

Теория автоматов

Лекция 7: ДКА = регекс

Дьулустан Никифоров

Кафедра ИТ
Северо-Восточный Федеральный Университет

Осень 2024

- Теперь мы знаем, что такое конечные автоматы и как с ними работать;
- Мы знаем, что такое регулярные выражения и как с ними работать;
- Эти две конструкции выглядят *совершенно* по разному и работают по непохожим принципам.
- Однако, удивительный факт в том, что оба вида “машин” описывают одно и то же множество языков — регулярные языки!

Theorem

Язык регулярный тогда и только тогда, когда какое-то регулярное выражение распознает его.

Утверждение теоремы состоит из двух частей:

- регулярный язык \Rightarrow регекс;
- регекс \Rightarrow регулярный язык.

Theorem

Язык регулярный тогда и только тогда, когда какое-то регулярное выражение распознает его.

Утверждение теоремы состоит из двух частей:

- регулярный язык \Rightarrow регекс;
- регекс \Rightarrow регулярный язык.

Поэтому доказательство тоже будет состоять из двух частей:

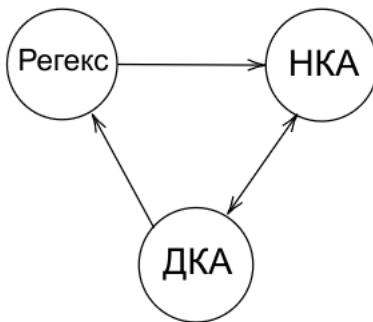
- **Регекс \Rightarrow регулярный язык.**

Для этого мы придумаем алгоритм превращения [регекс \rightarrow НКА].

- **Регулярный язык \Rightarrow регекс.**

Для этого мы придумаем алгоритм превращения [ДКА \rightarrow регекс].

Автоматы = Регексы



Доказательство позволяет увидеть глубокую фундаментальную связь между НКА и регексами.

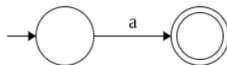
Proof: [регекс \rightarrow НКА]

Пусть нам дан регекс R . Тогда, по определению, он выглядит как один из этих случаев:

- a для некоторого $a \in \Sigma$,
- ε ,
- \emptyset ,
- $(R_1 + R_2)$, где R_1 и R_2 регулярные выражения,
- $(R_1 R_2)$, где R_1 и R_2 регулярные выражения, или
- R_1^* , где R_1 регулярное выражение.

Proof: [регекс \rightarrow НКА]

- a для некоторого $a \in \Sigma$:



- ε :



- \emptyset :



Proof: [регекс \rightarrow НКА]

- $R_1 + R_2$:
строим НКА для R_1 и $R_2 \Rightarrow$ далее строим НКА для $(R_1 + R_2)$;
- $R_1 R_2$:
строим НКА для R_1 и $R_2 \Rightarrow$ далее строим НКА для $R_1 R_2$;
- R_1^* :
строим НКА для $R_1 \Rightarrow$ далее строим НКА для R_1^* .

Пример 1: $(ab + a)^* \rightarrow$ НКА.

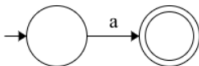
По алгоритму:

Регекс \rightarrow НКА: примеры

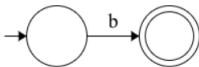
Пример 1: $(ab + a)^* \rightarrow$ НКА.

По алгоритму:

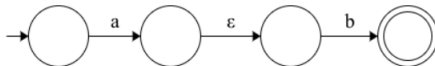
a :



b :

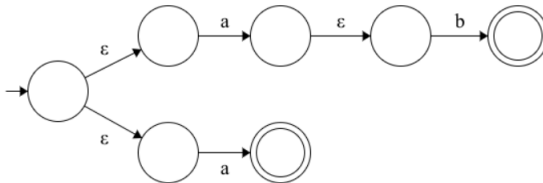


ab :

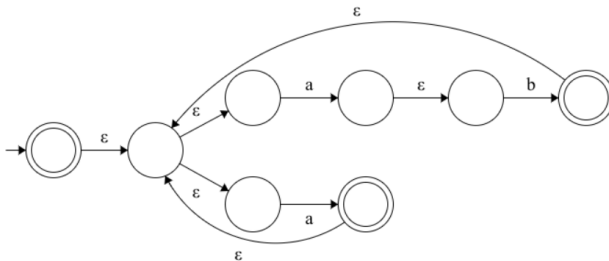


Регекс \rightarrow НКА: примеры

$ab + a$:

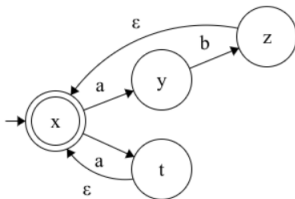


$(ab + a)^*$:



Регекс \rightarrow НКА: примеры

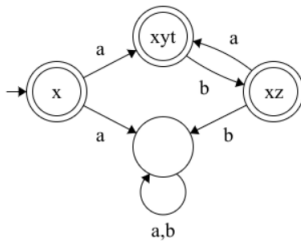
Полученный НКА достаточно очевидно можно сократить:



Можно было такой НКА построить сразу из регекса (юзать мозг).

Регекс \rightarrow НКА: примеры

Далее, можем также НКА превратить в ДКА:

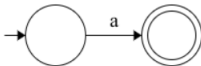


Таким образом, можно регекс превратить в эквивалентный ей ДКА!

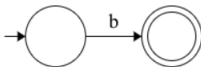
Регекс \rightarrow НКА: примеры

Пример 2: $(a + b)^*aba \rightarrow$ НКА

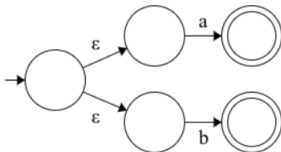
a :



b :

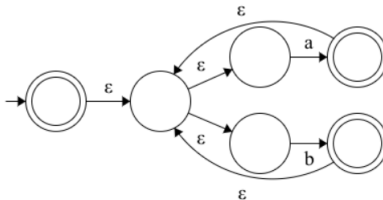


$a + b$:

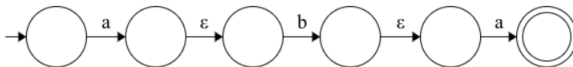


Регекс \rightarrow НКА: примеры

$(a + b)^*$:

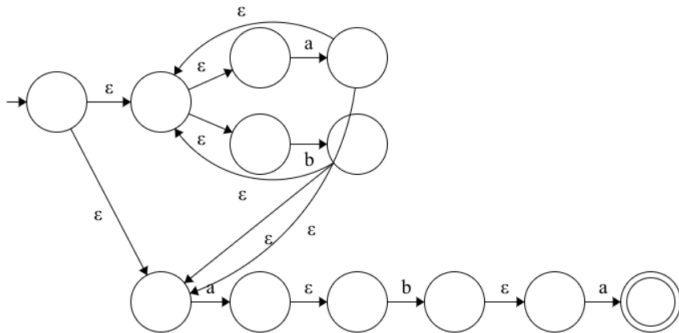


aba :



Регекс \rightarrow НКА: примеры

$(a + b)^*aba$:



Proof: [ДКА \rightarrow Регекс]

Вот это более сложное доказательство и алгоритм!

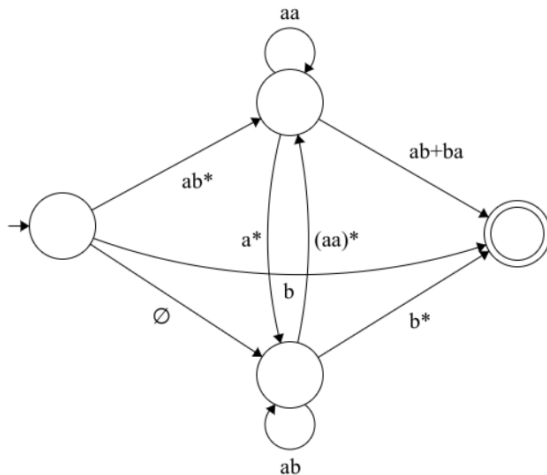
Конструкцию будем делать в два этапа:

- ДКА \rightarrow **ОНКА** (обобщенный НКА).
- ОНКА \rightarrow регекс.

ОНКА — обобщение НКА, которому мы разрешаем в переходах кушать не просто символы, но и целые регексы.

ДКА \rightarrow ОНКА

Пример ОНКА:

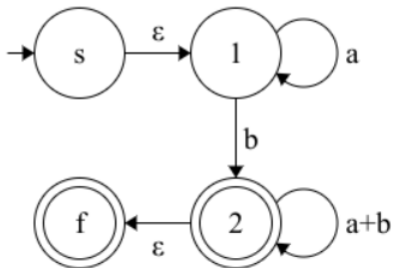
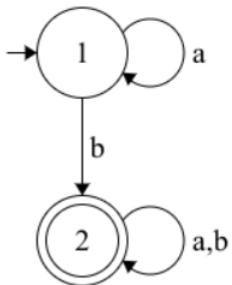


Алгоритм превращения ДКА в ОНКА (1-й этап построения регекса):

- Добавить новое начальное состояние s (старое нач. состояние перестает быть начальным). Направить от s ε -переходы в старое начальное состояние.
- Добавить новое принимающее состояние f (остальные перестают быть принимающими). Направить от всех старых принимающих состояний ε -переходы к f .
- Для любых двух состояний i, j , все переходы $i \rightarrow j$ объединяем в один переход (например, переходы по буквам x, y, z объединяются в переход $x + y + z$).

ДКА \rightarrow ОНКА

Пример:

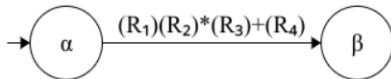
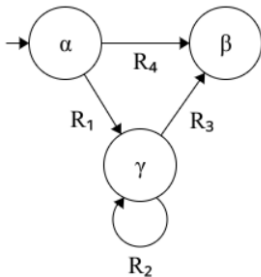


ОНКА \rightarrow Регекс:

- ОНКА с k состояниями — убрать одно из состояний, кроме начального состояния s и принимающего состояния f .
- Починить оставшийся автомат так, чтобы скоменсировать потерю состояния (как это сделать — на следующем слайде).
- Получили ОНКА с $k - 1$ состояниями: repeat until $k = 2$.
- В конце остается ОНКА только с двумя состояниями s и f . На переходе $s \rightarrow f$ написан желаемый регекс!

ОНКА \rightarrow Регекс:

Если убрали состояние γ , то необходимо для *всех* пар состояний α, β дополнить переход $\alpha \rightarrow \beta$ (точнее, его регекс) всеми возможными путями через γ :

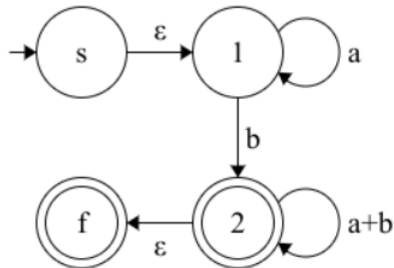
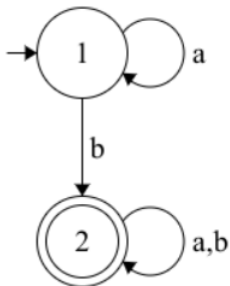


Это надо проделать для **всех** пар состояний α, β таких, что есть путь $\alpha \rightarrow \gamma \rightarrow \beta$.

ДКА \rightarrow Регекс: Примеры

Пример 1: ДКА \rightarrow ОНКА \rightarrow Регекс

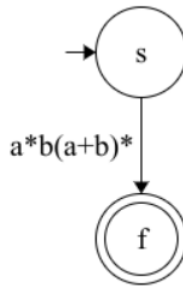
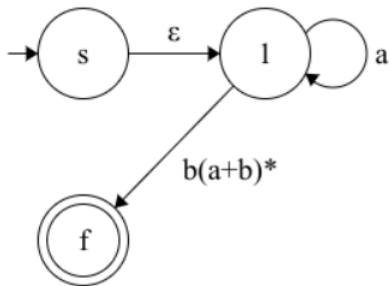
Сначала превращаем в подходящий ОНКА.



ДКА \rightarrow Регекс: Примеры

Пример 1: ДКА \rightarrow ОНКА \rightarrow Регекс

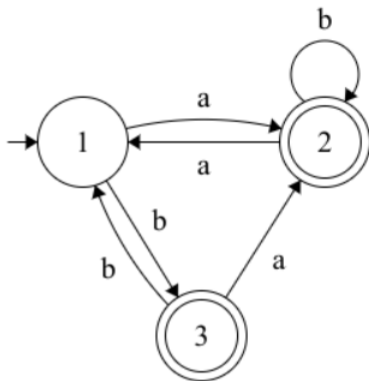
Сначала избавились от состояния 2, потом от состояния 1.
Можно и в другом порядке.



Ответ: $a^*b(a+b)^*$

Можно было сразу получить ответ, юзая мозг.

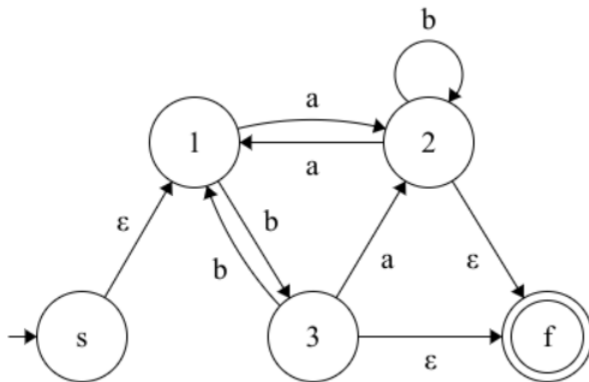
Пример 2: ДКА \rightarrow ОНКА \rightarrow Регекс



ДКА \rightarrow Регекс: Примеры

Пример 2: НКА \rightarrow ОНКА \rightarrow Регекс

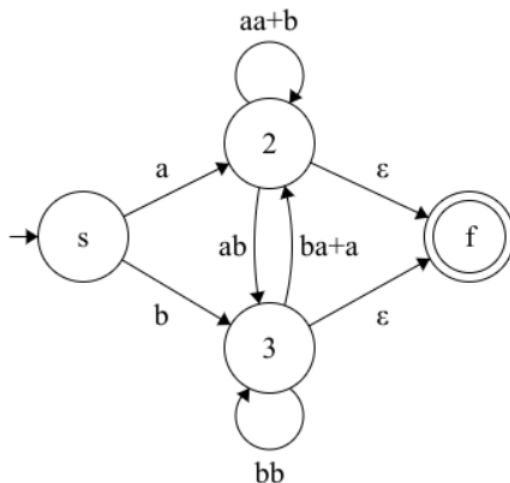
Сначала превращаем в подходящий ОНКА.



ДКА \rightarrow Регекс: Примеры

Пример 2: НКА \rightarrow ОНКА \rightarrow Регекс

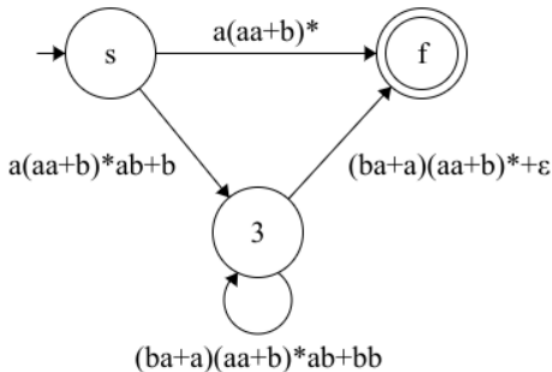
Избавились от состояния 1.



ДКА \rightarrow Регекс: Примеры

Пример 2: НКА \rightarrow ОНКА \rightarrow Регекс

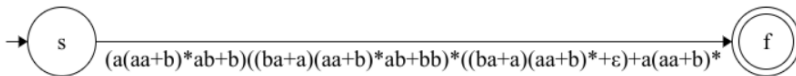
Избавились от состояния 2.



ДКА → Регекс: Примеры

Пример 2: НКА → ОНКА → Регекс

Избавились от состояния 3.



Получили отличный регекс!

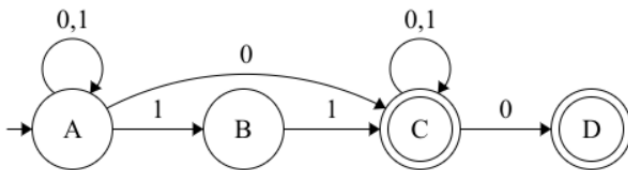
Ответ:

$$(a(aa+b)^*ab+b)((ba+a)(aa+b)^*ab+bb)^*((ba+a)(aa+b)^*+\epsilon)+a(aa+b)^*$$

Без алгоритма вряд ли получилось бы!

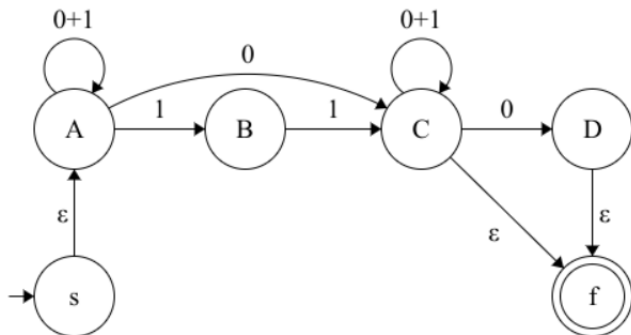
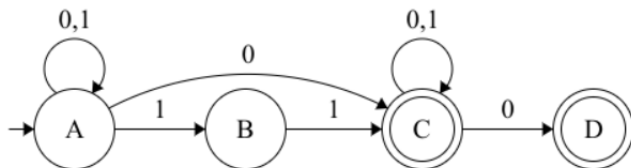
ДКА \rightarrow Регекс: Примеры

Пример 3: НКА \rightarrow ОНКА \rightarrow Регекс

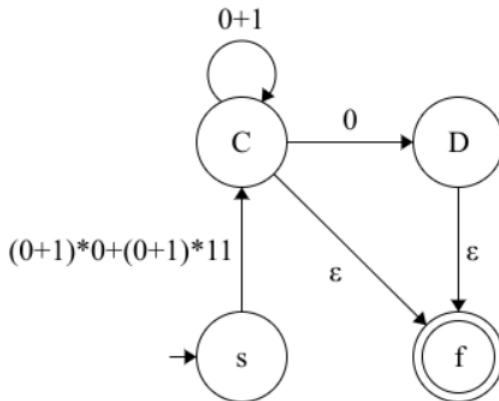


ДКА \rightarrow Регекс: Примеры

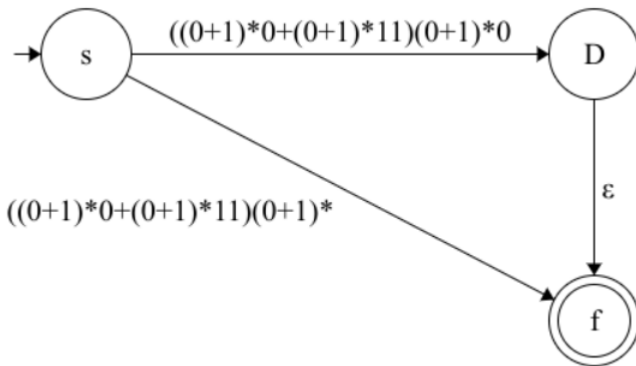
Пример 3: НКА \rightarrow ОНКА \rightarrow Регекс



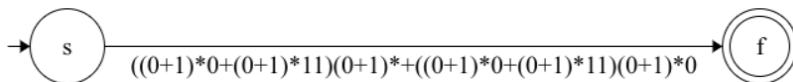
Пример 3: НКА → ОНКА → регекс



Пример 3: НКА → ОНКА → регекс



Пример 3: НКА → ОНКА → регекс



Видим, можно полученный регекс сократить. Лучше было сокращать сразу в ходе удаления состояний (естественным образом).

Ответ: $(0+1)^*(0+11)(0+1)^*(0+\varepsilon)$.

Тоже несложно получить напрямую мозгом, глядя на исходный автомат.