

11.520: A Workshop on Geographic Information Systems

11.188: Urban Planning and Social Science Laboratory

Lab Exercise 8: Address Matching & Geocoding

Distributed: Lab 8

Due: Lecture 13

In this lab, you will learn how to geocode mailing list style addresses. Quite often, planners have mailing lists or locate Web pages that list the names and addresses of sites, clients, activities and the like. GIS packages such as ArcGIS provide tools that facilitate converting these addresses into X,Y coordinates so that the spatial patterns of the addresses can be studied.

Part I: Finding the Addresses to be geocoded

To obtain a list of addresses for this exercise, we will use an online database of the Bureau of Waste Site Cleanup in the Massachusetts Department of Environmental Protection. This website: <http://www.mass.gov/dep/cleanup/index.htm> lists information about reportable releases of hazardous material.

We have saved a snapshot of the Bureau of Waste Site Cleanup list as of October, 2004, in the spreadsheet, **ma_release_04oct25.xls**, stored in the class data locker: M:\data. The main table is in the worksheet labeled 'RTN' and the column headings are explained in the 'metadata' worksheet. There are 381 sites in Cambridge that involve the chemical type = 'oil'. (Specifying type = 'oil' will select both the 'oil' and the 'oil and hazardous material' cases. That's fine.) These 381 rows have been copied and pasted into the worksheet labeled 'RTN-oil' in the spreadsheet. To facilitate moving these data into MS-Access, we have replaced the original first row with short, one word attribute names: RTN TOWN ADDRESS PLACETYPE REPORTCAT REPORTDATE CSTATUS STATUSDATE PHASE RAO CTYPE RTN2. Save a copy of this spreadsheet in your personal workspace.

Open a new database (let's call it **lab8_oil.mdb**) in MS-Access and import the Excel table you just copied (let's call that table, 'RTNoil'). Design and save an MS-Access query that selects the 81 cases involving both oil and hazardous material and that have a 120 day reporting category.

Part II: Mapping the Location of Cases Involving Oil and Hazardous Material

Open ArcMap, add the Arc portion of the **cambtigr** coverage layer. Also add your MS-Access table containing the 81 cases involving both oil and hazardous material with a 120 day reporting category. Use the ArcGIS geocoding tools to locate these 81 cases. (When you create the geocoding service using the **cambtigr** coverage be sure to set the datatype to coverage rather than shapefile or the cambtigr coverage will not be visible.) Only 40 of the 81 cases are address-matched automatically using the default settings. Hand match two more of the cases: the Ames and Amherst St case that shows up first in the unmatched list and the 367-371 FRANKLIN ST case that is a few rows down.

The steps involved in geocoding the addresses are just like those discussed in lecture (see the lecture notes) using the examples in Chapter 17 of the Ormsby textbook, "Getting to Know ArcGIS". The datasets for these Chapter 17 exercises are available in the 'chapter17' sub-directory of the class data locker (M:\data). By referring to the online lecture notes and the online ArcMap help files, you should be able to figure out the geocoding steps without having Chapter 17 in front of you.

Create a map layout showing Cambridge streets along with the location of the 42 matched addresses. Change the symbols so that the two hand-matched addresses are distinguishable from the rest. Also, draw a 500 meter buffer from the intersection of Mass Ave and Vassar St. How many of the 42 address-matched addresses fall within this 500 meter buffer? Add this number as an annotation on your map.

***Turn in** your map layout showing the Cambridge streets, the matched addresses (with different symbols for the two interactively-matched addresses), the 500 meter buffer around Mass Ave and Vassar, and your annotation stating the number of matched addresses that fall within the 500 meter buffer.*
