Second semester Test
GIS 310 – Advanced GIS
Department of Geography, Geoinformatics and Meteorology

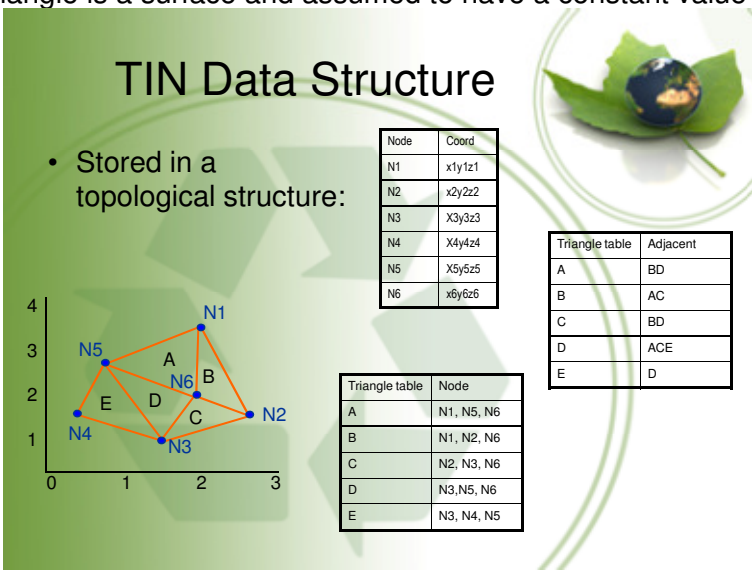
11 May 2011 Time: 50 min

1. STUDY UNIT 3: Spatial Data Models

1.1 Name and discuss the two models used to model surfaces in a GIS.
Surface – entity with a continuous change in value.
Heights, pollution, rainfall
Calculated from sample points taken from the field or terrain
More points – more accurate model
Vector related model used to represent terrain heights:

Triangulated Irregular Network (TIN Model)

Very useful for irregularly spaced points
Data points are connected to form triangles
Triangles does not overlap
Tin surfaces are created by performing delaunay triangulation of the points
Smallest triangle is formed from any three points
Lines from one triangle does not cross the lines from another triangle
Identify the convergent circle for a set of 3 points
Each triangle is a surface and assumed to have a constant value



Node	Coord
N1	x1y1z1
N2	x2y2z2
N3	X3y3z3
N4	X4y4z4
N5	X5y5z5
N6	x6y6z6

Triangle table	Adjacent
A	BD
B	AC
C	BD
D	ACE
E	D

Triangle table	Node
A	N1, N5, N6
B	N1, N2, N6
C	N2, N3, N6
D	N3,N5, N6
E	N3, N4, N5

Advantages of TIN

- They incorporate the original sample points
- Variable densities of triangles means it is an efficient way of storing surface representations
- Easy to calculate elevation, slope , aspect and line of sight

Raster model used to represent surfaces:

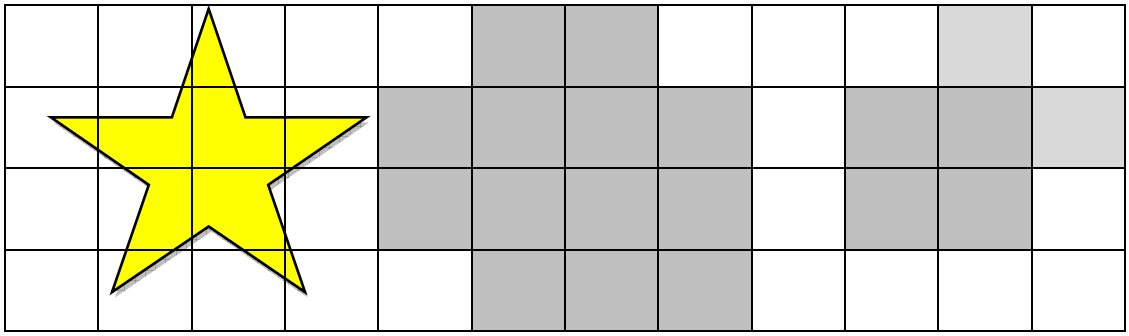
Digital Elevation Model (DEM)

Wide application sphere
Most software packages incorporate it
Also called DTM :digital terrain model
Used to model topographic surfaces
A surface representing height data

Models a continuous surface
Using a finite number of points
Can be calculated from points or lines
Resolution →frequency of observations used
Optimise the resolution data storage ratio
Geographical coordinates must be known:
X,Y & Z (elevation)
In practice Z value can be any observed attribute with characteristics of a continuous surface
E.g. pollution, ranfall, humidity, soil moisture

(10)

1.2 Illustrate how the following feature will be converted to a raster data model when using the 50% rule and the present or absent rule.



(6)

1.3 Explain how vector and raster data models are linked to attribute tables.
Vector
- Table is used to organise the attributes.
- Linkage between rows in the table and the spatial data in topological data layer.
- Most common:
  - One to one relationship – each entry in table is linked to one entity in data layer
- Occasionally:
  - Many to one relationship (Try to avoid)
- Linked through unique identifier in spatial and attribute data

Raster:
- Table the same as in vector data
- Results in large attribute tables
- Many to one relationship often allowed
- Reduces flexibility of the table

(6)

2. STUDY UNIT 4: Geodatabases

2.1 Define the following terminology:

2.1.1 Data - are the raw facts and observations that are stored in a database.

2.1.2 Database - is body of related data stored in a structured manner

2.1.3 Information - is the pay-off that results from successful analysis of data.

2.1.4 Database Management System - a general purpose computer program which makes manipulating a database more convenient

(4)

2.2 Name and discuss five advantages and/or disadvantages of the relational database model.

1. Simple logical model (tables and only tables)
2. - Considerable data independence
3. - High level declarative language (no programming knowledge necessary)
4. - Greater productivity.

Disadvantages

1. **Complex Objects:** The strength of the Relational model, its simple tabular data-model, is also in some circumstances its weakness, since it is difficult to compress some of the complex relationships that exist in the real world into tables. For example, in a GIS we might wish to define objects that are comprised of other objects. For example we might have an object 'Property' which is composed of objects called 'building', 'land', and 'fence'. Modelling such complex, nested entities in a Relational model is difficult.
2. **Lack of Semantic Knowledge:** 'Semantic knowledge' is knowledge about what data MEANS, i.e. how the data can be interpreted, and what are the legitimate processes for which the data can be used?
3. **Limited Data Types:** Related to the two previous limitations, the inability of a pure RDBMS
4. system to recognise anything other than simple atomic data types such as integers, characters, etc. is a serious limitation. To give trivial examples, it would be nice to be able to declare a type 'Voter' which recognised that only persons over 17 years old can vote, or to declare a type 'Day' which would accept only values 'Monday' to 'Sunday'.
5. **Difficulties with Time:** It is difficult with relational databases to satisfactorily model the temporal aspects of data. Time has a natural sequence, events happen 'before' or 'after' each other, but modelling such sequences in a relational database poses difficulties.
6. **Lack of Efficiency:** Strictly this is an implementation issue, rather than a criticism of the logical adequacy of the theoretical model, but it is convenient to mention it here. Relational databases have been criticised for their lack of speed compared to other DBMS systems, and for computationally intensive applications, such as GIS this may present a problem for large applications. (But see Section 6 where recent developments in processing power are discussed).

(10)

6.1 Name the steps that you will follow when designing and EAR diagram

1. Identify entities
2. Identify relationships between entities
3. Identify attributes of entities
4. Derive tables.

(4)

6.2 Explain what SQL is and also name the advantages of SQL that has led to its present popularity.

Structured Query language – Practical language present users with suitable means of managing databases.

1. **Completeness:** SQL contains statements that implement most of the structural, manipulative, and integrity requirements of the Relational model. It thus provides a comprehensive database language, which allows databases to be created, used, and deleted.
2. **Simplicity:** Despite its comprehensive nature, there are less than thirty SQL commands, so that
3. learning the language is not an overwhelming task. New users can rapidly learn sufficient SQL to
4. be able to carry out most operations. On the other hand the flexibility of SQL allows experienced
5. users to carry out complex operations.

6. **Declarative:** SQL is a declarative ***language*** in the sense that SQL statements tell the computer what is required but leave the computer to determine how to achieve the required results. This is in contrast to the lower level approaches of ***third generation languages***, such as COBOL, which require the user to be able to write a program to tell the computer how to retrieve information. Whereas third generation languages tend to operate on one record at a time, SQL statements, like relational algebra, operate on sets of records at a time. The inputs to SQL statements are sets of records, and logical conditions that the rows must satisfy, and the outputs are new sets of rows.
7. **Pseudo English command structure:** An early prototype for SQL was the 'Structured English Query Language' (SEQUEL), which IBM introduced in 1974, and SQL has retained an English-like syntax. All SQL statements can be thought of as a very restricted and structured English sentence. SQL commands always begin with a verb which specifies the action which the software is to take
8. **Widely Used:** SQL has become a very widely used database language. Mainframe RDBMS have used SQL for several years, and increasingly SQL is now incorporated into most PC DBMS packages. Windows based RDBMS, such as MS ACCESS, which present a friendly graphical interface to the user commonly rely on SQL 'under the skin'.

(10)

Total [50]