

Школа лингвистики, 2023-24 уч. год**Дискретная математика для лингвистов****Регулярные языки и автоматы (9 – 19 января 2024 года)**

В. В. Кочергин, Ю. Г. Кудряшов, А. В. Михайлович, И. В. Щуров, И. А. Хованская

Задача 1. Являются ли следующие языки регулярными? Если да, то постройте как можно более простую диаграмму, задающую следующие языки над алфавитом $\{a, b\}$.

- Все возможные слова, состоящие из букв a и b .
- Все слова, начинающиеся с буквы a .
- Все слова, начинающиеся с буквы a , и такие, что после любой буквы b должна следовать буква a .
- Все слова, начинающиеся с буквы a и заканчивающиеся буквой b .
- Слово $abbbababb$ (и никакие другие).
- Слово $abbbababb$ или слово $bbabbaa$ (и никакие другие).
- Слово $ababb$, повторённое сколько угодно раз, то есть все слова вида «», « $ababb$ », « $ababbababb$ », « $ababbababbababb$ » и т.д.
- Все слова, получающиеся следующей процедурой. Запишем произвольное слово из букв A и B . Затем заменим каждое вхождение буквы A на $abba$, а каждое вхождение слова B на aab .
- Все слова, содержащие чётное число букв.

Задача 2. Являются ли следующие языки регулярными? Если да, то постройте диаграмму для них.

- Непустые слова, содержащие только букву « a » (одну или много) и никаких других букв.
- Слова, содержащие хотя бы три буквы « a » и не содержащие никаких других букв.
- Слова, содержащие чётное количество букв « a » (и, возможно, какое-то количество букв « b »).
- Слова, длина которых делится на 3.
- Слова, начинающиеся и заканчивающиеся на букву « b ».
- Слова, в которых встречаются буквы « a » и « b », причём первое вхождение « a » встречается раньше, чем первое вхождение « b ».
- Слова, в которых запрещено ставить три буквы « a » подряд.
- Слова, в которых встречается хотя бы раз подслово « aba ».
- Слова в алфавите из трёх букв $\{a, b, c\}$, в которых встречается хотя бы раз подслово « $abac$ ».

Задача 3. Пусть $A = \{a, b\}$, L_1 — некоторый язык над алфавитом A . Пусть $B = \{c, d, e, f\}$, α, β — слова над алфавитом B . Язык L_2 получен из языка L_1 заменой буквы a на слово α , а буквы b на слово β . Язык L_1 является регулярным. Является ли регулярным язык L_2 ?

- Задача 4.** (a) Построить диаграмму, задающую язык L_1L_2 , если $L_1 = \{ab, cd\}$, $L_2 = \{e, f\}$;
- Построить диаграмму, задающую язык L^* , если $L = \{aba\}$;
 - Построить диаграмму, задающую язык L^* , если $L = \{aba, abb, aab\}$;
 - Построить диаграмму, задающую язык L_1L_2 , если $L_1 = \{ab\}^*$, $L_2 = \{ad, cd\}$.

Задача 5. Для языков из задач 1, 2, 4 построить регулярные выражения.

Задача 6. Постройте диаграмму, реализующий тот же язык, что и регулярное выражение.

- (a) $(\{a\}\{a\} \cup \{b\}\{b\})^*$;
- (b) $(\{a\} \cup \{b\})^*\{a\}$;
- (c) $(\{a\} \cup \{b\})^*\{a\}(\{a\} \cup \{b\})(\{a\} \cup \{b\})$;
- (d) $(\{a\} \cup \{b\})^*\{a\}(\{a\}\{a\} \cup \{b\}\{b\})$.

Задача 7. Пусть $A = \{a, b\}$, $Q = \{q_0, q_1, q_2, q_3, q_4\}$, $Q' = \{q_3, q_4\}$, функция $G : A \times Q \rightarrow Q$ задаётся следующим образом:

$$\begin{aligned} G(a, q_0) &= q_0; \\ G(b, q_0) &= q_2; \\ G(a, q_1) &= q_1; \\ G(b, q_1) &= q_4; \\ G(a, q_2) &= q_1; \\ G(b, q_2) &= q_3; \\ G(a, q_3) &= q_4; \\ G(b, q_3) &= q_1; \\ G(a, q_4) &= q_2; \\ G(b, q_4) &= q_0. \end{aligned}$$

Нарисовать диаграмму для автомата (A, Q, G, q_0, Q') и написать регулярное выражение для языка L , представляемого этим автоматом. Построить автомат для языков L^* , CL (дополнение языка L)

Задача 8. Построить конечный автомат, как можно более простую диаграмму и регулярное выражение, описывающее язык L_1 над алфавитом $A = \{a, b, c\}$, который содержит только те слова, которые не содержат одновременно все три буквы (в любом порядке) и язык CL_1 .

Задача 9. Построить регулярные выражения, задающие те же языки, что и диаграммы на рисунке 1.

Задача 10. Доказать, что язык, содержащий только слова, в котором букв «а» больше, чем букв «b», не является регулярным.

Задача 11. Является ли регулярным язык над алфавитом $A = \{a\}$, содержащий все слова, длина которых выражается простым числом, и только их.

Задача 12. Является ли регулярным язык над алфавитом $S = \{(,)\}$ (символы алфавита — открывающая и закрывающая скобки), содержащий все слова, в которых выполняется «баланс скобок», то есть слова, содержащие одинаковое количество символов (и), причём в любом префиксе этого слова число символов (не превосходит число символов (, и только их.

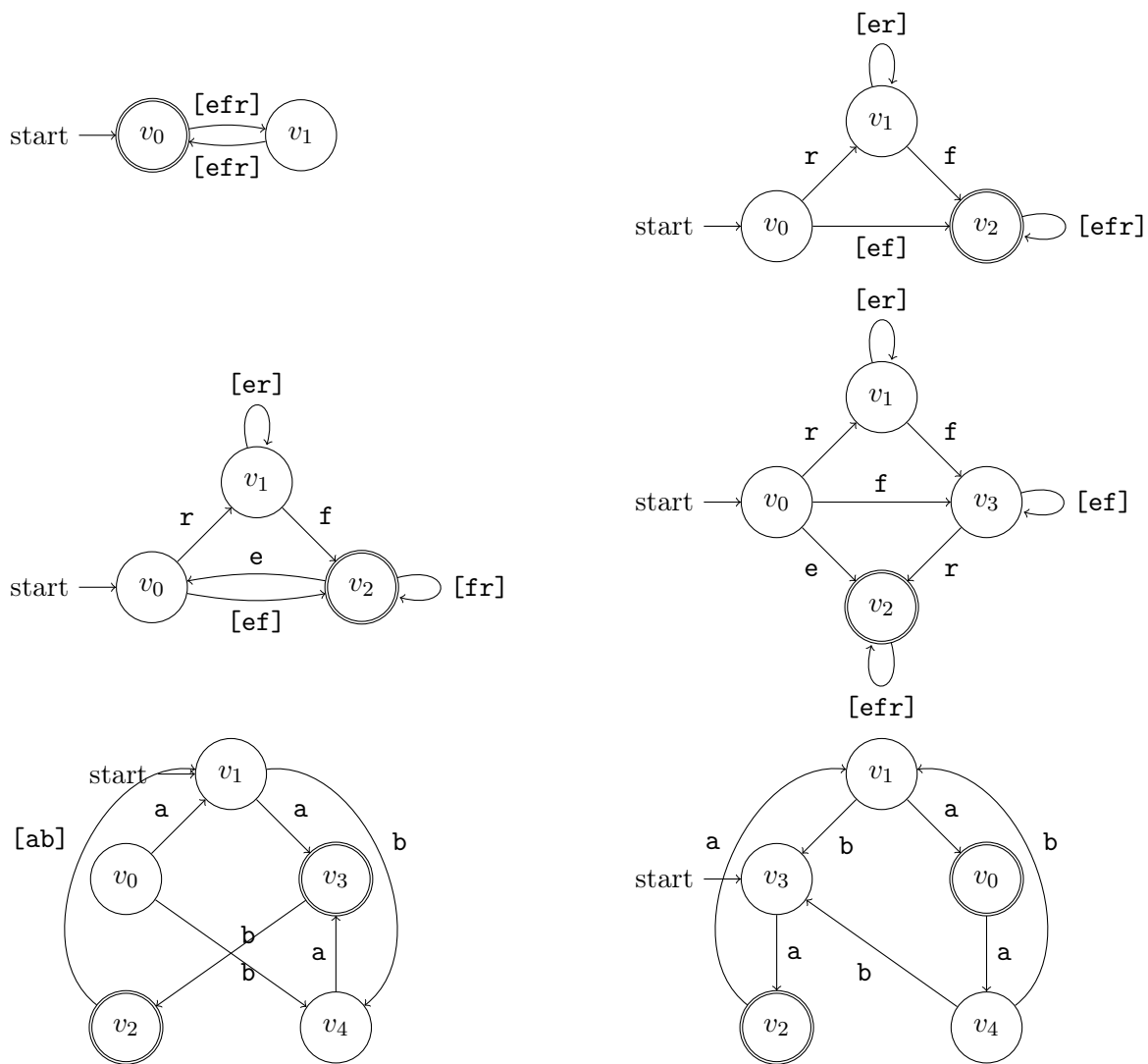


Рис. 1: К задаче 9