

Chapter 10. Geocoding

Objectives

- Understanding the geocoding process
- Knowing the different styles of geocoding
- Setting up an address locator service
- Geocoding using several styles of locators

Mastering the Concepts

GIS Concepts

What is geocoding?

As everyone knows, a street address tells a postal worker where to deliver mail. It contains a type of spatial information. However, the postal worker requires additional knowledge to get the letter to the house. Unlike a latitude-longitude point, which uniquely locates a house on the globe, a street address requires information about the city and state, the location of the street, and the sequence of house numbers. A newly hired postal worker delivering a piece of mail in a strange city would need a good map in addition to the address.

Likewise, geocoding combines map information with street addresses in order to locate a point uniquely. It enables someone to convert a list of addresses into points on a map (Fig. 10.1). Geocoding has many uses.

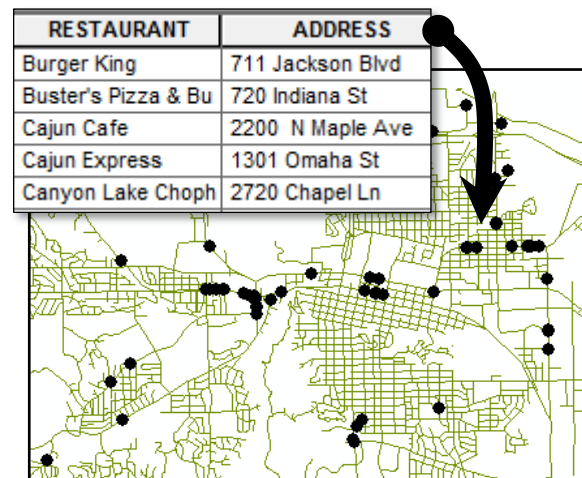


Fig. 10.1. Geocoding can convert restaurant addresses in a table to points on a map.

911 addressing

One critical use being developed by many governmental agencies concerns finding street addresses given over the phone during an emergency call in order to quickly route response vehicles to the location. Routing of vehicles is accomplished by providing a map or a global positioning system (GPS) location in vehicles equipped with GPS units. Because lives are at stake, this application requires highly accurate address information, and it is a challenging application to implement.

Driving directions

Many people have gone to one of the popular web sites, such as MapQuest or Google Maps, to find directions from a starting address to a destination. The results include step-by-step driving directions and a map showing the route. Other people may have tried the GPS systems installed in cars, such as OnStar or Magellan, which guide people to their destination from their current position as identified by GPS. Both of these applications are built upon geocoding technology.

Crime analysis

Police records of crime incidents include the address, and these locations can be converted to points on a map through geocoding. Analysts could use these data to help identify “hot spots” in order to help plan patrol routes for officers or to target efforts to develop neighborhood watch programs. Police departments can glean information about spatial patterns for different types of crime or target the time of day that particular crimes tend to occur. Such information can be critical in developing strategies to reduce crime and to protect residents and businesses.

Marketing

Many companies collect address information about their customers as a regular part of conducting business. Converting addresses to locations and analyzing their distribution might help a company provide better service by opening new stores in underserved areas. Mail-order businesses can locate areas with low sales for targeted marketing efforts.

Sales and service companies can better plan where to place agents to serve their clients. Recently, a study helped a university in South Dakota to understand the national distribution of its students, and it is helping them to plan marketing strategies to increase national enrollment (Fig. 10.2).

A whole host of business applications is available once information about customer locations is known, and many companies now routinely use GIS in their operations.

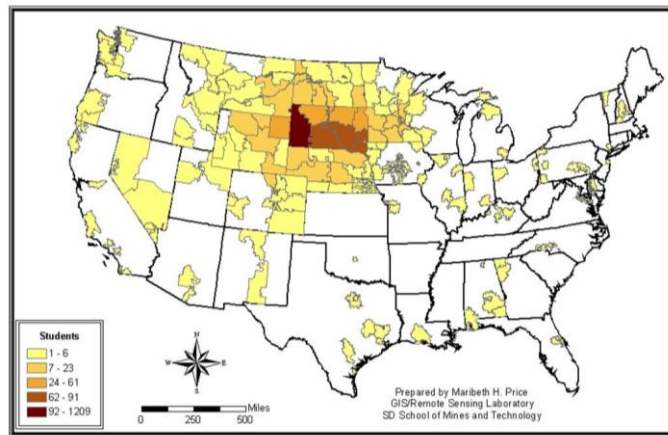


Fig. 10.2. Distribution of students attending the South Dakota School of Mines and Technology in 2004

Public Land Survey System locations

Many parts of the United States, particularly out west, utilize the Public Land Survey System, or PLSS, as a means of locating points of interest. In the days before GPS units, field workers often located themselves on a topographic map and marked field notes or sample locations by means of the township, range, section, and section quarters, rather than units of latitude or longitude. Property descriptions also use this format, and they are still common terms in western agriculture.

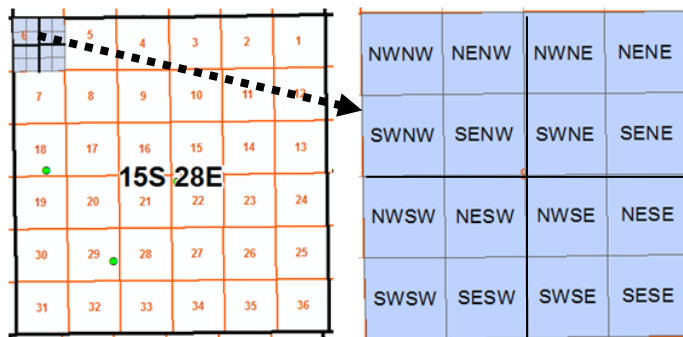


Fig. 10.3. A township, sections, and quarter-quarter sections

The basic unit of PLSS is the township, a 6-mile by 6-mile square divided into 36 one-mile square sections (Fig. 10.3). The township is referenced by its position relative to a surveyed east-west baseline and north-south meridian, defined for the state or region. Figure 10.3 shows a township that is the 15th one south of the baseline and the 28th one east of the meridian for Oregon. Notice that the sections are numbered in alternate rows: 1–6 from east to west on the first row and then 7–12 from west to east on the second. The section in the upper-left corner has been

divided once into quarters (quarter-sections NW, NE, SW, and SE), and those quarters divided into quarters again (quarter-quarter sections also designated NW, NE, SW, SE). The full designation for the upper-left quarter-quarter section is T15S, R28E, S6, NWNW. The quarter-quarter section designation comes first, then the quarter-section.

A GIS practitioner may need to convert PLSS locations from old index cards or field notebooks or legal descriptions to points in a geodatabase. Geocoding provides a quick and straightforward method to convert descriptions to points once the descriptions are entered in a spreadsheet.

However, the geometric accuracy of these methods is much lower than even the cheapest GPS. A section is a mile on the side, a quarter section is a quarter-mile, and a quarter-quarter section, the smallest unit typically used, is a 16th of a mile, or 330 feet. Geocoding finds the right polygon unit and places a point at its center. So if you locate a fossil or a rock sample to the quarter-quarter section, you still have a large area to hunt through for its source (40 acres!).

TIP: If you are not already familiar with PLSS history and terminology, a practical review can be found at [Wikipedia.org](https://en.wikipedia.org).

How does geocoding work?

Just as a stranger requires a map to find an address in a city, the computer requires map data in order to turn an address into a location. This spatial data is called the **reference layer**. Geocoding typically uses a street layer composed of polylines with specific attributes indicating each street's name and the address ranges present on that street. However, other types of geocoding can use point or polygon reference data, such as matching store franchises to points representing cities or matching a list of owner addresses to parcel polygons.

Although humans automatically recognize a variety of address styles, a computer requires specialized software to interpret addresses. An address might be written 123 Maple Street, or 123 Maple St, or 123 N. Maple St, or even 123 Maple, and the postmaster would have no difficulty finding the house. To computers, however, each of these addresses is different. Geocoding programs must interpret different styles of addresses and manage typographical and other errors, such as someone entering Maple Rd instead of Maple St. They accomplish this task by dividing the address into different components and then scoring how well they match the reference components. A high score (> 80) indicates a good match, a medium score (60–80) indicates a tentative match, and a low score (< 60) suggests a poor match.

To interpret an address, a geocoder begins by parsing the address into individual fragments of information. The address "235 W. Main Str" would be divided into four pieces: the address number (235), the direction prefix (W.), the street name (Main), and the street type designation (Str). After separating the components, each fragment is standardized so that it has the best chance to match the information in the reference layer. All letters may be capitalized (MAIN), periods removed (W), and common abbreviations reduced to a single standard (ST). This process is called **address standardization**. This step is necessary because a computer, confronted with the problem of comparing (Str) in the address table to (ST) in the reference layer, will conclude that the two do not match. A human would hardly notice the difference, but computers can be tiresomely literal. An address may include these components:

Prefix direction. A direction indicator that comes before the street name, such as *E Maple St* or *North Main St*. Some towns and cities use prefixes for directions; others use suffix directions (e.g., *Main St North*). This part may have any of the standard directions, N, E, S, W, NE, SW, SE, and NW. The complete spellings North, East, South, and West are also recognized.

Prefix type. An indicator of the road type that comes before the street name, such as *Highway 85*. In this case, 85 is the street “name” and Highway, or Hwy, is the prefix type.

Number. The address number, such as *123 Maple St*

Street name. The name of the street, such as *123 Maple St* or *Mount Rushmore Road*

Street type. A type indicator that comes after the street name, such as *Main St* or *Maple Ave*. The software recognizes a variety of street types (Street, Lane, Place, Road, etc.) and understands most common abbreviations, such as St, St., Rd, Av, and Ave.

Suffix direction. A direction indicator that comes after the street name, such as *Florman St E* or *87th St North*. This part may have the same values as the prefix direction.

Zip code. The standard U.S. zip code or, optionally, the Zip+4 nine-digit codes

Table 10.1. Examples of address standardization.

Address	Prefix dir	Prefix type	Number	Street	Type	Suffix dir
123 Maple Road			123	MAPLE	RD	
15 Center St East			15	CENTER	ST	E
314 Hwy 85 N		HWY	314	85		N
234 E. St. Patrick St	E		234	SAINT PATRICK	ST	

Table 10.1 shows some examples of addresses and their standardizations. Note that, in addition to separating the address into components, the text is converted to capital letters to facilitate comparisons in case-sensitive databases. The complete words in the directions and types are also converted to standard abbreviations, so that Road becomes RD and the periods are omitted.

Once the address is standardized, the components are compared to attributes in a spatial data layer, such as a line shapefile containing the city streets. This reference layer must contain certain attribute fields to compare with the standardized components. The score for each component match is averaged to give an overall match score for the address, in terms of percentage. A score of 100 indicates a perfect match. A zero indicates that none of the components match. Usually, a score of 80 or better is considered a good match, and a point will be created for the address. When the score is lower than 60, the geocoding software looks for other possible candidates. Typically, candidates have scores above 30. The user can examine the candidates individually to find out if any of the candidates might provide a reasonable match.

Figure 10.4 shows an example of a street feature class used for geocoding. Notice that most of the fields described are present, although the names are not the same. Also notice that, instead of a single address, the street has two fields (LEFT_FRADD and LEFT_TOADD) that describe the range of addresses that fall along the left side of the street, and another pair of fields for addresses on the right side. This reference layer can place an address on either the left or right of the street. The FRADD field is the From address and contains the lower range value; the TOADD field is the To address and contains the higher range value.

Streets									
PREFIX_DIR	PRE_TYPE	STREET_NAM	STREET_TYP	SUFFIX_DIR	LEFT_FRADD	LEFT_TOADD	RIGHT_FRAD	RIGHT_TOAD	
		NORTHLAND	DR		3301	3345	0	0	
		CEDAR BEND	DR		1605	1617	1610	1616	
		NORTHLAND	DR		3051	3225	0	0	
		BRECOURT MANOR	WAY		7621	7637	7620	7632	
		CARENTAN	DR		7701	7709	7700	7708	
N	IH	35	SVRD	NB	0	0	3213	3223	
N	IH	35	SVRD	NB	0	0	3201	3211	
		E 32ND ST TO N IH 35	RAMP		0	0	0	0	
		DECKER	LN		10714	10806	10715	10807	
		DESSAU	RD		10206	10214	0	0	
		DECKER	LN		10808	11016	10809	11017	

Fig. 10.4. The attribute table of a shapefile that could be used as the basis for geocoding

To create a point feature based on an address and street reference layer, the following logic is used. The number of the address is tested to see if it falls within the range of addresses on that street section. If it does, the point locating the address is placed along the street in proportion to its value between the range endpoints. In Figure 10.5, for example, Meadowlark St has an address range of 700 to 799. If matching the address 750 Meadowlark St, the point feature would be placed halfway between the street ends because 750 is halfway between 700 and 799. Likewise, 725 Meadowlark St would be placed a quarter of the distance along the street. This method of placing points is called the **One Range** method because a single address range is used for each street.

For more advanced geocoding, a **Dual Ranges** method can be used. In this method each street has four attributes defining its address ranges, a Left From and Left To address range for the left side of the street and a Right From and Right To address range for the right side. When using the Dual Ranges method, the addresses will be located on the left or the right side of the street, as dictated by the address ranges. In addition, the user can specify an offset distance such that the actual point is placed slightly to the side of the street instead of directly on top.

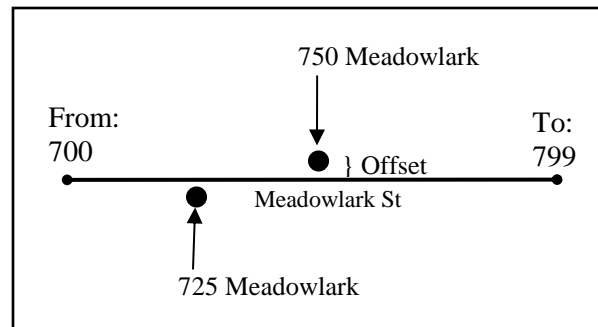


Fig. 10.5. Placing a located point along the street based on its proportional distance from the ends

The assumption that a house or other structure is located along a street segment at a distance proportional to its address number does not always work as well as it should. Parcel sizes may vary along the street, or some address numbers may be skipped entirely. This problem occurs most severely in rural areas where the spacing of addresses may be very irregular. If Meadowlark

St in Figure 10.5 represents a mile-long section line and all the houses are clumped together at one end, then the actual location and geocoded location may be quite different. Geocoded locations should always be considered approximations.

One also encounters great differences in the accuracy of the reference data used for geocoding. Large national data sets, such as TIGER, are generally less accurate and up to date than locally produced data intended for city governments and 911 applications. High-growth areas may experience difficulty keeping reference layers up to date. Users need to be aware of these potential pitfalls when geocoding. Creating and maintaining reference layers is a time-consuming and expensive process. Fortunately, the advent of cloud-based geoprocessing services has provided access to some national and international geocoding services through ArcGIS Online. These resources may not be as accurate as high-quality local data, but they are free and cover large areas in a standard way.

One Range and Dual Ranges are examples of different geocoding styles. The style chosen will depend on the way in which the **streets** layer has been set up for geocoding. Other styles are also available, as described in the next section. The user must ensure that the attributes in the reference layer support the desired geocoding style.

Although geocoding is most often associated with placing addresses, it applies to many types of data. Points can be located by city and state, by zip code alone, or even by township/range information. These different styles of geocoding have two common themes: they convert attributes to point locations using a reference layer, and they allow the user to link data that are similar but are not necessarily an exact match.

Available geocoding styles

Geocoding takes a variety of approaches, from finding a specific house location on one side of the street to simply finding a city and country on a world map. The available options depend on the attributes of the reference layer, the setup in the address table, and the type of matching to be done. A **geocoding style** specifies what type of matching will be done and defines a set of required attributes in the reference layer. To choose an appropriate style, the user must be familiar with the organization of the reference data.

Single Field

A single field locator uses one field to match to the reference layer, which typically contains points or polygons, although lines could also be used if desired. Any single field, such as the zip code, a state name, or a place name, can be used as the key to match the locations. This method expands geocoding from addresses to *any* type of information. For example, a table might contain information about petroleum wells keyed to specific well IDs, such as “TenRun Hole No. 6.” Such IDs may change slightly from database to database, depending on who typed them in (Ten Run Hole #6 or Ten Run Num 6). Performing an attribute join to link the table with a point layer of well locations would yield many missing values because the matches are not exact for each ID. Geocoding would work better because it can link close matches, not just exact matches. Single field locators are also used for township-range locations.

U.S. Alphanumeric Ranges

Many people are familiar with using grid zones to locate streets on a map by consulting a list of streets, finding its zone (such as H7), locating the zone by means of letters and numbers on the map margin, and scanning the streets in the zone to find the one in question. This geocoding style uses this approach to reduce the search time while geocoding. Each address is attributed with its

zone identification, and the geocoding algorithm need only search inside that zone. The required fields are the grid zone, the street name, and the From and To ranges for both sides of the street.

U.S. Cities with State

This style accommodates national scale matching of location based on the city and state names only. Cities may be matched to points or polygon features. This style could create a map showing the locations of customers across the United States. The required fields are the city name and a state name or state abbreviation.

U.S. Single House

This style uses points or polygon reference data set up so that each address corresponds to one point or polygon (Fig. 10.6). For example, it could be used with points representing buildings or polygons representing parcels. The required fields include street number and street name. The user may also include prefix directions, suffix directions, and other address components.

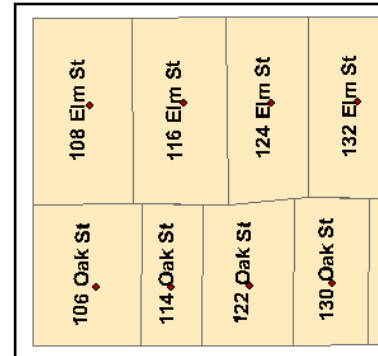


Fig. 10.6. The US Single House style codes to point or polygon layers, such as these parcels.

U.S. Streets—One Range

This style uses one range of addresses along a street rather than having ranges for both the left and right sides of the street. The required fields are the street name, the From address, and the To address. Prefixes, suffixes, and street types may also be used if desired.

U.S. Streets—Dual Ranges

This style relies on a reference layer with dual ranges of addresses, one for each side of the street. The required fields include the street name, a From and To address for the left side of the street, and a From and To address for the right side. Prefixes, suffixes, and street types may also be used if desired.

U.S. Hyphenated

This style follows conventions in locations, such as Queens, New York, where housing locations are indicated by a cross-street designation followed by the number and the street of the house. For example, the address 72-40 38th Ave specifies a location at 40 38th Ave near the cross street 72nd.

World Cities with Country

This style allows matching of a city and a country on a world map. The reference layer may contain cities as points or as polygons. The required fields include the city name and the country name or abbreviation.

Zip 5-Digit and Zip+4

The Zip style matches zip codes to points or polygons in a zip code reference layer. The only field required is the 5-digit zip code. The Zip+4 style works the same way but has an additional separate field containing the 4-digit extension.

With Zone

Many of the above styles offer a zone option. A zone is an additional piece of information in the reference layer, such as state, city, or zip code, which is added to the other required fields for that style. The addition of the zone allows matching over larger areas. For example, the US Streets style can only find addresses within a single city because only the street address is included; there

Address locators are created using ArcCatalog. First, the geocoding style must be designated. Figure 10.8 shows the settings for creating a Dual Ranges locator. A single streets reference layer is assigned and must be designated the Primary table. Supporting information in an Alias table may also be added at this time. The fields preceded by an asterisk (*) are required; the others are optional.

Next, the user must enter the fields from the reference data that correspond to the address components needed for the geocoding style. The fields shown in the setup will vary depending on the style of locator requested. This Dual Ranges style requires many different fields. The Single Field style requires only one.

Finally, the user specifies a name and location to store the locator. Locators may be stored in a folder or inside a geodatabase.

Geocoding

Geocoding starts with a table of addresses to be matched. The table may come from a dBase file, a geodatabase, or even Excel. At minimum it must have an address field and any other required fields for the style of locator being used. Geocoding may be performed live in ArcMap or using the Geocoding tools from ArcToolbox.

First, the user specifies the fields in the address table that will be used for geocoding (Fig. 10.9). If the common names are used for the table, as shown, then they will be recognized and included by default, so the user only has to change them if they are incorrect.

The user must specify a shapefile or feature class to contain the output points and table. Usually, the default static option is used.

Geocoding options

Next, the user has the chance to set various geocoding options that control how the locator decides on candidates and how it structures the output (Fig. 10.10). These options may also be adjusted later if necessary. Most users will be happy using the defaults at first.

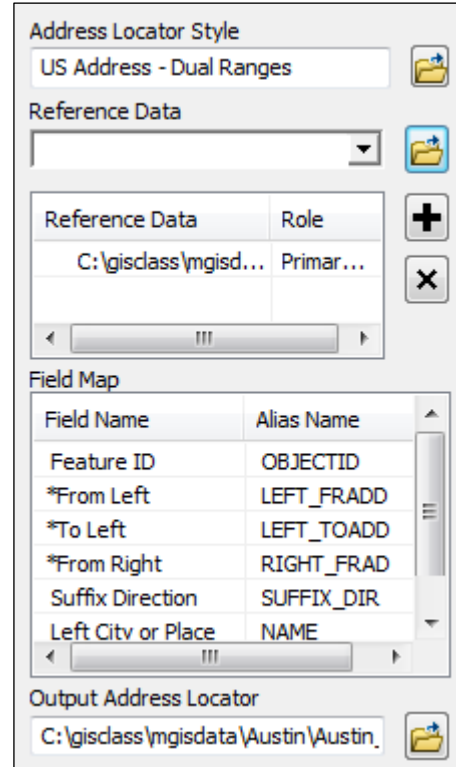


Fig. 10.8. Defining the reference layer for a new address locator

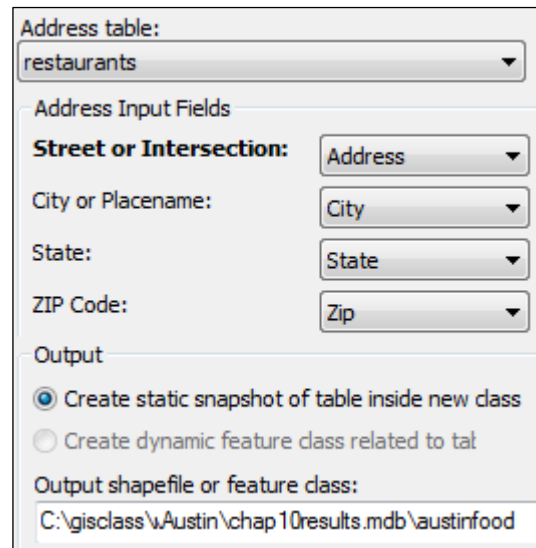


Fig. 10.9. Starting to geocode

At this time, an Alias table may be specified if one was not provided with the locator. The **spelling sensitivity** specifies how closely a street name must match to be a valid candidate. With this option, small misspellings or typing mistakes do not prevent a match from being made. Thus, Pierce St would still be recognized even if it were spelled Peirce St.

The **minimum candidate score** is used to determine how closely a reference feature must match the address to be considered a possible candidate for a match. Generally, a low value is desirable to aid in finding candidates for mistyped addresses, but if the user finds that too many candidates are being generated, the score can be raised.

The **minimum match score** is the lowest score that will constitute a match. In applications in which an occasional mismatch is not important, it is acceptable to lower the default in order to match as many records as possible. To a company analyzing customer locations, a few wrong addresses are not likely to significantly affect the analysis. In other cases, the minimum score might be raised. To an emergency management team, a poorly matched address may be a matter of life and death. In this case, poor matches should be identified and individually considered by the dispatcher to avoid sending the team out to a wrong location.

Geocoding can also match intersections, such as the corner of 5th St and St. Joseph St. This option specifies the symbols that are considered to be connectors so that the address locator can properly interpret the address as an intersection rather than a regular address. The default values are **&**, **|**, and **@** characters, but additional characters may be used. However, the separation character should not be one that might ordinarily appear in an address.

The output options control parameters related to the output file. First, for the Dual Ranges style, the user can set an offset distance to place the locations slightly off the street to the left or to the right. A setback from the street endpoints may also be specified so that no point ends up at the end of a street. Finally, when an address matches more than one street with the same score, marking the check box dictates that the address will be matched to the first street found.

The user can request additional fields in the output file for more information about the locations. These fields include the *x-y* coordinates of the point, the reference data ID (the feature-ID of the street it was matched to), the standardized address, and the percentage distance along the street where the point was placed.

After modifying the options, actual geocoding begins.

Fig. 10.10. Geocoding options

Matching addresses

In the first pass through the address table, the locator matches all of the addresses it can automatically. When finished, it issues a report of how many addresses were matched, how many were tied, and how many remain unmatched. If all addresses were matched (rare), you can close the window and the geocoding is done. Usually, though, the next step is to review and match addresses using the Interactive Rematch window (Fig. 10.11).

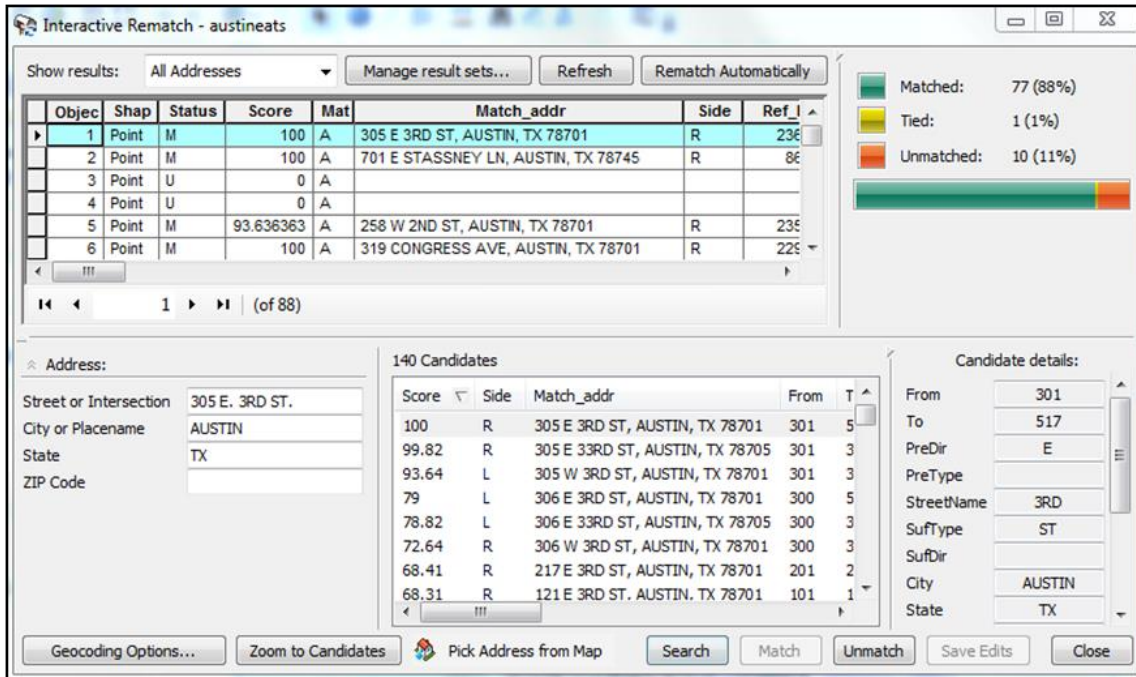


Fig. 10.11. The Interactive Rematch window

The top part of the window shows the records from the address table with additional fields added by the geocoding, such as the matched status (U for unmatched or M for matched), the score, and the address from the reference layer to which it was matched. A Show Results filter allows you to see all the addresses, just the unmatched ones, just the tied ones, and so on.

At this point you have two choices. You can modify the geocoding options from this window and then rematch automatically, or you can review each unmatched address individually and try to find matches that way. Before deciding, it is helpful to examine the unmatched addresses to see what types of problems are occurring.

You might notice that the unmatched addresses have multiple references to place names, such as City Hall or Capitol Building. In that case, it would be wise to create or locate an Alias table and specify it using the geocoding options. Then you would rematch automatically.

You might see that the address table is fully of sloppy typing mistakes, and that candidates are being generated that are mostly acceptable but do not quite meet the minimum match score. In that case, you would probably use the geocoding options to lower the spelling sensitivity and minimum match score and rematch automatically.

If you don't see any obvious repeated issues, then interactive matching is the next step. Each unmatched address is examined. Suspected typing mistakes can be corrected on the fly followed by a new search for candidates. Tied candidates or those with lower scores can be evaluated individually to see if a match is acceptable. If no candidates appear, you can try lowering the spelling sensitivity to generate more. However, it is not always possible to match every address. The address may be located beyond the extent of the locator reference layer, or errors in either the address table or reference layer may be present. Once you have matched all the addresses that can be matched, the Interactive Rematch window is closed and geocoding is complete.

Occasionally, the address table or the geocoding layer contains errors or inconsistencies. In this case the user may decide to halt geocoding, fix the problems, re-create the address locator, and try again to match the addresses. For example, the original reference layer prepared for this textbook had many streets named for saints, such as Saint Patrick St. In the reference layer as well as the address table, the street names were abbreviated to St. Patrick, St. Anne, and so on. However, during geocoding the software automatically converted *St* to *Saint* during the address standardization. As a result, dozens of addresses remained unmatched. To streamline matching, the author edited the streets layer to replace all the St. abbreviations with Saint.

In another example, students were analyzing crime patterns using street addresses and the StreetMap USA product. The police reports used the standard address prefixes for Rapid City (*West Main St*), but StreetMap uses suffixes (*Main St West*). In this case, the easiest solution involved editing the spreadsheet containing the addresses so as to place them in the address file with suffixes instead of prefixes.

These examples demonstrate that geocoding is not a simple push-button operation but may require some investigation and ingenuity on the part of the user. A thorough understanding of the geocoding process, address formats, and reference data being used helps in detecting and solving barriers to matching.

Summary

- Geocoding converts information in a table to points on a map, using reference feature classes that contain the relevant information tied to feature locations.
- Although geocoding most often refers to addresses, other types of information, such as zip codes, city-state pairs, or township-range, are also possible data on which to match table records to features in a reference layer.
- Geocoding requires an address locator that specifies the style of information to be used and stipulates the fields in the reference layer to be used for matching. Locators may come ready-made or may be created by a user for a specific purpose.
- National and international locators are available as cloud-based geoprocessing services on ArcGIS Online.
- Address geocoding breaks down addresses into standard components, compares these to the attributes in the reference layer, and scores how closely the components match possible candidates in the reference layer.
- Once a match is found, the address location is placed along the street at a distance proportional to the address value between the street address ranges.

- Geocoding styles vary on a number of factors, such as whether zip codes are used and whether the address ranges are given for the street as a whole (One Range) or for both the left and the right sides of the street (Dual Ranges).
- The process begins by matching as many addresses as possible. Then the user may change the geocoding options and try again, or interactively review individual addresses and attempt to match them. The process may be repeated until as many addresses as possible have been matched.
- Locators store a snapshot copy of the reference layer. If changes are made to the reference layer, a new locator should be created.

TIP: If you edit the features or attributes of a reference feature class, you must rebuild or re-create the address locator.

Important Terms

address locator	geocoding style	One Range
address standardization	minimum candidate score	reference layer
Dual Ranges	minimum match score	spelling sensitivity

Chapter Review Questions

1. What two data objects are required for geocoding?
2. Describe the difference between One Range and Dual Ranges geocoding.
3. If the 1200 block of Maple Street is 1000 ft long, how many feet from the 1200 corner would the address 1275 Maple be located during geocoding?
4. What is a street type? Give examples.
5. Which components are present in the address 1298 Bear Valley Road NW?
6. How does automatic matching differ from interactive matching?
7. What is an address locator and what does it include? If you don't have one, where might you look?
8. What is an Alias table and what is its purpose?
9. What is an Alternate Name table and what is its purpose?
10. Discuss the factors that would impact the geometric accuracy of a feature class derived from geocoding.

Mastering the Skills

Teaching Tutorial

The following examples provide step-by-step instructions for doing basic tasks and solving basic problems in ArcGIS. The steps you need to do are highlighted with an arrow →; follow them carefully. Click on the video number in the VideoIndex to view a demonstration of the steps.

- Start ArcMap and open ex_10a.mxd in the MapDocuments folder.
- Use Save As to rename the document and remember to save frequently as you work.

Setting up an address locator

For our first example, we will geocode restaurants in Austin, Texas. The addresses were obtained from an Internet search, copied or typed into a spreadsheet, and converted to a geodatabase table. However, before we can start, we need to decide what style of geocoding we will use, and we need to create an address locator.

- 1 → Open the Streets attribute table and examine the fields.

1. For each of the following address components, write down the name of the field in the table that contains the information.

Prefix direction	_____
Prefix type	_____
Street numbers	_____ and _____
	_____ and _____
Street name	_____
Street type	_____
Suffix direction	_____
City	_____
State	_____
Zip code	_____

2. Based on the fields just listed, which geocoding style is best to use? _____

- 1 → Open the restaurants table and examine the fields and records. (If the table is not visible, switch to the List by Source option in the Table of Contents.)

3. Which field contains the restaurant addresses? _____
4. Are there any intersections for addresses? _____ Which connector symbol is used? _____

- 1 → Close the Table window.

- 1 → Open the Properties of the Streets layer and examine the Source tab contents.

5. What is the source layer for Streets and where is it located? _____

Now we have all the information we need to set up an address locator using the Streets as the reference layer.

- 2➔ Close the Properties window for Streets.
- 2➔ Open the Catalog tab in ArcMap.
- 2➔ Open the Folder Connections, if necessary. Right-click the mgisdata\Austin\Austin geodatabase and choose New > Address Locator.
- 2➔ Click the Browse button for the Style and choose US Address – Dual Ranges (Fig. 10.12). Click OK.
- 2➔ Choose Streets as the Reference Data. Ignore the warnings for a moment.
- 2➔ Click in the Role field for Streets and use the drop-down to change it to Primary Table, if it does not already say it.

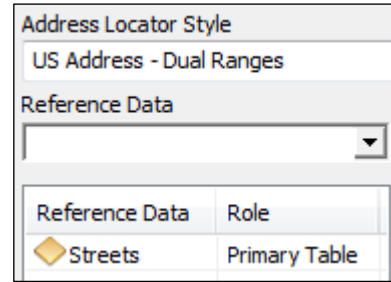


Fig. 10.12. Set the style and reference layer.

Next, we specify the fields to be used for the address components. Note that the required fields are preceded by an asterisk (*).

- 3➔ For the Feature-ID field name, click in the box under Alias Name, and use the drop-down to select the OBJECTID field (Fig. 10.13).
 - 3➔ For the From Left and To Left components, select the LEFT_FRADD and LEFT_TOADD fields.
 - 3➔ For the From Right and To Right components, choose the RIGHT_FRAD and RIGHT_TOAD fields.
 - 3➔ For the Prefix Direction and Prefix Type components, choose the PREFIX_DIR and PRE_TYPE fields.
 - 3➔ For the Street Name component, choose STREET_NAM.
 - 3➔ For the Suffix Type and Suffix Direction components, choose the STREET_TYPER and SUFFIX_DIR fields.
-
- 4➔ For both the Left City and Right City components, choose the NAME field.
 - 4➔ For both the Left and Right ZIP code components, choose ZIPCODE.
 - 4➔ For both the Left and Right State components, choose STATE.
 - 4➔ Leave the remaining components set to the <None> value.
 - 4➔ Store the locator in the Austin geodatabase and name it **Austin_Streets**.
 - 4➔ Click OK. Wait for the tool to finish; it may take a few minutes.

Field Name	Alias Name
Feature ID	OBJECTID
*From Left	LEFT_FRADD
*To Left	LEFT_TOADD
*From Right	RIGHT_FRAD
*To Right	RIGHT_TOAD
Prefix Direction	PREFIX_DIR
Prefix Type	PRE_TYPE
*Street Name	STREET_NAM
Suffix Type	STREET_TYPER
Suffix Direction	SUFFIX_DIR
Left City or Place	NAME
Right City or Place	NAME
Left ZIP Code	ZIPCODE
Right ZIP Code	ZIPCODE
Left State	STATE
Right State	STATE
Left Street ID	<None>
Right Street ID	<None>
Left Additional Field	<None>
Right Additional Field	<None>
Altname JoinID	<None>

Fig. 10.13. Setting the locator fields

Geocoding

Now we can geocode the restaurants. As always, we will create a practice geodatabase to hold the results of this chapter's work.

- 5 → Open the Catalog tab and navigate to the Austin folder. Right-click it and choose New > Personal Geodatabase. Name it **chap10results**.
- 6 → Right-click the Restaurants table in the Table of Contents and choose Geocode Addresses.
- 6 → Choose the Austin_Streets locator from the list and click OK.

TIP: If the desired locator is not already listed, click Add, navigate to the place where it is stored, and select it.

- 6 → Check the address input fields. The field names in the table are recognized by the software and added automatically, but check them to be sure (Fig. 10.14).
- 6 → Name the output feature class **austinfood** and save it in chap10results.
- 6 → Click the Geocoding Options button and examine the options.
- 6 → Note that the connector symbol you identified in the Restaurants table, &, is present in the list.
- 6 → Change the *Minimum match score* to 80 and click OK.
- 6 → Click OK to start geocoding.
- 6 → Examine the progress bar. When it is finished, you will find that you have some unmatched addresses.

Fig. 10.14. Starting to geocode

6. How many restaurants were matched? _____ Tied? _____ Unmatched? _____
(Your results may differ slightly.)

At this point, you have several ways to proceed. You could change the geocoding options and try to rematch. However, it is usually instructive to examine the unmatched addresses in interactive mode to find out why they won't match.

- 7 → Click the Rematch button.
- 7 → Examine the Rematch window. Note that it is currently set to Show All Addresses, both matched and unmatched.
- 7 → Examine the fields in the upper part of the window, showing the table from the geocoding output feature class.

Figure 10.15 shows the upper table, with some of the fields resized so that you can see more of them than in your window. The Status field contains an M for Matched or U for Unmatched. The match score is shown in the Score field. The Match_addr field shows the final matched address. In this table, the first two fields were matched and the second two were not matched.

Show results: All Addresses							
Manage result sets... Refresh Rematch Automatically							
	Objc	Shape	Status	Score	Match_type	Match_addr	Side
▶	1	Point	M	100	A	305 E 3RD ST, AUSTIN, TX, 78701	R
	2	Point	M	100	A	701 E STASSNEY LN, AUSTIN, TX, 78745	R
	3	Point	U	0	A		
	4	Point	U	0	A		
	5	Point	M	90.97	A	258 W 2ND ST, AUSTIN, TX, 78701	R
	6	Point	M	100	A	319 CONGRESS AVE, AUSTIN, TX, 78701	R

Fig. 10.15. All Addresses results from the first geocoding attempt, with field widths adjusted

Examine the lower half of the window. The area to the left shows the address for the currently highlighted field. A list in the middle shows potential candidates, sorted by score. The area to the right shows the details for the currently highlighted candidate. Take a moment to examine the various candidates and scores, to get a feel for how the locator scores and matches the addresses.

8➔ Change the Show Results box from *All Addresses* to *Matched Addresses with Candidates Tied*.

8➔ Examine the lower panel for the first entry, 201 W 3RD ST, seeing the actual address on the lower left, and the candidates on the right.

The address in the table did not specify E or W 3rd Street, so there are two equally likely candidates. A little detective work might help.

8➔ Scroll the upper table to the right until you can see the restaurant name.

The restaurant's name is Catina Laredo (probably a misspelling of Cantina). You could whip out your Austin telephone directory (or yellowpages.com) and look up the restaurant, discovering that it is in fact on West 3rd Street.

8➔ Click on the 201 W 3RD ST candidate to highlight it, and click Match.

8➔ If you see more tied addresses, examine them and choose the top candidate. Click Match for each one if you can find a match.

9➔ Change the Show Results box to *Unmatched Addresses*.

9➔ Resize the first few fields to make them smaller.

9➔ The Match_addr field is empty and very wide, preventing you from seeing the addresses. Right-click the Match_addr field heading and choose Turn Field Off.

9➔ Highlight the first row in the table by clicking on the open gray box to its left.

TIP: You can adjust the size of each window section, or the size of the fields within a section, as needed. Place the cursor on the right side of the lower-left Address window, if necessary, and resize it so that you can see the entire address.

The first unmatched address should be 13343 HWY 183N. No candidates appear at all for this one. Let's try adjusting the geocoding options.

- 9➔ Click the Geocoding Options button at the bottom left of the window.
- 9➔ Relax the spelling sensitivity to **60** and click OK.
- 9➔ Click Search to look for more candidates. No new candidates appear.
- 9➔ Set the spelling sensitivity back to the default of **80**.

Sometimes the commonly written address does not match the reference layer. The next step is to search the reference layer manually to see if this road is being portrayed differently.



- 10➔ Minimize (do not close!) the Interactive Rematch window.
- 10➔ Click the Find tool on the Standard toolbar and enter the text to search as **183**.
- 10➔ To refine the search, choose to search only in the Streets layer and only in the STREET_NAM field (since that is what you used in the locator). Click Find.

The search shows that there are many different ways that this reference layer stores HWY 183. This problem is not uncommon in reference layers. It is hard to be consistent when assembling data sets—with troublesome results, as you can see.

- 11➔ Click a few of the entries to flash them and see where they are located.
- 11➔ Click the first entry, scroll to the end, hold down the Shift key, and click the last record, to highlight them all.
- 11➔ Right-click on one of the entries and choose Select, to select them all.

With all of the segments highlighted, it appears that most designate the same road, although a few streets form spurs off the main road. You note these points as issues with the reference layer that may need fixing later. In the meantime, there are one or two tricks to try to get a match right now.

TIP: If you closed the Interactive Rematch window by mistake, or if you want to rematch a geocoded layer at any time, you can open the Interactive Rematch window by right-clicking the layer in the Table of Contents and choosing Data > Review/Rematch Addresses.

- 12➔ Close the Find window and clear the selection.
- 12➔ Restore the Interactive Rematch window.
- 12➔ In the Address section, edit the address field, placing a space between 183 and N.
- 12➔ Click Search.
- 12➔ If any reasonable candidates appear, click Match. If not, leave it unmatched and go on.

Address:	
Street or Intersection	13343 HWY 183 N
City or Placename	AUSTIN
State	TX
ZIP Code	

TIP: Between the initial release of Version 10.1 and the first service pack, the behavior of the geocoding changed substantially. Your results may not correspond exactly with the instructions. You may find new unmatched addresses or fewer than described below. For an unmatched address, you may find different candidates or no candidates at all. Don't worry about it; just be flexible and try to match all the unmatched addresses that you can.

13➔ Highlight the row of the next unmatched candidate, 3107 S. IH35.

A couple of candidates may appear, but with low scores. Keeping in mind the issue with HWY 183, let's try a couple of alternate road names to see if better candidates appear. IH35 probably refers to Interstate-35. The roads look reasonable, but the address ranges are not satisfactory.

13➔ Edit the address, trying variations such as **IH 35** or **I 35**. Try removing the period after S. Click Search after each try.

13➔ It might find the correct road, but not the correct address range. Try editing the City to AUSTIN instead of ROUNDROCK. Click Search.

13➔ Now the top candidate looks reasonable. Highlight it and click Match.

TIP: If you have room on your computer screen, move the Interactive Rematch window off the map. Potential candidates are highlighted on the map so you can examine them spatially if needed. The highlighted candidate is shown in yellow, the rest in blue.

14➔ Click on the next unmatched address, 500 NORTH I-35. Note that some candidates refer to IH 35. This designation seems to be standard for this Austin dataset.

14➔ Edit the address, replacing I-35 with **IH 35**, and click Search. Now the top candidate is a good match with a score over 90. Click Match.

15➔ Click on the next unmatched address, 605 West St. The top candidate is barely below the minimum match score of 80, and it looks reasonable, so click Match.

15➔ Click on the next unmatched address, 2900 N RM 620. The top candidate appears acceptable. RM is a mistyping of FM, perhaps. Click the top candidate and Match.

15➔ Click the next, 3506 RR 620. Match the top candidate if any appear.

16➔ Click the next, 4894 HWY 290 WEST. Match the top candidate if any appear.

16➔ Click the next, 5001 EAST BEN WHITE. Match the top candidate if any appear.

16➔ Click the next, 4715 HWY 290 WEST. Match the top candidate if any appear.

17➔ Click the next, 6500 BEE CAVES ROAD. Match the top candidate if any appear.

At this point you might have about 86 matched or tied and 2 unmatched, although your values may differ depending on the service pack you have. (Service Pack 1 yielded 84 matched and 4 unmatched.) In any case, you are finished.

18➔ Click Close to finish rematching.

18➔ Examine the geocoded austinood layer.

- 18➔ Zoom in to the cluster of restaurants in downtown Austin and examine the restaurants. Identify a few if you wish.
- 18➔ Return to the Full extent of the map.
- 18➔ Save your map document.

Geocoding with the Single House style

Instead of using street features, points or polygons can be used if they represent houses or parcels. A small section of Austin will serve to demonstrate, as these reference layers can be very large.

Imagine that a political group has come to City Hall with a set of 100 signatures regarding a neighborhood issue. You have been asked to verify whether the addresses on the petition are real.

- 19➔ Right-click the Crestview data frame name and choose Activate.
- 19➔ Zoom in to a block of houses and examine the parcels with the address points.
- 19➔ Return to the previous extent.
- 19➔ Turn off the Address Points layer.

- 20➔ Open the Parcels attribute table and examine the fields. Note the address component fields, similar to the ones we used in the Streets layer.
- 20➔ Open the Address Points attribute table and examine the fields.
- 20➔ Open the petition table and examine the fields.
- 20➔ Close the Table window.

We could use either the Address Points layer or the Parcels layer for this style of geocoding, but we will use the Parcels.

- 21➔ Open the Catalog tab, right-click on the Austin geodatabase, and choose New > Address Locator.
- 21➔ Click the Browse button for the Address Locator Style and choose the US Address – Single House style. Click OK.
- 21➔ Set the Reference Data to Parcels and make sure that it is the Primary Table.

- 22➔ Enter the reference fields (Fig. 10.16).
- 22➔ Name the locator **Austin_Parcels** and save it in the Austin geodatabase.
- 22➔ Click OK and wait for the locator to build.

Field Name	Alias Name
Feature ID	OBJECTID_1
*House Number	ADDRESS
Side	LEFT_RIGHT
Prefix Direction	PREFIX_DIR
Prefix Type	<None>
*Street Name	STREET_NAM
Suffix Type	STREET_TYP
Suffix Direction	SUFFIX_DIR
City or Place	CITY
ZIP Code	ZIPCODE
State	STATE
Longitude	<None>
Latitude	<None>
Street ID	<None>
Additional Field	<None>
Altname JoinID	<None>

Fig. 10.16. Single House

It's time to geocode to the parcels.

- 23➔ Right-click the petition table and choose Geocode Addresses.
- 23➔ Highlight the Austin_Parcels locator and click OK.
- 23➔ Set the Street or Intersection field to FULL_NAME.
- 23➔ Leave the City, State, and ZIP Code fields set to the defaults found in the table.

- 23➔ Name the output **petitionparcels** and save it in the chap10results geodatabase.
- 23➔ Click OK. Examine the results when it finishes.

You should have about 94 matched and 6 unmatched petition names.

- 24➔ Click Rematch.
- 24➔ Change Show Results to Unmatched Addresses.
- 24➔ Examine each of the first five unmatched addresses. None of them come up with good candidates, so leave them unmatched.
- 25➔ The last one, 1006 CHESAPEEK DR, has no candidates, but it might be a misspelling of CHESAPEAKE.
- 25➔ Edit the address and search for new candidates. Match the top one.

You should now have 95 matched addresses and 5 unmatched. The unmatched ones are either not real addresses or lie outside this neighborhood area. Just to be sure, though, you will check the reference layer.

- 26➔ Close the Interactive Rematch window, completing the geocoding session.
- 26➔ Notice the geocoded points that have been placed inside the parcel polygons.
- 26➔ Open the attribute table for Parcels and sort on the STREET_NAM field.
- 26➔ Scroll down and look for the missing street names.

Oertli is there, although the address number does not match any of the parcels. The other streets are not listed at all. You can be sure that these signatures on the petition are invalid.

- 27➔ Open the Geocoding Result: petitionparcels table.
- 27➔ Use Table Options > Select by Attributes to select the unmatched addresses using the expression [Status] = 'U'.



- 27➔ Click the Show selected records button in the table and scroll right to find the unmatched addresses.

You could export or screen-capture this list for the report you will prepare for your supervisor, indicating that these five signatures should not be counted.

- 27➔ Close the Table window and save the map document.

Address
212 OERTLI LN
8617 N LAMAR BLVD
8311 SWALLOW CT
8409 GOLDFINCH CT
8413 GOLDFINCH CT

Geocoding U.S. cities

Next, we will practice our geocoding skills by matching a table of precipitation data for U.S. cities to a feature class of U.S. cities. When finished, we'll make a map of monthly normal precipitation totals for the United States. Now let's prepare to do our geocoding and examine the table we'll be using for the "addresses."

- 28➔ Open a new, blank map document and set the home geodatabase to the mgisdata\Usa\usdata.
- 28➔ Add the cities and states feature classes from the usdata geodatabase. Also add the table normal_precip.

- 29→ Zoom in to the conterminous United States.
- 29→ Open the normal_precip table and examine the fields. The precipitation values are given in inches for each month; a yearly total is also given.
- 7. Which style of geocoding will we use this time? _____
 - 29→ Open the attribute table of the cities layer, which will be the reference layer.
- 8. Which field will we use for the city name? _____ Which field do we use for the state abbreviation? _____
 - 29→ Close the Table window.
 - 30→ Open the Catalog tab and navigate to the mgisdata\Usa folder.
 - 30→ Right-click the usdata geodatabase and choose New > Address Locator.
 - 30→ Click the Address Locator Style browse button and select the General – City State Country style.
 - 30→ Specify cities as the reference layer. Make sure it is the Primary Table.
 - 30→ Select the NAME field for the City Name component and the ST for the State component.
 - 30→ Name the locator **US_City_State** and save it in the usdata geodatabase.
 - 30→ Click OK to create the locator.

All right, let's give it a try.

- 31→ Right-click the normal_precip table and choose Geocode Addresses.
 - 31→ In the Choose an Address Locator window, select the US_City_State locator just created. Click OK.
 - 31→ In the Geocode Addresses window, specify CITY for the City field and STATE for the State field.
 - 31→ Put the output in the chap10results geodatabase and call it **cityprecip**.
 - 31→ Click OK to start geocoding.
 - 31→ Examine the results window.
9. How many matched/tied records are there? _____ How many unmatched? _____

That is not a bad effort. Let's look at the unmatched records.

- 32→ Click Rematch and change the Show Results to Unmatched Addresses.
- 32→ Highlight the first row, and examine the Address entry on the lower left of the window. It says NOME, AK. No candidates have appeared.

You've probably heard of Nome, Alaska, and think that the spelling is correct. Perhaps this smaller city is just not within the cities feature class?

- 32→ Open the Find tool on the Standard toolbar.
- 32→ Set it to search only in Cities, and type **Nome** as the Find text. Make sure it is searching in all fields, and click Find.

It did not find Nome, so apparently this city was not in the feature class.

TIP: If you are not certain Find is working, search for a major city, like Boston, to check.

- 33→ Leave the Find window open in case you need it again, moving it aside to see the Interactive Rematch window.
- 33→ Leave NOME unmatched and highlight the next unmatched record.
- 33→ The next unmatched record is WINSLOW, AZ. Use Find to search for it in the cities layer. Nothing is found, so leave it unmatched.
- 34→ The next several cities are not in the reference layer. Check each one to be sure and then leave them unmatched, until you get to MINNEAPOLIS-ST.PAUL.

When Find turns up nothing for Minneapolis-St. Paul, a major city, we are immediately suspicious. It must be in the cities layer.

- 35→ In the Find window, search just for **Minneapolis**.

Now we see the problem. The precipitation station is named for the combined city of Minneapolis-St. Paul, but in Cities they are listed separately.

- 35→ Edit the address in the City Names address box, changing it simply to **MINNEAPOLIS**.
- 35→ Click Search. The right candidate appears. Match it.

⌵ Address:	
City Names	MINNEAPOLIS
State	MN

- 35→ Click the next unmatched record, OMAHA EPPLEY.

Omaha, Nebraska, you have heard of. Eppley is actually the airport. In this case, the climate station is named not just for the city, but also for its location at the airport. It can be matched.

- 36→ Edit the address to read just OMAHA, click Search, and match it.
- 36→ The next unmatched one, ATLANTIC CITY AP, appears to have the same issue. Edit the address and match it to Atlantic City, NJ.
- 36→ The next unmatched one is a second climate station for Atlantic City. Edit the address and match it.

Some large cities have more than one climate station, like Atlantic City and New York. Simply match all the climate stations to the same city. A point will be created for each station, stacked on top of each other.

All of the remaining unmatched records have one of the issues we've already discussed.

37→ Keep moving through the table, examining each address in turn. Match them if you can, or leave them unmatched if you cannot.

37→ Also examine the tied candidates and check that they have been assigned the best match.

TIP: The initial release of Version 10.1 caused inappropriate ties to occur when matching some cities. For example, a climate station in Little Rock, AK, was matched to North Little Rock instead of Little Rock (both had scores of 100).

When finished, you should have about 222 matched stations and 4 unmatched ones.

37→ Click Close to close the Interactive Rematch window. Also close the Find window.

Once matched, the data can be used to analyze the distribution of rainfall across the United States.

38→ Turn off the Cities layer to see the precipitation locations better.

38→ Create a graduated symbol map of the Geocoding Result layer, showing the annual precipitation (field ANN) in symbols that increase in size with precipitation (Fig. 10.17).

38→ Save the map document.

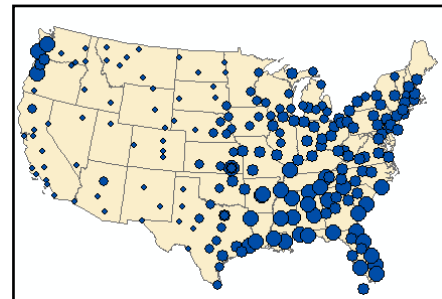


Fig. 10.17. Annual precipitation of the matched cities

10. Can you think of two reasons the geocoding method works better to accomplish this task, as opposed to joining the precipitation data to the Cities layer based on a common CITY_NAME attribute?

Geocoding township-range data

Geocoding provides an effective way to convert township-range-section locations to points on a map. Imagine that you have some rock samples that a retired professor collected from Grant County, Oregon, located by township, section, and quarter-quarter section. You want to produce a map showing the locations of the samples. Someone has already typed the locality information into a spreadsheet, as shown in Figure 10.18.

STATE	MERID	TSHP	RANGE	SECTION	QQSEC	QSEC
OR	33	T0150S	0270E	023	SE	NW
OR	33	T0170S	0260E	016	NW	SE
OR	33	T0160S	0270E	010	SE	SW
OR	33	T0150S	0290E	018	SW	SE
OR	33	T0170S	0280E	029	NE	SE

Fig. 10.18. Rock sample localities

39→ Open the map document ex_10b.mxd and save it under a new name.

39→ Zoom in to one of the gray squares, a township. Township labels and sections will appear.

- 39→ Zoom in to one of the red sections. The quarter-quarter section outlines and labels appear.
- 39→ Zoom to the full extent of the map.

The townships, sections, and quarter-quarter section feature classes have been included just as they come from the BLM web site, except that they have been converted from shapefiles to geodatabase feature classes.

- 40→ Open the QQsection table. Notice the lndkey, sectn, and qqsection fields and examine their contents carefully for the first row in the table.

Figure 10.19 shows the relevant attributes of the QQsection table. The lndkey field contains a concatenated value representing the state, zone, and township. The first row refers to the NE corner of the NE quarter of Section 19 of Township 17S-26E. Because these fields are formatted with additional zeros to keep the codes a consistent width, your data tables must include the zeros, as has been done in Figure 10.18. Note that the spreadsheet columns must be specifically formatted as text to retain the leading zeros.

The geocoding style that you must use is the General – Single Field style. Because you have only one field to match on, you must contrive a way that both your location data and your reference data have a single unique field that identifies the polygon and sets up the match. You can get this by concatenating the strings into a single key field, in both the reference layer and the spreadsheet.

lndkey	qsection	sectn	qqsection
OR33T0170S0260E	NE	019	NENE
OR33T0170S0260E	NW	020	NWNW
OR33T0170S0260E	NW	020	NENW
OR33T0170S0260E	NE	020	NWNE
OR33T0170S0260E	NE	020	NENE

Fig. 10.19. Relevant QQsection attributes

- 40→ Choose Table Options > Add Field.
- 40→ Type **KEY_QQSEC** as the field name, and make it a text field with a length of 25. Click OK.
- 40→ Right-click the empty KEY_QQSEC field and choose Field Calculator.
- 40→ Enter the expression [lndkey] & [sectn] & [qqsection]. In this window, the & symbol is a string concatenation operator. Click OK.

The KEY_QQSEC field in the first row should now read OR33T0170S0260E019NENE, including all the information needed to uniquely identify the polygon. (You may need to expand the field width to see it all.) Next, you must edit the spreadsheet to create matching identifiers.

TIP: If you don't have access to Excel™, simply use the rocksampleskey.xls spreadsheet, which already has the key added for you, and skip to Step 50.

- 41→ Close the Table window.
- 41→ Start Excel™, and open the mgisdata\Oregon\rocksamples.xls spreadsheet.
- 41→ Create a new column in the spreadsheet named KEY.
- 41→ In the top row, enter the formula =concatenate(B2,C2,D2,E2,F2,G2,H2).
- 41→ Copy the formula down to the other rows.

ArcMap cannot read spreadsheets with formulas, so you must convert the formulas to text.

- 42➔ Select the cells containing the formulas and click Ctrl-C to copy them.
- 42➔ Right-click inside the highlighted cells and choose Paste Special.
- 42➔ Fill the button to Paste Values and click OK.
- 42➔ Save the spreadsheet and close it. (ArcMap will not open the file while it is open in Excel™).

Now you build the address locator.

- 43➔ In ArcMap, open the Catalog tab.
- 43➔ Right-click on the mgisdata\Oregon folder and choose New > Address Locator.
- 43➔ Click the Browse button for the Style and choose General – Single Field.
- 43➔ Select QQsection as the reference data and make sure it is set as the Primary Table.
- 43➔ Set the Key field to KEY_QQSEC.
- 43➔ Name the locator **QQsection_locator** and store it in the mgisdata\Oregon folder.
- 43➔ Click OK to start building the locator.

Now you can geocode the rock sample locations.

- 44➔ Click Add Data and navigate to the mgisdata\Oregon folder.
- 44➔ Double-click the rocksamples.xlsx file to open it.
- 44➔ Select Sheet1\$ and click Add.
- 44➔ Open the table to make sure it looks all right, and then close it.

- 45➔ Right-click Sheet1\$ and choose Geocode Addresses.
- 45➔ Select the QQsection_locator and click OK.
- 45➔ Set the Key field to KEY.
- 45➔ Name the output **samplelocalities**, and save it in the GeolProject geodatabase. Click OK.

All of the records should be matched, so you shouldn't need to do any rematching.

- 46➔ Click Close to close the matching window.
- 46➔ Change the samplelocalities layer symbol to a large green circle so it shows up better against the geological map (Fig. 10.20).
- 46➔ Zoom in to one of the townships containing sample points.
- 46➔ Zoom in until you can see the quarter-quarter sections, and the sample points placed in the center of one.

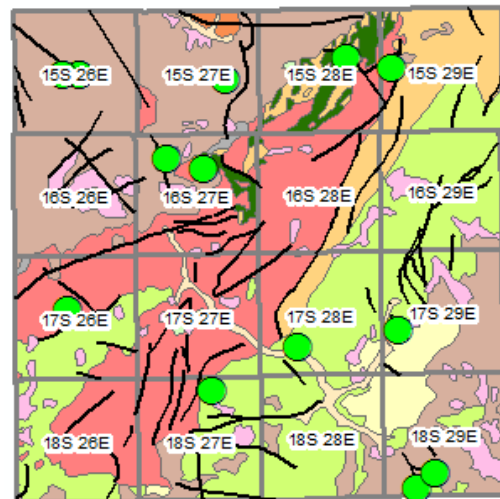


Fig. 10.20. Rock sample localities

46 → Return to the full extent of the map.

This geocoding technique can also be applied to locating just townships, sections, or quarter-sections, depending on the level of detail in the data you need to geocode. Simply modify the key field and use the appropriate reference layer.

TIP: GIS files of PLSS units are freely downloadable as shapefiles from the site http://www.geocommunicator.gov/geocomm/lsis_home/home/index.shtm.

This is the end of the tutorial.

→ Exit ArcMap and save the map document.

Exercises

1. Search on the Internet for 10 to 15 hotels in Austin, Texas. Type the names and addresses into a spreadsheet and geocode them. **Capture** a map showing the hotels.
2. Geocode the Austin hotels using the online geocoding service. Compare them with the results of Exercise 1. **Capture** a map showing both sets of hotels.
3. In the mgisdata\Rapidcity folder, you will find a table of gas stations called gas_loc.dbf and an Alias table called rc_alias.dbf. Create a One Range locator using the roads feature class in the rapidnets geodatabase, and geocode the gas stations. **Capture** a map showing the stations.
4. In the mgisdata\Usa\usdata geodatabase, you will find a table called customers. These data come from the test marketing of a catalog for natural landscaping to minimize watering of plants (xeriscaping). Create a map showing the location of the customers. **Capture** your map.

Challenge Problem

Search the Internet for a data set that interests you, is geocodable, and contains at least 25 data points. Create a spreadsheet that can be read by ArcMap. Then geocode it. With your answer, include the source of the data, the style of locator used, and the service or reference layer used.

Skills Reference

Creating an address locator

An address locator specifies the style and reference layer used for geocoding.

1. Examine the reference layer to ensure that all the fields required for the geocoding style are present, and identify the names of the fields.
2. In ArcCatalog or the Catalog tab, right-click the name of the folder or geodatabase in which you want to store the locator and choose New > Address Locator.
3. Click the Browse button to choose the address locator style (Fig. 176).
4. Specify the name and location of the reference feature class using the drop-down or the Browse button.
5. Specify Primary Table as the role for the reference layer.
6. Add additional table(s), if desired, and specify their role(s).
7. Set the required fields (asterisk) and the optional fields to the appropriate fields in the attribute file of the reference layer.
8. Specify the name and location where the address locator will be saved.
9. Click OK to create the locator.

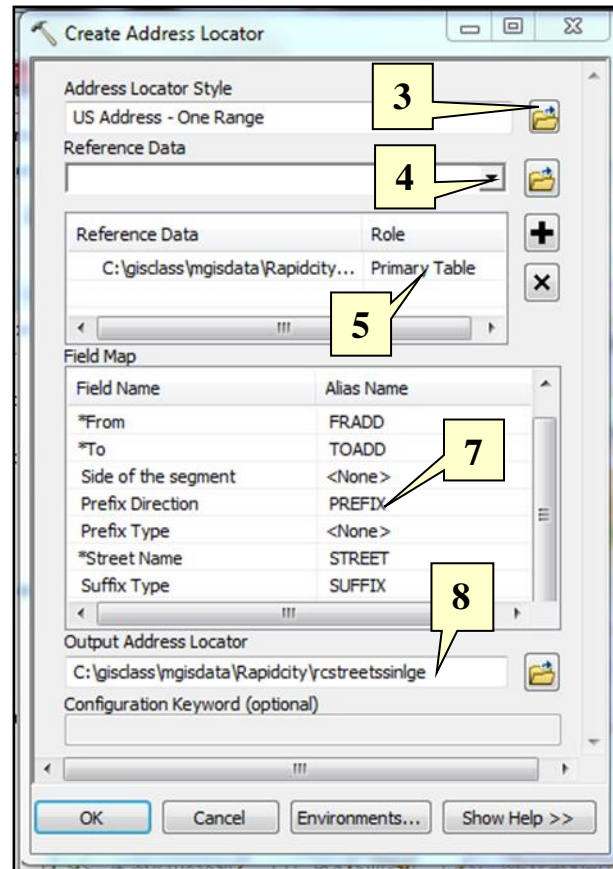


Fig. 176. Creating an address locator

TIP: Although theoretically optional, the City, State, and Zip fields must be used in the address locator for it to work. This is a bug in the initial release of ArcGIS 10.

Geocoding addresses

1. Add the address table and the reference layer (optional) to ArcMap.
2. Click the List by Source icon at the top of the Table of Contents.
3. Right-click the address table in the Table of Contents and choose Geocode Addresses.
4. If the desired locator does not appear in the menu, click Add (Fig. 177).

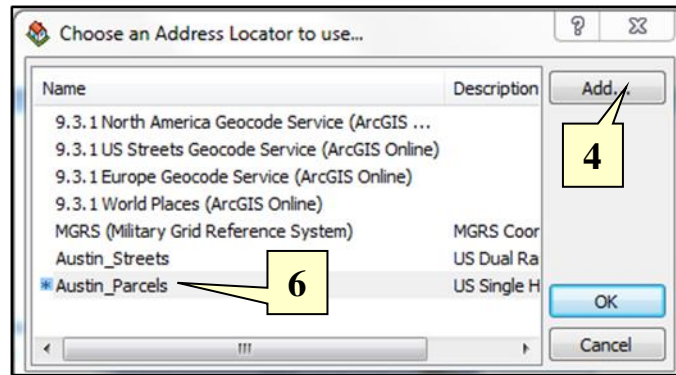


Fig. 177. Choosing an address locator

5. Navigate to the folder where the address locator is stored. Select the locator and click Add to close the menu and add the locator to the list.
6. Click the desired locator to highlight it and click OK.
7. Specify the table containing the addresses (Fig. 178).
8. Specify the fields containing the address information to be matched.
9. Specify the name and the location of the output file.
10. Click the Geocoding Options button to change the spelling sensitivity, minimum candidate score, or minimum match score.
11. Click OK to begin matching.
12. A progress bar will appear. When it is complete, click Close to finish matching or click Rematch to try matching more addresses.

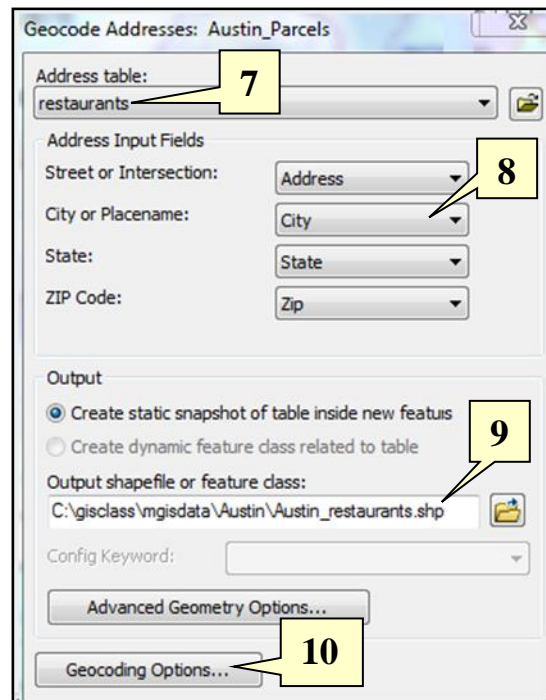


Fig. 178. Preparing to match

Rematching a geocoding result

You can reopen a completed matching session to rematch more addresses.

1. It is best to terminate any editing before rematching addresses. Choose Editing > Stop Editing from the Editor menu and save any changes, if necessary.
2. Right-click the geocoded layer in the Table of Contents and choose Data > Review/Rematch Addresses.
3. Rematch the addresses. See *Interactive rematching*.

Interactive rematching

Interactive rematching lets you examine each unmatched address, edit errors or inconsistencies in the address, and search for candidates.

1. Change the Show Results box to Unmatched Addresses (or another group) (Fig. 179).
2. Click on one of the unmatched addresses to highlight it.
3. If a list of candidates appears in the bottom box, click the best match to choose it, and click the Match button. The address will be matched, and the U in the upper panel will change to an M. The match score will also be shown.

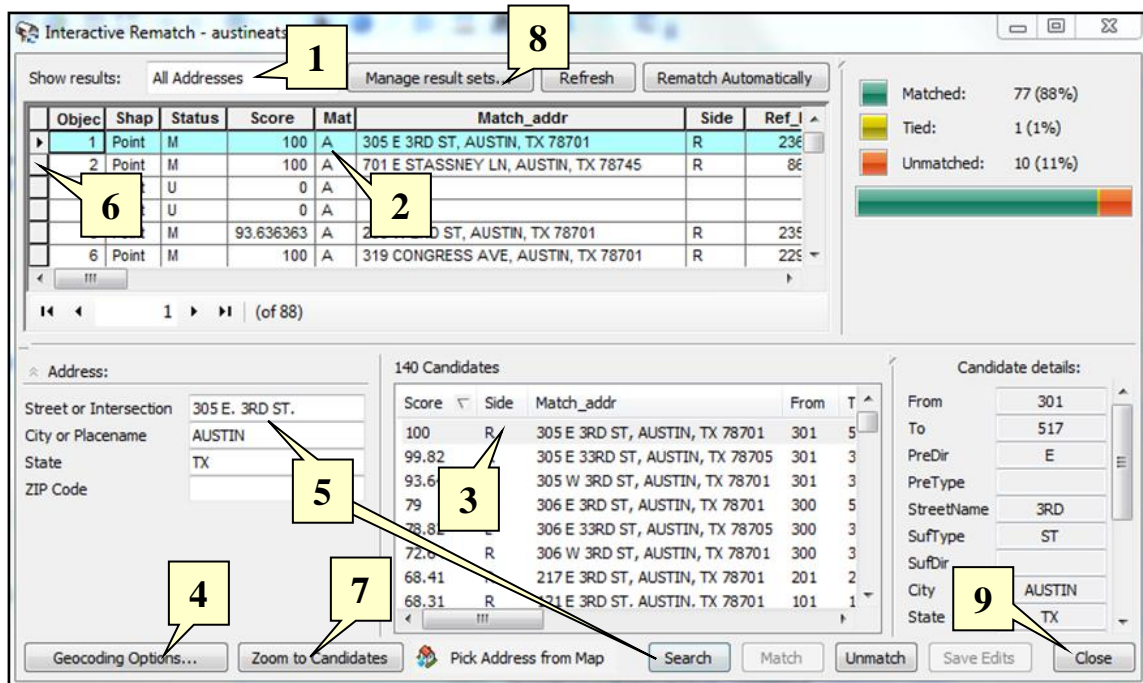


Fig. 179. The Interactive Rematch window

4. If no candidates appear, try clicking the Geocoding Options button to modify the spelling sensitivity or the minimum candidate score.
5. If still no candidates appear, examine the Address section to see if the address is being interpreted correctly. Edit the address to try alternatives, and then click Search.
6. If still no candidates appear, the address is probably unmatchable. It may be incorrect, or the reference layer may be incomplete. Click another address and begin again.
7. If you are trying to decide among several candidates, click the Zoom to Candidates button to view them on the map. Click Original Extent to go back to the previous extent.
8. You can use the *Manage result sets* button to save groups of records to work on later.
9. When you are finished rematching, click Close.