Regular expressions over Σ

$\emptyset$ is an r.e.

ε . . . . r.e.

a . . . . . for each $a \in \Sigma$

if R_1 & R_2 are r.e.s,
then so are

$(R_1 \cup R_2)$

$(R_1 \circ R_2)$

(R_1^*)

---

the language denoted by R, $L(R)$
is :

$L(\emptyset) = \emptyset$

$L(\varepsilon) = \{\varepsilon\}$

$\vdots$

$L(R_1 \cup R_2) = L(R_1) \cup L(R_2)$

Theorem:

$\forall$ regular expression R $\exists$ an NFA M_R st $L(R) = L(M_R)$

Proof:

By induction on K, the # of $\cup, \circ, *$ operators in R

Base cases ($K=0$):

Then R is "ϕ", "ε", or "a" for $a \in \Sigma$.

Explicitly give simple NFA's recognizing ϕ, $\{\varepsilon\}$, and $\{a\}$ for each $a \in \Sigma$. (details omitted)

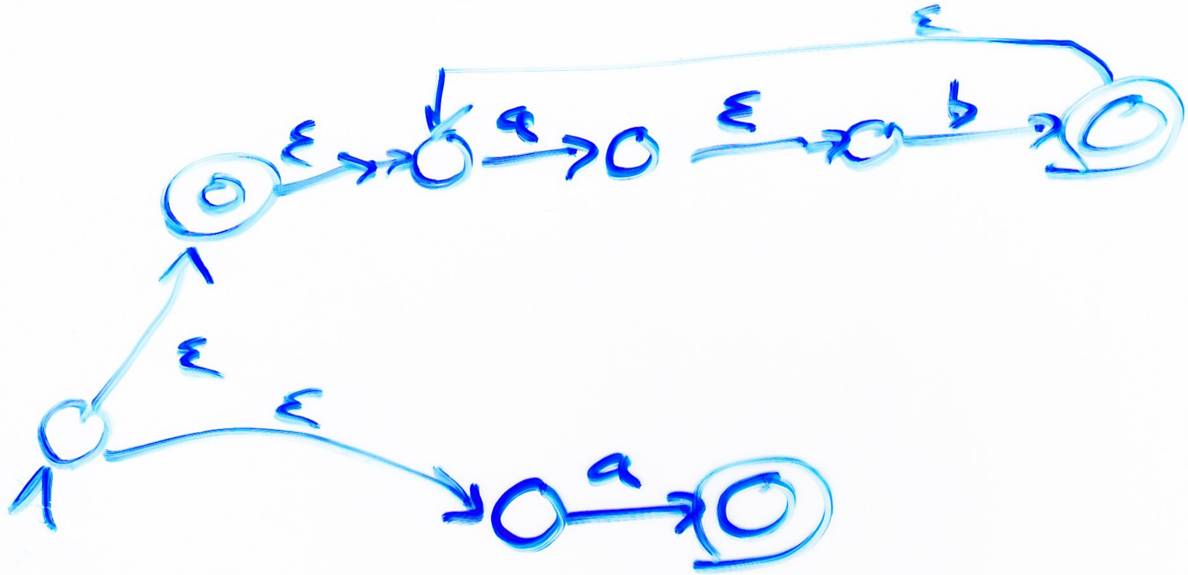
Induction Step (R has $K > 0$ operators)

> I.H.: assume that for all regular expressions R' with $< K$ operators, $\exists$ NFA $M_{R'}$ recognizing $L(R')$

R has $K > 0$ operators. So R is $(R_1 \cup R_2)$ or $(R_1 \circ R_2)$ or $(R_1)^*$ where R_1 ($\& R_2$ if any) have $\leq K-1$ operators. By I.H., $\exists M_{R_1}$ ($\& M_{R_2}$) st. $L(R_i) = L(M_{R_i})$, $i=1,2$. Modify/join it/them as in previous proofs of closure under $\cup, \circ, *$ to get M_R st. $L(R) = L(M_R)$.

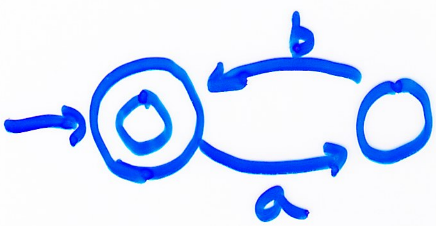
Example

$(ab)^* \cup a$



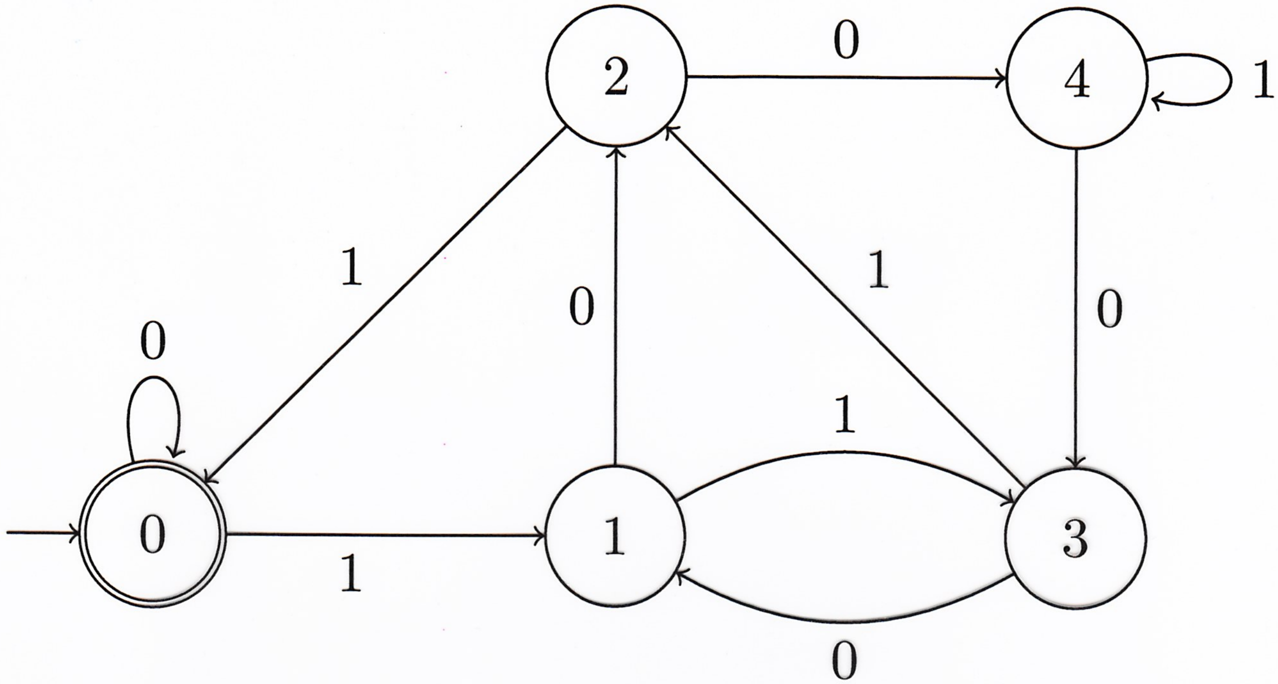
Converse?

For every D/NFA $\exists$ reg expr defining the same language



$\rightarrow$ $(ab)^*$

$\xrightarrow{?}$



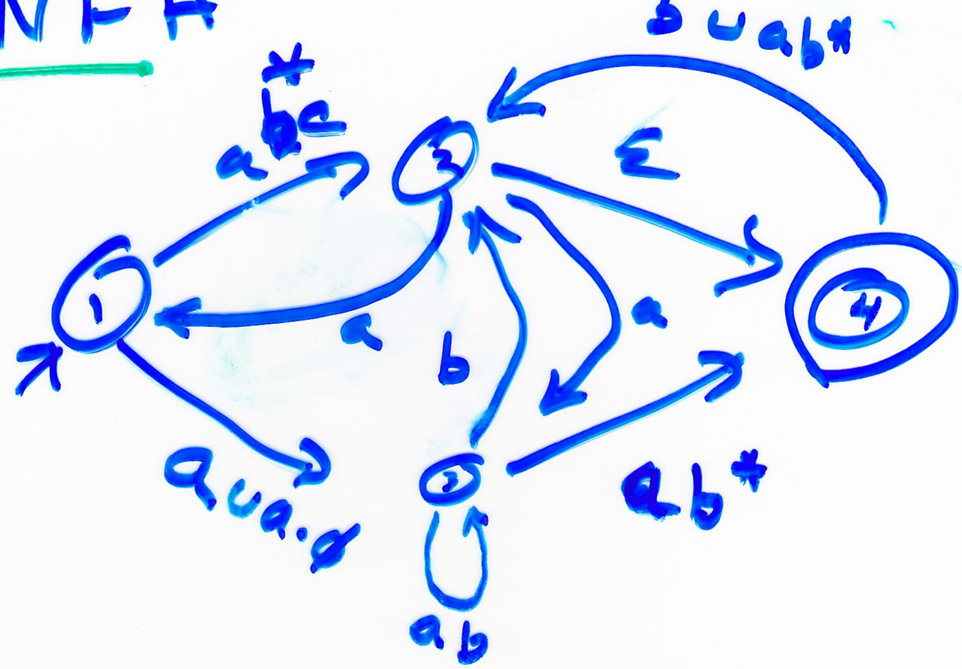
divisible by 5

pattern?

0	0
101	5
1010	10
1111	15
10100	20
11001	25
11110	30
100011	35
101000	40
101101	45
110010	50
110111	55
111100	60
⋮	

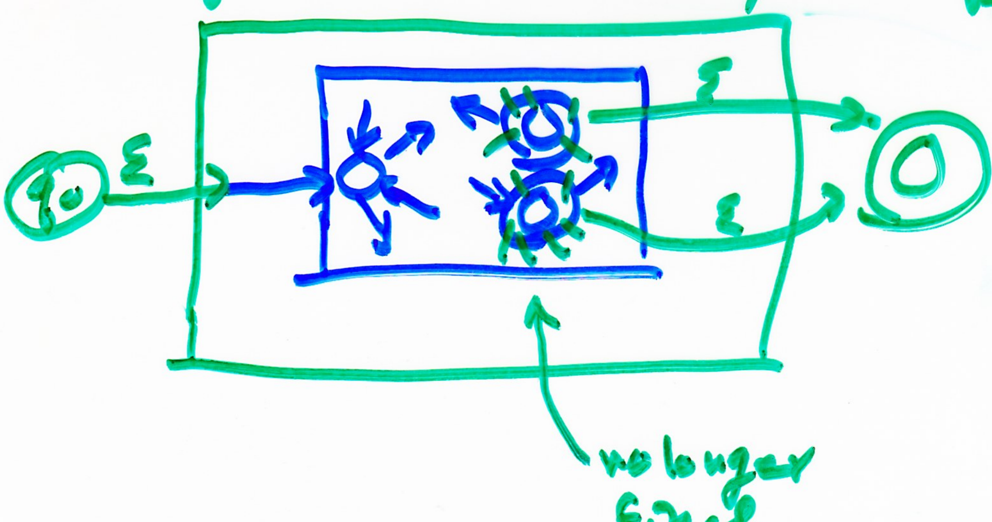
Every ~~FA~~ Regular language can be described by a regular expression.

GNFA



$_1\, a\, bbc\, _2\, a\, _3\, ab\, _3\, ab\, _3\, abb\, _4$

Note: No loss in assuming no edges into q_0 / out of F / only one $q_f \in F$



GNFA

$G = (Q, \Sigma, \delta, q_0, q_f)$

$Q, \Sigma, q_0, q_f \in Q$ as usual

Regular expressions over Σ

$\delta : (Q - \{q_f\}) \times (Q - \{q_0\}) \to R_\Sigma$

Defn

- G can be in state $q \in Q$ after reading $x \in \Sigma^*$ if $\exists k \geq 0$,

  $\exists\ r_0, r_1, \cdots, r_k \in Q$

  $\exists\ x_1, \cdots, x_k \in \Sigma^*$

  such that

  (i) $x = x_1 \cdot x_2 \cdot \cdots \cdot x_k$

  (ii) $r_0 = q_0$

  (ii) $r_k = q$

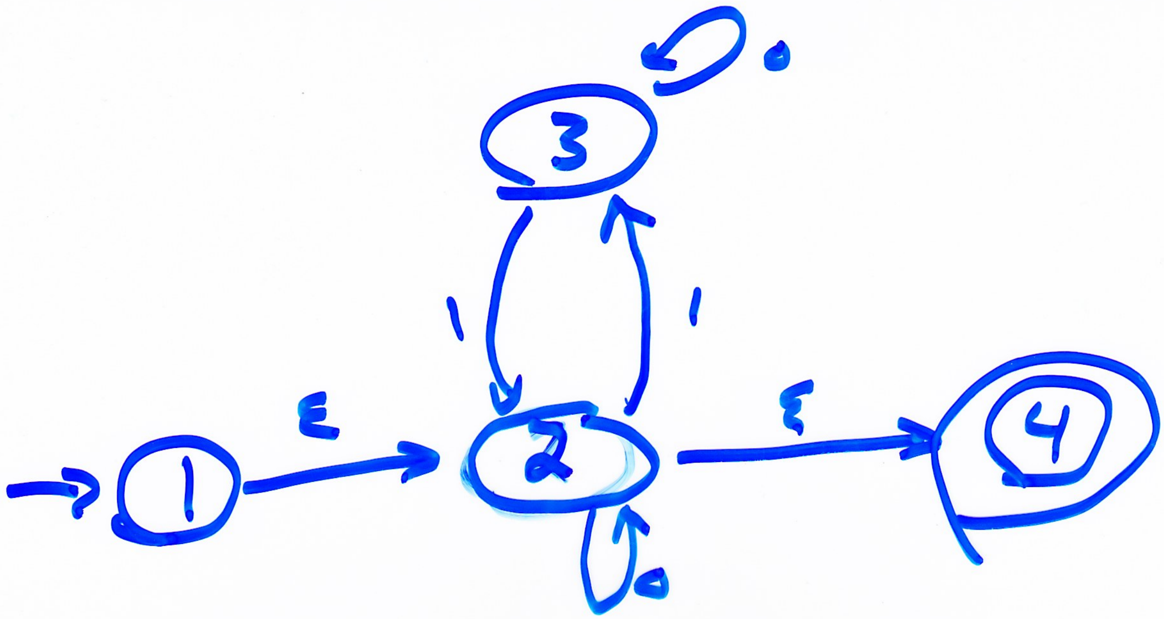
  (iii) $\forall\ 1 \leq i \leq k,\ x_i \in L(\delta(r_{i-1}, r_i))$
- $L(G) = \{ x \mid G$ can be in state $q_f \cdots \}$

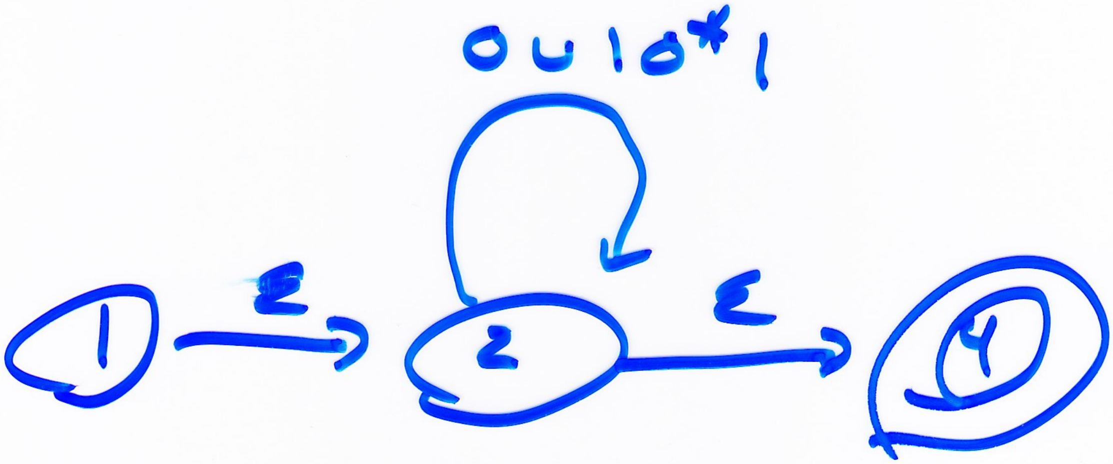
Note: δ syntax a little different;
maps state pair to label (reg. exp.)
rather than state $\times$ symbol $\to$ new state

Theorem

If L is accepted by a GNFA, then L is regular

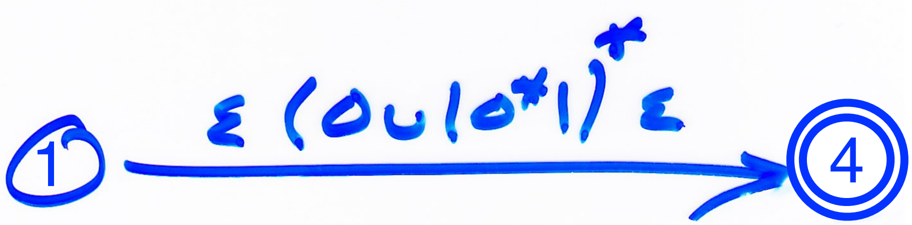
Pf sketch:
replace edge labeled "r" by NFA equivalent to r based on previous theorem.

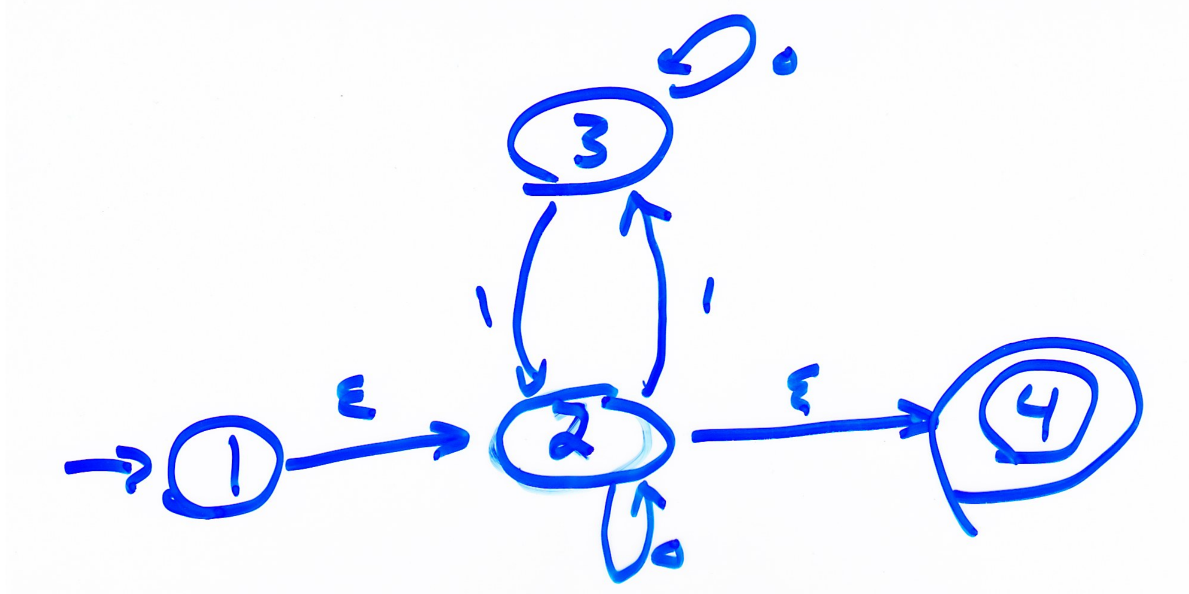




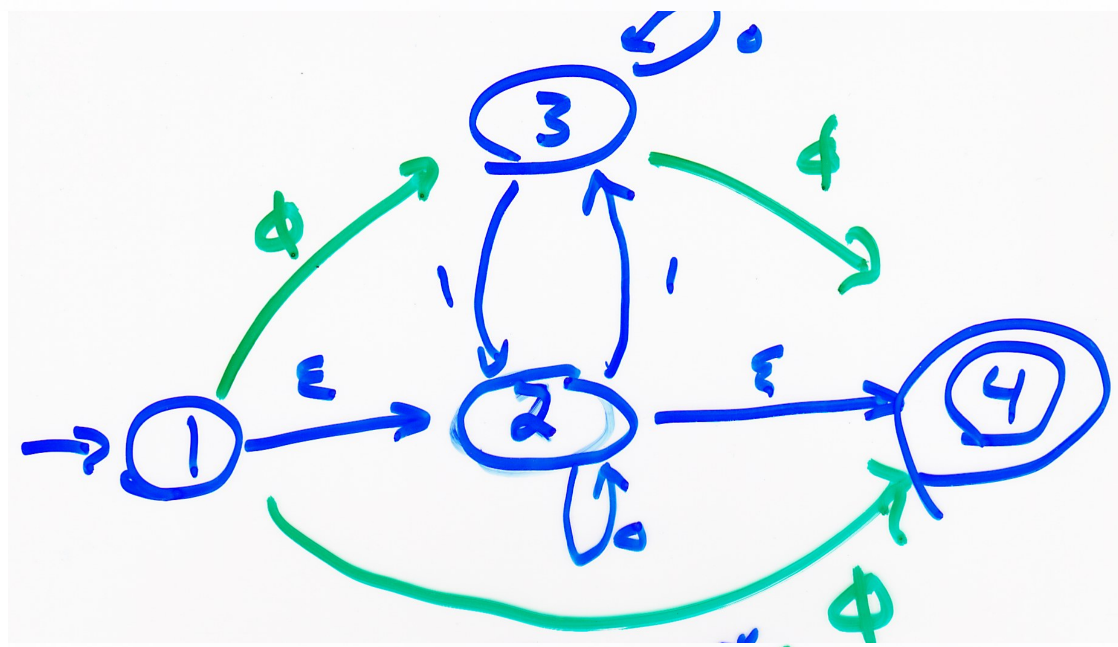
$${}^{1}\varepsilon^{2}\,0^{2}\,\underbrace{1^{3}\,0^{3}\,0^{3}\,1}^{2}\,0^{2}\,\varepsilon^{4}$$

$$1 \quad 2 \quad 2 \qquad\qquad 2 \quad 2 \quad 4$$

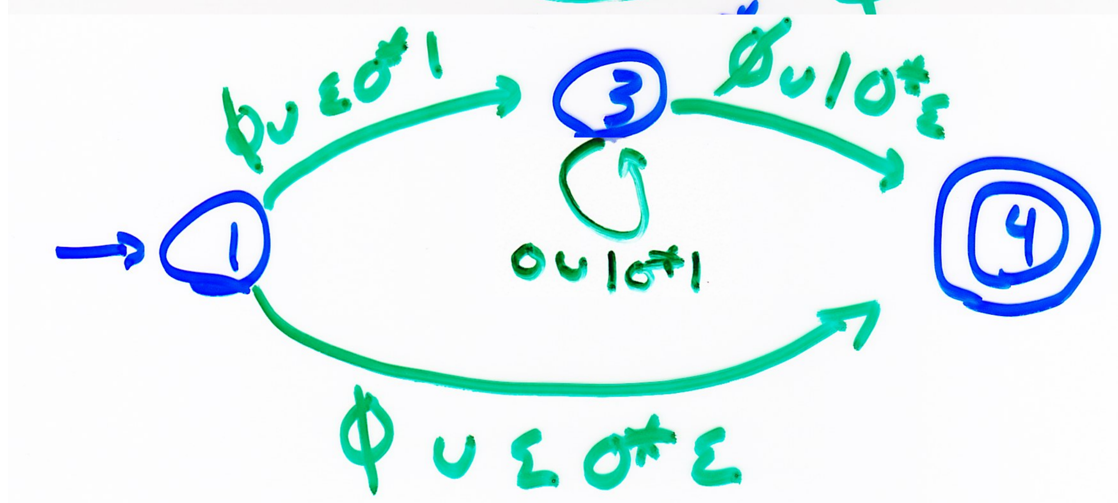




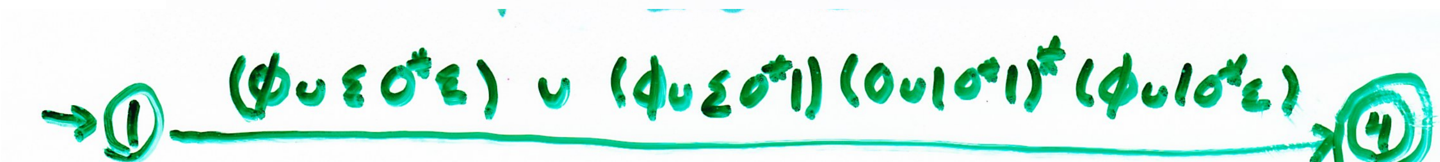
3
0
1
1
ε
ε
1
2
4
0



0
3
ϕ
ϕ
1
1
ε
ε
1
2
4
0
ϕ



ϕ∪ε0*1
3
ϕ∪10*ε
1
4
0∪10*1
ϕ∪ε0*ε



1
(ϕ∪ε0*ε) ∪ (ϕ∪ε0*1)(0∪10*1)*(ϕ∪10*ε)
4