GST 101: Introduction to Geospatial Technology Lab Series

Lab 5: Creating Geospatial Data

Document Version: 2013-07-30

Organization: Del Mar College
Author: Richard Smith

The development of this document is funded by the Department of Labor (DOL) Trade Adjustment Assistance Community College and Career Training (TAACCCT) Grant No. TC-22525-11-60-A-48; The National Information Security, Geospatial Technologies Consortium (NISGTC) is an entity of Collin College of Texas, Bellevue College of Washington, Bunker Hill Community College of Massachusetts, Del Mar College of Texas, Moraine Valley Community College of Illinois, Rio Salado College of Arizona, and Salt Lake Community College of Utah.

Contents

Introduction

This lab is part of a series of lab exercises designed through a grant initiative by the National Information, Security & Geospatial Technologies Consortium (NISGTC), funded by the United States Department of Labor in partnership with the Department of Education under the Trade Adjustment Assistance Community College and Career Training Grant Program (TAACCCT).

In this lab, students will learn how to digitize information from a source document, into a digital geospatial data model.

Your instructor may require that you provide screen captures and/or exported files. Please check with your instructor for the requirements specific to your class.

This lab includes the following tasks:

1. Creating a new geodatabase
2. Transforming a coordinate system of source data
3. Heads-up digitizing
4. Adding attribute information
5. Creating a layout

Objective: Digitize Information from a Scanned Hard-Copy Source

While there is a large amount of digital information readily available to users of GIS, there is also much information that has not been converted to digital format. It is possible to extract the information from hardcopy sources through a process called digitizing. In this lab, you will use heads-up digitizing to digitize counties, lakes, roads, and houses from a scanned map. This will be accomplished through a five-step digitizing process:

1. Create a data set to store the digitized data
2. Load source in the program
3. Register source
4. Digitize
5. Save

In this lab, you will complete steps 1, 3, 4, and 5. The hardcopy data sources have already been loaded for you, but have not been referenced to a coordinate system yet.

Lab Settings

Required Virtual Machines and Applications

Windows Machine User Account	Train
Windows Machine User Password	##########

1 Creating a New Geodatabase

The first step in digitizing is creating a data set to store the digitized information. In this task, you will create four feature classes inside a file geodatabase to store the following digitized features in an appropriate vector type.

Feature Class Name	Feature Type
Counties	Polygon
Lakes	Polygon
Roads	Line
Houses	Point

1. Copy the *Lab 5* Data folder from the *Shared_Drive\GST101* into the folder *C:\GST 101\Lab 5*
2. Start ArcCatalog, browse to the lab folder. Click **Connect to Folder** button on the Standard Toolbar; navigate to your lab folder. Then click **OK**. Your lab folder now becomes an entry in the catalog tree. Highlight it so that you lab folder is open in the display area.
3. You need to create four new (and empty) feature classes, corresponding to the county boundaries, lakes, roads, and houses. In ArcCatalog. **right click** in the display area, choose **New … File Geodatabase**. Name the geodatabase *myCity.gdb*.
4. **Right-click** on the newly created geodatabase, then choose **New … Feature Class**. Type **Counties** for the name of the new feature class. In the dropdown list next to Feature Type, select **Polygon Features**. Click **Next**.
5. The source data falls within UTM Zone 17N, so you have to set the coordinate system for the feature class to match that coordinate system. Expand the **Projected Coordinate Systems** folder, then navigate to **UTM →NAD 1983 → NAD 1983 UTM Zone 17N**. Click **Next**.
6. Leave the XY tolerance to its default value. Click **Next**. We will not specify any configuration options, so click **Next**. We will not add any fields, so click **Finish**.
7. **Repeat the same procedure (steps 4 to 6)** to create the other three feature classes. Make sure that correct feature types (as specified above) are selected. The coordinate systems of the three feature classes should be set the same as the first feature class.
8. When you have finished creating all 4 feature classes, **Close ArcCatalog.**

2 Transform Coordinate System of Source Data

Now that you have created empty feature classes to store the digitized information, you will perform a coordinate transformation (also known as registration or rectification) on the source data set so that it is in an Earth-based coordinate system. In this case, the coordinate system will be NAD 1983 UTM Zone 17 North.

1. **Start ArcMap**, and open a new, empty map document.
2. **Add *myCity.bmp*** to the map document (either drag from the Catalog tab, or use the 'Add Data' button).
3. An Unknown Spatial Reference warning will appear. This is because no spatial reference information has been assigned to the bitmap (but will be soon). Dismiss this dialog.
4. **Save your current workspace** inside the lab directory. (From the File Menu, choose Save). In ArcGIS, we call such a workspace file an **ArcMap document (*.mxd)**. Navigate to your lab folder and name the saved file ***Digitizing.mxd***. You need to save the workspace periodically. This is an important precaution to take to avoid losing your work accidentally.

Now you have a scanned image of features on screen to digitize from. You will digitize selected features on the image (raster) and save the features into corresponding feature datasets (vector). However, the scanned image is not on the UTM coordinate system, so we would need to **transform** the current coordinate system of the image to the UTM 17N coordinate system. We will use the **Georeferencing** tools in ArcGIS to achieve this.

5. Choose **Customize → Toolbars → Georeferencing**. (alternatively, right click on the gray area around your toolbars and add the georeferencing toolbar).
6. Place the georeferencing toolbar where you want. Click the **View Link Table** button to open a table. **Turn off** (uncheck) **the Auto Adjust** option in the table. Now let's check out the 'Add Control Points' button with the green and red plus signs  . Hover your mouse over the button and note what the tool tip has to say.
7. **Click the Add Control Points button**.
8. **Move the cursor** on the image to the **point marked "1"**. Observe the current X and Y coordinates of the location. These coordinates represent the location in the data source's (not Earth-based) coordinate system.
9. **Click** with the cursor on the center of the point as precisely as possible.
10. **Right-click** and select **Input X and Y** from the resulting menu. Put 201000 for X (Easting) and 2455000 for Y (Northing). **Click OK**. You just added a control point. Look at the link table to be sure that both the source and map coordinates have been recorded.
11. **Repeat the same procedure in steps 8-10** and add three additional control points. If at any time you get in trouble with any control point, highlight the control point entry in the Link table, and then click the **Delete Link** button  on the top left corner in the table.
    The UTM coordinates for the four control points are provided in the table below:

	X	Y
Point 1	201000	2455000
Point 2	201000	2461500
Point 3	211400	2461500
Point 4	211400	2455000

12. Once you have all four control points are input properly, **check the Auto Adjust** box in the Link table. The map (image) coordinates are transformed (moved to a different coordinate system) and your map will seem to have disappeared. Note that the Link Table shows Total RMS error – probably of a few meters (it may show no error value). It is the measure of how much error the transformation may bring. If the figure is larger than 10 meters, you need to redo the transformation by re-entering the control points. If not, you have transformed the image in the current workspace. To view where the image went to, click the "**Full Extent**" button.
13. **In the Link table, click the Save button**. Save the control points into your lab folder as **ControlPoints.TXT** file. We do this so that we do not have to re-enter these coordinates next time when you need to do the same transformation.

The transformation is only effective within the session (i.e. the original image will remain intact). So, each time you close the image and re-load it to ArcMap, you need to repeat the transformation. You can do so by loading the control point file in the Link Table window.

An alternative method is to export the transformed data into a new raster dataset by clicking Georeferencing → Rectify on the Georeferencing toolbar and saving the rectified source data into a new file. Doing this will make the rectification permanent. The last step would be defining the projection of the newly saved file to NAD 1983 UTM Zone 17 North.

3 Heads-up Digitizing

1. Add the four empty feature classes you created in Task 1 into ArcMap using Add Data.
2. In ArcMap, display the **Editor Toolbar** if it is not displayed.
3. It is always a best practice to edit with a defined coordinate system for the data frame. In the Table of Contents, **right-click** on the data frame named **Layers**. Then open **Properties**.
4. Click the **Coordinate System** tab. Expand the **Layers** folder, then click on **NAD_1983_UTM_Zone_17N**. Click **OK**.
5. On the Editor toolbar, click the **Editor** menu and click **Start Editing**.
6. On the Editor toolbar, click **Editor → Editing Windows → Create Features**.

For each feature class, you will trace the boundaries of corresponding features in the scanned image to create vector features.

3.1 Digitizing Lakes

1. In the **Create Features** table symbol, **click on Lakes**.
2. Move the cursor to any point on a lake's boundary and click (to record the point)**.** Circumnavigate the lake and click at every turning location to record a change in the direction. Bring the cursor back to the starting point and double click (double click means to it is the ending point of the polygon). Repeat this process for the other lake.

It is impossible to trace the lakes perfectly. Instead, record a reasonable number of points to represent the lakes accurately, but without recording a large number of points.

3. When you are done for the lake layer, click the **Editor** menu and click **Save Edits**. We will follow this routine when you create the other layers.
4. The last lake you digitize may be highlighted because this polygon is now selected by ArcMap. If this is the case, click **Selection menu → Clear Selected Features**.

3.2 Digitizing Houses

1. Click on **House** in the **Create Features** window.
2. Move the cursor to each house and click. Because it is a Point feature class, this one click is enough for a feature. Repeat this for all houses. When you are done, save edits. Clear selected features, if any.

3.3 Set the Snapping Environment and Use Topology

Snapping is important as it ensures that features snap to each other, and that dangles, overshoots, gaps, or slivers are avoided. This will help to reduce errors that cause confusion about topological relations.

To specify the snapping environment, two things need to be defined: 1. the types of features (geometries) to be snapped to; and 2. the snapping tolerance (distance).

1. To specify the snapping geometries, click **Editor → Snapping → Snapping Toolbar.** Here for each layer, you can toggle the boxes  to determine the types of features they will "snap" to – e.g. points, ends, vertices, or edges of existing features.
   a. When digitizing counties, enable all boxes.
   b. When digitizing roads, enable end, vertex, and edge only.
2. Snapping tolerance is another important control in the environment. The **Snap Tolerance** tool will allow you to interactively define a snapping distance. Add the **Snap Tolerance** tool to the editing toolbar. To do this, select **Customize → Customize Mode**. Click the **Commands** tab. Find the **Editor** Category and the **Snap Tolerance** Command (shown in Figure 1).

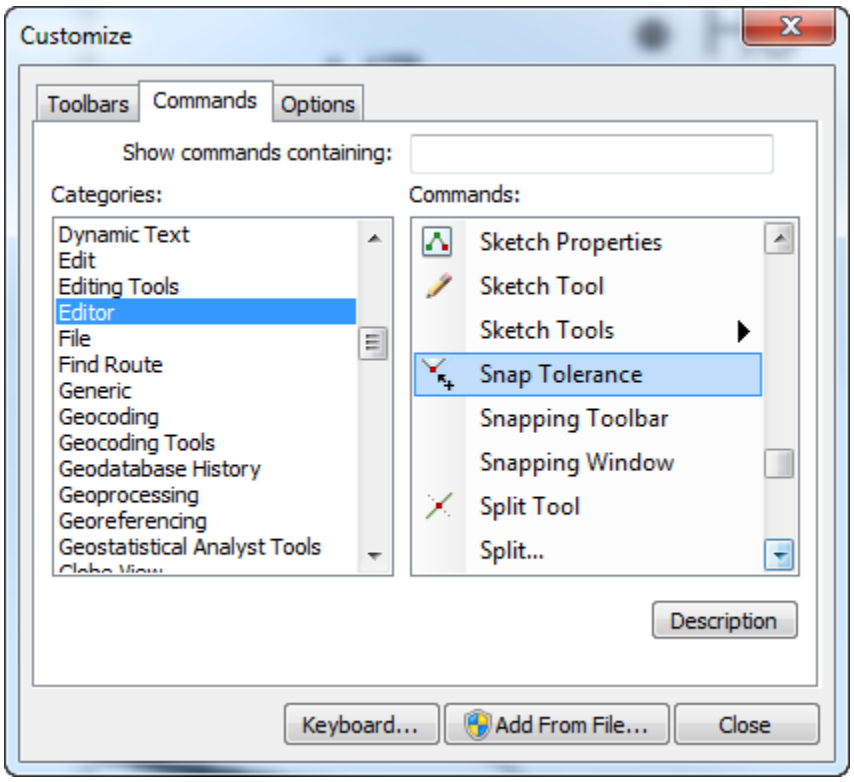

Figure 1: Snap Tolerance Tool

3. Click, hold, and drag the tool onto the Editor toolbar. Close the Customize dialog. Click the **Snap Tolerance** tool, move the cursor to the map area, click and drag a distance to define it as your snapping tolerance. When features are added that come within this distance of existing features, snapping will occur. **Make the snapping tolerance fairly small, otherwise you will snap to features when too far away.** You can check to see the current snapping tolerance by looking at **Snapping > Options** from the Snapping toolbar.

It is important to set the snapping tolerance before you start editing line or polygon features. You can always reset snapping tolerances at any time, based on the precision you need.

For this lab, the recommended snapping tolerance is about 10 pixels.

3.4 Digitize the County Layer

1. Click on **Counties** in the **Create Features** window. Check to make sure all boxes in the Snapping Toolbar are enabled
2. You will first draw the boundary of any of the three counties. Move the cursor to a corner point and click**.** This starts the polygon. Move to every turning point or intersection point in sequence (clockwise or anti-clockwise), click. You will need to digitize the county boundaries that frame the map as well. If necessary, you may want to use the Zoom and Pan tools to click precisely. If you use the Zoom or Pan tools, you will need to re-select the Counties button on the "Create Features" window.

Be sure to digitize at each intersection point because it will be used to "snap" (taken as the identical point) when you digitize other polygons that share the point.

3. At the last point, right before returning to the starting point, double click. This will close the polygon. When you are done, click the **Full Extent** tool. You should now see that a whole polygon has been created.

4. The next step is to add a new polygon with the auto-complete option (which aims to automatically detect and duplicate common edges between polygons). You should never attempt to digitize over existing edges, because it is likely you will be creating gaps or slivers. Auto-complete is part of the topology tools. To access these, select **Customize → Toolbars → Topology**.
5. From the **Topology** toolbar, click the **Select Topology** button  . Click the checkbox for **Counties**. In this editing session, this will be the only layer participating in the topology. Click **OK**.

6. On the Create Features window, in the Construction Tools section, click on the **Auto Complete Polygon** button. If you do not see the Construction Tools section, you may need to expand the window. See Figure 2 for an illustration of this step.

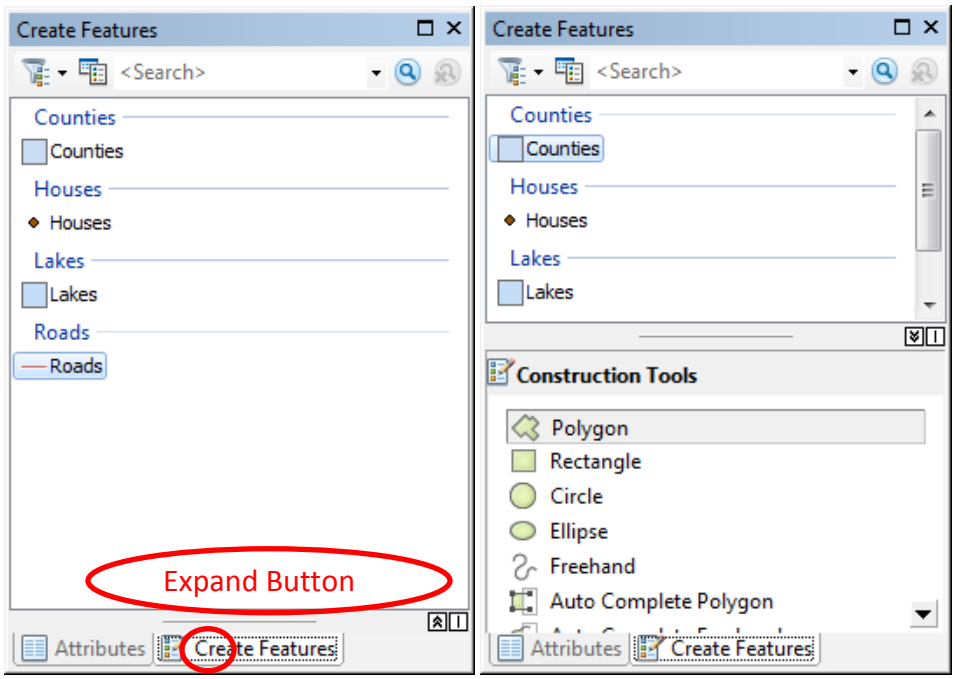

Figure 2: Finding the Auto Complete Polygon Tool

7. Before you start creating another polygon (county boundary), indentify the common edges that have already been created in other digitized polygons.
   a. First, **click on one end-node of the common edges**, and then move outside the existing polygon, **adding vertices as needed**.
   b. To finish the new adjacent polygon, move the cursor on the other end-node of the common edges, double click to complete the polygon. Do not attempt to return to the starting node of the polygon, as the common edges have already been identified and duplicated to be part of for the new polygon. See Figure 3 for clarification.

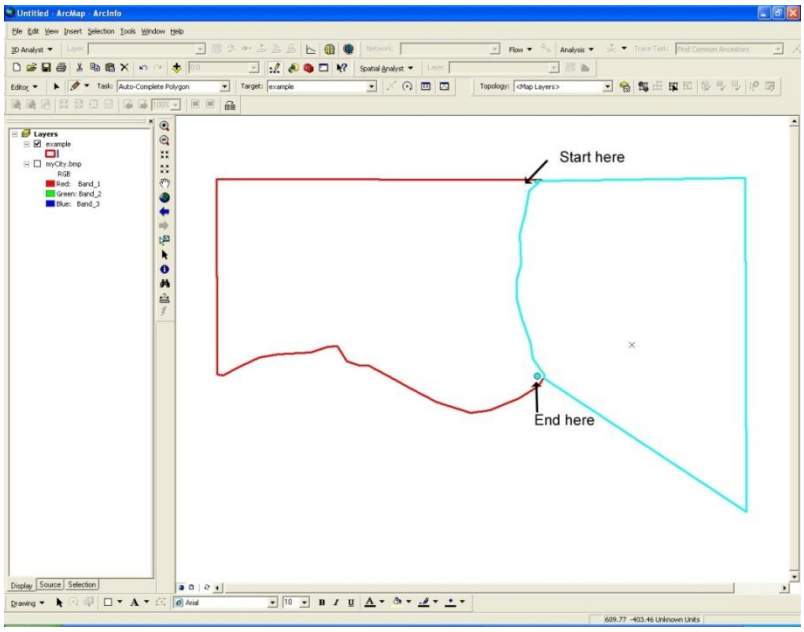

Figure 3: Start and End Nodes when Editing with Topology

8. Repeat this process for the third polygon. This time, there will be two common edges that are identified and duplicated.
9. When you are done with all three counties, click the **Editor** menu and click **Save Edits**.

3.5 Digitize the Roads Layer

1. Click on **Roads** in the **Create Features** window.
2. You may need to turn the visibility of the digitized counties layer off to see the Roads layer.
3. Set the snapping options (End,Vertex, and Edge only) and snapping tolerance. Be sure that the Line option is selected in Construction Tools.
4. **Move the cursor to where Highway 10 starts and click**. Travel along the road. Double **click when you reach the end of the road**. Because Highway 29 connects into Highway 10. The same as with Counties layer, you need to **capture all the intersection points** (for topology).
5. If the point where you click falls within a snapping tolerance of a boundary line in Counties, the point will be "snapped" onto the boundary line because we have enabled the Edge snapping.
6. Follow the same procedure to digitize **Highway 29**, remember to snap to the existing node at the intersection and snap to the boundary line at the end (or beginning).
7. Click the **Editor** menu and choose **Save Edits** when you are done.
8. We are done digitizing. Click **Editor → Stop Editing**

4 Adding Attribute Information

You will now add attribute information to each of these four layers (Counties, Lakes, Roads, and Houses). Attribute information (*e.g.* county name, area, population) describes the characteristics of the spatial features and thus is an integral part of a GIS. Without attribute information, it is not possible to differentiate between the lines that represent Highway 10 and Highway 29, nor can we carry out further GIS analyses. In ArcGIS, columns are referred to as Fields (or attributes) and rows are referred to as records (each correspond to a feature).

1. In the **Table of Contents**, right click on the **Counties** layer.
2. Click **Open Attribute Table**. You will see a table with 3 records. In the table, click the Table **Options** button  and click **Add Field**. (If the Add Field command is disabled, you may not have concluded the edit session. Click the **Edit** Menu and click **Stop Editing**). (If you get a Schema lock error, you need to close ArcCatalog if it is open.)
3. Type **County_Name** in the name textbox. Click the Dropdown arrow and select **Text**. **Change the Length property to 10**. Click **OK** to dismiss the **Add Field** dialog. You will see that the County_Name field has been added to the table.
4. **Add one other field** named **Population** and choose **Short Integer** for the type. Click OK
5. **Select the first county in the attribute table** (by clicking the gray tab at the left edge of the first record). The county in the Map Display area and the equivalent record in the table will now become highlighted (This means that this county and its record are selected). You have to examine the base map to see which county has been selected so that you can fill in the corresponding county name in the table.
6. Click the **Editor** menu and click **Start Editing**. Then come back to the table and enter the county name in the County_Name field and the population in the population field as follows. Repeat this until you have entered the names and population figures for all the counties. When you are done, clear the selected features.

County Name	Population
Oconee	30000
Green	20000
Clarke	10000

7. **Repeat the same process for the Lake features**. You will need to Stop Editing to Add Fields to the attribute table. For the lakes, just create and populate one field named **Lake_Name** (Text, Length 16).
8. **Repeat the same process for the Roads features**. Create two fields: **Road_Name** (Text, Length 16) and **Road_Class** (Short Integer) and populate them with the information below.

Road Name	Class
10	1
29	2

9. **Repeat the same process for House features**. Create two fields: **House_Name** (Text, Length 16), and **Household_Size** (Short Integer) and populate them with the information below.

House Name	Household Size
H1	2
H2	6
H3	4
H4	3
H5	8
H6	10
H7	6
H8	12

10. Click **Editor → Save Edits.**
11. Click **Editor → Stop Editing** as we are finished entering attributes.

5 Creating a Layout

Now that we have created a spatial database and digitized features, we are ready to create a map. In addition to the actual map, most maps have a **legend**, **scale bar**, **north arrow**, and **descriptive text**. We can create such maps using the **Layout** View of ArcMap. Before we create the layout, however, we will change the colors and symbols for the layers so that they are more in line with the map display standards commonly used.

1. In the Table of Contents, right click the **rectangular symbol** below counties layer. In the Symbol Selector dialog, click on the rectangle with a white background (Hollow). (You can also open the Symbol Selector by right clicking the counties layer in the Table of Contents, select **Properties**, select **Symbology** tab, then click **symbol button** in the Symbol Frame).
2. Right click the **Lakes** layer. Select **Properties,** select the **Symbology** tab**,** then click **Symbol** button in the Symbol Frame. In the **Symbol Selector window**, click on the **Lake** rectangle (the second one in the second row). Click OK.
3. **Click the Labels tab**. Under label field, select **Lake_Name**.
4. Check **Label features in this layer**. Click OK.
5. Open the Properties dialog for **Roads** layer. Click **Symbology** tab. In the show area, click **Categories**, then click **Unique Values.** In the Value Field Drop-down list, Select **Class** (This means that we are going to use a unique symbol for each road class). At the bottom of the dialog, click **Add All Values**. In the symbol column, **uncheck** the check box next to <all other values>. **Click** on the **Label** for value **1** and change it to **Hwy 10.** Double Click the **Symbol** for Value 1 and select the **symbol with "highway" (the first one).** Change the **width** to 1.5. Click **OK.**
6. **Repeat the same procedure for Value 2**. (label **Hwy 29**). Make the symbol dot-dot-dash with "Boundary, State" text below, Red color and Size 1.5). Click **OK**.

7. Open the Properties dialog for **houses** layer. Click **Symbology tab**. In the show area, click **Quantities**, and then choose **Graduated Symbol,** and for **Value Field**, choose **Household_Size** (the size of the house symbol will depend on the Household_Size). **Do not change the Label**. Click the **Template** button on the right of the dialog to access Symbol Selector window. Change the **Symbol** to a circle with a dot inside it (the symbol with the text "Circle 3" below). **Choose a color**. Then click **OK.** Click **OK** again to close the properties.

Now that you have a symbolized and labeled map, you will create a map layout and add additional map elements.

8. Turn off the visibility of the source data set (myCity.bmp).
9. From the main menu bar, **click View → Layout View**
10. From the mail menu bar, **click File → Page and Print Setup**. In the Page and Print Setup dialog, click the radio button next to either Landscape or Portrait depending on the orientation you want. **Click OK** to set the orientation.
11. Using the **Insert** menu on the main menu bar, insert the following map elements:
    a. Title
    b. Scale bar
    c. North arrow
    d. Legend
    e. Text (for metadata)

Conclusion

In this lab, you have successfully digitized information using the five-step digitizing process. Additionally, you have recreated the original source data (scanned as a raster) in the digitized vector format. Digitizing can be a time-consuming and tedious process, but can yield useful geographic information.

Discussion Questions

1. What other vector types (point/line/polygon) would be appropriate for digitizing a road? In which instances would you use one vector type over another?

2. What are all of the possible field types? Explain what each field type contains, and provide an example of a valid entry in the field.

3. The note at the end of the Task 2 mentioned defining the projection of the rectified source data set. Open ArcToolbox and expand Data Management Tools → Projections and Transformations. Open the Define Projection tool. What does this tool do? Why do you think it is important to define a projection even though the source data set has already been registered?

4. In this lab, fields were created by opening the attribute table, then clicking Table Options → Add Field. Where else in this lab could we have created fields? List the steps required to do this.