7.4.2 Sequences of regular definitions

Suppose we are given a sequence of regular definitions.

$$\begin{aligned} A_1 &= R_1 \\ A_2 &= R_2 \\ &\vdots \\ A_n &= R_n \end{aligned}$$

The generation of a scanner for this sequence involves the following steps:

(1) Generation of NFAs $M_i = (Q_i, \Sigma, \Delta_i, q_{0,i}, F_i)$ for the regular expressions R'_i (obtained by replacement), where the Q_i are pairwise disjoint.

(2) Construction of the NFA $M = (Q, \Sigma, \Delta, q_0, F)$ with
$Q = \bigcup_{i=1}^{n} Q_i \cup \{q_0\}, \quad q_0 \notin \bigcup_{i=1}^{n} Q_i, \quad F = \bigcup_{i=1}^{n} F_i$
$\Delta = \bigcup_{i=1}^{n} \Delta_i \cup \{(q_0, \varepsilon, q_{0,i}) \mid 1 \leq i \leq n\}$.
Thus, we obtain the NFA M for the sequence, by providing the new initial state q_0 with a ε transition to all initial states $q_{0,i}$.

(3) Application of the algorithm **NFA → DFA** (see Figure 7.5) to M, result DFA M'.

(4) Possible minimization of M'. We start with the **MinDFA** algorithm with a partition $\Pi = \{F'_1, F'_2, \ldots, F'_n, Q' - \bigcup_{i=1}^{n} F'_i\}$, where the F'_i are the final states of M' belonging to the symbol class A_i. More precisely: $F'_i = \{S \mid S \cap F_i \neq \emptyset \text{ and } S \cap F_j = \emptyset \ (1 \leq j < i)\}$ (this means that a symbol that takes the automaton M' into two former final states of the M_i is assigned to the class occurring earlier in the definition list).