

# UCS1602: COMPILER DESIGN

Recognition of tokens



# Session Objectives

---

- To learn concepts recognition of tokens

# Session Outcomes

---

- At the end of this session, participants will be able to
  - Recognize the tokens during lexical analysis

# Outline

---

- Recognition of tokens
- Transition diagrams

---

# Recognition of Tokens

# Recognition of tokens

---

- Starting point is the language grammar to understand the tokens:

```
stmt -> if expr then stmt  
      | if expr then stmt else stmt  
      |  $\epsilon$   
expr -> term relop term  
      | term  
term -> id  
      | number
```

# Recognition of tokens (cont.)

- The next step is to formalize the patterns:

*digit* -> [0-9]

*digits* -> digit+

*number* -> digit(.digits)? (E[+-]? Digit)?

*letter* -> [A-Za-z\_]

*id* -> letter (letter|digit)\*

*if* -> if

*then* -> then

*else* -> else

*Relop* -> < | > | <= | >= | = | <>

- We also need to handle whitespaces:

*ws* -> (blank | tab | newline)+

# Transition diagrams

---

- We represent the patterns in the form of flow charts, called as transition diagrams.
- Nodes: states, conditions that could occur during the process of scanning the input looking for a lexeme that matches one of several patterns
- Edges: directed from state to state
  - Labeled by a symbol or set of symbols
- Deterministic: there's never more than one edge out of a given state with a given symbol among its labels



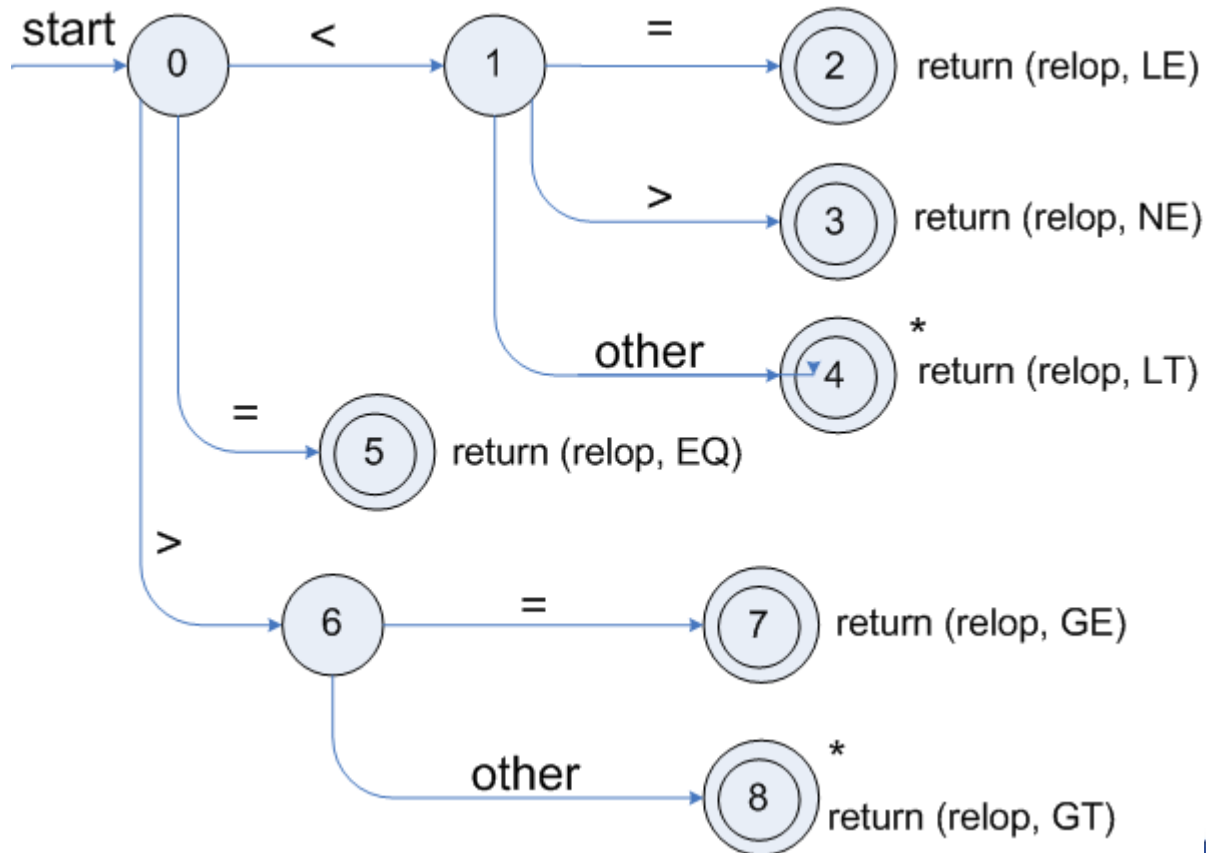
# Transition diagrams

---

- Certain states are accepting or final: a lexeme has been found Double circle
- If it's necessary to retract the forward pointer, we shall additionally place a \* near that accepting state
- Start state, or initial state, is indicated by an edge, labeled “start”, entering from nowhere start

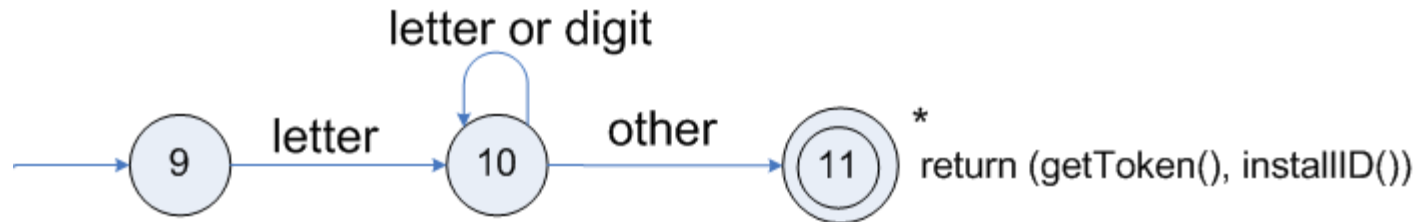
# Transition diagrams

- Transition diagram for relop



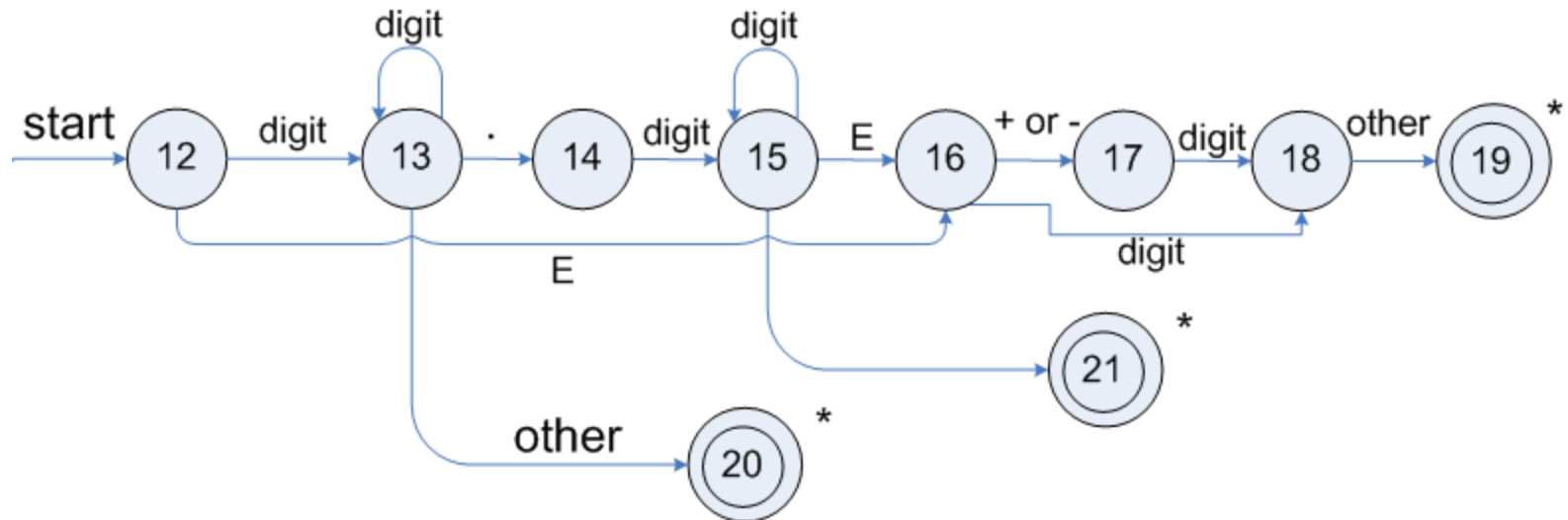
## Transition diagrams (cont...)

- Transition diagram for reserved words and identifiers



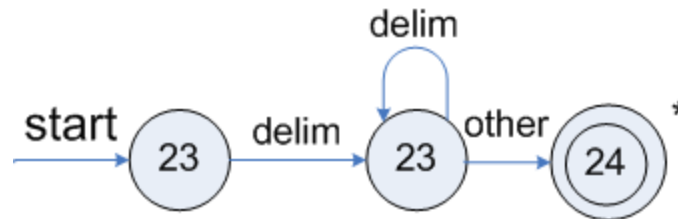
# Transition diagrams (cont.)

- Transition diagram for unsigned numbers



# Transition diagrams (cont.)

- Transition diagram for whitespace



# Summary

---

- Transition diagram
- Transition diagram for relational operator, identifier, reserved word and white spaces

# Check your understanding?

---

1. Draw the transition diagram for the following
  - (a) 3.456
  - (b) 3.45E -5
  - (c) 3456