Regular Expressions

If A and B are languages

- A or B is the union of A and B
- AB is the concatenation of A and B: contains all strings that are concatenation of string in A and string in B
- A^* is all strings that can be built by arbitrarily concatenating strings in A

Example: A = {ant, cow} B = {at, hash}

A or B = {ant, cow, at, hash}

AB = {antat, anthash, cowat, cowhash}

A^* includes ε, ant, cow, cowcowant, ...

A regular expression (RE) generates a language using these 3 operations and recursion.

Example Give RE for all binary strings:

(a) starting with 11

(b) ending with 11

(c) alternates 0s and 1s

(d) has even number of 0's.

(a) $11\underbrace{(0+1)^*}_{\text{any string}}$

(b) $\Sigma^* 11$ (We use Σ for the alphabet)

(c) $(1+\varepsilon)(01)^*(0+\varepsilon)$

(d) $(1^*01^*01^*)^* + 1^*$