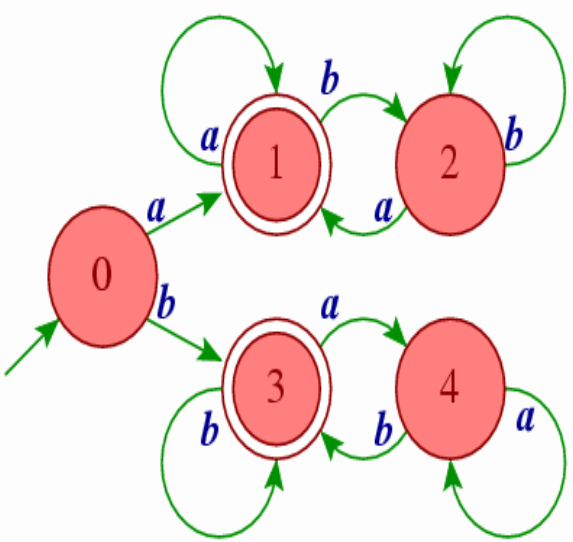
THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

UCS701
Theory of Computation

Chapter 1: Regular Languages & Finite Automata



Dr. Shashank Sheshar Singh

✉ Shashank.sheshar@thapar.edu

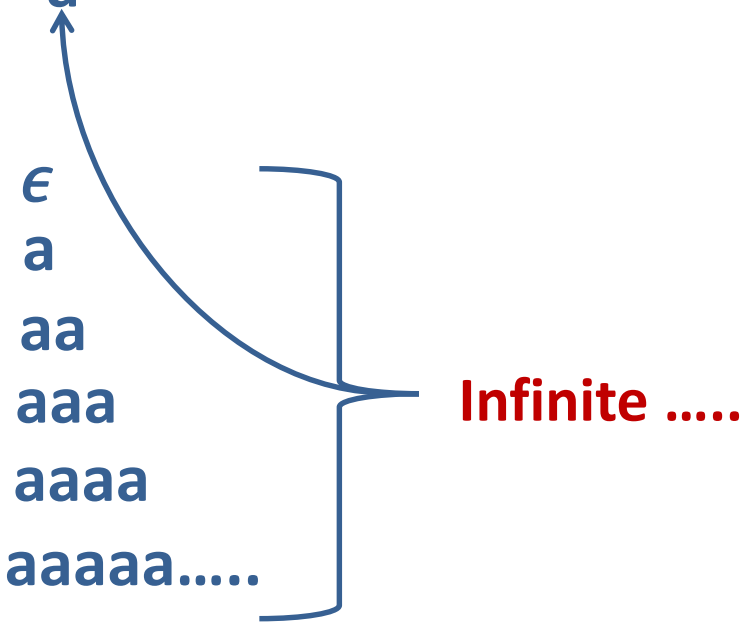
Regular Expression

Regular expression

- A regular expression is a sequence of characters that define a pattern.
- Notational shorthand's
  1. One or more occurrences: $+$
  2. Zero or more occurrences: $*$
  3. Alphabets: Σ

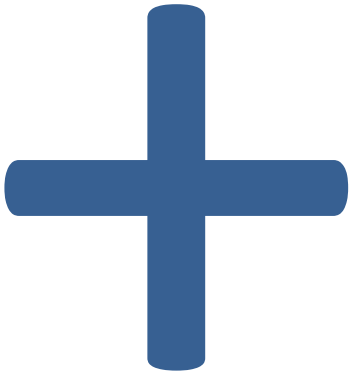
Regular expression

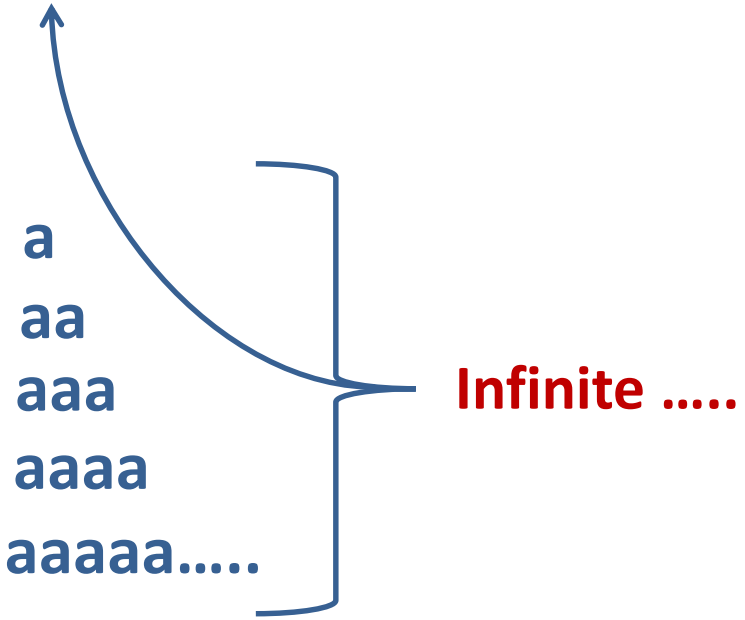
L = Zero or More Occurrences of a = a*



Regular expression

L = One or More Occurrences of a = a^+





Precedence and associativity of operators

Operator	Precedence	Associative
Kleene *	1	left
Concatenation	2	left
Union |	3	left

Regular expression examples

1. 0 or 1

   $\boldsymbol{Strings}: \mathbf{0}, \mathbf{1}$ $\qquad\qquad$ $\boldsymbol{R.E.} = \mathbf{0} \mid \mathbf{1}$

2. 0 or 11 or 111

   $\boldsymbol{Strings}: \mathbf{0}, \mathbf{11}, \mathbf{111}$ $\qquad\qquad$ $\boldsymbol{R.E.} = \mathbf{0} \mid \mathbf{11} \mid \mathbf{111}$

3. String having zero or more *a*.

   $\boldsymbol{Strings}: \epsilon, \boldsymbol{a}, \boldsymbol{aa}, \boldsymbol{aaa}, \boldsymbol{aaaa} \ldots\ldots$ $\qquad\qquad$ $\boldsymbol{R.E.} = \boldsymbol{a}^{*}$

4. String having one or more *a*.

   $\boldsymbol{Strings}: \boldsymbol{a}, \boldsymbol{aa}, \boldsymbol{aaa}, \boldsymbol{aaaa} \ldots\ldots$ $\qquad\qquad$ $\boldsymbol{R.E.} = \boldsymbol{a}^{+}$

5. Regular expression over $\Sigma = \{a, b, c\}$ that represent all string of length 3.

   $\boldsymbol{Strings}: \boldsymbol{abc}, \boldsymbol{bca}, \boldsymbol{bbb}, \boldsymbol{cab}, \boldsymbol{aba} \ldots.$ $\qquad\qquad$ $\boldsymbol{R.E.} = (\boldsymbol{a}|\boldsymbol{b}|\boldsymbol{c})\ (\boldsymbol{a}|\boldsymbol{b}|\boldsymbol{c})\ (\boldsymbol{a}|\boldsymbol{b}|\boldsymbol{c})$

6. All binary string.

   $\boldsymbol{Strings}: \mathbf{0}, \mathbf{11}, \mathbf{101}, \mathbf{10101}, \mathbf{1111} \ldots$ $\qquad\qquad$ $\boldsymbol{R.E.} = (\mathbf{0} \mid \mathbf{1})+$

Regular expression examples

7. 0 or more occurrence of either a or b or both

   $Strings: \epsilon, a, aa, abab, bab \ldots$ $\quad R.E. = (a \mid b)*$

8. 1 or more occurrence of either a or b or both

   $Strings: a, aa, abab, bab, bbbaaa \ldots$ $\quad R.E. = (a \mid b)+$

9. Binary no. ends with 0

   $Strings: 0, 10, 100, 1010, 11110 \ldots$ $\quad R.E. = (0 \mid 1)*0$

10. Binary no. ends with 1

    $Strings: 1, 101, 1001, 10101, \ldots$ $\quad R.E. = (0 \mid 1)*1$

11. Binary no. starts and ends with 1

    $Strings: 11, 101, 1001, 10101, \ldots$ $\quad R.E. = 1\,(0 \mid 1)*1$

12. String starts and ends with same character

    $Strings: 00, 101, aba, baab \ldots$ $\quad R.E. = 1\,(0 \mid 1)*1$ or $0\,(0 \mid 1)*0$
    $a\,(a \mid b)^*\,a$ or $b\,(a \mid b)^*\,b$

Regular expression examples

13. All string of a and b starting with a

    $Strings: a, ab, aab, abb\ldots$ $\quad R.E. = a(a \mid b)*$

14. String of 0 and 1 ends with 00

    $Strings: 00, 100, 000, 1000, 1100\ldots$ $\quad R.E. = (0 \mid 1)*00$

15. String ends with abb

    $Strings: abb, babb, ababb\ldots$ $\quad R.E. = (a \mid b)*abb$

16. String starts with 1 and ends with 0

    $Strings: 10, 100, 110, 1000, 1100\ldots$ $\quad R.E. = 1(0 \mid 1)*0$

17. All binary string with at least 3 characters and 3rd character should be zero $Strings: 000, 100, 1100, 1001\ldots$ $\quad R.E. = (0|1)(0|1)0(0 \mid 1)*$

18. Language which consist of exactly two b's over the set $\Sigma = \{a, b\}$

    $Strings: bb, bab, aabb, abba\ldots$ $\quad R.E. = a^* b\, a^* b\, a^*$

Regular expression examples

19. The language with $\Sigma = \{a, b\}$ such that 3rd character from right end of the string is always

    $\boldsymbol{Strings: aaa, aba, aaba, abb\ldots}$ $\qquad$ $\boldsymbol{R.E. = (a \mid b) * a(a|b)(a|b)}$

20. Any no. of a followed by any no. of b followed by any no. of c

    $\boldsymbol{Strings: \epsilon, abc, aabbcc, aabc, abb\ldots}$ $\quad$ $\boldsymbol{R.E. = a^{*} b^{*} c^{*}}$

21. String should contain at least three 1

    $\boldsymbol{Strings: 111, 01101, 0101110\ldots.}$ $\qquad$ $\boldsymbol{R.E. = (0|1)^{*}1\ (0|1)^{*}1\ (0|1)^{*}1\ (0|1)^{*}}$

22. String should contain exactly two 1

    $\boldsymbol{Strings: 11, 0101, 1100, 010010, 100100\ldots.}$ $\qquad$ $\boldsymbol{R.E. = 0^{*}10^{*}10^{*}}$

23. Length of string should be at least 1 and at most 3

    $\boldsymbol{Strings: 0, 1, 11, 01, 111, 010, 100\ldots.}$ $\quad$ $\boldsymbol{R.E. = (0|1) \mid (0|1)(0|1) \mid (0|1)(0|1)(0|1)}$

24. No. of zero should be multiple of 3

    $\boldsymbol{Strings: 000, 010101, 110100, 000000, 100010010\ldots.}$ $\quad$ $\boldsymbol{R.E. = (1^{*}01^{*}01^{*}01^{*})^{*}}$

Regular expression examples

25. The language with $\Sigma = \{a, b, c\}$ where a should be multiple of 3

    $\boldsymbol{Strings: aaa, baaa, bacaba, aaaaaa..}$ $\boldsymbol{R.E. = ((b|c)^* a (b|c)^* a (b|c)^* a (b|c)^*)^*}$

26. Even no. of 0

    $\boldsymbol{Strings: 00, 0101, 0000, 100100....}$ $\boldsymbol{R.E. = (1^* 0 1^* 0 1^*)^*}$

27. String should have odd length

    $\boldsymbol{Strings: 0, 010, 110, 000, 10010....}$ $\boldsymbol{R.E. = (0|1)\ ((0|1)(0|1))^*}$

28. String should have even length

    $\boldsymbol{Strings: 00, 0101, 0000, 100100....}$ $\boldsymbol{R.E. = ((0|1)(0|1))^*}$

29. String start with 0 and has odd length

    $\boldsymbol{Strings: 0, 010, 010, 000, 00010....}$ $\boldsymbol{R.E. = (0)\ ((0|1)(0|1))^*}$

30. String start with 1 and has even length

    $\boldsymbol{Strings: 10, 1100, 1000, 100100....}$ $\boldsymbol{R.E. = 1(0|1)((0|1)(0|1))^*}$

Regular expression examples

31. All string begins or ends with 00 or 11

    $Strings: 00101, 10100, 110, 01011 \ldots$ $R.E. = (00|11)(0 \mid 1) * \mid (0|1)^{*} (00|11)$

32. Language of all string containing both 11 and 00 as substring

    $Strings: 0011, 1100, 100110, 010011 \ldots$

    $R.E. = ((0|1)^{*}00(0|1)^{*}11(0|1)^{*}) \quad | \quad ((0|1)^{*}11(0|1)^{*}00(0|1)^{*})$

33. String ending with 1 and not contain 00

    $Strings: 011, 1101, 1011 \ldots.$ $R.E. = (1|01)^{+}$

34. Language of C identifier

    $Strings: area, i, redious, grade1 \ldots.$ $R.E. = (\_ + L)(\_ + L + D)^{*}$

    $where\ L\ is\ Letter$ & D is digit

Finite Automata

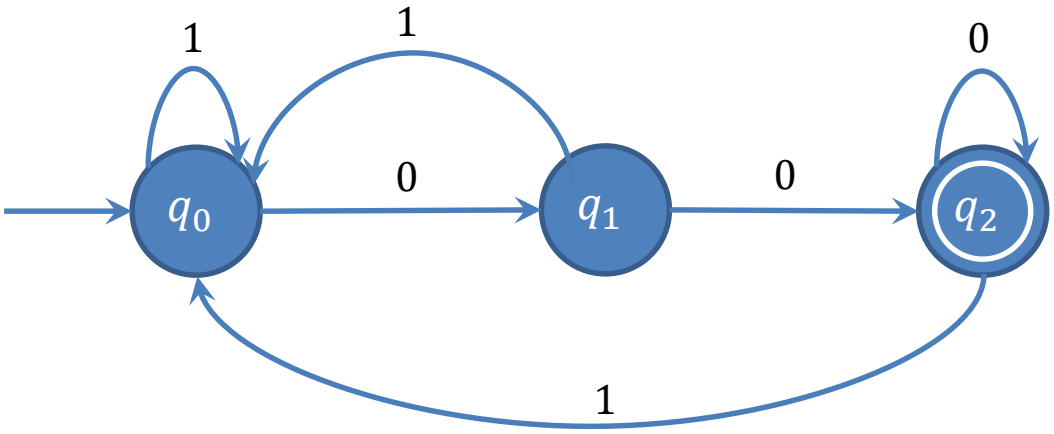
Finite Automata

- A finite automaton, or finite state machine is a 5-tuple $(Q, \Sigma, q_0, A, \delta)$ where
  - Q is finite set of states;
  - Σ is finite alphabet of *input symbols*;
  - $q_0 \epsilon Q$ (*initial* state);
  - $A \subseteq Q$ (the set of *accepting* states);
  - δ is a function from $Q \times \Sigma \; to \; Q$ (the *transition* function).
- For any element q of Q and any symbol $a \epsilon \Sigma$, we interpret $\delta(q, a)$ as the state to which the FA moves, if it is in state q and receives the input a.

Example: Finite Automata

- $M = (Q, \Sigma, q_0, A, \delta)$
  - $Q = \{q_0, q_1, q_2\}$
  - $\Sigma = \{0,1\}$
  - $q_0 = q_0$
  - $A = \{q_2\}$
  - δ is defined as

δ	Input	
State	0	1
q_0	q_1	q_0
q_1	q_2	q_0
q_2	q_2	q_0

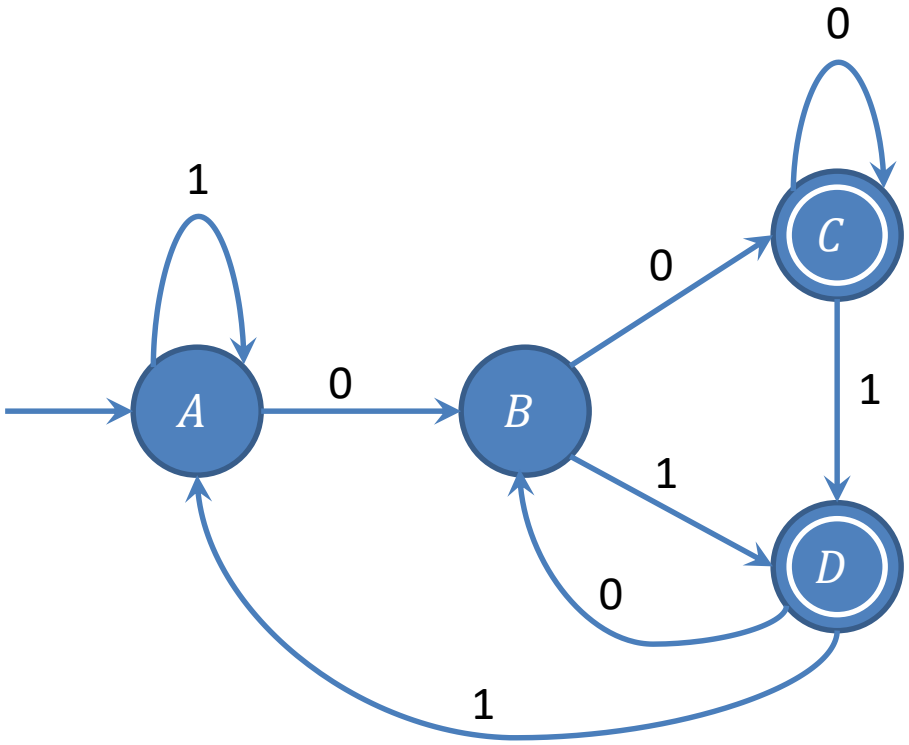


Applications of FA

- Lexical analysis phase of a compiler.
- Design of digital circuit.
- String matching.
- Communication Protocol for information exchange.

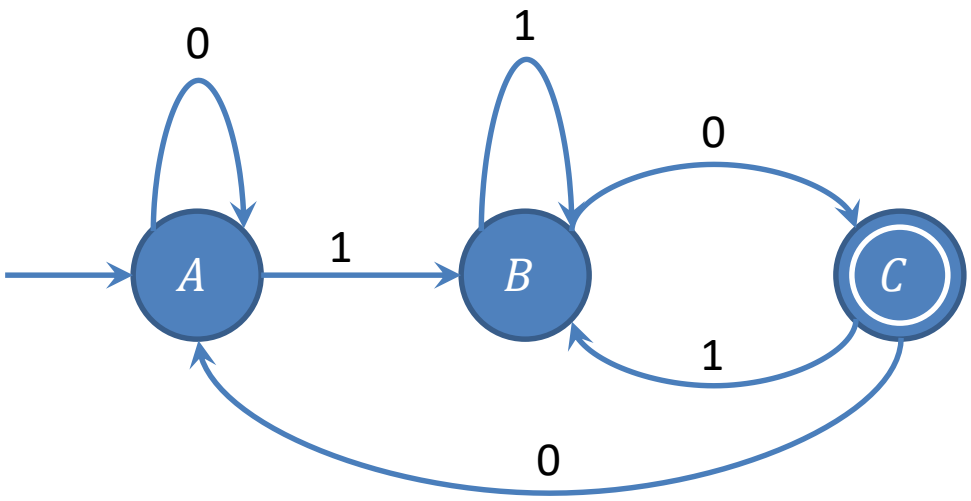
FA Examples

- The string with next to last symbol as 0.



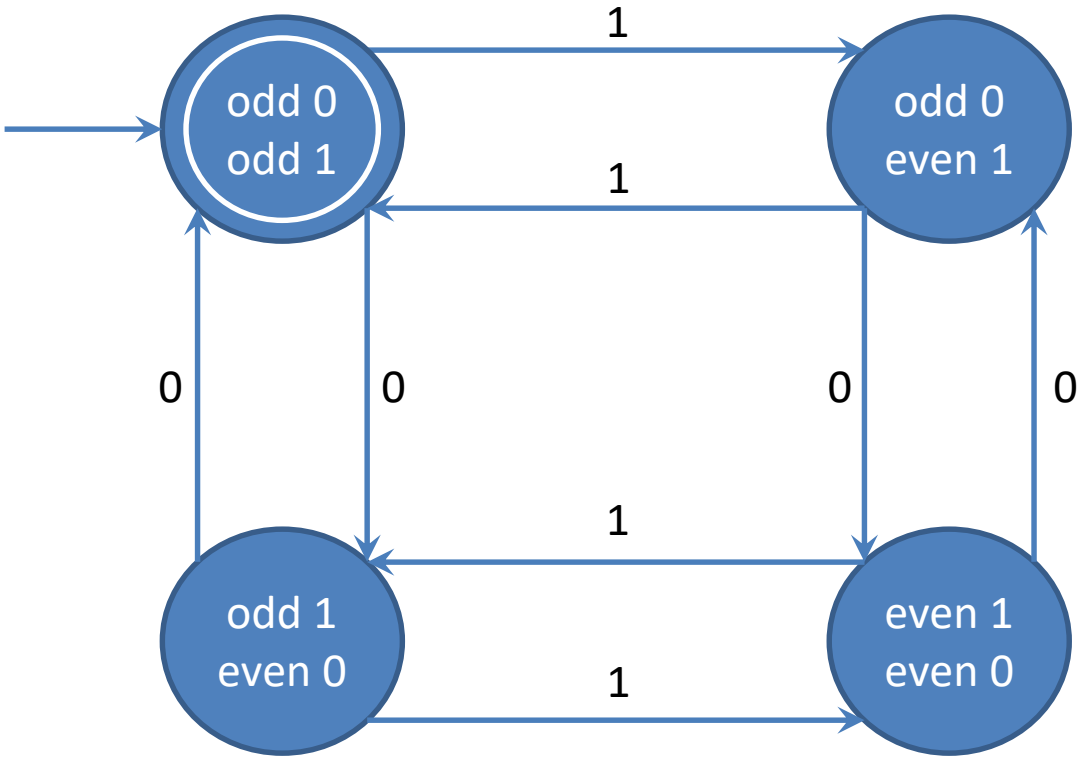
FA Examples

- The strings ending with 10.



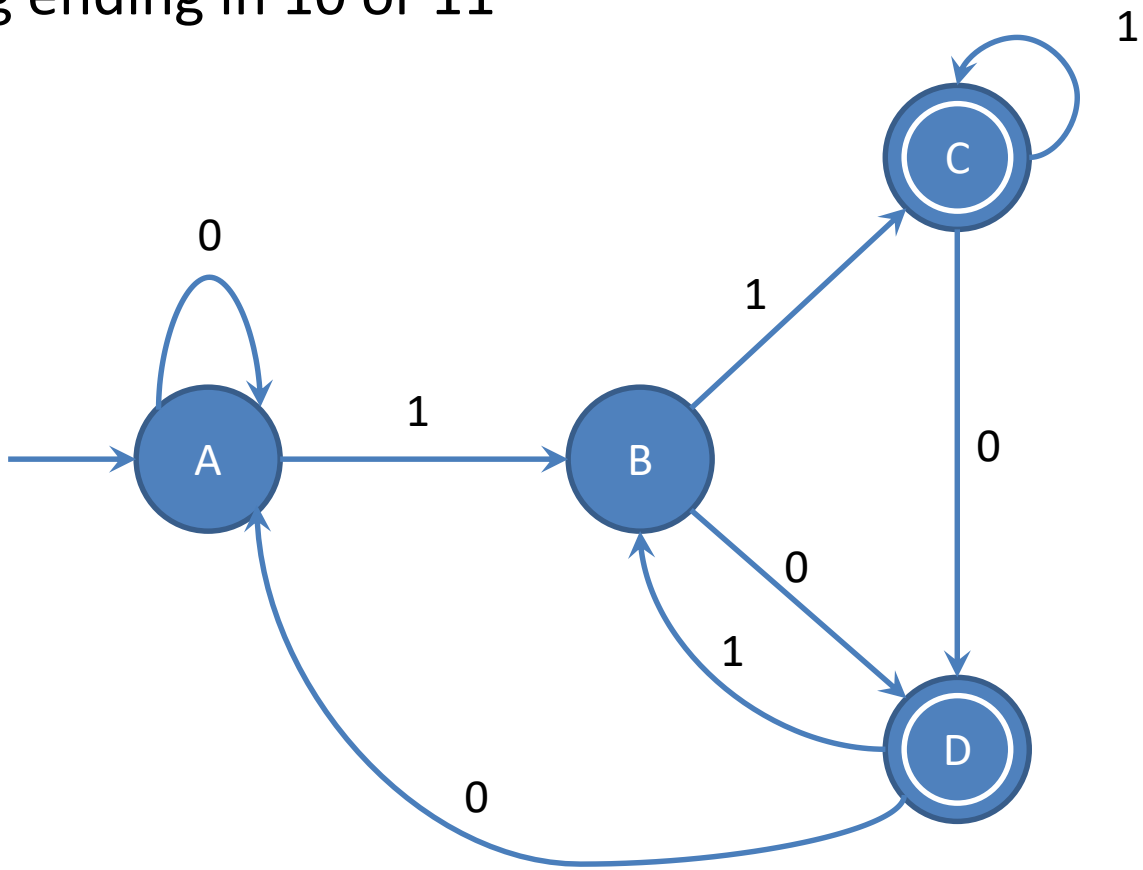
FA Examples

- The string with number of 0's and number of 1's are odd



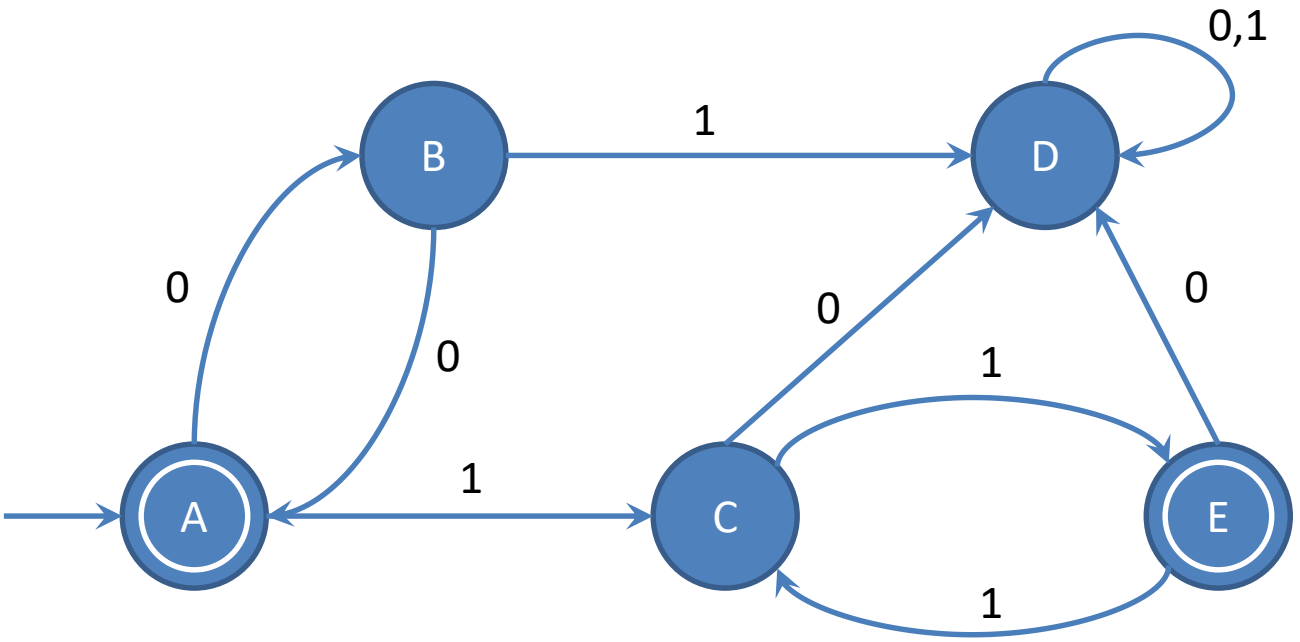
FA Examples

- The string ending in 10 or 11



FA Examples

- The string corresponding to Regular expression {00}*{11}*



FA Examples

- (a+b)*baaa