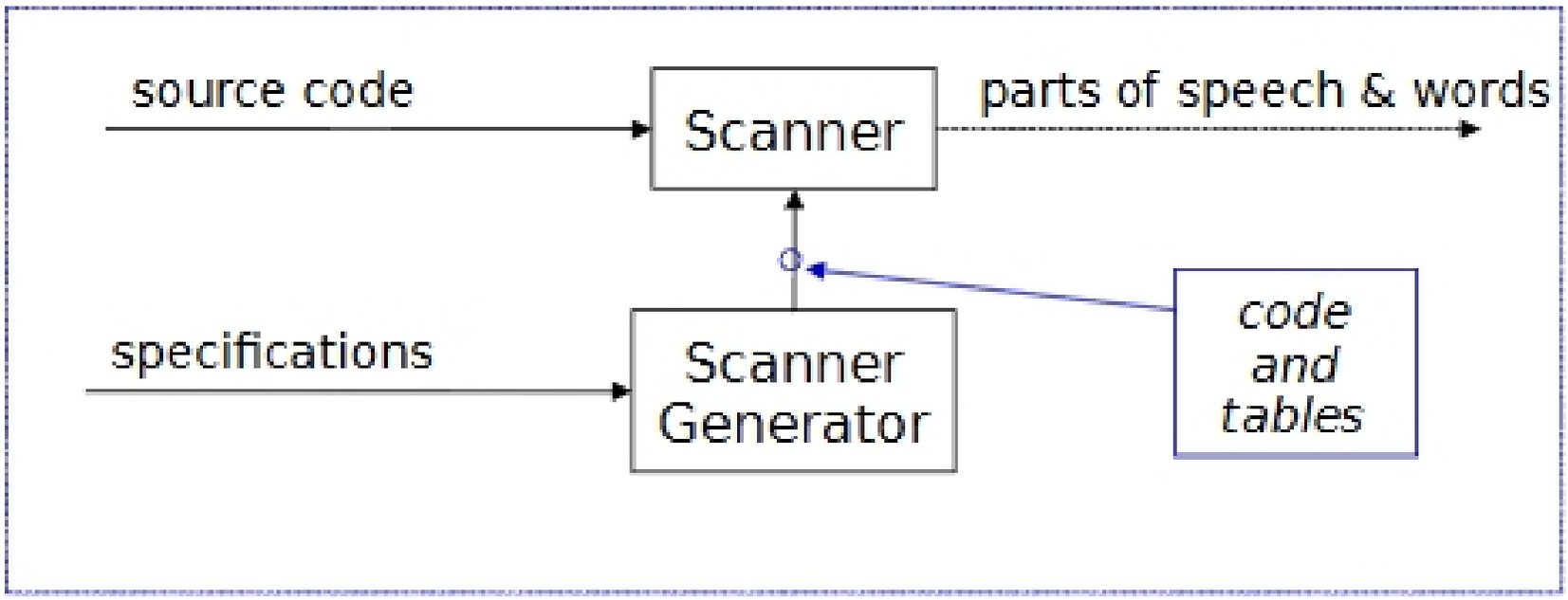
Lexical Analysis — Part II: Constructing a Scanner from Regular Expressions

Quick Review



Previous class:

- → The scanner is the first stage in the front end
- → Specifications can be expressed using regular expressions
- → Build tables and code from a DFA

Goal

- We will show how to construct a finite state automaton to recognize any RE
- This Lecture
  - → Convert RE to an **nondeterministic finite automaton (NFA)**
    - Requires **ε-transitions** to combine regular subexpressions
  - → Convert an NFA to a **deterministic finite automaton (DFA)**
    - Use Subset construction

Next Lecture

- → Minimize the number of states
  - Hopcroft state minimization algorithm
- → Generate the scanner code
  - Additional code can be inserted