Notes

CSC222 Geographic Information Systems

04 November 2014

1. equivalence of programs: same inputs → same outputs

2. equivalence of computers: if a program can be run on computer A, an equivalent program can be run on computer B

3. if all computers are equivalent, why bother designing new computers?

    (a) faster

    (b) more compact

    (c) use less power

    (d) more reliable

    (e) easier to use

4. if all programming languages are equivalent, why bother inventing new languages?

    (a) allow programmers to express themselves more concisely

    (b) allow programmers to express themselves more clearly

    (c) give programmers better ways to divide task into parts, to collaborate, to organize parts, to detect/correct errors, to describe their designs to other programmers, to modify and build upon previous work (reuse code)

    (d) make it easier for people to write programs more quickly and with fewer errors

5. largest computer programs are among the most complex objects that human beings have ever designed and built

6. data types

    (a) integer

    (b) floating point

    (c) character/string

    (d) boolean

7. operators

   (a) arithmetic—$+$ $-$ $\times$ $/$
   (b) relational—$<$ $\leq$ $=$ $\neq$ $\geq$ $>$
   (c) logical—**AND**, **OR**, **NOT**, **XOR**

8. shapefiles

   (a) vector data
   (b) attribute values
   (c) no topological information
   (d) not a single file—several files
      i. *.shp—vector data
      ii. *.shx—association of geometric data with attributes
      iii. *.dbf—database table
      iv. *.prj—information about projection

9. object-oriented programming

   (a) class—a blueprint for the construction of objects
   (b) object—a bundle of related data and methods for retrieving, updating, and combining data
   (c) inheritance—a derived class (subclass) has all methods and attributes of a parent class (superclass)
   (d) polymorphism—behavior of object depends upon the class to which it belongs
      i. suppose that we have an Employee class with SalariedEmployee and HourlyEmployee subclasses
      ii. all kinds of employees have a method for computing pay
      iii. pay is computed differently for hourly and salaried employees
      iv. given a collection of employees, system automatically uses correct formula to compute pay of each
   (e) references—one object can point to another

10. database—peristent data

11. relational database

   (a) data stored in tables
      i. each row contains a record—information about a single thing
      ii. each column contains values of a single attribute
      iii. each record includes a unique identifier—a primary key
   (b) database may contain several or many tables

(c) tables can be linked—a record in one table can contain the key of a record in another table

12. design of databases

   (a) common sense, formal methods and notations, plus experience

   (b) identify relationships among "entities"

      i. one-to-one
      ii. one-to-many
      iii. many-to-many

   (c) normalization

      i. eliminate redundancy
      ii. reduce opportunities for introducing inconsistencies
      iii. trade-off with performance

13. SQL—Structured Query Language

   (a) pronounce the letters ("S", "Q", "L") or say "sequel"

   (b) common to many vendors, many database management systems

   (c) many dialects

   (d) a declarative language—describe properties of result, not how to produce result

   (e) commands include: SELECT, INSERT, DELETE, UPDATE

   (f) SELECT *attributes* FROM *table* WHERE *attributes have specified values*

14. XML–eXtensible Markup Language

   (a) tags enclosed in angle brackets
       `<   >`

   (b) values between start tag and end tag
       `<name> Isaac Asimov </name>`

   (c) attributes in tags
       `<a href="http://www.slowgerman.com"/>German lessons</a>`

   (d) document with the structure of a tree—nested tags

   (e) schemas describe permitted order and nesting of tags and attributes allowed within tags

   (f) used to exchange data

   (g) used to store data

   (h) used to describe the configuration of programs

15. JSON—JavaScript Object Notation

(a) "dictionary"

    (b) key–value pairs

    (c) value can be another dictionary (another list of key–value pairs)

16. big data

    (a) volume, variety, velocity

    (b) NoSQL databases

    (c) map-reduce—distribute parts of database among many machines, query in parallel, collect partial results to process and format