Arrays, Records and Unions

Alark Joshi
Definitions

• *data type*
  - collection of data objects
  - a set of predefined operations

• *descriptor* : collection of attributes for a variable

• *object* : instance of a user-defined (abstract data) type
Structured Data Types

• Built out of other types
  o usually composed of multiple elements.
  o homogeneous: all elements have the same type
  o heterogeneous: elements have different types
Structured Data Types

• Arrays
  o aggregate of **homogeneous** data elements indexed by its position

• Associative arrays
  o **unordered** collection of key-value pairs

• Records
  o **heterogeneous** aggregate of data elements indexed by element name

• Unions
Array Types

• An array is
  o an aggregate of homogeneous data elements in which
  o an individual element is identified by its position in the aggregate, relative to the first element.
  o Example: int fruits[5];
    • fruits[0] is used to access the first element
    • fruits[2] is used to access the third element
Array Operations

• Whole array operations:
  o Assignment
  o Catenation
  o Slices

• Indexing: mapping from indexes to elements
Array Design Issues

• What types are legal for subscripts?

• Are subscripting expressions in element references range checked?

• When are subscript ranges bound?
Array Design Issues

• When does allocation take place?

• What is the maximum number of subscripts?

• Can array objects be initialized?

• Are any kind of slices supported?
Array Indexing

- **Indexing** (or subscripting) is a mapping from indices to elements
  
  \[
  \text{array\_name (index\_value\_list)} \rightarrow \text{an element}
  \]

- **Index Syntax**
  - FORTRAN, PL/I, Ada use parentheses
    - Ada explicitly uses parentheses to show uniformity between array references and function calls because both are *mappings*
  - Most other languages use brackets
    - One set of brackets for each index
    - One set of brackets for all indexes
Arrays Index (Subscript) Types

• FORTRAN, C: integer only
• Ada: integer or enumeration (includes Boolean and char)
• Java: integer types only
Index range checking

• C, C++, Perl, and Fortran do not specify range checking
• Java, ML, C# specify range checking
• In Ada, the default is to require range checking, but it can be turned off