Computerized Information System For Poverty Monitoring

USER'S MANUAL

November 1998

CIRDAP

Centre on Integrated Rural Development for Asia and the Pacific

Computerized Information System
For Poverty Monitoring

USER'S MANUAL

November 1998

CIRDAP

Centre on Integrated Rural Development
for Asia and the Pacific (CIRDAP)

Bangladesh Space Research and Remote
Sensing Organization (SPARRSO)

This work was carried out with the aid of a grant from the International Development Research Centre (IDRC), Ottawa, Canada which included a contribution from the Canadian International Development Agency (CIDA).

Computerized Information System for Poverty Monitoring: User's Manual

MAP Technical Paper Series No. 8

November 1998
CIRDAP

The manual has been prepared by SPARRSO, under the 'Monitoring Adjustment and Poverty (MAP) in Bangladesh' Project. The MAP project is being implemented by CIRDAP to institutionalize mechanism for regular monitoring of poverty and impact of structural adjustment (SA) policies by relevant national institutions of Bangladesh. The CIS and PMS components of the project are being conducted by Bangladesh Space Research and Remote Sensing Organization (SPARRSO) and Bangladesh Bureau of Statistics (BBS), respectively.

Jointly Published by:

Centre on Integrated Rural Development for Asia and the Pacific (CIRDAP)
Chameli House, 17 Topkhana Road
GPO Box 2883, Dhaka 1000
Bangladesh
Tel: 9568379, PABX 9558751,9559686
Fax: 880-2-9562035
E-mail: rescir@citechco.net

Bangladesh Space Research and Remote Sensing Organization (SPARRSO)
Mahakash Biggyan Bhaban
Agargaon, Sher-E-Bangla Nagar
Dhaka 1207, Bangladesh
Tel : 323994, 9113329, 815863
Fax: 880-2-813080
E-mail: sparrso@bangla.net

Contents

	Page
CHAPTER 1 GEOGRAPHIC INFORMATION SYSTEM (GIS)	1
1.1 WHAT IS GEOGRAPHIC INFORMATION?	1
1.2 WHAT IS GIS	3
1.3 THE COMPONENTS OF GIS	4
1.4 TOPOLOGY IN GIS	15
CHAPTER 2 GETTING SPATIAL DATA INTO ARC/INFO	26
2.1 DATA CAPTURE THROUGH DIGITIZING	26
2.2 SPATIAL DATABASE STRUCTURE	27
CHAPTER 3 MAKING SPATIAL DATA USABLE	34
3.1 CONSTRUCTION OF TOPOLOGY	34
3.2 IDENTIFYING DIGITIZING ERRORS	35
3.3 FIXING OF ERRORS	36
3.4 RECONSTRUCTION OF TOPOLOGY	39
CHAPTER 4 GETTING ATTRIBUTE DATA INTO ARC/INFO	41
4.1 PROCEDURE OF ADDING ADDITIONAL DATA TO THE FEATURE ATTRIBUTE TABLES	41
CHAPTER 5 MANAGING THE DATABASE	45
5.1 CREATION OF A TIC COVERAGE CONTAINING REAL-WORLD COORDINATES	45
5.2 COORDINATE TRANSFORMATION FROM DIGITIZER TO REAL-WORLD	46
5.3 JOINING OF TWO COVERAGES	48
CHAPTER 6 GEOGRAPHIC ANALYSIS AND QUERY	49
6.1 TABULER ANALYSIS	52
CHAPTER 7 DISPLAYING DATA	54
7.1 WHAT IS A MAP?	54
7.2 DESIGNING AND CREATING A MAP	55
7.3 MANIPULATING MAP COMPOSITION ELEMENTS	61
CHAPTER 8 INTEGRATION OF DATA FROM GENERAL DATABASE MANAGEMENT SYSTEM (DBMS) WITH GIS	62
8.1 GIS DATABASE INTERFACE WITH OTHER DATABASE	62
8.2 PREPARATION OF THE DATABASES FOR INTEGRATION WITH GIS	63
8.3 STEPS IN DATA INTEGRATION	64

Chapter 1

Geographic Information System (GIS)

1.1 What is Geographic Information?

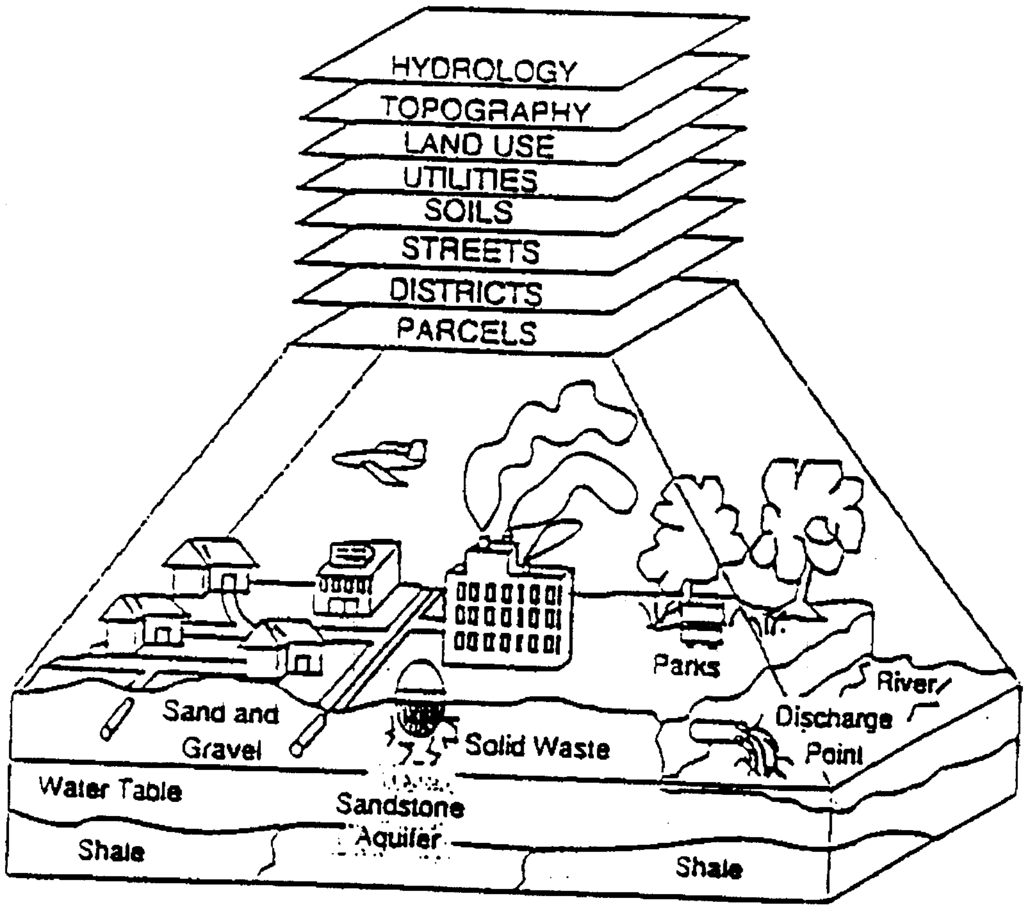
- Any information at any geographic location or region (may be the whole globe) of the earth is known as geographic information.

- The geographic information may be spatial, graphical and textual / tabular.

- The spatial information of the earth or of a region on the earth is represented on maps. The maps are the abstract representation of the reality of the earth on plain paper containing one or more thematic information. In a map, the spatial extent of the real world is scaled down many times for accommodating the real world distance on papers of standard sizes. This reduction is represented by the term SCALE of the map, which is represented by the expression 1: x where x is the number of times the real world has been reduced during representation on the map. The scale of the map is also represented by a horizontal scale bar representing the real world distance.

  The map is characterized by a number of map elements

  - Title
  - Border
  - Spatial Coordinates (geographic/projective)
  - Features: Points, Lines, Polygons, Texts
  - Symbols
  - Legends describing the features
  - Map scale
  - North Arrow
  - Projection Type
  - Other details of the source of information, time reference and ownership/ authorship.



The real world consists of many geographic features which can be represented a: number of resulted map / data layers.

Why geographic information is important?

The geographic information provides with the knowledge about the environme: disasters, resources, topography, infrastructure, demography, health, education a other socioeconomic variables of the planet where we live and the spatial a temporal variation of these features at different scales. The informations a presented in the form of maps and tables. These informations are extreme important for our daily life, resource management, disaster management a mitigation and short and long term economic development and planning.

1.2 What is GIS ?

The hard copy maps and tabular information on a number of themes of a geographic area may also constitute very important analog tool constituting a geographic information system for a required application. Such an information system of analog nature has been used for a long time until the modern digital technique for handling spatial / tabular data has been evolved in the late 1970s and early 1980s. At present, by 'Geographic Information System (GIS)' we understand a computerized tool for handing the spatial data integrated with tabular information in a widely flexible manner which can be used for a large number of applications. A more complete description of the GIS may be given as:

An organized collection of computer hardware, software, geographic data and personnel designed to efficiently capture and store, update, manipulate, analyze, and display all forms of geographically referenced information.

The GIS may be described in two ways:

⇨ Through formal definitions
⇨ Through the ability to carry out spatial operations and link the descriptive data sets with the spatial data using location ID as the common key, answer various queries and perform arithmetical, logical, modeling and other operations. Keeping away from any particular applications, it can be seen that there are five generic questions that a sophisticated GIS can answer:

- Location : What exists at a particular location (cropland, pond or road)

- Condition: In stead of finding what exists at certain location, we would like to find the location which satisfies certain conditions. Say, we would like to find a place within 100 meters from the road and with soils capable of supporting construction of buildings and is not inundated during the high floods and does not have forests. Such a site is suitable for planning a residential area without destroying the trees which are highly important elements of nature.

- Trends: The time differences of a parameter within an area: change of forest cover, water shed areas, etc.

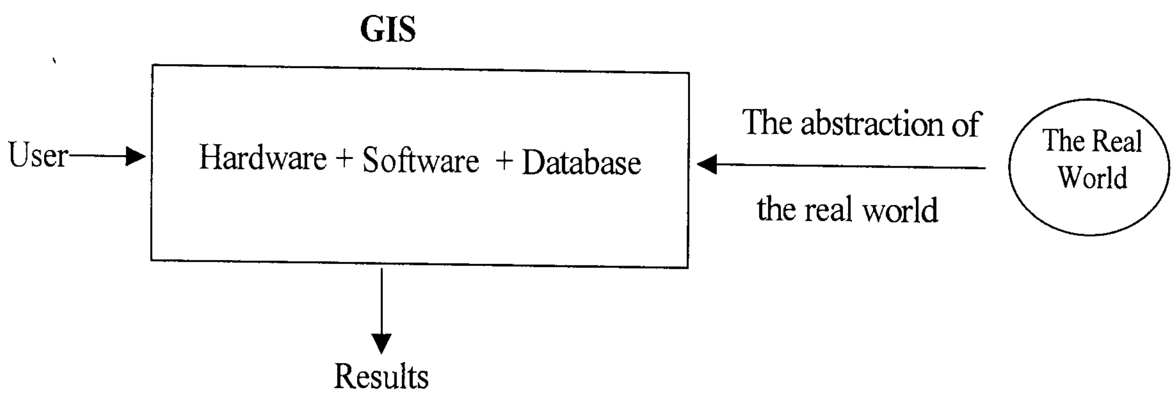
- Patterns: What spatial patterns exist? We may observe that, areas that are affected by frequent flooding have severe poverty situations. Also we may locate areas that do not follow such a pattern. Or we might ask the question whether cancer is the major cause of death among residents near the vicinity of a nuclear power station. Also, we may find out the exceptions which do not fit the pattern.

- Modeling: What will happen if the toxic waste is slipped into the local ground water? What will happen to the surface water condition/drainage capacity of the ecology if a road is constructed across a region.

The GIS is capable of performing more than the above. Measurement, mapping, monitoring, manipulations of environmental and other geo-referenced features / data and processes can be enhanced using the GIS. The use of GIS helps the planning and implementation process of economic development activities.

1.3 The Components of GIS

The components that constitute a GIS are the following:



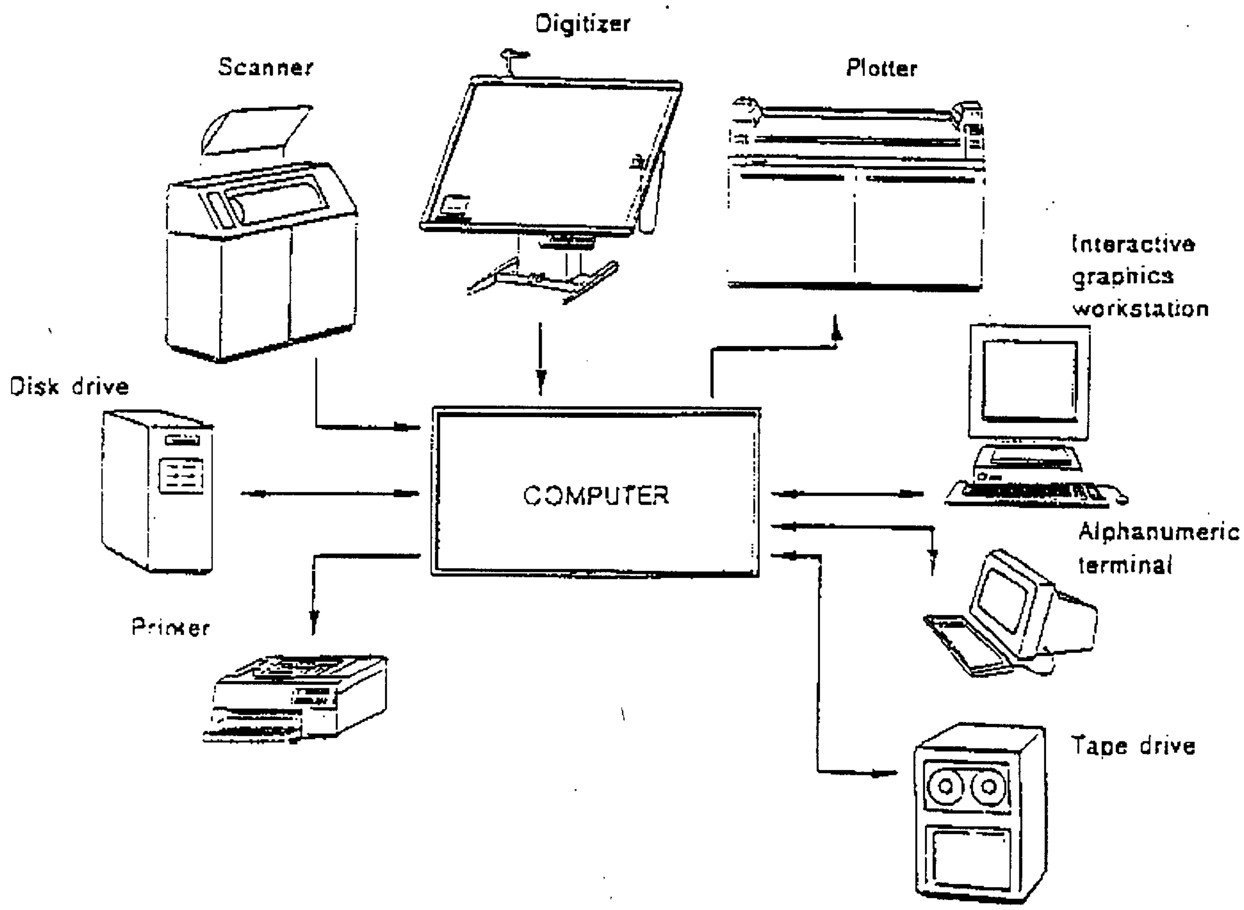
GIS: Main Hardware and Software:

Hardware

- Computer: CPU and Main memory
- Storage device : Disk Drive, Tape drive, CD drive/write
- Input Device: Scanner, Digitizer, Keyboard, Floppy drive, CD drive
- Output Device: Printer, Terminal, Plotter

Software

UNIX, VMS, Pc DOS/Windows versions of GIS ARC/Info (several modules), Arc/View, Map/Info, IDRISI, ILWIS, AutoCAD, etc.



All GIS has two components: **Spatial** component and **DBMS** component. The first one handles spatial data management and processing and the DBMS handles the tabular data processing while the tabular data may be linked with the spatial data through common key parameters.

User

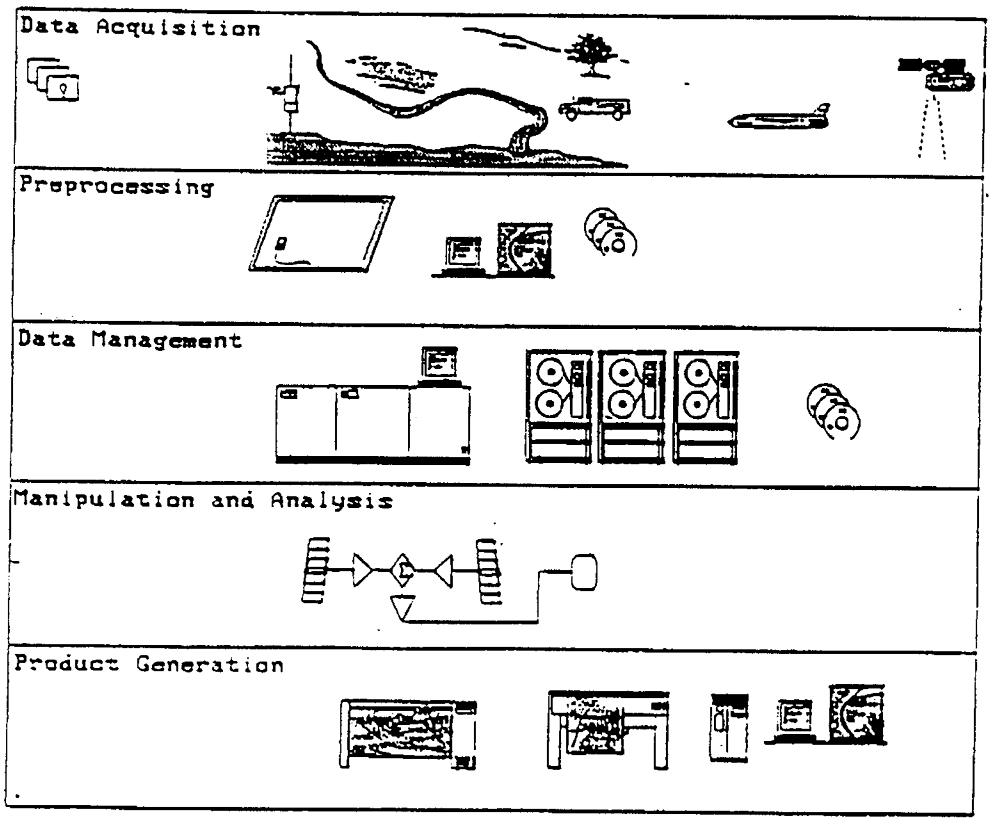
GIS is not only a map making tool. It is an analytic tool and the 'user' is an important component of the system, who uses the tool to perform the complicated analysis needed to solve a problem.

The efficient use of GIS depends on the user, how he focuses on the problem and plans the solution mechanism and his skill of selecting and using tools from GIS tool box required for the solution.

GIS Functional Elements

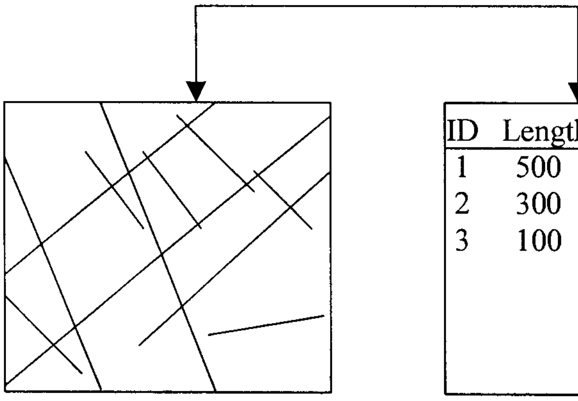
- Data acquisition
- Preprocessing
- Data Management (getting data in, editing and management)
- Manipulation and Analysis
- Product Generation

GIS Project Sequence



Characteristics of the GIS Database (GIS Data Model)

- The GIS has two basic types of geographic information: spatial and descriptive.

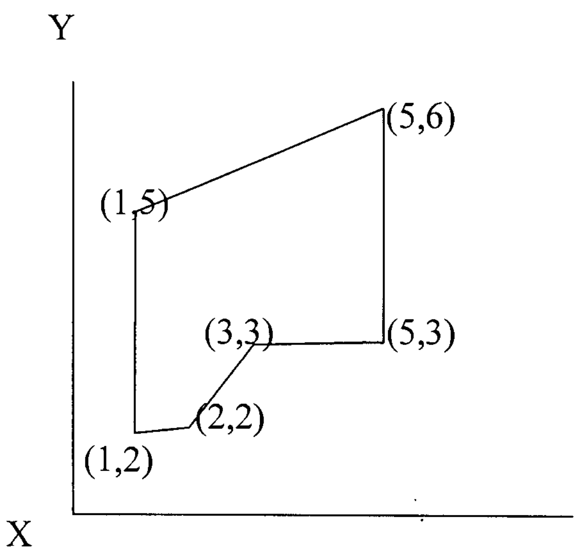


LINE (ARC)

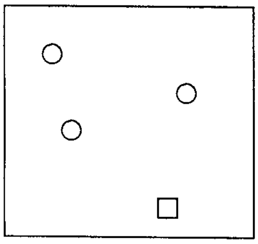
ID	Length	Class	Surface
1	500	1	paved
2	300	2	gravel
3	100	1	paved

FEATURE DESCRIPTION (INFO)

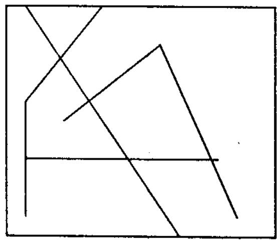
The GIS has specific data model for representing maps in the computer. It does not store the map, but the x, y position of the data points.



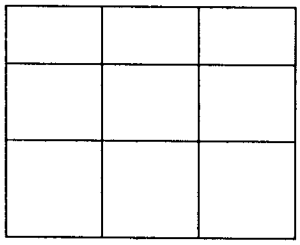
- Major geographic feature types: points, lines and polygons.



Points



Lines

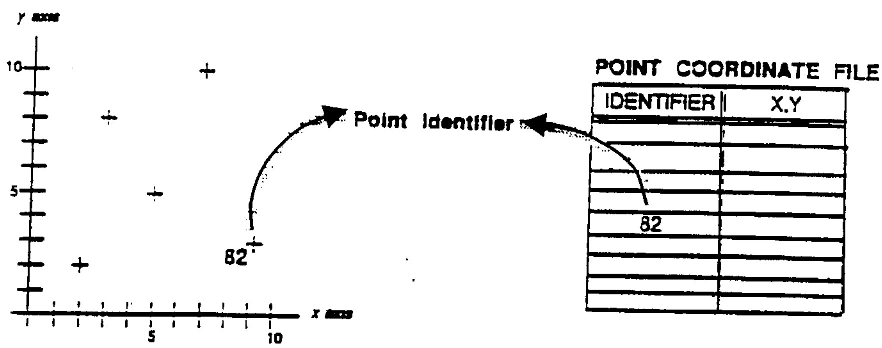


Polygons

❖ **Point Features**

- A single pair of X, Y coordinate represents point.
- The point has an identification number.
- There is no length and area of a point.

Feature Type: Points



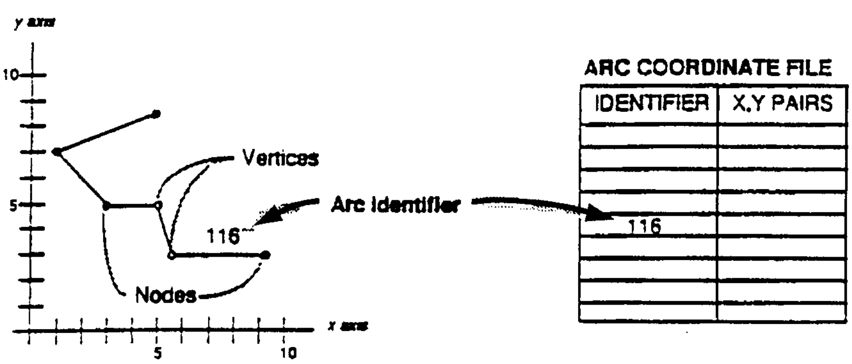
■ Spatial locations

■ Coordinate data

❖ Line Features

- The linear features are called the arcs.
- A string of X, Y coordinates, each point is known as vertex.
- Length is defined by the coordinate.
- Each arc has an **ID**.
- An arc may have maximum 500 vertices.

Feature Type: Arcs



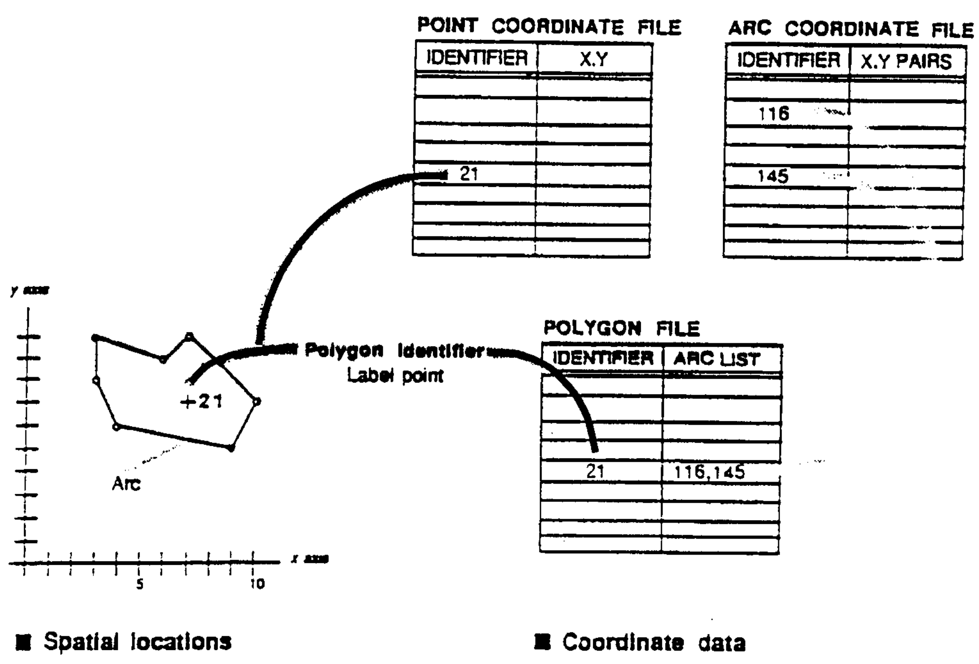
■ Spatial locations

■ Coordinate data

❖ Polygons Features

- Geographic feature having a homogeneous area enclosed by a boundary line forms a polygon.
- A polygon is composed of arcs and label points.
- One or more arcs outline an arc.
- A simple label point is located inside the area and serves to identify each polygon. The label points have an ID.

Feature Type: Polygons

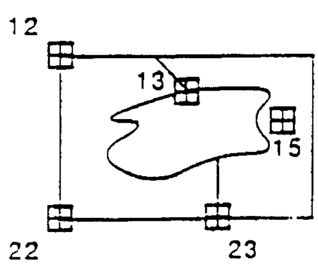


Tics, Boundaries, Annotations, Nodes, Links

◇ <u>Tics</u>: Geographical control points for real-world locations in a coverage are called tics.

- All coverage must have a set of tics.
- Tics represent a single pair of x, y coordinate.
- Identify a tic with a symbol **IDTIC**.
- X- coordinate of tic location measured in coverage units by **XTIC**.
- Y- coordinate of tic location measured in coverage unit by **YTIC**.
- All coverages have minimum four tics.
- Tic location must be known in real-world coordinates.
- Each tic has a unique value.

Tics



TIC

IDTIC	XTIC	YTIC
12	3.920.000	7.350.000
13	3.970.000	7.340.000
22	3.920.000	7.300.000
23	4.000.000	7.300.000
15	4.100.050	7.335.000

Tic identifier

■ Spatial data

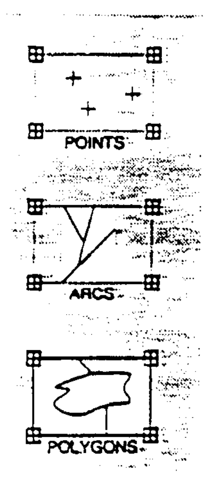
■ Tic table (TIC)

Additional features of a coverage:

◇ **Boundaries (BND):** The extent of all geographic data within a coverage is called its boundary (BND).

- An area defined by two pairs of x, y coordinates.
- Minimum x-coordinate and y- coordinate data within the coverage represent **XMIN** and **YMIN**.
- Maximum x-coordinate and y-coordinate data within the coverage represent **XMAX** and **YMAX**.

Boundary (BND)



■ Extent of spatial data

BND

XMIN	YMIN	XMAX	YMAX
3,950.000	7,320.000	3,990.000	7,340.000

BND

XMIN	YMIN	XMAX	YMAX
3,940.000	7,300.000	3,980.000	7,350.000

BND

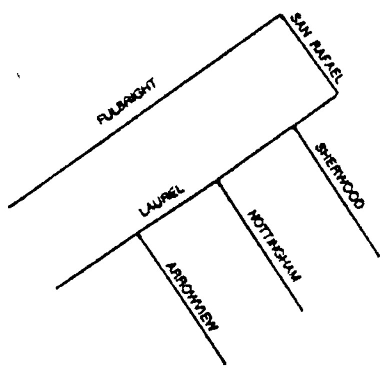
XMIN	YMIN	XMAX	YMAX
3,920.000	7,300.000	4,000.000	7,350.000

■ Boundary tables (BND)

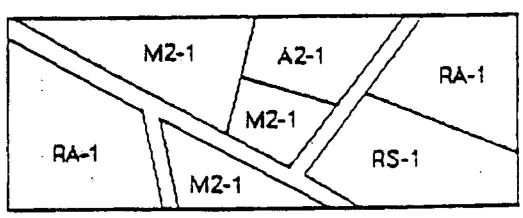
◇ **Annotation**: Textural information stored with a coverage that enhances the cartographic output is called annotation.

- Located in a single x, y coordinate or several x, y coordinates.
- Can be splinted along an arc
- Various sizes, colors or fonts
- Multiple levels of annotation can be created and stored for different applications
- It can be selected, copied, rotated, moved and deleted.
- Stored in coverage units
- It has no connection to points, lines or polygons or their attributes.

Annotation



■ Annotated arcs



■ Annotated polygons

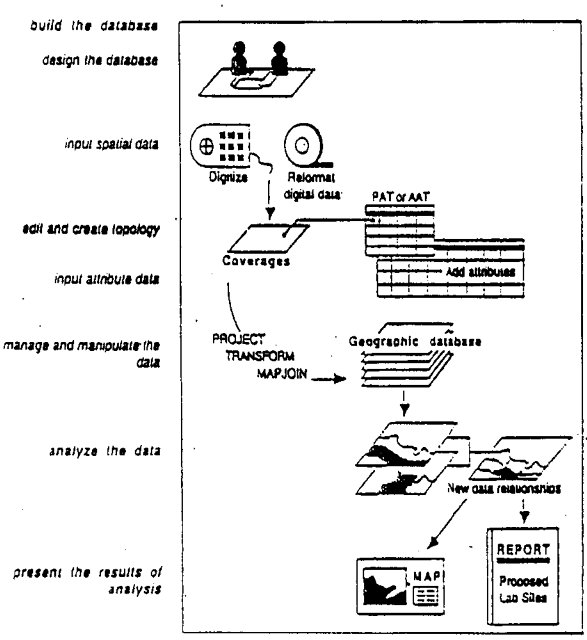
◇ **Nodes:**

- The beginning and ending location of an arc.
- Arcs connect at nodes.

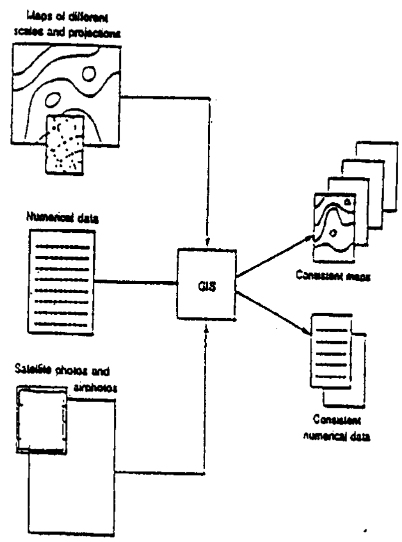
Feature	What they represent	Examples	IDs
Points	Single isolated location	Wells	1
		Markets	2
		Fire station	3
Arcs	Linear features	Roads	201
		Streams	300
		Railways	210
		Power lines	500
Polygons	Homogeneous area enclosed by boundary	Cropland	410
		Forest	420
		Water body	310

Integration of Satellite Data into GIS

GIS Project Sequence



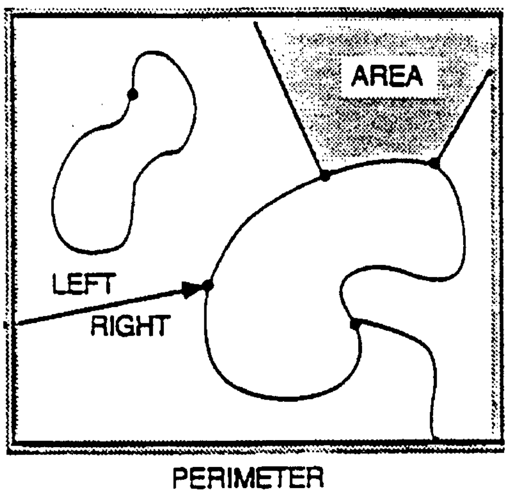
Satellite data are integrated in typical maps and tabular data



1.4 Topology in GIS

- The spatial data base is known as coverage which is characterized by a specific topology. The topology makes the GIS data base intelligent.
- The descriptive map information is stored in the computer in a tabulate database
- The spatial data and the tabular data are linked by a common identifier.

Topology: Defining Spatial Relationships



■ Spatial properties of topology

Topology

We know that the map features are represented by points, lines and polygons. By seeing a map, we interpret the spatial relationships between the features using our mind. For example, we can trace a route along a city map to find the way from an airport to a hotel. In digital map, such spatial relationships are depicted using topology. Topology is a mathematical procedure for explicitly defining spatial relationship. For a map, topology defines connection between features that identifies adjacent polygon, and can define one feature, such as an area, as set of others (i.e., lines).

Creating and storing topology has a number of advantages:

- ✓ Data are stored more efficiently when topology is used.
- ✓ You can process the data faster and can process larger data sets.
- ✓ You can perform analysis function, such as modeling flows through the connecting lines in a network.
- ✓ You can combine adjacent polygons with similar characteristics.
- ✓ You can overlay geographic features.

Three major topological concepts of **ARC/INFO** are:

- ⋆ Arcs connect each other at nodes (connectivity)
- ⋆ Arcs that connect to surround an area, define a polygon
- ⋆ Arcs have direction and left and right sides (contiguity).

Topology uses spatial relationships to define spatial properties.

Spatial Relationships	Spatial Properties
Each arc has a beginning node and an ending node	Length of arc Directionality
Arcs connect with other arcs at nodes	Connectivity
Connected arcs form polygon boundaries	Area of Polygon Perimeter of polygon
Arcs have polygons on their left and right sides	Adjacency or contiguity

Arc-node Topology

Arc- node topology is known as the relationship between the arc features and the nodes. The construction of these relationships allows to determine the important properties of direction and connectivity.

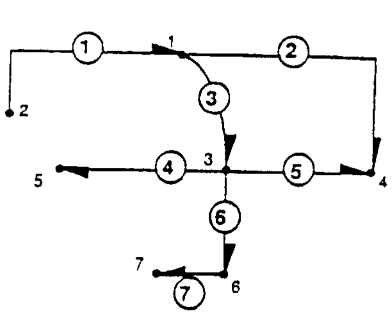
- From-node and To-node:
  - ◇ These numbers are based on how the data was digitized.
  - ◇ Except in specific applications, the direction of digitizing is insignificant
  - ◇ To do model, direction of digitizing is important, e.g. the flow of rivers and street traffic.

- Direction:
  - ◇ In an AAT the form-node is defined by **FNODE#** and to-node is **TNODE#.**

- Connectivity:
  - ◇ If the node number of one arc (FNODE# or TNODE#) matches the node number of another arc (FNODE# or TNODE#)

- Left and Right Polygons:
  - ◇ These can be determined by starting along an arc, beginning at the from-node and ending at the to-node. The enclosed area of the arcs left side is left polygon and the right side is right polygon.

Arc-node Topology



■ **Spatial Data**

ARC#	FNODE#	TNODE#	LENGTH	ID NO.
1	2	1		
2	1	4		
3	1	3		
4	3	5		
5	3	4		
6	3	6		
7	6	7		

ARC#	X,Y PAIRS
1	X,Y X,Y X,Y
2	
3	
4	
5	
6	
7	

ARC COORDINATE FILE

■ **Topologic Data**

Polygon-arc Topology

Polygon-arc topology is known as the relationship between the arc features and the polygon features for which they create boundaries. To determine the important properties of area definition and adjacency for constructing the relationship between these features.

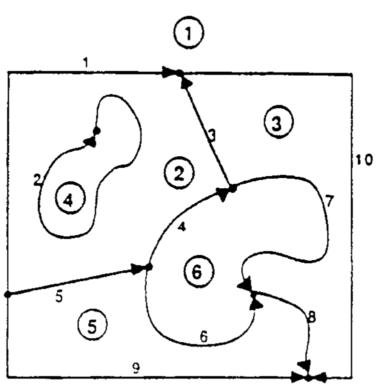
Adjacency:

Arcs form a boundary shared by two adjacent polygons **(LPOLY#, RPOLY#)**.

Area definition:

Area of the polygon is a set of arcs connected together to form a loop for integration of satellite data into GIS

Polygon-arc Topology



■ Spatial data

POLYGON-ARC LIST

POLY#	ARC#
1	1,9,10
2	1,0,2,3,4,5
5	5,6,8,9
4	2
3	3,7,8,10
6	4,6,7

ARC#	LPOLY#	RPOLY#
1	1	2
2	2	4
3	2	3
4	2	6
5	2	5
6	6	5
7	3	6
8	3	5
9	5	1
10	1	3

LEFT/RIGHT LIST

ARC#	X,Y PAIRS
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

ARC LIST

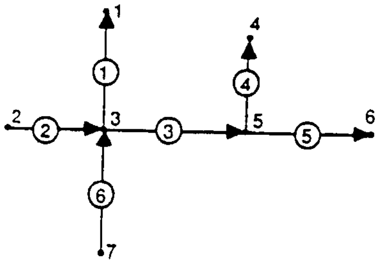
■ Topologic data

CONSTRUCT AN ARC-NODE TOPOLOGY

The map of roads below shows seven nodes which have been numbered. It also shows the arcs, with their numbers circled. Use the table to list the from-node and to-node of all the arcs. Note that one or more arcs can share a node; but by definition, a node cannot exist without an arc.

Now list the arcs along which you would traverse to get from node 6 to node 1. Also indicate the direction to travel across each arc (i.e., beginning at the from-node and traveling to the to-node, or vice versa). This demonstrates arc connectivity.

Roads



Arc	From node	To node
1	3	1
2	2	3
3	3	5
4	5	4
5	5	6
6	7	3

Path from Node 6 to Node 1

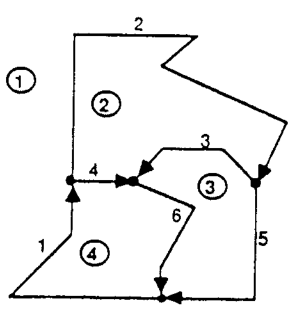
Arc #	5	3	1		
Direction	–	–	+		

\+ = from- to to-node (Forward direction)
– = to- to from-node (Reverse direction)

CONSTRUCT A POLYGON-ARC TOPOLOGY

This step will illustrate adjacency for the polygon map below. Using the first table, define each polygon (circled numbers) by listing the arcs that connect to create each polygon. Record the number of each arc.

Then, in the second table, for each arc, list the polygons on the left and right sides. The arrows indicate the direction of each arc.



Polygon	# of arcs	List of arcs
1	3	1 2 5
2	3	2 3 4
3	3	3 5 4
4	3	1 4 6

Arc	Left polygon	Right polygon
1	1	4
2	1	2
3	3	2
4	2	4
5	1	3
6	3	4

Feature Attribute Tables

* ⋆ Arc attribute table - **AAT**
* ⋆ Polygon attribute table - **PAT**
* ⋆ Point attribute table – **PAT**

Feature Attribute Tables

When constructed a topology for a coverage, attribute table is generated for the feature of the coverage.

- The attributes created define the topological relationships and the geometry o the features.
- The feature attribute table also includes the feature number assigned b **ARC.Cover#,** And an identifier assigned by the user, **Cover-ID**.
- The polygon attribute table and the point attribute table have the sam template, but the area and perimeter items of point features will always b zero.
- There may be two feature attribute tables for one set of spatial data.

Coverage	**Feature Types**	**Examples**
Point coverage	Points only	Wells
Arc coverage	Arcs only	Streets
Polygon coverage	Polygons only	Parcels
Network coverage	Arcs and polygons	Streets and blocks
Link coverage	Points and arcs	Water lines and meters

PAT and AAT

AREA	PERIMETER	COVER#	COVER-ID

FNODE#	TNODE#	LPOLY#	RPOLY#	LENGTH	COVER#	COVER-ID

Identifiers

ARC/INFO provides two different types of identifiers.

- Cover#, numbers assigned and manipulated by ARC/INFO
- Cover-ID, numbers assigned and manipulated by the user

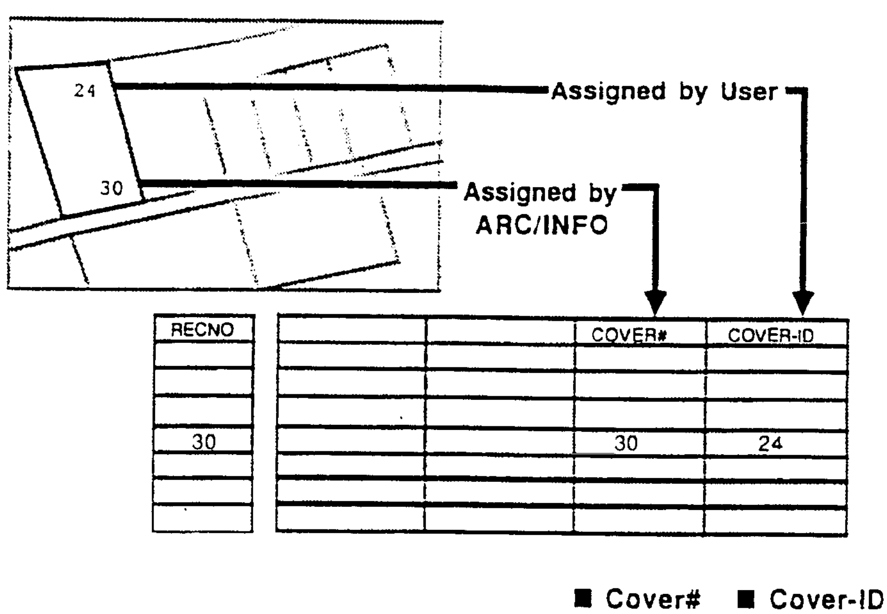
Cover# : This is a feature identifier whose number is defined by the software .

- Each feature values are sequential starting at 1 and incremented by 1
- Values change each time data is added or deleted from the coverage
- Values always match the record number in each feature attribute table.

Cover-ID : To control the feature identifier provided by the user

- Numbers can be any positive integer
- Assigned numbers do not have to be sequential but should be unique
- Numbers can be based on a coding scheme and used during da automation as a temporary attribute.

Identifiers



Arc Attribute Table (AAT)

Descriptive data about geographic features are stored in the feature attribute table. The standard attribute table will contain some components of the topological relationships.

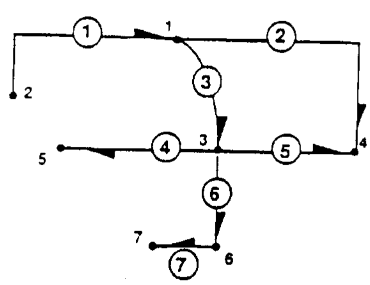
Standard attributes for AAT are:

- ◇ The internal number of the node where the arc begins is **FNODE#**
- ◇ The internal number of the node where the arc ends is **TNODE#**
- ◇ The internal number of the polygon on the left side of the arc is **LPOLY#**
- ◇ The internal number of the polygon on the right side of the arc is **RPOLY#**
- ◇ The length of the arc using the same units as the coordinates is **LENGTH**
- ◇ Internal feature number is **Cover#**
- ◇ User defined feature identifier is **Cover-ID**

Some Important Rules are:

- ◇ ARC/INFO's attributes are always to the left of the Cover-ID
- ◇ Additional user-defined attributes may be added to the right of the Cover-ID
- ◇ For a coverage with arcs but no polygons, the values for the LPOLY # and RPOLY# are always **zero#**

Arc Attribute Table



■ Spatial data

FNODE#	TNODE#	LPOLY#	RPOLY#	LENGTH	Cover#	Cover-ID
2	1	0	0	36.8	1	1
1	4	0	0	38.2	2	2
1	3	0	0	25.5	3	5
3	5	0	0	41.1	4	4
3	4	0	0	41.9	5	3
3	6	0	0	13.6	6	6
6	7	0	0	19.7	7	7

■ Arc attribute table

Polygon Attribute Tables (PAT)

Descriptive data about geographic features are stored in the feature attribute tab The standard attribute table will contain some components of the topologic relationships.

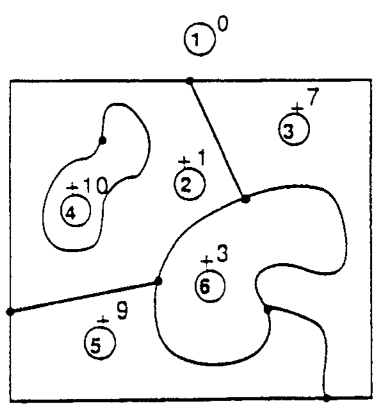
Standard attributes for PAT are :

- The area of each polygon using the same units as the x, y coordinates **AREA**
- The perimeter of each polygon measured using x, y coordinate units **PERIMETER**
- The internal feature number is **Cover#**
- The user defined feature identifier is **Cover-ID**

Some Important Rules are:

- Point and polygon features may not be stored in the same coverage
- Each polygon must have one, and only one, label point
- Each label point should have a unique value.
- ARC/INFO's attributes are always to the left of the Cover-ID
- Additional user-defined attributes are always added to the right of the Cove ID.

Polygon Attribute Table



■ Spatial data

AREA	PERIMETER	Cover#	Cover-ID
-685.6	325.6	1	0
137.4	45.2	2	1
141.3	52.3	3	7
120.4	41.2	4	10
140.8	51.1	5	9
145.7	60.0	6	3

■ Polygon attribute table

Point Attribute Tables (PAT)

Descriptive data about geographic features are stored in the feature attribute table. The standard attribute table will contain some components of the topological relationships.

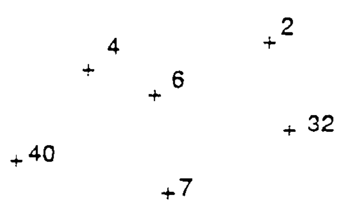
Standard attributes for PAT are:

- **AREA**: the point have no area
- **PERIMETER**: points have no perimeter
- **Cover#:** The internal feature number
- **Cover-ID**: User defined feature identifier

Some important rules are:

- Point and polygon features may not be stored in the same coverage
- Each label point should have a unique value
- ARC/INFO's attributes are always to the left of the Cover-ID
- Additional user-defined attributes are always added to the right of the Cover-ID.

Point Attribute Table



■ Spatial data

AREA	PERIMETER	Cover#	Cover-ID
0	0	1	4
0	0	2	6
0	0	3	2
0	0	4	40
0	0	5	32
0	0	6	7

■ Point attribute table

Chapter 2

Getting Spatial Data into ARC/INFO

The initial task in a GIS work is to create a spatial database by transferring data from map to the computer memory. A spatial database in ARC/INFO is created by converting features in a map to a digital format on the computer. The conversion tasks may be performed using any of the following input devices:

Digitizer: Data on a map may be captured by digitizing each feature, one by one using a digitizer table. This is a manual process of spatial data capture and is used widely. It has the simplicity of creating thematic data-base, but requires hours and days of digitizing and editing works for creating usable database. Accuracy of captured data is largely dependent on the quality of the source map as well as on the skills of digitizing.

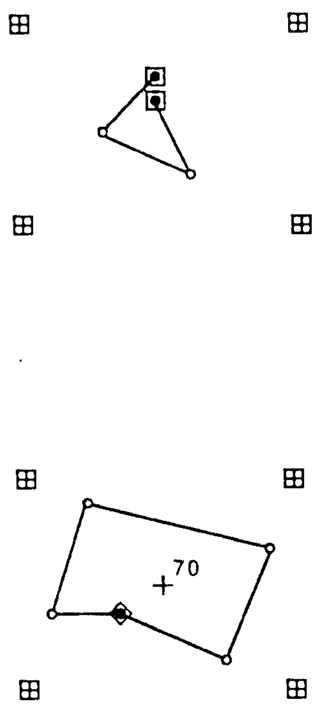
Electronic Scanner: This device may be used to capture data on an entire map sheet automatically. This procedure has the advantage of easy initial data capture, but the tasks of creating appropriate thematic database are much difficult and time consuming.

Keyboard: Data in the form of known coordinate values, can be captured by typing in the exact x- and y- co-ordinates. Data capture using keyboard is rare due to the unavailability of data in the required form.

2.1 Data Capture through Digitizing

In digitizing procedure, point, line and area features that form a map are converted into x, y co-ordinates and are transferred to the database stored in the computer. A point is represented by a single coordinate, a line is represented by a string of coordinates, and when combined, one or more lines with a label point inside outline, identify an area (polygon).

In ARC/INFO, there are some terms that should be understood before proceeding to the details of the data capture. The key terms needed for digitizing and the symbols used to present them are shown in the following:

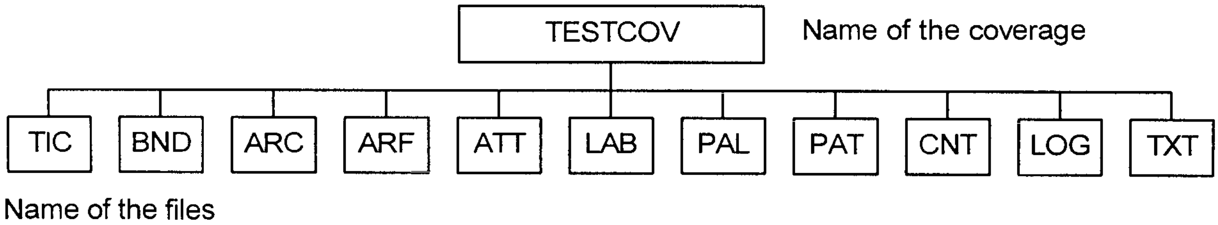


Terms	Symbol	Description
Arc	——	A line feature or the border of a polygon. An arc is defined by a from-node and a to-node and additional vertices in between.
Node	•	An arc endpoint. In many cases, the end of one arc marks the beginning of another, so it can be said that nodes also mark intersections.
Vertex	o	A point within an arc, providing shape.
Pseudo node	◇	The point at which an arc connects to itself (a loop), or to only one other arc.
Dangling node	◻	An arc endpoint not connected to another arc.
Label point	+	Used to either represent a point feature or to identify a polygon.
User-ID	70	A number assigned to each feature. These values should be unique. Once established, these values can be altered by the user, as needed.
Tic	⊞	A registration or geographic control point. These allow all coverage features to be registered to the same coordinate system.
RMS Error	0.004	The calculated difference between recorded and specified tic locations, expressed as a residual of the means squared. Values are usually just slightly greater than 0.000 – the higher the value, the greater the error.

2.2 Spatial Database Structure

In ARC/INFO, spatial database is created as coverage, which is a digit analogue of a single map sheet forming the basic unit of data storage. In coverage, map features are stored as primary features, such as arcs, node polygons and label points, and secondary features, such as tics, extent ar annotation. Map feature attributes are described and stored independently feature attribute tables. However, in spatial database management, a coveraş usually contains data for a single theme or layer, such as soils, roads, land us etc.

In ARC/INFO, a coverage, created during data capture procedure, is stored as set of files where each file contains information about a particular feature clas In order for a coverage's features to be stored, referenced and managed as a se the coverage's files are organized in a directory whose name is the name of tl coverage. This forms the primary ARC/INFO database structure, and is show below:



Among the number of files in a coverage, the user needs only a few to access. Polygon Attribute Table (PAT) and Arc Attribute Table (AAT) are the files most frequently accessed by the user for analysis of GIS data. These two tables contain attribute data for the polygon and arcs respectively and can be accessed through any database management system. The structure of these tables is similar to that of a general database created using softwares, like DBASE. However, these database contain some standard items, like area, perimeter, length, etc., which are automatically created and filled-up during digitizing procedure.

Steps for Capturing Data by Digitizing

Since data capture through digitizing often involve large amount of work, proper steps should be followed for accurate and efficient data capture. The following steps should be considered:

i) *Use of Good Source Map:* The accuracy of the digital data is directly affected by the quality of the map manuscript from which features are to be digitized. The map should be in good condition, clean and readable, not torn or folded. Map materials such as paper, are affected by climatic conditions, and introduce distortion in the digitized data. To minimize this, map should be copied onto a stable material, such as MYLAR, that minimizes stretching and shrinking.

ii) *Definition of Procedures:* How digitizing will be performed should be determined. Some guidelines are:

- A standard sequence procedures should be established. As an example: arcs may be digitized first and then points.

- ⋆ A sequence for digitizing features and map sheets should be establishec so that one can easily track which portions of the database have alread} been digitized.

- ⋆ Standard naming conventions should be established. As an example: i: the point features are being digitizing, PT could be used as a standarc suffix for all coverages being digitizing.

iii) *Preparation of Maps:* The goal of preparing the map is to minimize the number of questions the person digitizing will have to face and resolve Preparation of map mainly includes:

- ⋆ Locating tic points on the map and assignment of a unique number tc each of them.

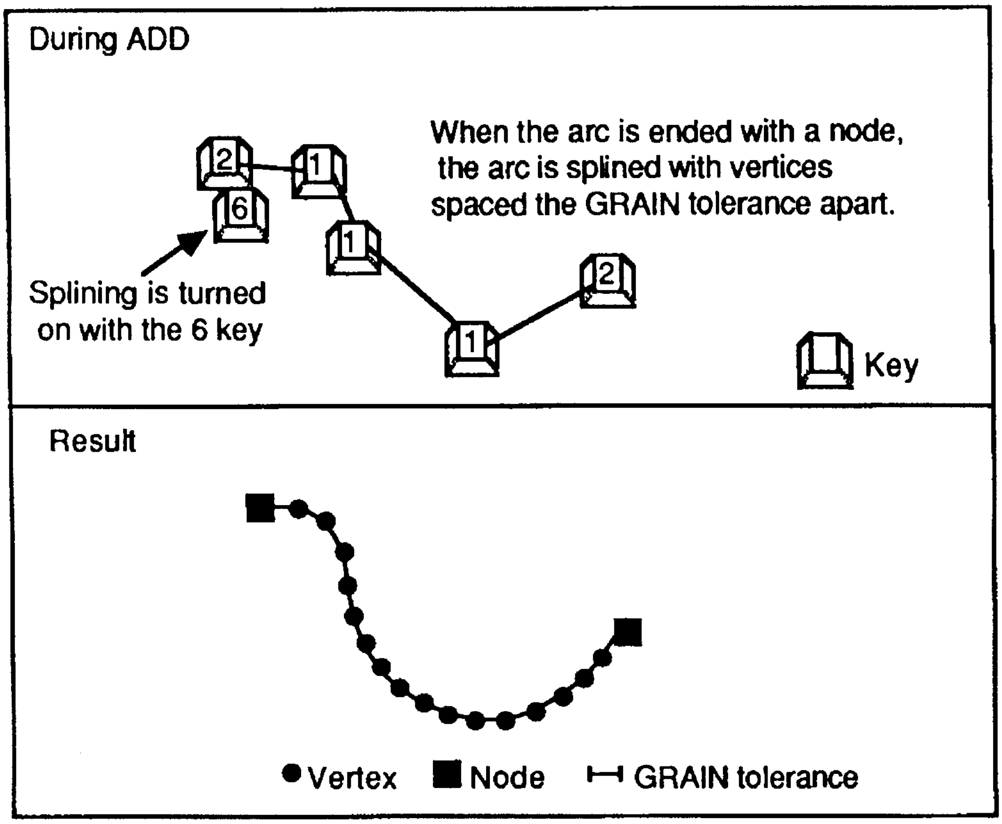
- ⋆ Making a new boundary, slightly larger than the real one.

- ⋆ Indicating start nodes for island polygons.

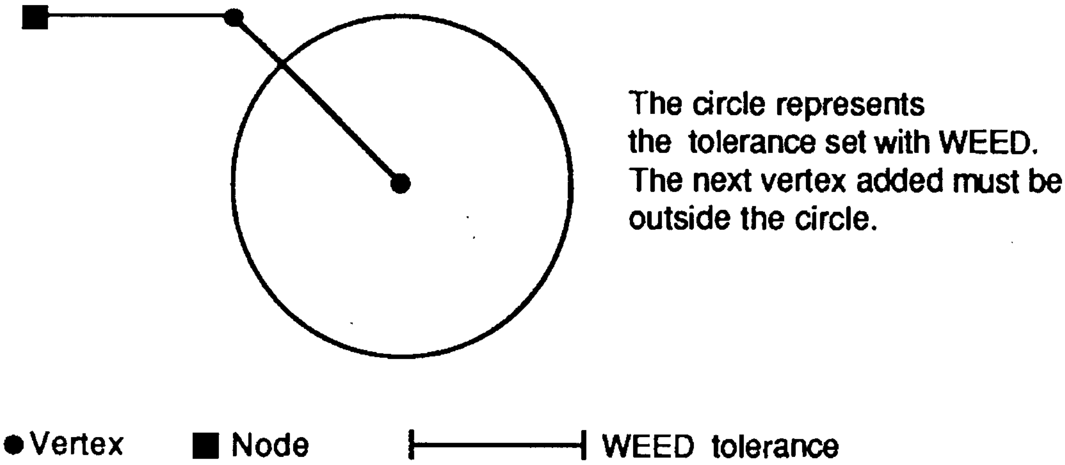
Digitizing of Maps in ARC/INFO :

In ARC/INFO, there are two modules for performing the tasks of digitizing These are ARC Digitizing System (ADS) and ARCEDIT. However, whateve may be the module used, the digitizing environment should be set up befor starting digitizing. In ARC/INFO, the digitizing environment is set up using th following parameters:

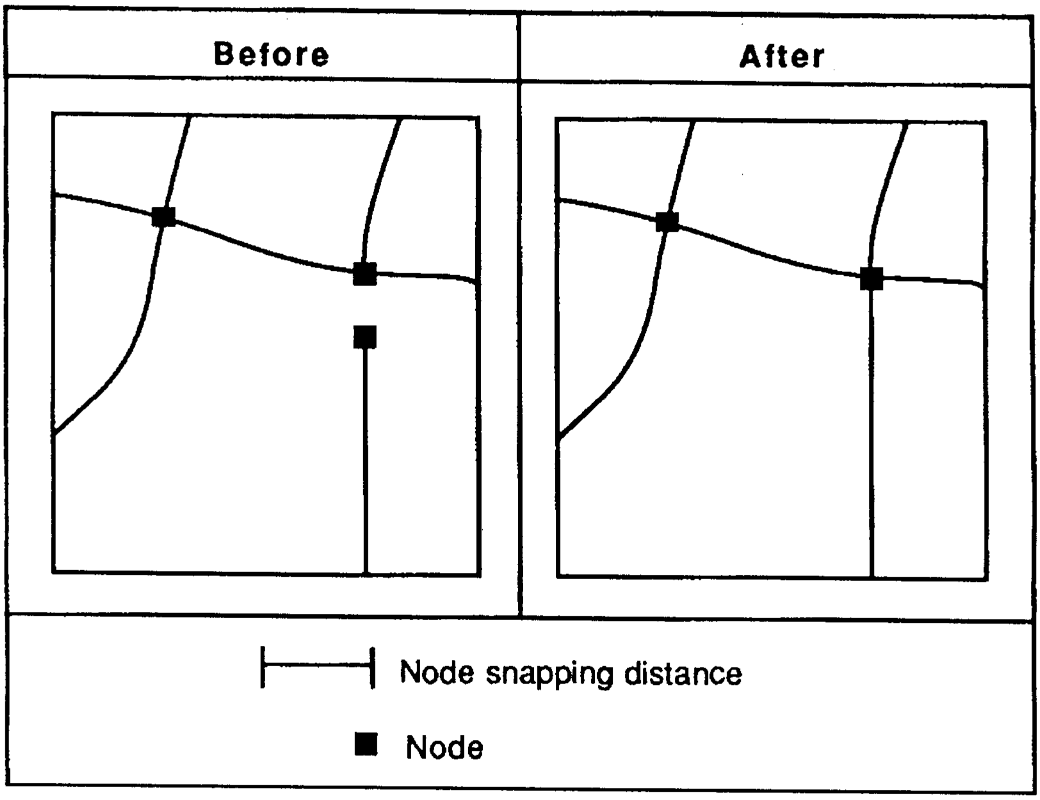
Grain tolerance: This tolerance controls the distance between vertices when a arc with a curve in it is added, a circle is added, or an arc is splined durin digitizing. This tolerance is given in coverage units or indicated by using th cursor to specify two points on the coverage, the distance between which is th grain tolerance. This tolerance is set using GRAIN command:



Weed tolerance: This tolerance controls the minimum allowable distance between vertices on an added arc. A vertices cannot be added if the distance between it and the previous vertices is less than the distance specified with WEED command.



Snap distance: This is used to snap a node of an arc being added to an exist node.



Use of ADS

ADS is a ARC command and is used to create a coverage by copying tics fr the master tic coverage and, then, after entering the corresponding tic locati at a digitizing station, adding arcs and label points. This command is normą used to add spatial features in a coverage, and has limited capability of edit coverage features.

ADS begins with a dialog to identify the tic locations and the extent of coverages. After entering the tic and the initial boundary for the coverage, m menu will appear on the screen as shown below:

1. Add Arcs 3. Add Labels 5. Draw Coverage 7. Add tics
2. Remove Arcs 4. Remove Labels 6. Define window 8. Set Options
9. EXIT Enter function:

Once a function has been selected, a second sub-menu appears. As an example, the menu below is for 'Add Arcs' function of the main menu:

Adding arcs - **POINT mode**
Options : **1) Vertex 2) Node 3) New User-Id 4) Delete vertex**
5) Delete Arc 6) New User-Increment 9) Exit *) Stream on/off

Use of ARCEDIT

ARCEDIT is a specialized sub-system of ARC/INFO which provides advanced capabilities for interactive, sophisticated graphics editing for coverage creation and update, and for final map production. It provides all of the facilities for digitizing coverages with a more comprehensive set of graphic editing commands. One can edit feature attributes, add high quality test annotation, use other database layers as background displays or attribute transfer. One can move, copy, delete, reshape, and undelete arcs, label points, map annotation, tics and nodes. It has numerous function for arc coordinate, attribute and annotation editing.

In ARCEDIT one has to enter a series of commands for performing data capture tasks. As an example, creation of a new coverage and spatial data entry tasks can be performed using the following command sequence:

1. Specifying the method for coordinate input: Three methods are available for this task. These, alone with the command to be used, are shown below:

   : COORDINATE (CURSOR/KEYBOARD/DIGITIZER)

2. Creating a new coverage: If the name of the coverage to be created is TESTCOV, the following command should be used:

   : CREATECOVERAGE TESTCOV

   The software prompt for this command are as follows:

   Creating TESTCOV
   DIGITIZER TRANSFORMATION
   Digitize a minimum of 4 tics.
   Signal end of tic input with tic-ID = 0

   Tic-ID: 1*
   Tic-ID: 2*

Tic-ID: 3*
Tic-ID: 4*
Tic-ID: 5*

Enter the corner point of boundary
Enter opposite corner of boundary
The edit coverage is now TESTCOV

3. Addition of new features: The following command should be used,

   : ADD

   The software prompt for this command are as follows:

   1) Vertex 2) Node 3) Curve
   4) Delete vertex 5) Delete arc 6) Spline on/off
   7) Square on/off 8) Digitizing option 9) Quite *Stream on/off
   (Line) User Id: 1 Points = 0

Digitizing may be started by entering the first node on the arc with the 2 key c the digitizer cursor, followed by a series of 1 keys to enter points along the ar and then press 2 to enter the last node on the arc.

Chapter 3

Making Spatial Data Usable

Whenever a coverage is digitized, the initial tasks of spatial data capture are completed. However, this data are not usable for analytical processing until some specific aspects of data preparation are ensured. These are:

- ☑ There should not be any missing data.
- ☑ There should not be any extra data.
- ☑ Data should be accurate.
- ☑ Feature that should connect actually do.
- ☑ All features are within the outer boundary.
- ☑ All polygons have one, and only one, label point. A polygon should have a label point with a unique identifier so that, the code identifying its use (e.g., urban, forest, etc.) can be associated with it.

The important requirement of a coverage to be used is that the features in it should be in the right place, have the correct shape and connected to the outer boundary. These relationships are known as topological relationships. Once these relationships are made explicit, one can query, analyze and display the data contained in a coverage. To make spatial data usable for the coverage that has been digitized, some specific steps are to be performed:

1. ***Construct topology.***
2. ***Identify digitizing errors.***
3. ***Fix the errors.***
4. ***Reconstruct topology.***

3.1 Construction of Topology

Construction of topology establishes the relationship between the geographic features within a coverage. When topology is constructed, arc intersections are created, the arcs that make up each polygon are identified, and a label point is associated with each polygon. All these help to identify the digitizing errors.

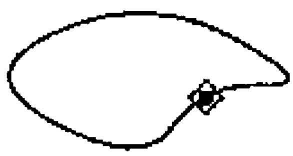
3.2 Identifying Digitizing Errors

Once topology is constructed, one can identify the digitizing errors exist in tl coverage. Common digitizing errors are mentioned below:

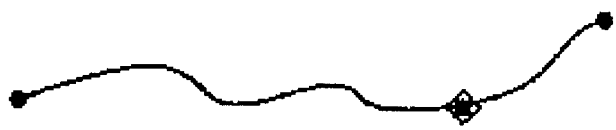
◇ Node Errors

ARC/INFO marks potential node errors with special symbols.

- *Pseudo nodes:* This, drawn with a diamond symbol, occurs when a sing line connects with itself (an island) or where only two arcs intersects.



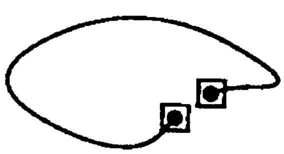
An island pseudo node



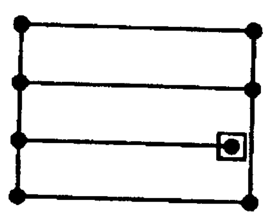
Pseudo node marking a two-arc intersection

This node do not necessarily indicate an error or a problem. Acceptab nodes may represent an island or the point where a road changes fro pavement to gravel.

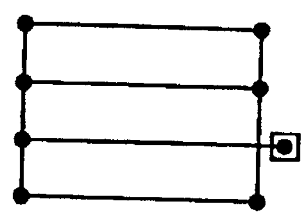
- *Dangling node*: Represented by square symbol, refers to the unconnecte node of a dangling arc. Every arc begins and ends at a node point. So if a arc does not close properly, or is digitized past an intersection, it w register as a dangling node.



An open polygon



An 'undershoot'

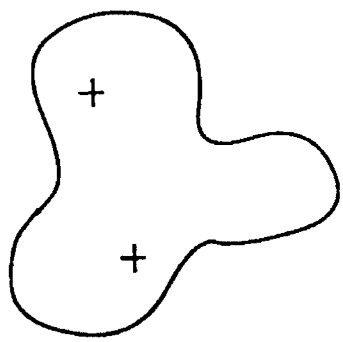


An 'overshoot'

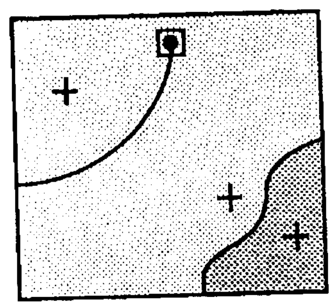
This node may be acceptable in some cases, like end (or start) of a road which may be presented by an arc and is not connected to other arc feature in the map.

◇ **Polygon Errors**

- *Missed polygon*: Arcs are assembled clockwise into closed polygon. If the arcs surrounding a polygon do not close, then the polygon does not exist.



Polygon with two label points



Polygon with two label points resulting from a dangling arc

- Label Errors: There may be no or more than one label point(s) for a polygon.

The errors mentioned above can be identified automatically by using the ARC commands **NODEERRORS** and **LABELERRORS** in majority of the cases. However, in other cases the digitized data should be compared with the source manuscript for identifying the errors.

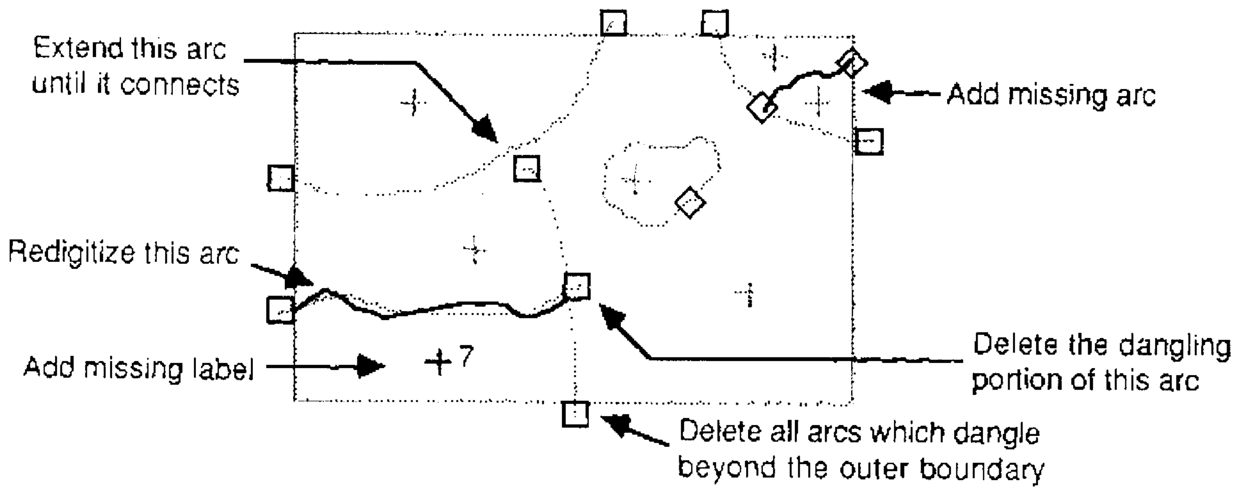
3.3 Fixing of Errors

It is evident that for geographic analysis, like calculation of areas, etc., fixing of errors is very important. Spatial database will not be usable without fixing the errors in the coverage. Fixing errors simply means that missing data are added, and bad data are removed and replaced by correct data. There is often more than one way to fix these errors. The following is a guideline for fixing errors:

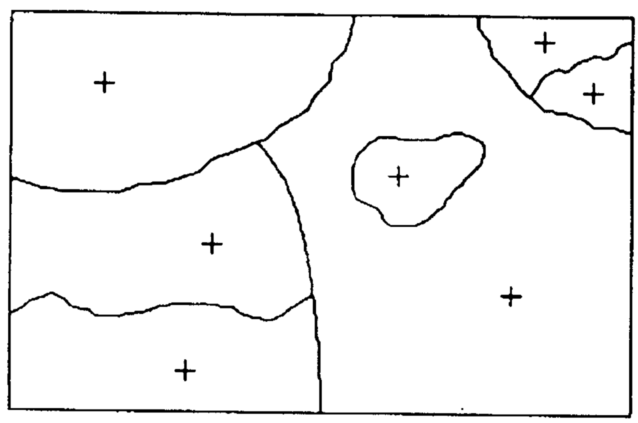
Errors	Action to be taken
Missing arc(s)	Draw them in
Missing label point(s)	Mark position and unique User-ID
More than one label points in a polygon	Identify which one(s) to delete
A gap between two arcs or an unclosed polygon	Identify which arc to extend or which node to move
An overshoot	Indicate whether it should be deleted
Incorrect User-ID values	Mark the correct value

The error fixing procedure and its result is shown below:

For example, here is a set of errors marked to be fixed.



Here is the same set of data after the spatial errors have been corrected.



3. 4. Reconstruction of Topology

Whenever the features of a coverage is edited, its topology is altered. Therefo topology of the coverage must be reconstructed to reestablish the spat relationships. However, after each reconstruction of topology, the coverɛ must be examined to see whether error(s) still exists in the coverage. If so, ı procedure of error fixing and topology reconstruction should be repeated.

ARC/INFO Commands to Construct Coverage Topology

There are two commands: **BUILD** and **CLEAN**

Both of these two commands are used to construct topology, but they somewl differ. The important difference is that BUILD processes all the three prim coverage features (points, lines and polygons), whereas, CLEAN processes oı lines and polygons. Another difference is that BUILD does not cre intersections of the arcs, whereas CLEAN does. If a coverage contains po feature along with the other features, BUILD should be used to construct topology initially. However, for identifying and fixing errors in the covera CLEAN should be used to reconstruct the topology of the coverage.

Feature Attribute Tables

Feature attribute tables are data files associated with each feature type. Wh topology for a coverage is constructed, **Polygon Attribute Table (PAT)** a **ARC Attribute Table (AAT)** are created, which contain the attribute data the spatial features of a coverage. The following standard items are created ɛ filled up automatically for the two tables:

standard items in an AAT

The AAT contains seven standard items created in the following order:

FNODE_	Internal node number for the beginning of an arc (from-node)
TNODE	Internal node number for the end of an arc (to-node)
LPOLY_	Internal number for the left polygon (the Cover_ in a corresponding PAT)
RPOLY_	Internal number for the right polygon (the Cover_ in a corresponding PAT)
LENGTH	Length of each arc, measured in coverage units
Cover	Internal arc number (values assigned by ARC/INFO)
Cover_ID	User-ID (values assigned by the user)

The AAT for a coverage containing only lines will have zero values for the left and right polygons.

standard items in a PAT

The PAT contains four standard items created in the following order:

If you're using Version 3.3, the standard item names use a pound sign (#) instead of an underscore (_) for items FNODE#, TNODE#, LPOLY#, RPOLY# and Cover#. Cover_ID becomes Cover-ID.

AREA	Area of each polygon, measured in coverage units
PERIMETER	Length of each polygon boundary, measured in coverage units
Cover_	Internal polygon number (assigned by ARC/INFO)
Cover_ID	User-ID (assigned by the user)

The PAT for a coverage of points will have zero values for both the area and perimeter.

These two tables, in their standard forms, contain the data regarding the geographic measurements. Another important thing is that, these tables are used for maintaining link between the spatial data in the coverage and the non-spatial data from other sources.

Chapter 4

Getting Attribute Data Into ARC/INFO

Whenever the topology of a coverage is constructed in ARC/INFO, two feature attribute tables are created:

- **Polygon Attribute Table (PAT).**
- **ARC Attribute Table (AAT).**

At this stage, the GIS database is usable for geographic analysis or output production. However, in majority of the cases, the production of the expected results through the analysis of the database still needs to add more non-spatial geographically referenced data to the database. As an example, it is necessary to add data regarding the land use types and cost per hectare value to the database in order to determine the land acquisition cost of each potential site.

In ARC/INFO, the non-spatial geographically referenced data are considered as attribute data, and are added to the feature attribute tables mentioned before. The link between each geographic feature in a coverage and a corresponding record in the feature attribute table is established automatically. When the additional attribute data contained in a data file are joined to the feature attribute tables, they too will automatically be linked to the geographic features. As we saw in the last chapter, among the standard items, the feature attribute tables have two items named **Internal_ID** and **User_ID**. Internal_ID is used by the computer to maintain link between the geographic feature in the coordinate database file and the record of the feature in the attribute table. User_ID is normally used to link the record in the feature attribute table with the record in the other database file(s) containing the additional data required to complete the project database.

4.1 Procedure of Adding Additional Attribute Data to the Feature Attribute tables

In ARC/INFO, there are two methods for adding additional attribute data to the feature attribute tables. The first one, allows direct access to the attribute tables, and the second one allows relating or joining data in a data file with the attribute tables.

⌘ Direct Access

Direct access to the attribute tables is performed using **TABLES** sub-system the ARC/INFO. Like other general database software, TABLES allows to a item(s) to the database structure of the attribute tables. It then allows to en data into the attribute tables using keyboard. The standard procedure of dir access is shown below:

TABLES is activated from **ARC** prompt. After loading of TABLES, t following software prompt appears,

Enter command:

The attribute table (say **TEST.PAT**), to which data will be added, should selected first,

Enter command: SELECT TEST.PAT

The item(s) to be added, say **Soiltype**, is (are) added to the database structure the attribute table using,

Enter command: ADDITEM

Software prompt will appear for Item Name, Item Width, Item Type, Decim places and Start-Item. Data for each item can then be entered using,

Enter command: ADD Soiltype

For data entry using **ADD** command of TABLES, no browse screen edit facili is available. Data can only be added sequentially to each item for each recor However, using **UPDATE** command of TABLES, data can be added modified for selected record.

⌘ Relating or joining data files to attribute tables

The following steps are to be completed for relating or joining data file to tl attribute tables:

- ➔ Create a new INFO data file to hold the attribute data.
- ➔ Add the attribute data to the newly created data file.
- ➔ Relate/join the new data file to the feature attribute tables.

New data file, say TEST.DAT, can be created using the command,

Enter command : DEFINE TEST.DAT

Software prompt for item name, item width, etc. will appear. In fact, the INFO data file structure is exactly similar to the structure of the data file created using general database software, like **DBASE**. Procedure of data entry to the INFO data file is similar as discussed earlier in this section.

Once the attribute data are added to the data file, one can attach them to the feature attribute table of a coverage using a common item as a key. Since the records in the feature attribute table can be linked to corresponding records in the new INFO data file, the new attribute data will also be associated with the geographic features. This makes it possible to perform queries and analysis on the data and create maps using attribute data in the feature attribute table.

In order to perform a relate or join operation, a relate item must be defined in the INFO data file, which is same as an item in the feature attribute table. The coverage User_ID is often used for this. However, none of the Internal_ID should be used for this, since it may change any time ARC/INFO process a coverage.

Relate: A relate temporarily associates two data files based on the relate item. The association between the files may be saved for future use. A relate can be perfomed using the **JOIN** command of the TABLES:

Enter command : JOIN [to_table] [join-item]

[**to_table**] is the database file to be related with the currently selected database file.

Join: A join is similar to a relate, except that it actually merges the two associated files into a new file. A join can be performed using the ARC command **JOINITEM**,

:JOINITEM <in_file> <join_file> <out_file> <relate_item> <start_item>

Since in ARC/INFO, an INFO data file is joined with the feature attribute table of a coverage to complete the database, both the in_file and the out_file in the JOINITEM command must be the name of the feature attribute table to which the join_file will be added.

Join vs. Relate:

- The advantage of a relate is that it allows various users in an organizatic to access the same coverage using their own descriptive data available in separate file. It may also make data updating easier.

- Operation on coverages with extended PAT or AAT files takes more tir and machine resources.

- Joining files may prove simpler for short term projects of small scope ar staff, because all the data exist in one table making access simpler.

Chapter 5

Managing the Database

In ARC/INFO, the aim of managing a GIS database is to finish building the project database and to ensure its functionality. A functional database may contain a number of associated coverages having the characteristics mentioned below:

- Each coverage contains clean topology.
- The accuracy of all feature locations has been verified.
- Feature attribute tables are present.
- The accuracy of the feature attribute values has been verified.
- A system of tics or ground control points exist.

The procedures of establishing these characteristics in ARC/INFO for a coverage has been presented in the previous chapters. However, there are some additional criteria in respect of database management that should be established in order to develop an usable geographic database. These are:

- Record all geographic features using real-world coordinates.
- Store all related coverages in one common coordinate system.
- Spatially reference the features of each coverage against features in associated coverages.

Meeting these criteria will ensure that the whole project database, having more than one adjacent coverages, is completed and its coverages are geo-referenced. The project database will then be fully functional, since queries and analysis can be done for the whole database, as well as for any part of the database. The practical steps needed to fulfil these additional criteria are:

- Create a tic coverage containing real-world coordinates.
- Perform coordinate transformation from digitizer to real-world.
- Edgematch adjacent coverages.
- Join adjacent coverages into one coverage.

5.1 Creation of a Tic Coverage containing Real-World Coordinates

In the chapter on Getting Spatial Data into ARC/INFO, it is seen that some tic points are to be digitized before starting digitizing the features on a map. Tics are reference points identified by known coordinates on a map and are used orient the coverage geographically. They are used in subsequent editing sessio for identifying and resistering the coverage boundary, and may also be used orient a coverage to relate with coverages captured at other times.

Typically a tic table, that contains the real-world coordinate values for the tic is developed manually using the commands of **ARCEDIT**. While developing tic table, the coordinate values of the tics must be entered using keyboard. Tl following steps should be followed in developing a tic table:

➔ Locate the tic points on the manuscript to be digitized.

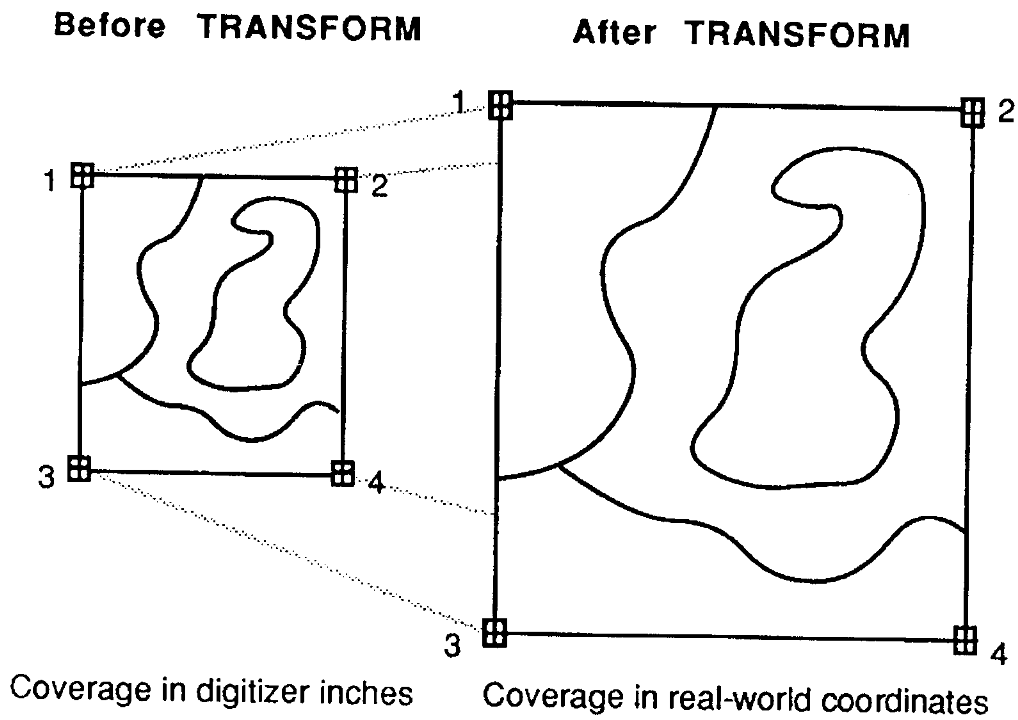
➔ Identify the x- and y- coordinates of the tics from the real- wor coordinate system to which the manuscript to be digitized will l transferred. If the real-world coordinates are not available, but tl geographic coordinates (Latitude/Longitude) are available, the latter can l taken for developing the tic table. In this case, the tics having geograph coordinates must be projected into real-world coordinates letter on.

➔ Use the ARCEDIT command **CREATECOVERAGE <COVERAG NAME>** to create a tic coverage. The software prompt will appear to ent the tic IDs and their coordinate values, as well as, the coordinate values two opposite corners. With all these entry, creation of tic coverage will l completed.

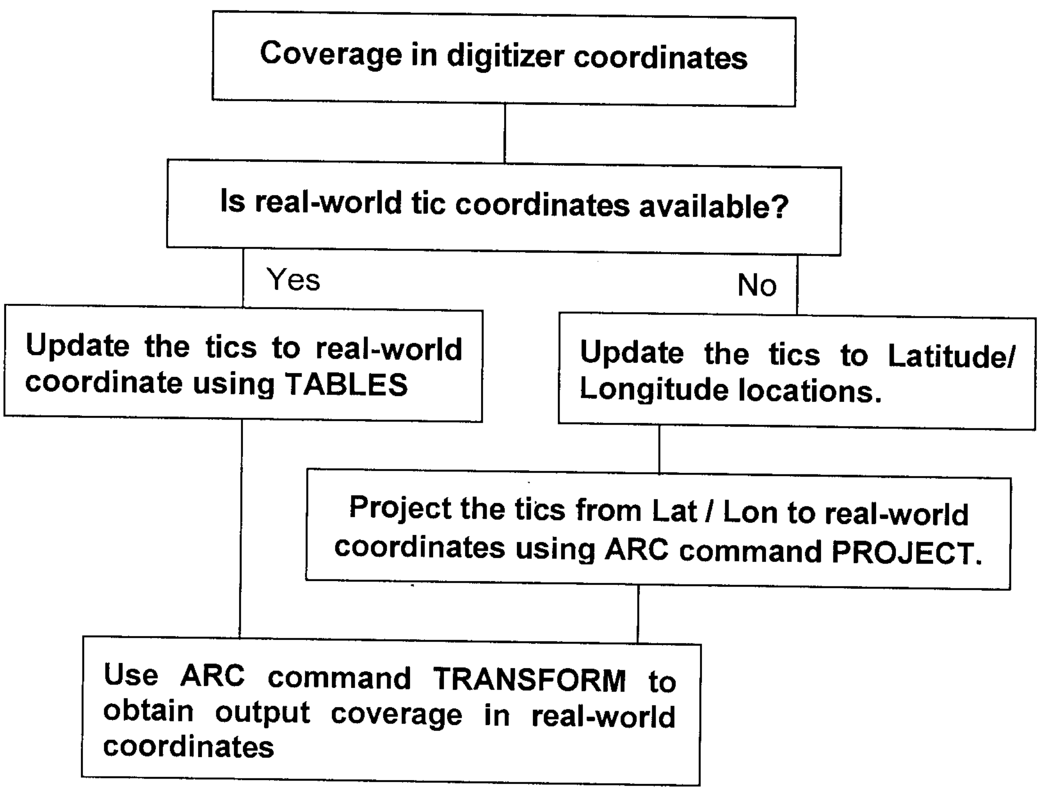
5.2 Coordinate Transformation from Digitizer to Real-World

Digitizer coordinates are the coordinate values of the digitizer itself. The valu of these coordinates are in inches. To make this information meaningful ai also to impose a scale factor, it is necessary to convert these measurements the real-world coordinate system having the same projection of the origir source map. Converting the digitizer coordinates to the real-world coordinates termed as **TRANSFORMATION** in ARC/INFO.

Real-world coordinates are those obtained from the map which has a specil map projection, like Universal Transverse Macerator (UTM). It may be in mel or feet depending on the type of projection of the map. In ARC/INFO, coverage can be digitized using the digitizer coordinate or the real-woi coordinate. However, when a map is digitized, each point on the map is initial converted to digitizer coordinates, and then transferred to the real-woi coordinates if the latter is enabled. The transformation can be performed durii digitizing or later on.



The transformation scheme from digitizer coordinates to real-world coordinates is shown below:



In the transformation process, the input coverage contains tics in digitize coordinates and the output coverage contains tics in real-world coordinates.

5.3 Joining of two Coverages

If the project area is so large that the map manuscript for it is not available in single map sheet, digitizing for the whole project area is to be performed froı separate map sheets. In that case, digitizing of each map sheet creates ε individual spatial coverage, and thus a number of coverages are obtained i order to cover the whole project area. The next task of creating the proje database is to merge all the coverages into a single one. The standard proceduı of joining the adjacent coverages are given below:

- ☑ Visually compare the adjacent coverages to determine that coveraç features match along adjacent borders. Generally, this process is performe by creating verification plots at the original map scale and laying them o on a table side by side. However, as another approach, various map displε can be created on the monitor screen to interactively compare adjacent mε edges.

- ☑ Use clip, split and other tools available in ARC/INFO to edit the coverag if they overlap each other.

- ☑ Verify that the item definitions of the attribute tables of the coverages to l joined match each other.

- ☑ Perform edgematching to assure that arcs and polygon boundaries mat across borders and that the map sheet corners match, using the AR command **EDGEMATCH.**

- ☑ Join the coverages using the ARC command **MAPJOIN**.

- ☑ Remove map sheet borders using ARC command **DISSOLVE**

With the performance of the above mentioned steps, the project database will completed. This database is now ready for analysis and mapping.

Chapter 6

Geographic Analysis and Query

Performing geographic analysis is what sets Geographic Information Systems apart from digital mapping systems. This chapter demonstrates four important geographic analysis concepts:

- ***Coverage Registration.*** Without the accurate registration of the **LANDUSE** and **SOILS** coverages, the accuracy of the results of the polygon overlays could not be assured.

- ***Spatial Joins of Geographic Features.*** The joining of features from various data layers uses the topological relationships inherent in the data model: the arcs comprising polygons in two data layers are known; the new arc intersections can be calculated when the layers are overlain; and the list of arcs comprising each new, resulting polygon can be identified.

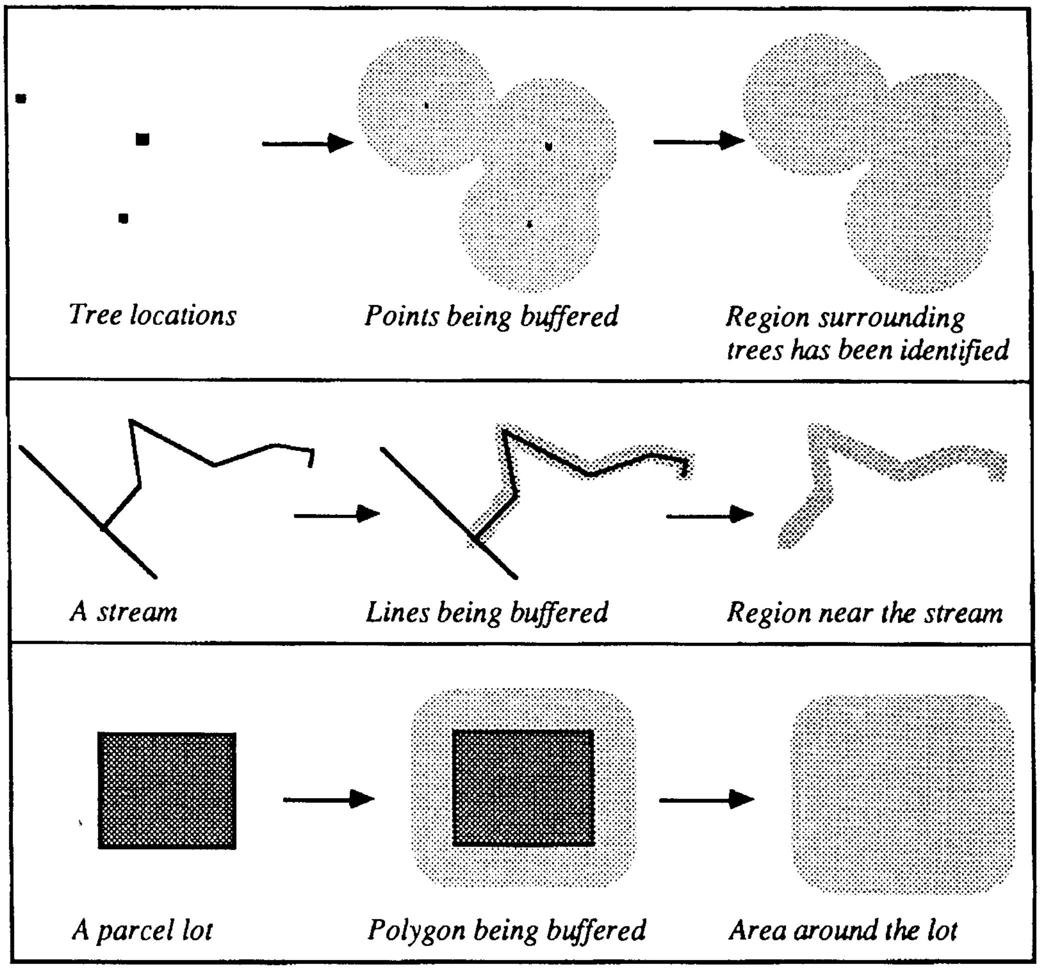
- ***Attribute Join through Coverage Overlay***. The joining of attributes from the various data layers uses overlapping coverage features to join feature attributes from multiple coverages.

- ***Attribute Manipulation (feature selection, adding items and values).*** The tabular database structure allows for logical record selection and calculation of new item values based on the values of attributes from the coverage feature attribute tables.

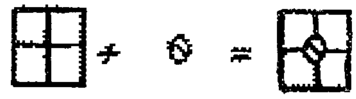
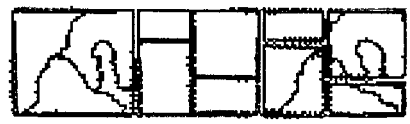
Before performing geographic analysis, it is necessary to define the problem and identify a sequence of operation to produce meaningful results. To better understand the kind of spatial operations required to generate results, geographic operations like creating buffer zones around features, manipulating spatial features and performing polygon overlay are performed.

Buffer generation is used when analysis requires that you identify the area surrounding geographic features. You can buffer any type of geographic features: points, lines or polygons.

You can buffer any type of geographic feature: points, lines, or polygons.

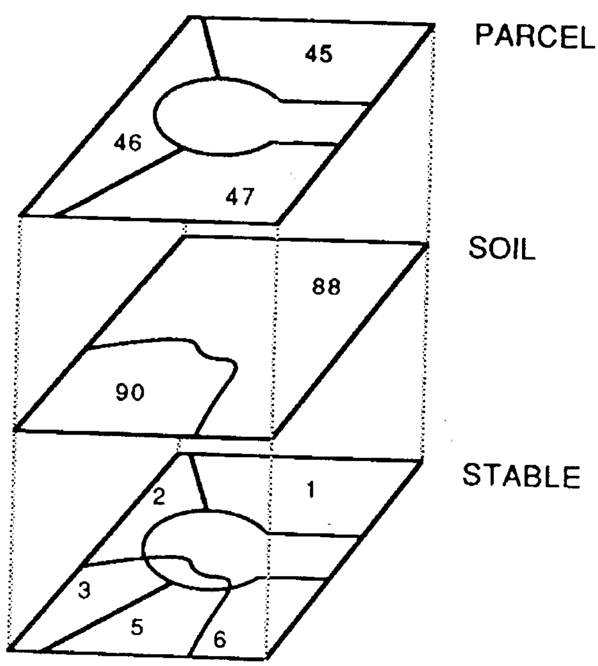


Geographic features can be identified and selected based on whether they are inside or outside the boundary of other coverages. Thus existing coverages can be overlaid or combined to remove, replace, cut, or merge geographic features. The **ERASE** command erases features from one coverage that overlaps with another coverage. Several other commands, including **UPDATE, CLIP** and **SPLIT** can also be used to manipulate spatial features based upon coverage boundaries.

Coverage Updating	Merge new features using a 'cut and paste'	UPDATE	
Feature extraction (subset a coverage or reduce it)	Cut out a piece of a coverage using a 'cookie cutter'	CLIP	
	Split a coverage into a number of smaller coverage	SPLIT	
	Remove part of the inside of a coverage	ERASE	

Overlay

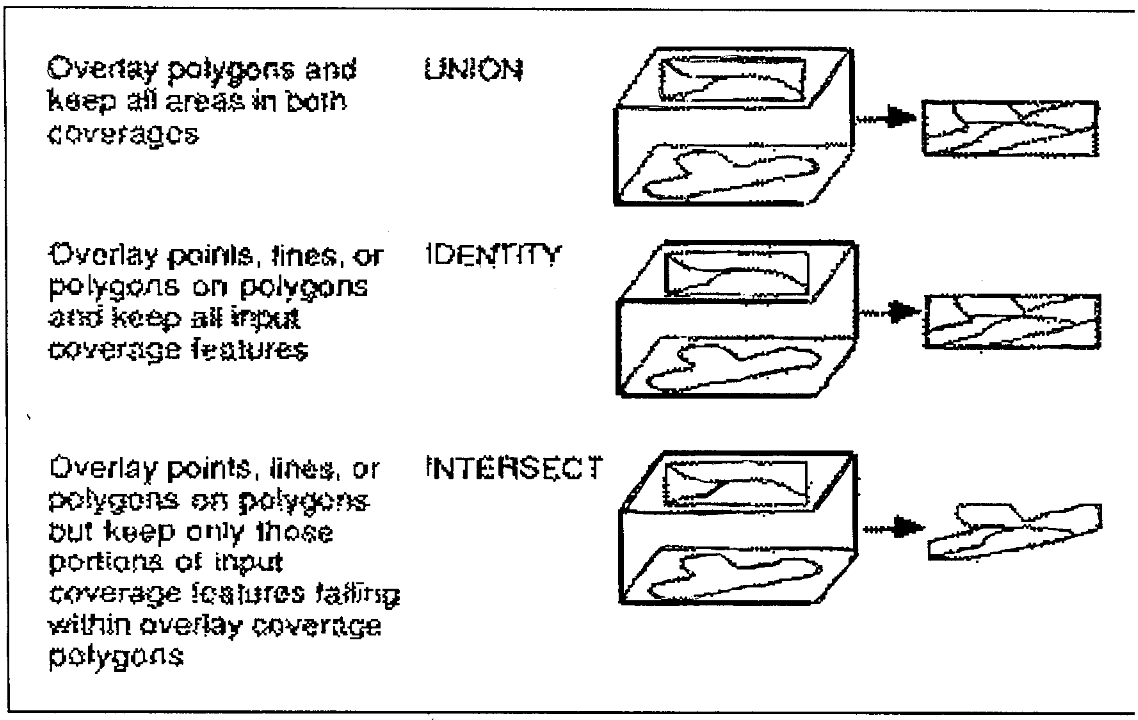
Polygon overlay is a spatial operation when one polygon coverage overlays onto another to create a new polygon coverage. The spatial locations of each set of polygons and their polygon attributes are joined to derive new data relationships. The overlay operation illustrated bellow lets you identify all parcel lots located on unstable soil types.



This topological overlay joins polygons from two coverages to establish the spatial relationships between parcels and soil types. The spatial results contain both sets of polygons.

Other overlay operations include line-in-polygon overlay, in which line featur assume the attributes of the polygon they lie within, and point-in-polygc overlay, in which point features assume polygon attributes.

Three ARC/INFO commands can perform polygon overlay. These comman are similar, differing only in the spatial features that remain in the outp coverage. The results of these commands are shown bellow:



6.1 Tabular Analysis:

The next part addresses how to perform tabular operations in terms of t feature attribute table using logical and arithmetic equations. Before performi tabular analysis, you need to make sure the feature attribute table contains the items needed to hold the new values you will create. The **ADDITE** command can be used to add additional items to an existing PAT or AA Other commands can be used to manipulate an existing PAT or AAT. The include **DROPITEM**, which removes an item from an INFO data file a

PULLITEM, which selects specified items from INFO data file and copies them to a new file.

Once the new items have been created, calculations for the site selection model can be done. This can be done using several tabular analysis operations available in INFO. Following are three of the most commonly used operations.

RESELECT selects a subset of the currently selected set of records that matches a logical expression. Features are selected based on a logical expression which contains **ITEMNAME, OPERATOR, VALUE**, etc. For example, the expression ` Enter command: Reselect area > 2000 ´ will select all polygons with area greater than given 2000.

ASELECT adds records to the selected set matching specified criteria.

Logical expressions are used in ASELECT and RESELECT to select records. Some of the available logical operators include:

EQ or = equal
NE or < > not equal
LT or < less than
GT or > greater than
GE or > = greater or equal
LE or < = less or equal
CN contains the specified character
IN is contains in the subset of values or values range.

CALCULATE assigns values to an item for all currently selected records. Some sample CALCULATE expressions are:

CALCULATE COST = VALUE * HA

CALCULATE ACRES = AREA / 43560

The computation is performed for every selected records. The result of the arithmetic expression is then assigned to the target item.

Chapter 7

Displaying Data

Geographic features are displayed on a map in the same form in which they are created and stored in the computer. The coverage features are translated from the computer to the display surface through a series of operations called map-to-page transformation. This includes positioning and scaling of map on the graphic page. Storing coverage features in real-world coordinates allows maps to be correctly scaled and ensures that measurements such as area and length are accurate. Associating attributes in the tabular database with the map features allows the features to be drawn and labelled based on their attribute values. Additional graphics can be added to the map to make it easier to read, such as lines and boxes, and are only stored as part of the map, not as geographic features. This chapter discusses the components of a map and the nature and use of symbols. It also presentes a step-by-step approach for creating a map using **ARCPLOT**.

7.1 What is a Map

A map is a graphic representation of a part of the earth's surface. The structure of a map is such that it conveys information easily and readily to the reader.

A map contains a series of layers or coverages which are often combined to form the final product. A map also includes descriptive information which helps the reader to interpret the information on the map. Here are some of the main components of a map:

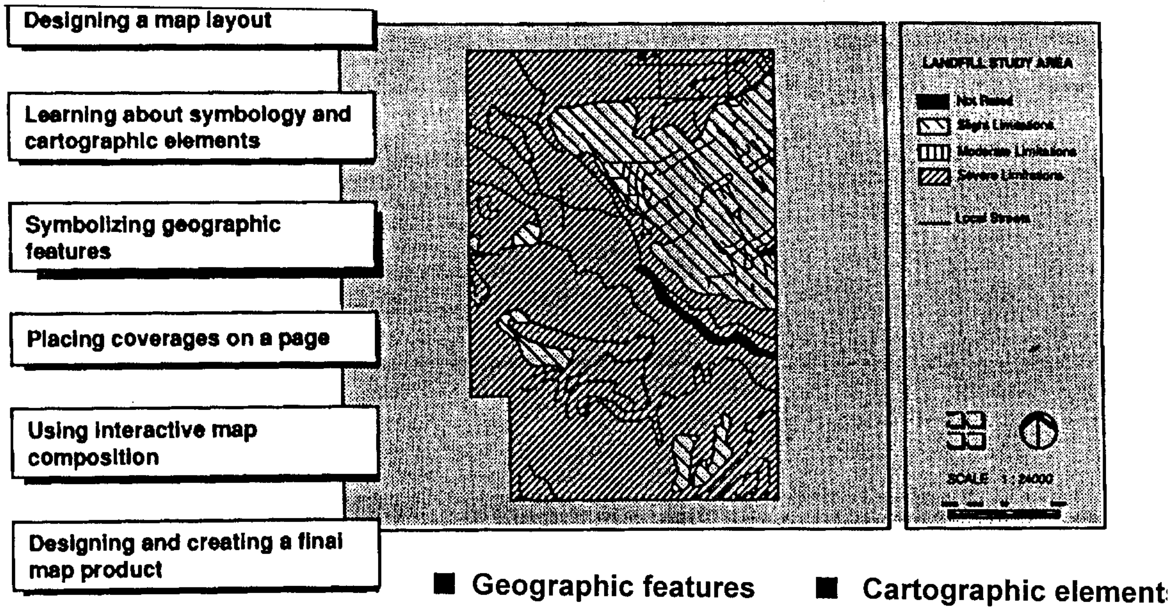
Geographic Features

Geographic map features include areas, lines and points drawn from various coverages in the geographic database.

- Areas are polygon features, such as land use areas. Polygon boundaries are drawn using line symbols. Polygons can be shaded based on attributes using colours, patterns, or both. They can also be labelled with attributes using text symbols.

- Line elements are arcs, such as roads or streams. Arcs are drawn using line symbols and labelled with attributes using text symbols.

Designing and Creating a Map



- o Point elements represent point features or polygon label points. They a drawn with marker symbols and labelled with attributes using text symbols

Cartographic Elements

Cartographic elements help to make the map easier to send and interpret.

◇ Titles and explanatory text describe the purpose of the map and are drav using text symbols.

◇ Neatlines create borders and partitions within a map and are drawn using li symbols.

◇ Legends describe the symbols used to represent the geographic elements a are drawn using line, shade or marker symbols and text.

◇ North arrows and scale bars describe the orientation and scale of the ma They are drawn using line, shade and text symbols.

7.2 Designing and Creating a Map

There are five basic steps for creating a map using **ARCPLOT**.

- Indicate the display method for the map - on the computer's screen or on a graphics plotter.

- Specify the portion of the earth's surface that you want to include in the map i.e., specify the map extent. This could be based on a coverage, a portion of a coverage, or several adjacent coverages.

- Specify certain parameters about the size and layout of the map. These include the final size of the map page, the position on the geographic features on the map, and the scale of the map.

- Draw the geographic features that you want displayed on the map, such as land use areas, roads and streams, and specify the symbols used to draw and label them.

- Add additional cartographic elements to make the map easier to read and understand. These include neatlines, titles, legends, a scale bar, a North arrow and so on.

Indicate the Display

You can create the map as a display on the computer's screen or you can create a digital file of the map known as a plot file, which can then be sent to a plotter to create hard copy version of the map. The map is drawn, it is also stored as a series of plot file instructions, which can then be redisplayed on screen or sent to a plotting device.

Specify the geographic area: Specify the area of the earth's surface that the map will display. This window is known as the map extent. The coverage's boundary (BND) file is used to set the map extent window.

Map Parameters (size, position, and scale)

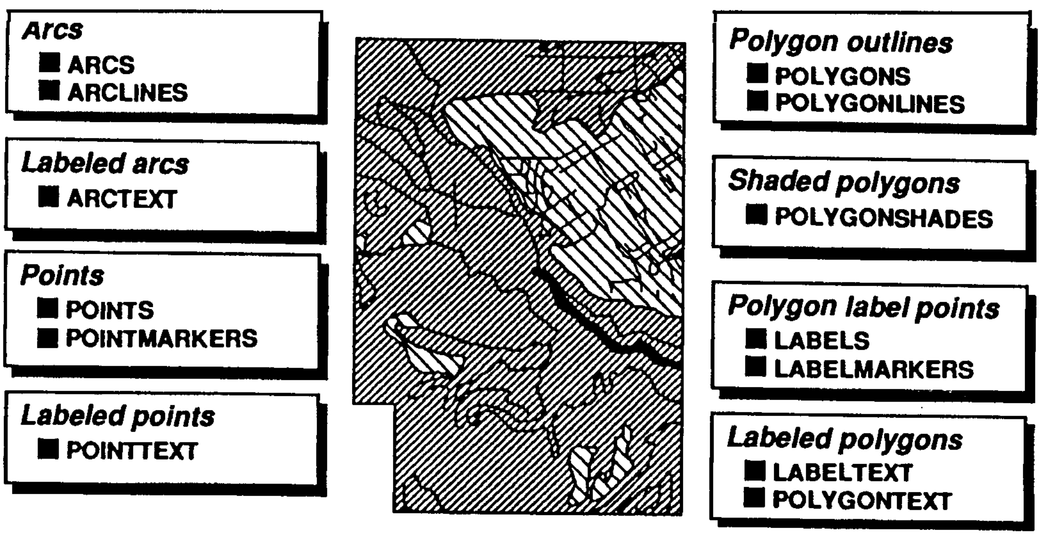
Pagesize: The map is displayed on a physical page. This can be a piece of plotter paper or the computer terminal's screen. The area on the physical page where the map elements can be drawn is known as the graphics page. In ARC/INFO, the graphics page is defined by the **PAGESIZE** command.

Map limits: An area on graphics page can be reserve for displaying those elements that represent coverage features. Limiting where the coverage features can be drawn allows to designate the space on the page for other map elements. The **MAPLIMITS** command specifies the coverage drawing area.

Map position: It is possible to specify the position of coverage features witl the map limits. For example, the **MAPPOSITION** command will let you cen the coverage features within the map limits by placing the center of the m extent at the center of the map limits.

Mapscale: **ARC/INFO** does not store a specific scale for a coverage. Rather, only stores coordinates in real-world units. You can display the coverage at a size by specifying a scale when you draw it. But before you do this, you t ARCPLOT what units your coverage is stored in (usually feet or meters), sc can perform the correct scaling. The **MAPUNITS** command does it.

Draw the coverage features:



ARCPLOT includes several commands for simple display of coverage featur When using these commands, the specified features are drawn using the curr line or marker symbol. Features may also be labelled with their IDs as an opti on these commands. Tics can be drawn with the **TICS** command but instead using the current marker symbol, they are always drawn using a square marl symbol.

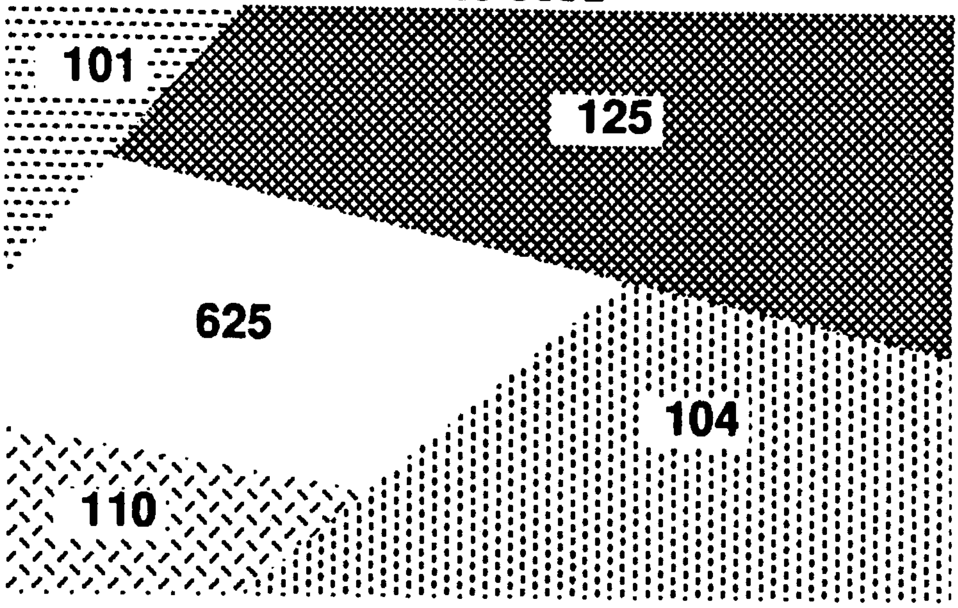
As well as using the current symbol, features can be drawn using a speci symbol. Symbol numbers can be specified explicitly on the command line, can be based on values stored in the feature attribute table or lookup table.

Look up tables

Lookup tables offer several advantages: First, you can use any symbol to represent the attribute- You are not limited to the value stored in the feature attribute table. Also, it can be easier to change symbols if you want to change how your map looks. You can simply specify another item in the lookup table storing different symbol numbers. Another advantage is that lookup tables can reduce data storage. Rather than storing the symbol number with each record in the attribute table, you need only store the symbol once for each item value in the lookup table.

One of the ways of using attributes to specify symbols is to refer to a lookup table. A lookup table is an INFO data file with an item defined the same as an item in the feature attribute table, and another item is called **SYMBOL**. When you draw the map you specify the item to be used to reference the symbol. Then ARCPLOT looks up the value of this item for each record to obtain the correct symbol number.

POLYGONSHADES LU LU-CODE



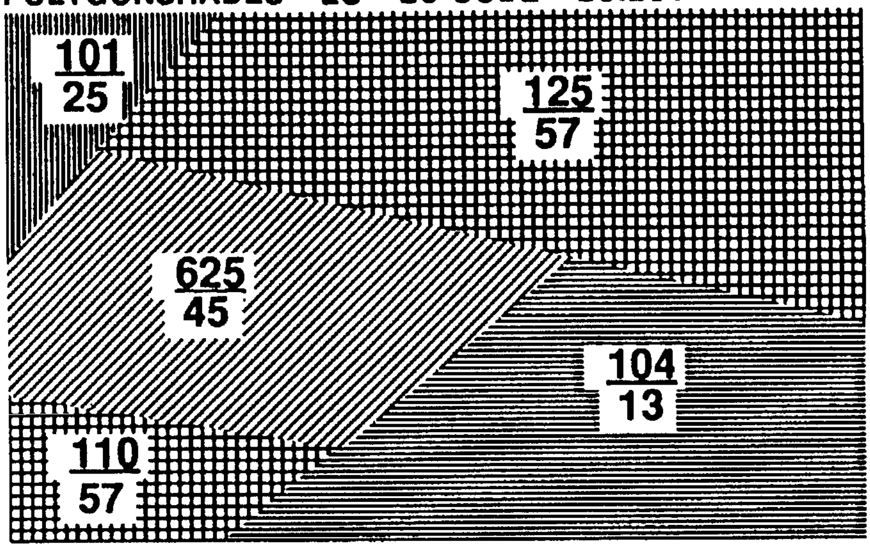
LU.PAT

	#	ID	LU-CODE
	1	0	0
	2	10	101
	3	20	125
	4	70	625
	5	35	104
	6	15	110

→ Universe Polygon

- **LU-CODE values in PAT used as shade symbol numbers**
- **Shade symbol 625 has not been defined, so no shade is drawn**

POLYGONSHADES LU LU-CODE LU.LUT



LU.PAT

	#	ID	LU-CODE
	1	0	0
	2	10	101
	3	20	125
	4	70	625
	5	35	104
	6	15	110

LU.LUT

LU-CODE	SYMBOL
101	25
104	13
125	57
625	45

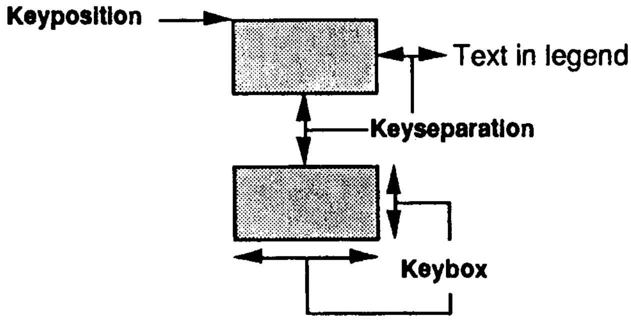
- **LU-CODE values reference symbol numbers stored in a related lookup table**

Add the additional cartographic elements- key legends: If you symbolize coverage features on the map, it will be helpful to explain what the symbols mean by adding a legend. To do this, you will need to create a key legend file specifying what will be in the legend. Key legends are created by combining the current ARCPLOT settings with the contents of specially prepared key legend file. These are text files written with the text editor. They contain symbol numbers and accompany text describing the symbol's representation.

SOILSPLT.KEY

```
.9
Not Rated
.31
Slight Limitations
.43
Moderate Limitations
.47
Severe Limitations
```

■ Key legend file



■ Key legend settings

7.3 Manipulating Map Composition Elements

A map composition is made up of a number of map elements. When using interactive map composition, each element can be manipulated any number of times.

- MINFO – lists the composition element number and the respective commands used to create each. Often used before MSELECT, if selecting by number.
- MSELECT, MASELECT, MUNSELECT, and MNSELECT are used to establish a selection set for the following operations. The use of these is similar to the SELECT command in ARCEDIT.
- MWHO – flashes or redraws the currently selected elements.
- MDELETE – deletes selected elements. Once deleted, elements cannot be restored.
- MGROUP – groups selected elements into a single element. Element may be given a name when grouped. Once grouped, elements cannot be ungrouped.
- MMOVE – moves selected elements from one location to another.
- MCOPY – copies selected elements from one composition to another.
- MROTATE – rotates selected elements.
- MSCALE – sizes selected elements.
- MFIT – moves and sizes selected elements. This technique is recommended for moving elements created from graphic primitives, not coverage features.
- MFRESH – refreshes the screen by clearing and redrawing all the elements.

Chapter 8

Integration of Data from General Database Management System (DBMS) with GIS

In the earlier chapters, the process of addition of attribute data to the GIS database has been discussed. In more advanced applications of GIS, the integration of attribute data with the GIS database is required not only for pure geographic analysis, but also for important socio-economic data analysis. Monitoring of the status of poverty based on geographic areas is an example of this type of application. Since these type of applications are based on geographic areas, the values of the parameters used for the analysis must be raised to the level of the geographic areas concerned. In such cases, integration of attribute data (representing the parameter value) with the GIS database is not simple like the process discussed earlier. A sequence of steps should be followed in order to integrate attribute data correctly, as well as, to maintain the validity of data integration at all spatial data processing level.

8.1 GIS Database Interface with other Database

For the type of application mentioned above, the attribute data must be added to the GIS database with reference to a geographic area which is called polygon in ARC/INFO. In earlier chapters, we have seen that a GIS database contains two attribute tables:

- **Arc Attribute Table (AAT)**
- **Polygon Attribute Table (PAT)**

These two attribute tables contain some standard items and have the structure similar to that of a general database file created by software like DBASE. The structure and the data of these tables can be accessed by a general database software, and thus, these two tables act as the interface of the GIS database with other databases.

8.2 Preparation of the Databases for Integration with GIS

Both the GIS and DBMS database should be in the proper state of integration.

GIS Database

- ✓ This database must contain the geographic areas concerned. As fo example, if the analysis is to be performed for thanas, the GIS databas should contain a coverage of thana boundary for all the thanas under th project.

- ✓ In order to maintain the validity of the integrated data, the database shoul contain the complete coverage before the integration is performed. W have seen that coverage topology should be reconstructed after the edit of coverage. If a coverage is edited after the integration of data reconstruction of topology may alter the attribute data assignment to th spatial features of the coverage. For example, data originally added to thana may be wrongly assigned to another thana.

- ✓ GIS database should contain only the primary geographic features (points arcs and polygons). This will ensure that coverage edit will not be require after the integration of data. For example, if annotation data exist in coverage, it may be required to be edited in order to produce differen maps using different attribute data in the database. This edit may cause th problems as mentioned above. All the annotations should be added to th final map product as separate map elements, but not as coverage elements.

- ✓ The AAT or the PAT should contain an item that is used to address eac geographic feature in the spatial data base, as well as, to relate or joi item(s) during automatic data integration process. This relate item shoul be as global as possible, so that the user can readily identify the spatia features addressed by the items. For example, BBS geocode for thana ca be used as relate item in a thana database. The data file that contain th DBMS data should contain the same relate item so that data integratio become possible.

DBMS Database

- ✓ Data to be integrated should be referenced geographically. For example roads in a GIS database may need to have attribute data, like width, roa type, etc. at their references. Again a geographic area may have numerou attribute data, like total population, cropping intensity, literacy rate, etc.

- ✓ Some of the data may be integrated with GIS database directly without any pre-processing. For example, road types, cost of land, etc. However, for some data pre-processing is required before adding with the GIS database. Population density calculated from the population and area data, average per capita calorie intake in a thana, etc. are examples of such data.

DBMS data pre-processing

DBMS data pre-processing is performed using the programming capabilities of any general database software. The ultimate aim of processing is to automatically develop an output data file having the processed data, through processing of data contained in a base data file. The processing techniques are specific of the requirement of the output, and are of the following two types:

- ➧ Processing based on single record: For example, calculation of the percentage of the poor population in thanas from the numbers of poor and total population in the thana contained in a base data file. Evidently, in this case, the base data file is a thana database.

- ➧ Processing based on group of records: If in the calculation of the percentage of poor population in thanas, the base data file is a household database that contains items of household size and parameters giving the status of poverty of households (like calorie intake), processing must be performed on all the records of each thana.

8.3 Steps in Data Integration

Once both the GIS and DBMS databases are prepared, the actual task of integration may be performed. This includes:

- ⌘ Verification of the relate item: As mentioned earlier, both the databases should contain the relate item in order to address the records to be added/merged. However, if there is any mistake in the values of this item, data integration will not be fully valid. When an application needs to integrate huge amount of attribute data with the GIS database, it become extremely difficult and time consuming to verify the integrated database for mis-integrated data. It is therefore, extremely important to verify the match of the relate item between the two databases. The said verification can most efficiently be performed by developing a software which will report the matching discrepancy.

- ⌘ Making backup copy of the GIS attribute data file: This is very important restart data integration process, if the GIS attribute data file becom corrupted due to any mistake in the data integration process.

- ⌘ Manual data entry: Manual data entry can be performed from DBMS da file to GIS attribute data file using either the ARC/INFO or DBMS da entry sub-system. However, as mentioned earlier, manual data entry is easi in DBMS than in ARC/INFO.

- ⌘ Automatic data integration: This may be performed by merging the DBM data file(s) with the GIS attribute data file(s). In the merging process, ca should be taken for the following aspects:

  - ◇ There should not be any item in both the data files, except relate ite having same name but different data. In this case, this item in the join fi will not be added to the GIS attribute file.

  - ◇ The 'start item' should not be placed within the standard items of the G attribute file. In this case, if the coverage topology is to be reconstruct for any reason, the name of standard item(s) will be changed according its position in the file. Apart from the naming discrepancy, this m result in loss of data.

  - ◇ Merging of databases should be performed using ARC/INFO comma discussed earlier. Merging command in DBMS software, like DBASE, w drop the record(s) from the GIS attribute files for which no match of t relate item exists. In many applications, attribute data may be collected or sample basis. For example, in the first round survey under the pove monitoring system of BBS, data for 75 thanas were collected. In this ca the join file contains only 75 records. On the other hand, a thana G coverage contains records for all the thanas in Bangladesh. Under t condition, if merging is performed using DBASE command, the result PAT will contain records for only 75 thanas. This will create massive m assignment of the records in PAT with the polygon features (thanas) in t coverage.

C I R D A P

The Centre on Integrated Rural Development for Asia and the Pacific (CIRDAP) is a regional, inter-governmental, autonomous institution, established in July 1979 at the initiative of the countries of the Asia-Pacific Region and the Food and Agriculture Organization (FAO) of the United Nations with support from several other UN bodies and donors. Its member countries include Afghanistan, Bangladesh (Host State), India, Indonesia, Lao PDR, Malaysia, Myanmar, Nepal, Pakistan, the Philippines, Sri Lanka, Thailand and Vietnam.

The main objectives of CIRDAP are to (i) assist national action; (ii) promote regional cooperation, and (iii) act as a servicing institution for its member countries for promotion of integrated rural development through research, action research, pilot project, training and information dissemination. Amelioration of rural poverty in the Asia-Pacific region has been the prime concern of CIRDAP. The Centre is committed to the WCARRD Follow-up Programmes. The programme priorities of CIRDAP are set under four areas of concern: (1) agrarian development; (2) institutional/ infrastructural development; (3) resource development including human resources; and (4) employment.

Operating through designated Contact Ministries and Link Institutions in member countries, CIRDAP promotes technical cooperation among nations of the region. It plays a supplementary and reinforcing role in supporting and furthering the effectiveness of integrated rural development programmes in the Asia-Pacific region.