CMPT 379
Compilers

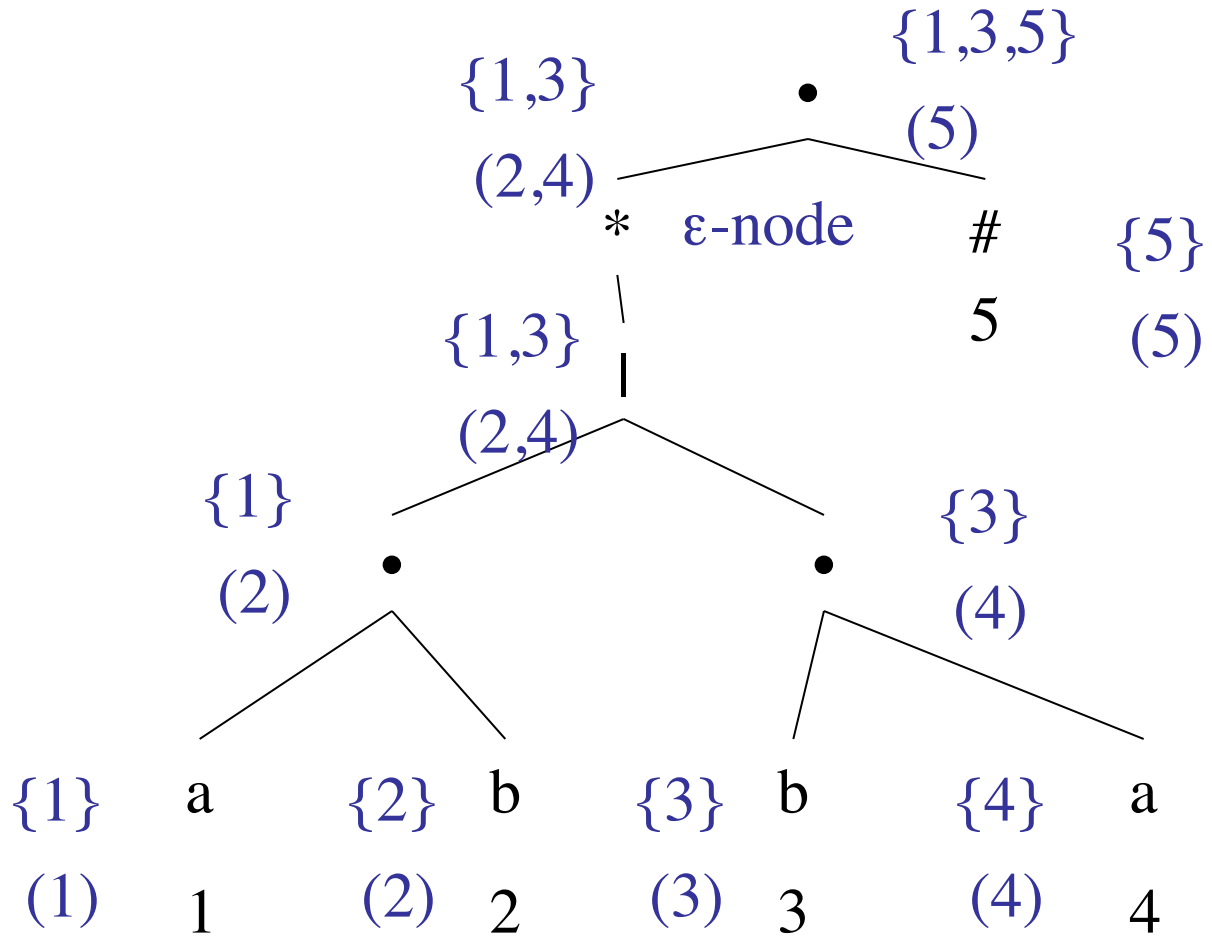
Anoop Sarkar

http://www.cs.sfu.ca/~anoop

Converting Regular Expressions directly into DFAs

This algorithm was first used by Al Aho in `egrep`, and used in `awk`, `lex`, `flex`

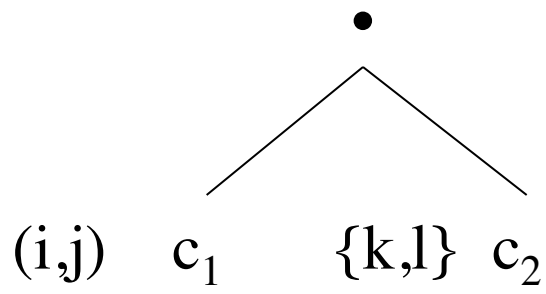
Regexp to DFA: `((ab)|(ba))*#`



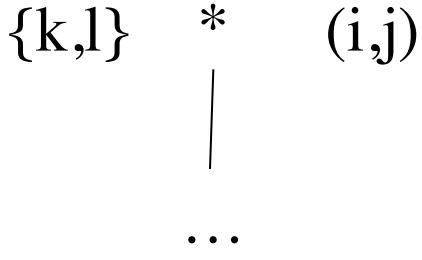
firstpos = {}
lastpos = ()

Regexp to DFA: *followpos*

- *followpos(p)* tells us which positions can follow a position *p*
- There are two rules that use the *firstpos* {} and *lastpos* () information

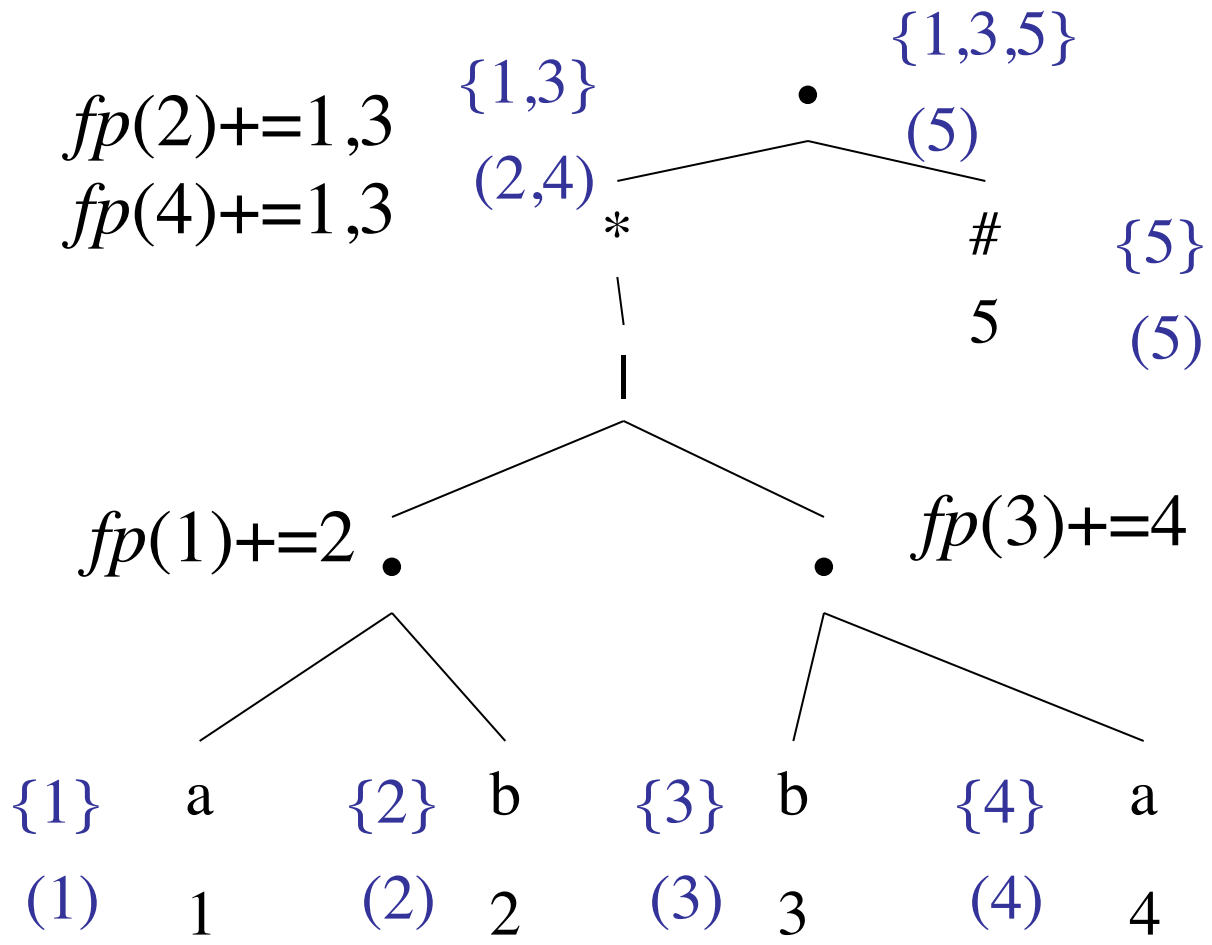


followpos(i)+=k,l
followpos(j)+=k,l



followpos(i)+=k,l
followpos(j)+=k,l

Regexp to DFA: `((ab)|(ba))*#`



root={1,3,5}
fp(1)=2
fp(3)=4
fp(2)=1,3,5
fp(4)=1,3,5

Regexp to DFA: ((ab)|(ba))*#

root={1,3,5}
fp(1)=2
fp(3)=4
fp(2)=1,3,5
fp(4)=1,3,5

{1,3,5} A

A: *fp*(1),a {2},a B,a

A: *fp*(3),b {4},b C,b

A: *fp*(5),# {},# E,#

B: *fp*(2),b {1,3,5},b A,b

C: *fp*(4),a {1,3,5},a A,a

1:a
2:b
3:b
4:a
5:#