

Лекція №8

Способи синтаксичного розбору

Контекстно-вільні граматики традиційно виступають основою для створення синтаксичних аналізаторів у компіляторах. Опис вхідної мови включає правила її граматики (які також називають синтаксисом мови), відповідно до яких складаються тексти програм. Тому послідовність правил граматики, яка була застосована для написання тексту якоїсь програми, визначає структуру цієї програми.

Варіантів виводу деякого ланцюжка α за граматиною G загалом може бути досить багато. Така ситуація виникає в такому випадку, коли деякий проміжний ланцюжок виводу містить більше одного нетермінального символу, і правила підстановки можуть застосовуватись до них в будь-якій послідовності. Якщо використати ті самі правила, а змінити лише порядок їх застосування, то буде отримано вже інший вивід, проте заключний термінальний ланцюжок та його синтаксична структура залишаться незмінними. Зазвичай множину усіх можливих виводів за КВ-граматиною не використовують, розглядаючи лише, так звані, лівосторонні чи правосторонні виводи або розбори.

Розглянемо лівосторонній та правосторонній виводи (розбори).

Нехай граматика визначена правилами:

1. $\langle \text{оператор} \rangle \rightarrow \langle \text{змінна} \rangle := \langle \text{вираз} \rangle$
2. $\langle \text{вираз} \rangle \rightarrow \langle \text{терм} \rangle$
3. $\langle \text{вираз} \rangle \rightarrow \langle \text{терм} \rangle + \langle \text{вираз} \rangle$
4. $\langle \text{терм} \rangle \rightarrow \langle \text{змінна} \rangle$
5. $\langle \text{терм} \rangle \rightarrow (\langle \text{вираз} \rangle)$
6. $\langle \text{змінна} \rangle \rightarrow a$
7. $\langle \text{змінна} \rangle \rightarrow b$
8. $\langle \text{змінна} \rangle \rightarrow c$

Для рядка $a := b + c$ зробимо лівосторонній вивід (розбір):

$$\begin{aligned} \langle \text{оператор} \rangle &\Rightarrow^1 \langle \text{змінна} \rangle := \langle \text{вираз} \rangle \Rightarrow^6 a := \langle \text{вираз} \rangle \Rightarrow^3 \\ a := \langle \text{терм} \rangle + \langle \text{вираз} \rangle &\Rightarrow^4 a := \langle \text{змінна} \rangle + \langle \text{вираз} \rangle \Rightarrow^7 a := b + \langle \text{вираз} \rangle \Rightarrow^2 a := b + \langle \text{терм} \rangle \Rightarrow^4 a := \\ b + \langle \text{змінна} \rangle &\Rightarrow^8 a := b + c \end{aligned}$$

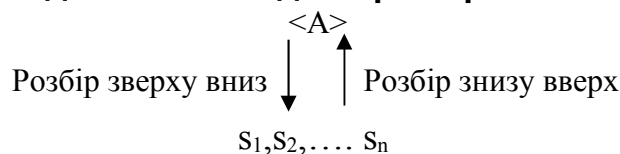
Лівий розбір – 1, 6, 3, 4, 7, 2, 4, 8

Для цього ж рядка зробимо правосторонній вивід (розбір):

$$\begin{aligned} \langle \text{оператор} \rangle &\Rightarrow^1 \langle \text{змінна} \rangle := \langle \text{вираз} \rangle \Rightarrow^3 \langle \text{змінна} \rangle := \langle \text{терм} \rangle + \langle \text{вираз} \rangle \Rightarrow^2 \\ \langle \text{змінна} \rangle &:= \langle \text{терм} \rangle + \langle \text{терм} \rangle \Rightarrow^4 \langle \text{змінна} \rangle := \langle \text{терм} \rangle + \langle \text{змінна} \rangle \Rightarrow^8 \\ \langle \text{змінна} \rangle &:= \langle \text{терм} \rangle + c \Rightarrow^4 \langle \text{змінна} \rangle := \langle \text{змінна} \rangle + c \Rightarrow^7 \langle \text{змінна} \rangle := b + c \Rightarrow^6 a := b + c \end{aligned}$$

Правий розбір – 1, 3, 2, 4, 8, 4, 7, 6.

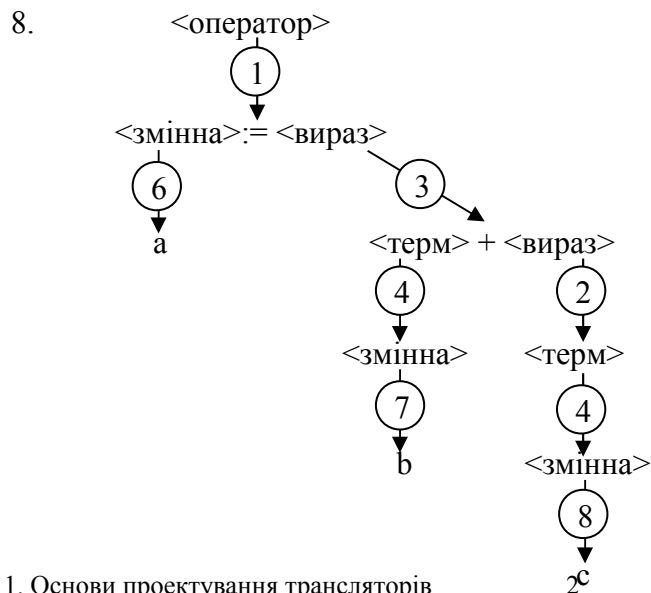
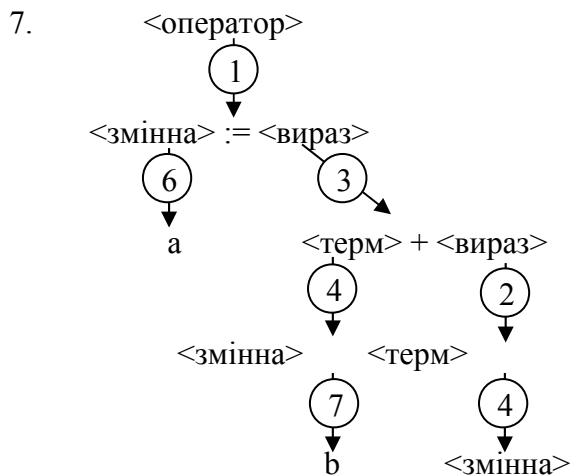
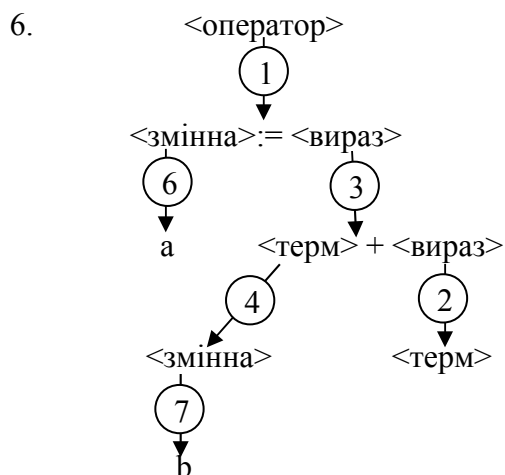
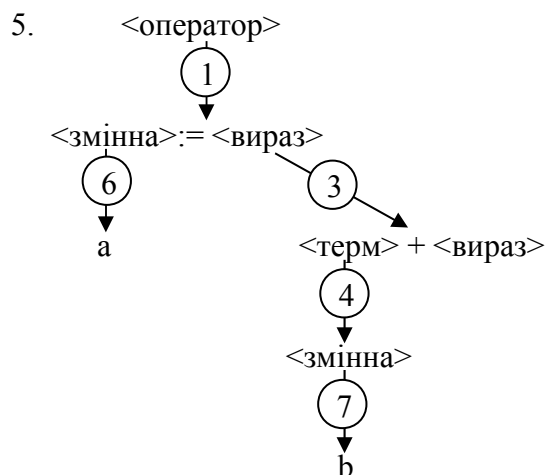
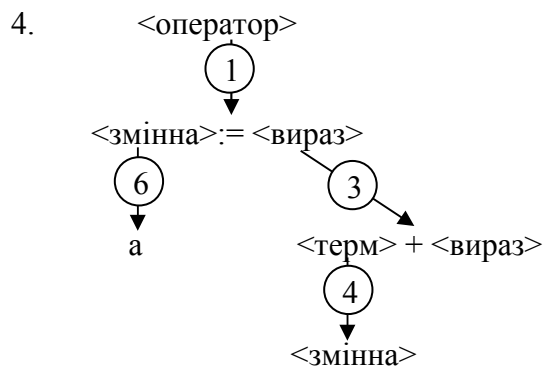
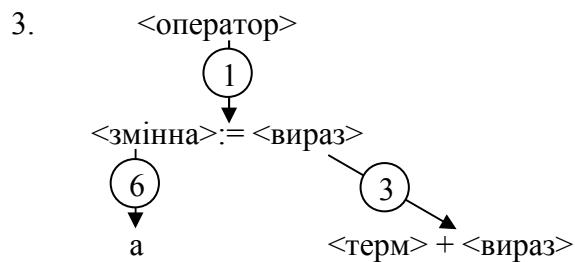
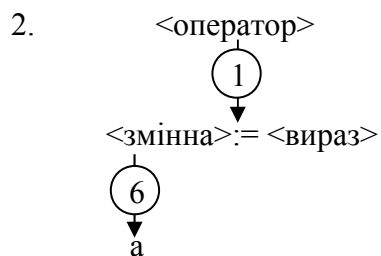
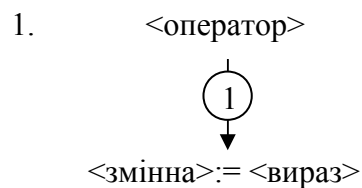
Низхідний та висхідний розбір або аналіз



$\langle A \rangle$ - аксіома граматики,
 S_1, S_2, \dots, S_n –рядок, який аналізується.

Низхідний аналіз

Розглянемо низхідний аналіз (розбір) для рядка $a:=b+c$



Висхідний аналіз

Висхідний аналіз базується на понятті «основа».

Якщо деякий рядок терміналів та/або нетерміналів (сентенціальна форма) прямо приводиться до нетерміналу, то він називається **безпосередньо приводимою фразою**.

Найлівіша безпосередньо приводима фраза заданої сентенціальної форми (рядка термінальних та/або нетермінальних символів) називається **основою**.

Суть алгоритму аналізу полягає у пошуку основи й заміни її нетерміналом, до якого ця основа безпосередньо приводиться. Процес такої заміни називається **згорткою**. Розбір завершиться тоді, коли нетерміналом, до якого застосовано згортку, стане аксіома граматики.

Розглянемо висхідний аналіз (розбір) рядка $a := b + c$ за такою граматику:

1. $\langle \text{оператор} \rangle \rightarrow \langle \text{змінна} \rangle := \langle \text{вираз} \rangle$
2. $\langle \text{вираз} \rangle \rightarrow \langle \text{терм} \rangle$
3. $\langle \text{вираз} \rangle \rightarrow \langle \text{вираз} \rangle + \langle \text{терм} \rangle$
4. $\langle \text{терм} \rangle \rightarrow \langle \text{змінна} \rangle$
5. $\langle \text{терм} \rangle \rightarrow (\langle \text{вираз} \rangle)$
6. $\langle \text{змінна} \rangle \rightarrow a$
7. $\langle \text{змінна} \rangle \rightarrow b$
8. $\langle \text{змінна} \rangle \rightarrow c$

Слід зазначити, що в цій граматиці, на відміну від граматики у прикладі низхідного розбору, у правилі 3 використовується ліва рекурсія, а не права. Це зроблено для того, щоб підкреслити, що при висхідному аналізі можна використовувати правила як з правою, так і з лівою рекурсією на відміну від низхідного аналізу.

Розглянемо процес висхідного аналізу (розбору) рядка $a := b + c$

