

lecture 3 Lexical Analysis

1. whitespace: blank, newline, tab and perhaps other characters that are used to separate tokens in the input.
2. Technically, the techniques used in parsers are still available for lexical analysis.
3. tokens also include comma, semicolon, parentheses, braces.

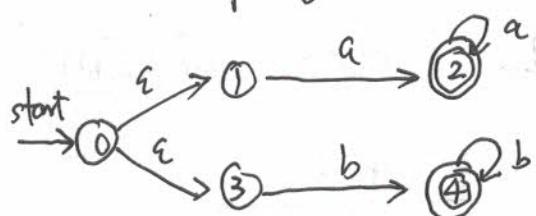
4. token name (abstract) $\xrightarrow{\text{lexeme, lexeme, lexeme}}$

5. lexical errors $f(a == \text{fox})$

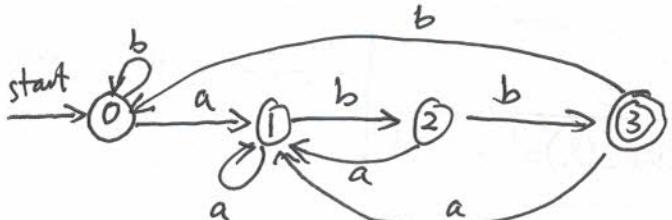
\downarrow if function name? (undeclared?)

6. More examples of NFA and DFA

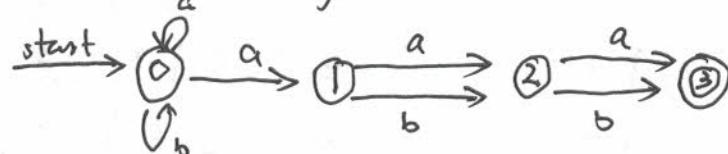
(1) NFA accepting $aa^* \mid bb^*$



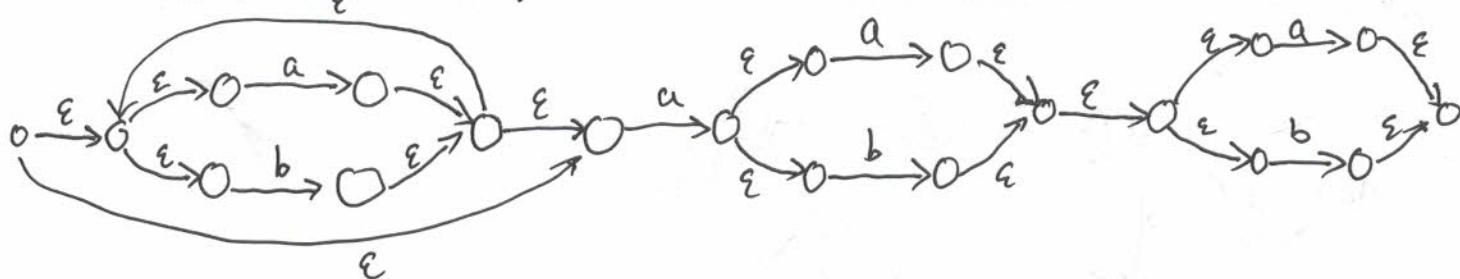
(2) DFA accepting $(ab)^*abb$



7. exg. NFA \rightarrow regular expression

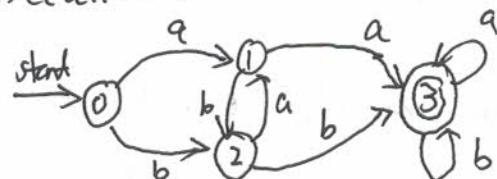


$$(a|b)^* a (a|b) (a|b)$$



8. Recognize strings on $\Sigma = \{a, b\}$ that include two consecutive a or two consecutive b.

DFA:



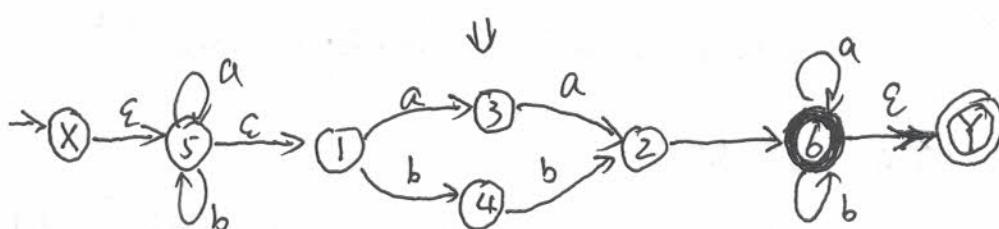
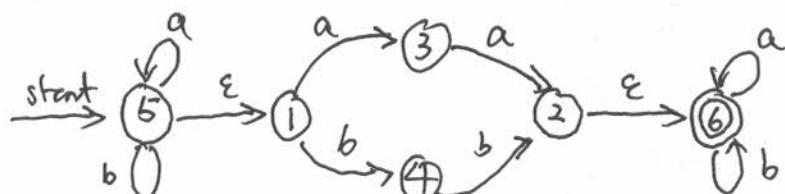
0: initial

1: we have found one a

2: we have found one b

3: we have found "aa" or "bb"

NFA



I	$I_a = \text{\textit{\epsilon}-closure}(\text{move}(I, a))$	$I_b = \text{\textit{\epsilon}-closure}(\text{move}(I, b))$
$\text{\textit{\epsilon}-closure}(X) = \{X, 5, 1\}^0$	$\{5, 3, 1\}^1$	$\{5, 4, 1\}^2$
$\{5, 3, 1\}^1$	$\{5, 3, 1, 2, 6, Y\}^3$	$\{5, 4, 1\}^2$
$\{5, 4, 1\}^2$	$\{5, 3, 1\}^1$	$\{5, 4, 1, 2, 6, Y\}^4$
$\{5, 3, 1, 2, 6, Y\}^{3*}$	$\{5, 3, 1, 2, 6, Y\}^3$	$\{5, 4, 1, 6, Y\}^5$
$\{5, 4, 1, 2, 6, Y\}^{4*}$	$\{5, 3, 1, 6, Y\}^6$	$\{5, 4, 1, 2, 6, Y\}^4$
$\{5, 4, 1, 6, Y\}^{5*}$	$\{5, 3, 1, 6, Y\}^6$	$\{5, 4, 1, 2, 6, Y\}^4$
$\{5, 3, 1, 6, Y\}^{6*}$	$\{5, 3, 1, 2, 6, Y\}^3$	$\{5, 4, 1, 6, Y\}^5$

