Лекция: Регулярные выражения и регулярные множества. Теорема Клини о совпадении классов автоматных множеств и регулярных множеств.

Лектор - доцент Селезнева Светлана Николаевна

Лекции по "Дискретной математике 2".
1-й курс, группа 141,
факультет ВМК МГУ имени М.В. Ломоносова

Лекции на сайте http://mk.cs.msu.su

Регулярные выражения

Определим по индукции **регулярные выражения** над конечным алфавитом A.

Базис индукции:
1) $\emptyset$ – регулярное выражение;
2) a, где $a \in A$, – регулярное выражение;
3) Λ, где $\Lambda \in A^*$ – пустое слово, – регулярное выражение.

Индуктивный переход: пусть s и t – регулярные выражения. Тогда
1) $s + t$ – регулярное выражение;
2) st – регулярное выражение;
3) $(s)^*$ – регулярное выражение.

Других регулярных выражений нет.

Множества, определяемые регулярными выражениями

Каждое **регулярное выражение** в алфавите A определяет некоторое **множество (язык)** $L \subseteq A^*$.

Базис индукции:
1) регулярное выражение $\emptyset$ определяет пустое множество $\emptyset$;
2) регулярное выражение a, $a \in A$, определяет множество $\{a\}$;
3) регулярное выражение Λ, где $\Lambda \in A^*$ – пустое слово, определяет множество $\{\Lambda\}$.

Индуктивный переход: пусть регулярные выражения s, t определяют соответственно множества $S, T \subseteq A^*$. Тогда
1) регулярное выражение $s + t$ определяет объединение $S \cup T$;
2) регулярное выражение st определяет произведение ST;
3) регулярное выражение $(s)^*$ определяет итерацию S^*.

Регулярные множества

Множество L слов в конечном алфавите A называется **регулярным множеством** (**языком**), если оно определяется некоторым регуляным выражением.

Заметим, что одно и то же регулярное множество может определяться **разными** регулярными выражениями.

Примеры

Пример 1. Пусть $A = \{0, 1\}$.

1. Регулярное выражение 01 обозначает множество $\{01\}$, содержащее одно слово $\alpha = 01$.

2. Регулярное выражение $(0+1)^*011$ обозначает множество всех слов из нулей и единиц, оканчивающихся на 011.

3. Докажем, что множество всех слов из нулей и единиц, которые содержат две единицы подряд, является регулярным. Заметим, что это множество задается регулярным выражением $(0+1)^*11(0+1)^*$.

Автоматность регулярных множеств

Теорема 1. *Пусть A – конечный алфавит. Каждое регулярное множество $L \subseteq A^*$ является также и автоматным множеством.*

Автоматность регулярных множеств

Доказательство. Пусть $A = \{a_1, \dots, a_n\}$.

Докажем сначала автоматность множеств $\emptyset$, $\{\Lambda\}$ и $\{a_i\}$, $i = 1, \dots, n$.

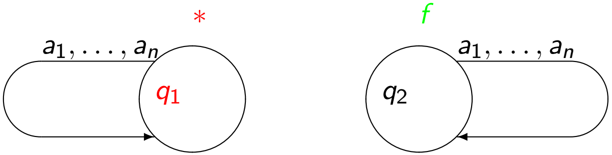
Для доказательства построим диаграммы переходов ДКА без выхода, которые принимают перечисленные множества.

Автоматность регулярных множеств

Доказательство (продолжение). ДКА без выхода, принимающий множество $\emptyset$:

Автоматность регулярных множеств

Доказательство (продолжение). ДКА без выхода, принимающий множество $\emptyset$:

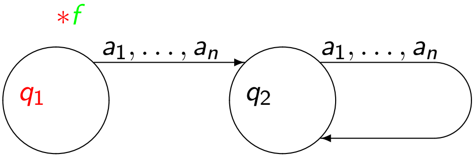


Автоматность регулярных множеств

Доказательство (продолжение). ДКА без выхода, принимающий множество $\{\Lambda\}$:

Автоматность регулярных множеств

Доказательство (продолжение). ДКА без выхода, принимающий множество $\{\Lambda\}$:

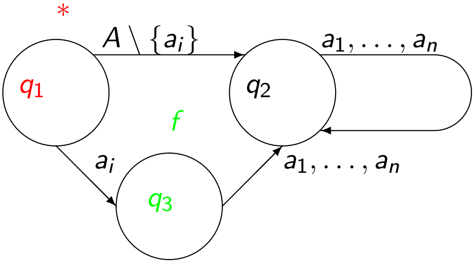


Автоматность регулярных множеств

Доказательство (продолжение). ДКА без выхода, принимающий множество $\{a_i\}$:

Автоматность регулярных множеств

Доказательство (продолжение). ДКА без выхода, принимающий множество $\{a_i\}$:



Автоматность регулярных множеств

Доказательство (продолжение). Кроме того, мы уже доказали, что операции объединения, произведения и итерации сохраняют автоматность множеств.

□

Регулярность автоматных множеств

Теорема 2. *Пусть A – конечный алфавит. Каждое автоматное множество $L \subseteq A^*$ является также и регулярным множеством.*

Регулярность автоматных множеств

Доказательство. Пусть $A = \{a_1, \dots, a_n\}$. Множество L – автоматное, значит найдется такой ДКА без выхода

$$\mathcal{A} = (A, Q, \psi, {\color{red}q_1}, {\color{green}F}),$$

что $L = L(\mathcal{A})$.

Регулярность автоматных множеств

Доказательство (продолжение). Пусть $Q = \{q_1, \ldots, q_r\}$. Не ограничивая общности рассуждений, пусть $F = \{q_r\}$.

Определим множества Z_{ij}^k, $Z_{ij}^k \subseteq A^*$, как множества слов в алфавите A, по которым КА без выхода $\mathcal{A}$ из состояния q_i переходит в состояние q_j, проходя **только** через состояния $q_1, \ldots, q_k$.

Т.е.

$$\begin{aligned} Z_{ij}^k &= \{\alpha \in A^* \mid \bar{\psi}(\alpha, q_i) = q_j, \bar{\psi}(\beta, q_i) \in \{q_1, \ldots, q_k\}, \\ &\qquad \alpha = \beta\gamma, \beta, \gamma \in A^*, \beta, \gamma \neq \Lambda\}. \end{aligned}$$

Докажем индукцией по k, что каждое из множеств Z_{ij}^k является регулярным.

Регулярность автоматных множеств

Доказательство (продолжение).

Базис индукции: $k = 0$. Множества Z_{ij}^0 состоят из всех слов в алфавите A, которые переводят КА $\mathcal{A}$ из состояния q_i в состояние q_j, **не проходя** через промежуточные состояния.

Если $i = j$, то Z_{ii}^0 содержит пустое слово Λ и все слова в алфавите $B \subseteq A$, где

$$B = \{a \in A \mid \psi(a, q_i) = q_i\}.$$

Если $i \neq j$, то

$$Z_{ij}^0 = \{a \in A \mid \psi(a, q_i) = q_j\}.$$

Значит, Z_{ij}^0 – регулярны.

Регулярность автоматных множеств

Доказательство (продолжение).

Индуктивный переход: Пусть множества Z_{ij}^{k} регулярны. Рассмотрим множества Z_{ij}^{k+1}. Оно содержит все слова в алфавите A, которые переводят КА $\mathcal{A}$ из состояния q_i в состояние q_j, проходя **только** через состояния $q_1, \ldots, q_k, q_{k+1}$.

Регулярность автоматных множеств

Доказательство (продолжение).

Пусть $\alpha \in Z_{ij}^{k+1}$, т.е. $\bar{\psi}(\alpha, q_i) = q_j$. Посмотрим, какие промежуточные состояния проходит КА $\mathcal{A}$ при чтении слова α:

1) если КА $\mathcal{A}$ **не проходит** через состояние q_{k+1}, то $\alpha \in Z_{ij}^{k}$;

2) если КА $\mathcal{A}$ **проходит** через состояние q_{k+1}, то слово α можно разбить на такие слова $\beta, \gamma_1, \ldots, \gamma_t, \delta \in A^*$, что $\alpha = \beta\gamma_1 \ldots \gamma_t\delta$, и

$$\beta \in Z_{i(k+1)}^{k}, \ \gamma_1, \ldots, \gamma_t \in Z_{(k+1)(k+1)}^{k}, \ \delta \in Z_{(k+1)j}^{k}.$$

Получаем, $Z_{ij}^{k+1} = Z_{ij}^{k} \cup Z_{i(k+1)}^{k}(Z_{(k+1)(k+1)}^{k})^* Z_{(k+1)j}^{k}$.

Т.е. Z_{ij}^{k+1} – регулярное множество.

Регулярность автоматных множеств

Доказательство (продолжение). Осталось заметить, что $L = Z^{r}_{1r}$. $\square$

Теорема Клини

Теорема 3 (Клини). *Пусть A – конечный алфавит. Класс автоматных множеств в алфавите A совпадает с классом регулярных множеств в алфавите A.*

Примеры

Пример 2. Пусть $A = \{0, 1\}$.

1. Доказать, что множество L_1 слов в алфавите A четной длины является автоматным множеством. Заметим, что множество L_1 задается регулярным выражением

$$(00 + 01 + 10 + 11)^*.$$

Значит L_1 – автоматное множество.

2. Доказать, что множество L_2 всех слов в алфавите A, кроме слова 01, является регулярным множеством. Заметим, что $L_2 = A^* \setminus \{01\}$. Множества A^* и $\{01\}$ – автоматные множества, разность автоматных множеств – также автоматное множество. Значит, L_2 – регулярное множество.

Можно найти регулярное выражение, определяющее множество L_2:

$$\Lambda+0+1+00+10+11+(000+001+010+011+100+101+110+111)(0+1)^*.$$

Задачи для самостоятельного решения

1. Какие множества в алфавите $A = \{0, 1\}$ определяют следующие регулярные выражения:

1) $(0 + 11) * 1$;

2) $10^*1(0 + 1)^*1$?

2. Найдите регулярные выражения, определяющие следующие множества:

1) множество слов из нулей и единиц, содержащих и подслово 00, и подслово 11;

2) множество слов из нулей и единиц, **не** содержащих ни подслово 00, ни подслово 11.

Литература к лекции

1. Марченков С.С. Конечные автоматы. М.: Физматлит, 2008.

Конец лекции