

# ПОБУДОВА ЛЕКСИЧНОГО СКАНЕРУ НА КІНЦЕВИХ АВТОМАТАХ

Квасенко О.В., Гавриленко С. Ю.

*Національний технічний університет*

*«Харківський політехнічний інститут», м. Харків*

При побудові лексичного аналізатору безпосередній опис кінцевого автомату (КА) незручний з практичної точки зору. Однак у випадках, коли потрібен аналізатор для лексики обмеженого обсягу, можна використовувати формалізм КА для безпосереднього проектування сканеру. Опис КА здійснюється за допомогою графу переходу.

Граф переходу кінцевого автомату – це навантажений односпрямований граф, в якому вершини представляють стани КА, дуги відображають переходи з одного стану в інший, а символи навантаження дуг відповідають функції переходу кінцевого автомату. Якщо функція переходу автомату передбачає перехід зі стану  $q$  в  $q'$  за декількома символами, то між ними будуватиметься одна дуга, яка позначається усіма символами, за якими відбувається перехід з  $q$  в  $q'$ .

Недетермінований кінцевий автомат незручний для аналізу ланцюжків, так як в ньому можуть зустрічатися стани, що допускають неоднозначність (стан коли дві або більше дуги помічені одним і тим же символом). Очевидно, що програмування роботи такого КА – нетривіальне завдання. Для простого програмування функціонування КА  $M(Q, \Sigma, \delta, q_0, F)$  він повинен бути детермінованим – у кожному з можливих станів цього КА для будь-якого вхідного символу функція переходу повинна містити не більше одного стану.

Будь-який недетермінований кінцевий автомат може бути перетворений в еквівалентний йому детермінований. Крім того детермінований КА може бути мінімізований – приведений до еквівалентного йому КА з мінімально можливою кількістю станів.

Алгоритм побудови детермінованого кінцевого автомату по НКА:

Вхід:  $M = (K, VT, \delta, H, S)$  – недетермінований кінцевий автомат.

Вихід:  $M' = (K', VT, \delta', H', S')$  – детермінований кінцевий автомат, що допускає ту ж мову, що й автомат  $M$ .

Метод:

- 1) Множина станів  $K'$  складається з усіх підмножин множини  $K$ . Кожний стан з  $K'$  будемо позначати  $[A_1, A_2, \dots, A_n]$ , де  $A_i \in K$ .
- 2) Відображення  $\delta'$  визначимо як  $\delta'([A_1, A_2, \dots, A_n], t) = [B_1, B_2, \dots, B_n]$ , де для кожного  $1 \leq j \leq m$   $\delta(A_i, t) = B_j$  для будь-яких  $1 \leq i \leq n$ .
- 3) Нехай  $H = \{H_1, H_2, \dots, H_k\}$ , тоді  $H' = [H_1, H_2, \dots, H_k]$ .
- 4) Нехай  $S = \{S_1, S_2, \dots, S_p\}$ , тоді  $S'$ , всі стани з  $K'$ , що мають вид  $[\dots S_i \dots]$ .  $S_i \in S$  для якого-небудь  $1 \leq i \leq p$ .

Якщо у множині  $K'$  можуть виявитися стани, які недосяжні з початкового стану, їх можна виключити.