

定義

有限アルファベットを Σ , 空文字列を ε とする. 任意の文字列 $p \in \Sigma^*$ と $t \in \Sigma^*$ に対して, 部分文字列パターン $\langle p \rangle$ は, $s \cdot p \cdot u = t$ となる $s, u \in \Sigma^*$ が存在するとき, テキスト t に位置 $|s|$ でマッチ (照合) するという.

Definition 1 (部分語オートマトン subword automata). ある文字列 $p \in \Sigma^*$ の部分文字列パターン $\langle p \rangle$ がマッチする文字列を受理し, そうでないものを却下する有限オートマトン M_p を p の部分語オートマトンという.

非決定性の $M_p = (Q, \delta, q_0, F)$ は, 以下のように構成できる. 状態の集合を $Q = \{0, \dots, |p|\}$ で状態数 $|p| + 1$, 初期状態を $q_0 = 0$, 受理状態の集合を $F = \{|p|\}$ とし, $0 \leq i \leq p, a \in \Sigma$ に対して遷移関係 $\delta \subseteq Q \times \Sigma \times Q$ を以下のように定義する.

$$\delta(i, a) = \begin{cases} 0 & (i = 0) \\ 1 & (i = 0, a = p[0]) \\ 0 & (0 < i < |p|, a \neq p[i]) \\ i + 1 & (0 < i < |p|, a = p[i]) \\ |p| & (i = |p|) \end{cases}$$