Overview of Digital Maps in the WHEAT Geographic Information System (GIS)

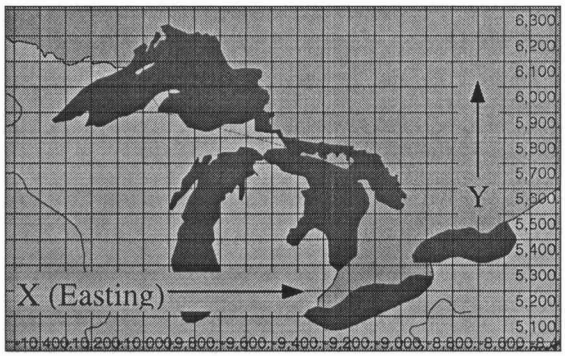
WHEAT is a public-domain, MSWindows-based Geographic Information System (GIS) developed for natural resource management and exploration. The centerpiece of WHEAT is EMAP, the Electronic Mapping program, whose primary functions are to display maps and perform spatial queries. This document provides a conceptual overview of how WHEAT "thinks" of maps and what is needed to create digital maps for use in WHEAT.

In WHEAT, a map is a view of a set of overlaid themes. A theme is a set of geographic features of the same geometric type (symbols, lines, polygons, textlabels...) that come from a table in a database: a theme includes coordinates of the features as well as plot-style instructions and serial numbers. A map includes not only what is shown but also where is shown (i.e., what area is visible). A theme definition instructs EMAP how to assemble a theme from information in a database.

Gregory W. Pouch
Kansas Geological Survey

MAPS AND THEMES

The maps shown below consist of several themes. The area shown is the Great Lakes Region of the United States and Canada. Each feature (an individual thing, like Canada, the Mississippi River, or Lake Ontario) belongs to a category (like countries, rivers, or lakes) depicted as a theme (like LandAreasPolygons, DrainageLines, or DrainageAreas).



FEATURES AND PROPERTIES

A feature has properties that depend on its category. For example, a lake has a name, a location,, a volume of water, a salinity, a water-surface elevation, a temperature, and so forth. A country has a name, a population, a location, a legal system, a head of state, citizens, residents, and so forth.

Some properties are easily represented in a database, while others are not. Some of these properties are relevant to a particular purpose, while others are not. Some

properties have a unique value, while others vary with time or have many values. (e.g., temperature (time-variant), language (possible multiple values).

The database author designs a database for a particular purpose by organizing the features into tables and queries so that the database user can get, and perhaps edit, information for a particular task. The database author is usually a user as well.

LOCATIONS AND COORDINATES

To produce maps with WHEAT, feature-locations must be stored in a form WHEAT can understand: as map coordinates. Map coordinates specify a location relative to a set of numbered lines forming a grid, with the X coordinate specifying distance east of the grid origin (easting) and the Y coordinate specifying distance north of the grid origin (northing). {Addresses in cities are often like this, where the street has a number indicating easting or northing, and an address indicates northing or easting: this system is often used on roads in rural areas that were surveyed on a grid-like system. Locations in atlases and road maps are often specified in a nearly-identical form, but with letters used for one coordinate and numbers for the other.}

For WHEAT to work well–show features properly and produce meaningful results to spatial queries– the grid system must meet an additional condition: distances in X and Y must be equivalent. That is, one unit in X must be the same distance as one unit in Y, or nearly the same, due to the curvature of the earth. Examples of systems that meet all of WHEAT's conditions include UTM (Universal Transverse Mercator) coordinates and State Plane Coordinates. In mathematical terms, WHEAT works with any cartesian, right-handed, isotropic coordinate system. {Although longitude-latitude coordinates can be converted to coordinates suitable for mapping, X and Y distances are equivalent only at the equator. Address systems are usually not suitable, because each unit in X is usually not the same distance as all other units (that is, not all blocks or sections are the same size).}

Feature coordinates can come from surveying or from measuring locations on a map. Measuring map locations can be done manually by reading coordinates or using a digitizer or scanner. EMAP can be used to "measure" map locations by positioning the cursor on a map location and recording the coordinates. Surveying can include use of a GPS (Global Positioning System receiver), traditional distance and angle measurements, and determining locations on airphotos or satellite images.

For points, WHEAT gets coordinates from two separate fields. For lines and polygons, WHEAT gets coordinates from a single field containing a chain of coordinates, a special data format known as an XYChain. For tiles (rectangles oriented parallel to the coordinate axes), WHEAT uses data from four separate fields describing the bounds in X and Y. WHEAT uses Microsoft Access format databases, for both locations and other properties. WHEAT includes an assortment of programs that convert to/from XYChain format. Most other data manipulation should be done in MSAccess or other relational database management software.

DATABASES, TABLES, AND QUERIES

A relational database stores information on a topic, such as water users or schools in tables. A table is an organized set of information about a features in a category. Each feature is described in a record, and each property is contained in a field in the table. In printed tables, records correspond to rows and fields correspond to columns. A field has a particular data type.

A table concerning schools might have records for each school in a county, and the fields might include the name (as text), date founded (as date), operating budget (as a floating point number), date closed (as date), number of buses owned (as integer number), principal's name (as text), and number of pupils (as integer number). Some of the fields listed, such as date closed and number of buses, will not occur for each individual school. In these cases, a relational database stores a null, indicating that no information is available.

While a relational database *stores* information in tables, it can retrieve information as virtual tables called queries. A query is the answer to a question asked about the datatable, such as show all schools with an operating budget above $500,000. A query can return data from several related tables by joining them. For example, suppose we had another table containing information about students. It might contain student's name, age, address, and name of school attended. We could then find out who a student's principal is by designing a query which relates the school name field in the student table to the name field in the school table. Queries can be used to combine information from several tables, select subsets of records, select subsets of fields, re-name fields in the virtual table, or any combination of these.

By allowing data to be stored in a single location (such as the principal's name in the school field) and related at a later time, relational databases can reduce storage requirements, and allow the updating of data to be considerably easier than if each record contained all data that might ever be needed.

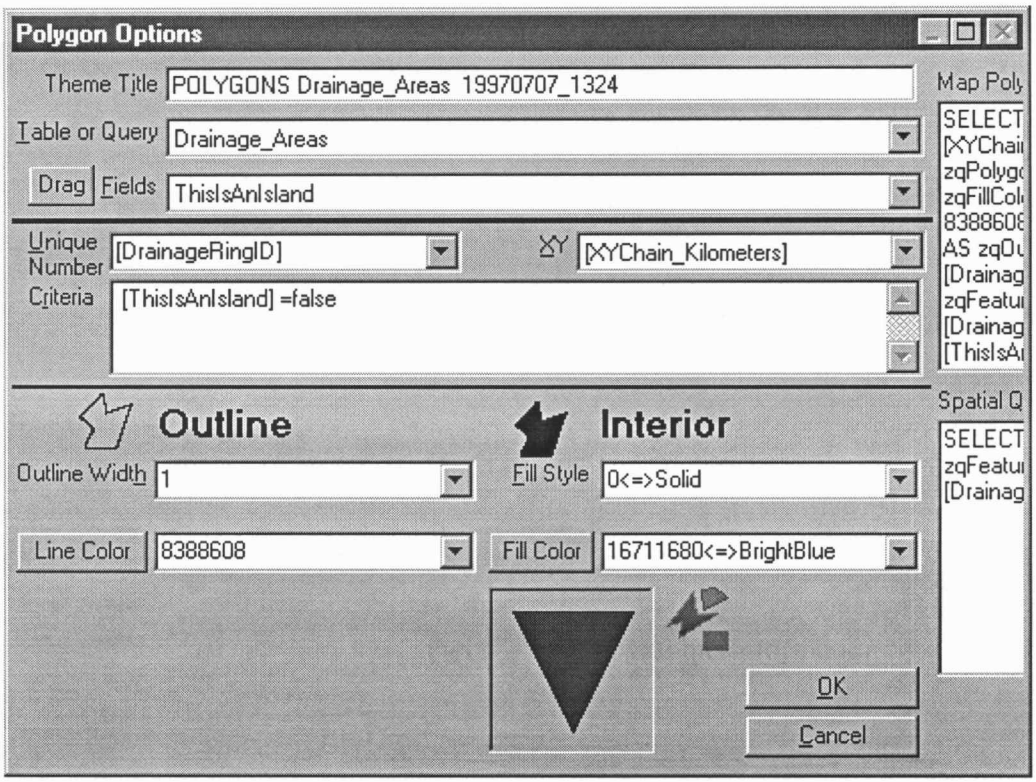
THEME DEFINITIONS

A theme definition tells EMAP how to load a theme from the database. You design a theme using dialog boxes that let you select options from dropdown lists, and you can store them in the database to use later.

The main part of a theme definition is an SQL statement that tells EMAP how to build a theme, including locations, serial numbers, plot styles, and any limitations on which data to load. (SQL is a database language.) This lets the theme author decide what to display and how to display, as well as control what information spatial queries retrieve. The SQL statement is built by the program based on user choices, and users don't need to use SQL or know about SQL to design themes, but it does help to know what is happening behind the scenes.

A theme definition includes a name, a serial number, author name, date of modification, geometric type, and an SQL statement used for retrieving information on features found by a spatial query.

This example shows how designing a theme looks to the user,



how it looks on the screen after it has been loaded (zoomed in to the Great Lakes),

and part of a theme definition, showing the SQL statements. (Users do not usually work with SQL statements directly.)

Name	Type	MapSQL	SpatialQuerySQL
LandAreas	Polygons	SELECT DISTINCTROW [XYChain_Kilometers] AS zqPolygonOutline , [CountryColorRANDOM] AS zqFillColor , 0 AS zqFillStyle , 0 AS zqOutlineColor , 1 AS zqOutlineWidth , [LandAreaSerialNumber] AS zqFeatureID FROM [LandAreas KM QDF] ORDER BY [CountryColorRANDOM] ;	SELECT *, [LandAreaSerialNumber] AS zqFeatureID FROM [LandAreas KM QDF]

VIEWBOUNDS

A map covers a particular area. In EMAP, the area covered by a map is always a rectangle oriented parallel to the coordinate axes (a tile) known as the ViewBounds. The ViewBounds can be adjusted by the user by zooming or shifting the view area. EMAP adjusts the ViewBounds so that screen distances are proportional to ground distances (it keeps the scaling isotropic).

SPATIAL QUERIES

A spatial query is a search through a theme for features meeting some spatial criterion, such as finding the river closest to a given a point, finding all wells within a five-mile radius of a given a point, or finding what geologic formation underlies a given a point. You can perform spatial queries interactively or in batches. Interactive spatial queries are most common: you use the mouse to point or enter coordinates and see results immediately.

SUMMARY

In WHEAT, a map is composed of a set of themes. Each theme has a particular geometric type and comes from one table or query in a database. A theme contains one or more features, each with map coordinates. The area seen on a map is defined by the ViewBounds. A spatial query allows the analyst (you) to search a theme for features matching a spatial criterion.

Kansas Geological Survey
Open-file Report

Disclaimer

The Kansas Geological Survey does not guarantee this document to be free from errors or inaccuracies and disclaims any responsibility or liability for interpretations based on data used in the production of this document or decisions based thereon. This report is intended to make results of research available at the earliest possible date, but is not intended to constitute final or formal publication.