликаний облегшити перевірку лабораторних робіт за рахунок простого порівняння результатів автоматичного аналізу (очікуваний) та аналізу, виконаного власноруч студентами (дійсний). Цей засіб реалізовано як веб-застосунок за допомогою мови програмування *JavaScript*, мови розмічення документів *HTML* та мови стилізації *CSS*. Аналогічно *Formal Grammar*, він здатний визначати різні типи помилок та конфліктів відношень передування.

**Висновки.** Апробацію розробленого засобу проведено на прикладах 30-ти протоколів студентів 4-6 груп потоку АС-21х за темою «Налагодження граматики паралельного передування». На основі отриманих показників можна стверджувати, що використання *FGPP* дійсно дозволяє пришвидшити перевірку лабораторних робіт. Варто зазначити, що виграш у часі дещо відрізняється між роботами. Було помічено, що він значно більший, якщо перевірена робота є правильно виконаною, коли в процесі не потрібно зазначати помилки та відслідковувати їх залежності. Загалом середній його показник склав 2,4 рази. Подальше випробування цієї розробки полягатиме в оцінюванні прискорення на лабораторних роботах із СА методом паралельного передування, складання яких розпочнеться за два тижні.

### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Колесніков, О. Є., Білощицький, А. О., Файзуллін, А. Р. Розробка теоретичних засад створення освітнього середовища навчального закладу. *Прикладні аспекти інформаційних технологій*. 2020. Т. 3, № 2. С. 32–43. DOI: <https://doi.org/10.15276/aa.02.2020.2>.
2. Komleva N., Liubchenko V., Zinovatna S. Improvement of teaching quality in the view of a resource-based approach. *CEUR Workshop Proceedings*, Vol. 2740. 2020. Pp. 262–277. <http://ceur-ws.org/Vol-2740/20200262.pdf>.
3. Komleva N., Liubchenko V., Zinovatna S., Kobets V. Decision support system for quality management in learning process. *CEUR Workshop Proceedings*, Vol. 2711. 2020. Pp. 430–442. <http://ceur-ws.org/Vol-2711/paper33.pdf>.
4. Комлева Н. О., Жиро Л. С. Конспект лекцій з дисципліни «Теорія формальних граматики і трансляції» для студентів спеціальності 121 – «Інженерія програмного забезпечення». Одеса: ОНПУ, 2018. 66 с.
5. Конструктор формальних граматики *Formal Grammar*. URL: <https://t.ly/tBB7> (дата звернення: 15.05.2023).

### A SOFTWARE TOOL FOR DEBUGGING FORMAL GRAMMARS OF PARALLEL PRECEDENCE AND SUBSEQUENT SYNTAX ANALYSIS

Vitalii Sychkov

PhD, Associate Professor and Head of the Department of SE Nataliia Komleva  
Odesa Polytechnic National University, UKRAINE

**ANNOTATION.** The paper describes some methods of syntax analysis and reviews the *Formal Grammar* constructor used in laboratory classes in the discipline “Theory of Formal Grammars and Translation”. Based on this, the *FGPP* software tool has been created, which allows to reduce the time of checking lab works of students of our department on certain topics of this discipline.