Алгоритм Мура

Допустим, у нас есть конечный автомат с n состояниями и алфавитом размера k. Вспомним, что такое алгоритм Мура (а именно – алгоритм минимизации DFA).

Пусть у нас есть некоторый автомат. На первом шаге разделим множество состояний на 2: принимающие и не принимающие. На каждом последующем шаге делаем проверку: если из одного множества по разным состояниям и одной и той же букве мы переходим в разные m множеств, то делим исходное множество на соответствующие m подмножеств. Продолжается алгоритм до тех пор, пока множества не перестанут делиться. Не сложно заметить, что в худшем случае время алгоритма $O(kn^2)$.

Определение. Структурой перехода называется четвёрка вида $(\Sigma, Q, \cdot, q_0)$.

Определение. Пусть E - непустое конечное множество. Для любого $\gamma \in (0, 1)$ вероятность r, определенная на подмножествах X из E, есть $r(X) = \gamma^{|X|}(1-\gamma)^{|E|-|X|}$, которая называется распределением Бернулли параметра γ на элементах E.

Определение. Для любого $n \geqslant 1$ вероятность p, определенная на D_n (множество DFA с n состояниями), является моделью Бернулли, когда существует вероятность q, определенная на T_n (множество структур перехода с n состояниями), и действительное число $\gamma \in (0, 1)$ такое, что для любого $A \in D_n$, $p(A) = q(T_A)r(F_A)$, где $A = (T_A, F_A)$ и r - распределение Бернулли по параметру γ на состояниях.

Определение. Пусть $\gamma \in (0, 1)$. Пусть $p : D \to [0, 1]$ - такое отображение, что для любого $n \geqslant 1$ ограничение p на D_n является моделью Бернулли параметра γ. Тогда p - модель Бернулли на D (параметра γ).

Теорема 1. *Для любой моделли Бернулли на D средняя временная сложность (ожидаемое количество итераций в алгоритме Мура для случайно выбранного автомата) алгоритма минимизации равна $O(kn \log n)$.*

То есть в среднем время работы алгоритма значительно лучше. Для доказательства этого факта мы находили условия для того, чтобы алгоритм Мура делал больше, чем l итераций и замечали, что вероятность данного события экспоненциально мала. Затем мы выбирали $l = \log n$.