A Crash Course in Modern Geography for HCI Researchers and Practitioners

Brent Hecht
Assistant Professor
University of Minnesota

Ayman Shamma
HCI Research Group
Yahoo! Research

“...Maps have the possibility of becoming as important a means of human computer interaction as our current GUI paradigm.”

— Tom Erickson (IBM)
Introduction to Modern Geography

1. Learn about the academic discipline of Geography.

Illustrative Deep Dives into Geography

2. Cartography, Space/Place, and Geostatistics.

Open Discussion

3. What are your geographic questions?

Learning Objectives

- Get the “lay of the land” of modern geography and its sub-fields.
- (Briefly) learn about the history of modern geography.
- Understand a small sample of key theories and frameworks from modern (human) geography.

Countries With Poor Decision-making Skills

- 13 Attendees
- 2 Attendees
- 1 Attendee
In many countries, geography education reduces to the null set (the U.S., in particular).  

“geographic vocabulary” (most countries besides UK, Germany)

Place names are to geography as dates are to history.

Place names are to geography as vocabulary words are to literature.

(Modern) Geography can be defined as:

"... a broad discipline that essentially seeks to understand and study the spatial organization of human activity [, the environment,] and people’s relationships with their environment. It is also about recognizing the interdependence among places and regions, without losing sight of the individuality and uniqueness of specific places.”

(Moseley et al. 2009)
Learning Objectives

- Get the “lay of the land” of modern geography and its sub-fields
- (Briefly) learn about the history of modern geography
- Understand a small sample of key theories and frameworks from modern (human) geography

Geography is old...
(A ridiculously condensed timeline)

<table>
<thead>
<tr>
<th>~300BCE</th>
<th>Ancient Greeks doing crazy modern stuff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Ages</td>
<td>Arab and Chinese explorers build up our geographic vocabulary</td>
</tr>
<tr>
<td>1700-1800</td>
<td>Kant and Humboldt lay groundwork for academic geography</td>
</tr>
<tr>
<td>Early 1800s</td>
<td>First geography degrees in Europe</td>
</tr>
<tr>
<td>1960s</td>
<td>Modern Computer Science Begins</td>
</tr>
</tbody>
</table>

Modern Geography Begins
Modern Geography
(A ridiculously condensed timeline)

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>early 1800s</td>
<td>First Geography Departments in Europe</td>
</tr>
<tr>
<td>late 1800s</td>
<td>UK Comes On the Scene</td>
</tr>
<tr>
<td>1905</td>
<td>AAG Formed</td>
</tr>
<tr>
<td>1950s</td>
<td>The decline</td>
</tr>
<tr>
<td>1960s</td>
<td>The quantitative revolution</td>
</tr>
<tr>
<td>2000s</td>
<td>The rebirth</td>
</tr>
</tbody>
</table>

Learning Objectives

- Get the “lay of the land” of modern geography and its sub-fields
- (Briefly) learn about the history of modern geography
  - Understand a small sample of key theories and frameworks from modern (human) geography

Key Question in Geography: What is where and why?

- Why is SF the way it is, and why is it here?
- Why are the Great Lakes what they are, and why are they here?
- Why is Toronto what it is, and why is it here?
- Why is NY what it is, and why is it here?
- Why is the U.S. / Canada Border here?

Key Theory / Framework #1
Central Place Theory

Explains the location of cities, towns, and hamlets (in a homogenous plain)

Theory published in 1933
Central Place Theory
Underlying Premise

“Each kind of product or service available to the dispersed population has its own threshold, or minimum number of potential consumers needed to support its supply.”

(Fellman et al. 2007)
The world is not a homogenous plane...

Reilly's Law of Retail Gravitation

\[
\frac{D_A^2}{D_B^2} = \frac{P_A}{P_B}
\]

\[
\frac{D_A^2}{D_B^2} = \frac{P_A}{P_B}
\]

\[
P_A = 1 \quad P_B = 1
\]

\[
P_A = 1 \quad P_B = 0.5
\]
Site v. Situation

**Site**
- McCarthy, AK

**Situation**
- Toronto, Ont.

Site and situation advantages/disadvantages can explain the prominence of a metropolitan area.

**Site Advantages**
- Moderate temperatures
- Protected harbor

**Situation Advantages**
- Nearby Gold Rush
- "Nearby" Asia
- Nearby Stanford
- Nearby Silicon Valley

**Site Disadvantages**
- Weather is not the best
- No room for expansion
- High real estate costs
- Earthquakes!

**Situation Disadvantages**
- Very far from Wall Street
- Very far from Europe

**Key Theory / Framework #2**

- Topography
- Climate
- Culture
- Religion

**Site**: refers to the specific properties of locale.

**Situation**: refers to the external relations of a locale.

To what and how well a place is connected.
Site v. Situation

What about your hometown?
Volunteers?

It is quite foolish to think advantages are permanent!

Gravity Models

Well-known geography theory most often rediscovered by folks in HCI and related fields

Explains interaction between places

Gravity Models

Goods
People
Social Ties
User-generated Content

Place A  Interaction  Place B
Key Theory / Framework #3

Gravity Models

\[ T_{ij} = k \frac{P_i P_j}{d_{ij}} \]

Basic form of the gravity model

Key Theory / Framework #3

Gravity Models

\[ T_{ij} = k \frac{P_i^\lambda P_j^\alpha}{d_{ij}^\beta} \quad T_{ij} = k \frac{P_i P_j}{d_{ij}} \]

More advanced form
Basic form

What do the other parameters do?

Learning Objectives

- Get the “lay of the land” of modern geography and its sub-fields
- (Briefly) learn about the history of modern geography
- Understand a small sample of key theories and frameworks from modern (human) geography
Introduction to Modern Geography
Learn about the academic discipline of Geography

Illustrative Deep Dives into Geography
Cartography, Space/Place, and Geostatistics

Open Discussion
What are your geographic questions?

Cartography
A super brief summary for people who make online maps all the time yet never took a course in cartography

Brent Hecht & David Ayman Shamma
CHI 2014

Even he uses maps…

Spatial Computing MOOC taught with Shashi Shekhar • Fall 2014
Learning Objectives

• Know the general idea of map projections and the plusses and minuses of Web Mercator

• Know some basic thematic cartography rules of thumb

(So many things, so little time!)

The World is 3D!
Maps are (generally) 2D!

Projection
Always introduce some distortions!

You can understand pretty much all projections with three things...

Globe
A Piece of Paper
Lightbulb
(Web) Mercator

Advantages
- Rectangular (Fits in Window)
- Great for sailing! (Designed for it!)

Disadvantages
- Drastically distorts shape as you move away from equator

How can we make things better?

Winkel Tripel projection

Balances and minimizes distortion across a number of different properties (area, angle, scale, etc.)
Local, local, local!

Lambert Conformal Conic
(Centered in North America)

Distorts Areas

Johann Heinrich Lambert
1728-1777
Where is this projection perfect?

Unprojected = :-(
Distorts everything (area, shape, angles) badly

UTM
Universal Tranverse Mercator

Lat, Lon != Linear Units!
Communication vs. Exploration


All this is made worse by the “McDonaldsification” of Reference Maps

Learning Objectives

- Know the general idea of map projections and the plusses and minuses of Web Mercator
- Know some basic thematic cartography rules of thumb

(St many things, so little time!)
The world of cartography can be split into two (overlapping) categories:

**Reference Maps**
- Purposes: Orientation and Navigation

**Thematic Maps**
- Purpose: Communicate nature of geographic distribution

- **Thematic Map Type**
  - **Choropleth**
    - **Spatial Distribution**
      - Discretized Population
      - Type: Categorial
    - Source: [FlowingData](http://flowingdata.com)

- **Isopleth**
  - **Spatial Distribution**
    - Type of Precipitation
    - Type: Categorial
  - Source: [Weather.com](http://www.weather.com)
Why do we map spatial distributions?
Why Thematic Cartography is Useful

[MacEachren 1983; Slocum et al. 2009]

Median Family Income
by U.S. state

1. Specific information
2. Regional information
3. Compare between maps

$47.3K - $56.2K
$56.2K - $61.4K
$61.4K - $65.6K
$65.6K - $71.0K
$71.0K - $86.3K

Data source: Department of Housing and Urban Development

Why Thematic Cartography is Useful

[MacEachren 1983; Slocum et al. 2009]

Obesity Rates
by U.S. state

1. Specific information
2. Regional information
3. Compare between maps
4. Compare between map and mental model

21.0% - 23.9%
24.0% - 26.0%
26.1% - 27.8%
27.9% - 30.4%
30.5% - 34.0%

Data source: Centers for Disease Control and Prevention

The "Map Trap"? An Evaluation of Map Versus Text-based Interfaces for Location-based Mobile Search Services

Karen Church, Joachim Neumann, Mauro Cheumbiri and Nuria Oliver
Telefónica Research
Via Augusta 177, 08028 Barcelona, Spain
karen@tid.es, joachim@tid.es, mauro@tid.es, nuria@tid.es

“...maps are not the interface of choice for certain information access tasks.”

Relatedness to concept “nuclear weapon”

Trad'l Geographic Reference System

Periodic Table reference system
Challenge #1: Choosing the correct visual variables

Data Source: U.S. Census 2010

Percent Latino Population

Data Type: Continuous (sequential) = Color Scheme: Categorical (qualitative)
Challenge #2: Choosing the correct class breaks

How did we determine these class breaks?

Data Source: U.S. Census 2010

Percent Latino Population

1.2-6.4%
6.4%-11.8%
11.8%-17.5%
17.6-29.6%
29.6%-46.3%

Quantile Classification

1.2-3.6%
3.6%-5.9%
5.9%-8.9%
8.9%-13.4%
13.5%-46.3%

Equal Interval

Data Source: U.S. Census 2010
Challenge #2: Choosing the correct class breaks

Thematic cartography is both an art and a science

Resources and tools...

ColorBrewer  QGIS  MapBox
Space and Place

*Space*

- position on surface of Earth
- distance and directions
- the basis of pretty much all spatial computing
- "GIS tends to favor features with precisely defined boundaries" (Elwood et al. 2013)
- permanent (roughly speaking)
Space and Place

Place

- planimetric control is not that important
- “humanized space” (Yi-Fu Tuan)
- emotional and personal
- often conceived of relationally
- temporary

Space and Place

CenturyLink Field

+ Anger, Frustration, Sad Memories
+ Happiness, Confidence, Comfort

Space and Place

Traditional Geographic Information Systems

“Volunteered Geographic Information” (VGI)

Space

“spatial information”

Place

“platial information”

(Elwood et al. 2013)
What's a Boundary? What's a Name?
Identity's in the Lines

What 'hood is this?
FINDING BOUNDARIES

La Lengua

http://www.flickr.com/photos/spierisf/8982696847
http://www.flickr.com/photos/telstar/5620478775

Burrito Justice
http://burritojustice.com/la-lengua/

La Lengua

https://www.flickr.com/photos/spierisf/8982696847
http://www.flickr.com/photos/telstar/5620478775

La Lengua

https://www.flickr.com/photos/spierisf/8982696847
http://www.flickr.com/photos/telstar/5620478775
We have a lot of geotagged photos

Almost ninety million, as I write this, and the numbers keep growing especially as nearly every new smart phone released to market has not only a camera but also the ability to capture location information with it.

For every geotagged photo we store up to six Where On Earth (WOE) IDs. These are unique numeric identifiers that correspond to the hierarchy of places where a photo was taken: the neighborhood, the town, the county, and so on up to the continent. This process is usually referred to as reverse-geocoding.

Over time this got us wondering: if we plotted all the geotagged photos associated with a particular WOE ID, would we have enough data to generate a mostly accurate contour of that place? Not a perfect representation, perhaps, but something more fine-grained than abounding box. It turns out we can.

So, starting today there are 150,000 (and counting) WOE IDs with proper (i.e.) shape data, available via the Places API. What kind of shapes, you ask?

---

<table>
<thead>
<tr>
<th>Place Type</th>
<th>Place Type Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continent</td>
<td>20</td>
<td>One of the major land masses on the earth. GeoPlanet is built on a seven-continent model: Asia (24865571), Africa (24865676), North America (24865672), South America (24865573), Antarctica (24865421), Europe (24865675), and Pacific/Australia, New Zealand, and the other islands in the Pacific Ocean – 24865674.</td>
</tr>
<tr>
<td>Country</td>
<td>12</td>
<td>One of the countries and dependent territories defined by the ISO 3166-1 standard.</td>
</tr>
<tr>
<td>Admin1</td>
<td>8</td>
<td>One of the primary administrative areas within a country. Place type names associated with this place type include: State, Province, Prefecture, County, Region, Federal District.</td>
</tr>
<tr>
<td>Admin2</td>
<td>9</td>
<td>One of the secondary administrative areas within a country. Place type names associated with this place type include: County, Province, Parish, Department, District.</td>
</tr>
<tr>
<td>Admin3</td>
<td>10</td>
<td>One of the tertiary administrative areas within a country. Place type names associated with this place type include: Commune, Municipality, District, Ward.</td>
</tr>
<tr>
<td>Town</td>
<td>7</td>
<td>One of the major populated places within a country. Place type names include: incorporated cities and towns, major unincorporated towns and villages.</td>
</tr>
<tr>
<td>Suburb</td>
<td>22</td>
<td>One of the subdivisions within a town. Place type names include: suburbs, neighborhoods, wards.</td>
</tr>
<tr>
<td>Postal Code</td>
<td>11</td>
<td>One of the postal code areas within a country. Place type names include: both full postal codes such as those in UK and US and partial postal codes. Examples include: SW1A 0AA (UK), 00101 (US), 095.0074 (FR).</td>
</tr>
<tr>
<td>Subname</td>
<td>19</td>
<td>A place that refers to a region consisting of multiple countries or an historical country that has been dissolved into current countries. Examples include: Scandinavia, Latin America, USSR, Yugoslavia, Western Europe, and Central America.</td>
</tr>
<tr>
<td>Colloquial</td>
<td>34</td>
<td>Examples are: France, French Riviera, Kansai Region, South East England, Pacific States, and Chubu Region.</td>
</tr>
<tr>
<td>Time Zone</td>
<td>31</td>
<td>A place that refers to an area defined by the Olson standard. Examples include: America/Los Angeles, Asia/Tokyo, Europe/Madrid.</td>
</tr>
</tbody>
</table>
Just like in Photoshop
Gaussian Blur

Saying too much with 1 slide

By analyzing check-in patterns throughout the city, we discover its structure.
Visual features can reveal 'aesthetics' and 'beauty'

Multimodal trip planning & analysis

OpenTripPlanner provides a range of passenger information and transportation network analysis features using our infrastructure for finding times for combining transit, pedestrian, bike, and car segments.

OpenTripPlanner   OTP Analyst
Time and Temporality

A question for discussion...

Historic Photos of the Western Addition in San Francisco.

All photos © Dave Glass https://flic.kr/p/41rpm1 https://flic.kr/p/dFg5J3 https://flic.kr/p/drhcW
Where were these photos taken?

Make All Regions as Space-Time Regions
Learning Objectives

- Geographic information can lie about **statistical independence**
- Geographic information can like about **precision**
- Geographic information can lie about **direction**
- Geographic information can lie about **time**

First Law of Geography

Aka “Tobler’s First Law”

“Everything is related to everything else, but near things are more related than distant things”

- Waldo Tobler (1970)
Quantitative manifestation of Tobler’s First Law

Spatial autocorrelation violates the independence assumptions of many statistical tests

A fundamental property of geographic information

Twitter Users Per Capita

Percent Urban Population

First Law of Geography

Quantitative manifestation of Tobler’s First Law

Spatial autocorrelation violates the independence assumptions of many statistical tests

A fundamental property of geographic information

rs = 0.46, p < 0.0001*

n = 3,143
Moran’s I Measure of Autocorrelation

\[ I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2} \]

Figure 3. Participants in a study found maps like those on the bottom (i.e. with high spatial autocorrelation) significantly more interesting than those on the top (i.e. with low autocorrelation).

Resources and tools...

GeoDa

R Packages

Geospatial Analysis

de Smith, Goodchild, Longley

Free online textbook!
Learning Objectives

- Geographic information can lie about statistical independence
- Geographic information can like about precision
- Geographic information can lie about direction
- Geographic information can lie about time

How would you visualize 1.2 Million Photos?

Ok here’s what you got

“Photo_Id”、“Long”、“Lat”

JUST put it on a map?

Answer
So where's the most popular point?

Command line Magic
GPS delivers a bit of noise.

HorseShoe Falls

DISCOVERING POINTS OF INTEREST MEASUREMENTS & COMPOSITION

http://www.flickr.com/photos/ayman/9205723524
DISCOVERING POINTS OF INTEREST
LOCALIZATION algorithm
Who’s time is it anyway?
GPS will tell you it loves you then lie to you in the following sentence.

• User’s ‘hometown’ may not reflect ‘where they are from’. (Contributor-contribution distance)

• Photos have several timestamps. Photo taken time, photo uploaded time, GPS timestamp. GPS stamps are often off or drift from location for various reasons (power, accuracy, etc.).

• The photo taken time (often set by camera phones automatically) and the GPS time differ by at least 10 minutes in 40% of all camera phone photos and as much as 12+ hours in 10% of the cases (from a sample of 10M public photos)