

Asking Spatial Questions to Identify GIS Functionality

Song Gao

Department of Geography and Center for Spatial Studies
University of California, Santa Barbara
CA 93106-4060
Email: sgao@geog.ucsb.edu

Michael F. Goodchild

Department of Geography and Center for Spatial Studies
University of California, Santa Barbara
CA 93106-4060
Email: good@geog.ucsb.edu

Abstract—Current desktop-GIS software cannot answer users' spatial questions directly. The GIS functionality is hard to identify and use without specific training of GIS skills because of the complex hierarchical organization and the gap between users' spatial thinking and systems' implement descriptions. In order to bridge this gap, we propose a semantic framework for designing a question-based user interface that integrates different levels of ontologies (spatial concept ontology, domain ontology and task ontology) to guide the process of extracting the core spatial concepts and translating them into a set of equivalent computational or operational GIS tasks. We also list some typical spatial questions that might be posed for spatial analysis and computation. The principle introduced in this paper could be applied not only to desktop-GIS software but also to web map services. The semantic framework would be useful to enhance the ability of spatial reasoning in web search engines (e.g. Google semantic search) and answering questions in location-based services as well (e.g. iPhone Siri assistant).

Keywords—GIS; functionality; spatial questions; ontology; semantic

I. INTRODUCTION

Spatial questions that relate to space, place, and location often reflect the concepts that humans use in spatial thinking and reasoning, e.g. "How far is it from city A to city B?" and "Which areas will be affected by Hurricane Sandy?" Compared with other systems, one of the distinct characteristics of geographic information systems (GIS)[1] is its ability to ask questions and get answers in a spatial context. However, the user interface of existing GIS software cannot support this directly. The company Esri defines desktop-GIS software as "mapping software that is installed onto and runs on a personal computer and allows users to store, display, query, update, and analyze data about geographic locations and the information linked to those locations."¹ Spatial analytic functionality is the nucleus of GIS software, but many existing analysis tools are hard to be identified and used for non-expert users without specific training of GIS software skills because of the complex hierarchical organization of functionalities. In addition, the service-oriented architecture (SOA) on the cloud is emerging as a new paradigm of computing, in which tasks are performed by chaining together on-line web services rather than by operations at the desktop. Popular web map services such as ArcGIS Server, MapServer, and Google Maps API, use SOA to offer fruitful functionalities such as mapping,

geocoding and hot-spot analysis for the user. Since not every user has the knowledge or the experience of translating spatial thinking questions into GIS functionality, it might be useful to design a friendly interface that captures users' spatial questions and identifies corresponding functional services based on their spatial purposes instead of a simple keyword-based search on the web browser. In general, keywords are good for matching the names of functions but not for matching connections between the questions and the functionalities.

One of the co-authors has served the "spatial help desk" that is open to the entire campus community and offers research consultation service at the Center for Spatial Studies at the University of California, Santa Barbara (spatial@ucsb)². This "spatial help desk" provides support for requests ranging from simple cartographic design to complex, multi-steps spatial information modeling and analysis. The most important step in this GIS consultation is to understand these spatial questions and then identify which analysis tools or functions in GIS software should be adopted. However, the problem is that current GIS interface is isolated from the user's conceptualization of the problem. The following is a collection of questions that have received technical support from the "spatial help desk":

- Overlaying a topographical basemap for cartographic design (Department of History and Religious Studies)
- Joining updated survey tables to existing spatial data (Department of Psychology)
- Assessing crime patterns near various categories of media sites (Department of Communication)
- Delineating information flow for disaster mapping (Department of Geography)
- Identifying dangerous bridges with respect to pedestrian traffic (Department of Political Science)
- Providing a batch process of converting geographic features in different folders into new geodatabase classes (Library)

In the past five decades of development the GIS field has made significant progress in standardizing data formats for interoperability but still lacks standardized or formalized functionalities that integrate well with the cognition process

¹<http://support.esri.com/en/knowledgebase/Gisdictionary/term/desktop-GIS>

²<http://www.spatial.ucsb.edu>

of the user. Goodchild (2011) has proposed to integrate the concepts of spatial thinking in the design of a new GIS user-interface and in organizing spatial analysis tools[2]. One of the alternative strategies to address this concern is to implement a question-based user interface that captures users' spatial questions and identifies the corresponding analysis tools as needed. This might also be useful to organize the help-documentation both for the end-user and for the developer.

It has been widely recognized that ontology plays an important role in knowledge organization and information integration. The use of ontologies in GIS development has been widely discussed in related research [3], [4], [5], [6], but relatively few works have addressed the design of the user interface based on semantic integration. Cuarino (1997) classified ontologies into four categories: top-level ontology (general concepts), domain ontology (vocabulary related to a domain), task ontology (a task or activity), and application ontology (a specific use or application)[7]. We consider that a task-ontology of spatial analysis should be an essential component of formalizing GIS functionality. Thus, we propose to integrate ontologies in the design of a new GIS interface to guide the process of parsing and extracting the core spatial concepts from users' questions and translating them into computational or operative tasks.

The remainder of the paper is organized as follows: in Section II, we briefly introduce some examples of mainstream organization structures of desktop-GIS functionality and web map services; in Section III, the taxonomies of GIS functions and operations are discussed; in Section IV, we propose a semantic framework for designing question-based user interface to identify GIS functionality as well as list some typical spatial questions that might be posed frequently for spatial analysis and computation; in Section V, we summarize the main conclusions and discuss future work.

II. EXISTING ORGANIZATION STRUCTURES OF GIS FUNCTIONALITIES AND WEB MAP SERVICES

We posit that the interfaces of GIS software and web map services provide us with a natural environment to identify fundamental spatial concepts. With the growing needs of users to explore more and more spatial concepts, developers have responded with various implementations, but without any overall design, any guiding principles of granularity, or any basic functional taxonomy. We use the examples of existing organization structures of desktop-GIS functionality and web map services to illustrate this argument.

A. ArcMap

Esri's desktop-GIS software ArcMap 10 includes hundreds of functional operations. Some of them are invoked through drop-down menus, while a much larger number (615) are invoked through the Toolbox (see Table I). Even though some of these headings in this hierarchical organization are distinct, it would be virtually impossible for a user to determine whether a given spatial concept would be regarded by Esri's designers as an instance of *3D analysis*, *analysis*, *geostatistical analysis*, *spatial analysis*, or *spatial statistics*. In addition, many categories overlap substantially (such as the same *Interpolation* tools exist both in *3D Analyst Tools* and in *Spatial Analyst*

Tools; and the same label *Overlay* contains different functions in *Analysis Tools* and in *Spatial Analyst Tools*).

TABLE I. ORGANIZATION OF GIS FUNCTIONS IN ARCMAP 10 TOOLBOX

Title	Count of functions
3D Analyst Tools	34
Analysis Tools	19
Cartography Tools	43
Conversion Tools	46
Data Interoperability Tools	2
Data Management Tools	178
Editing Tools	7
Geocoding Tools	7
Geostatistical Analyst Tools	22
Linear Referencing Tools	7
Multidimension Tools	7
Network Analyst Tools	21
Parcel Fabric Tools	4
Schematics Tools	5
Server Tools	14
Spatial Analyst Tools	171
Spatial Statistics Tools	26
Tracking Analyst Tools	2
Total	615

B. GRASS

GRASS is an open-source GIS software used for geospatial data management and analysis, image processing, graphics/maps production, spatial modeling, and visualization. The organization of GRASS functionality is driven by data format. The drop-down menus and the menu tree list different functions categorized by raster, vector, imagery, volumes, and database (see Fig.1). This interface relies on the assumption that the user has a good sense of data format and knows exactly which operation is suitable for their needs. However, this is not always the case for non-expert users who are interested in asking questions and finding spatial patterns from their datasets but have little knowledge about GIS data formats and the taxonomy of functionality. Therefore, we suggest that the software should offer a simplified user interface, remove the distinction between different operations in various categories, and allow the user to start to work at a higher conceptual plane. Note that GIS software could offer multiple user interfaces rather than replace the existing one.

C. Google Maps API

Google Maps API is a powerful web-map library that supports developers to embed the Google Map Object in web applications, to build location-based apps, and to visualize geospatial data. The structure of online documentation provides a view to look at the functionality of the map services. Take the Google Maps JavaScript API as an example, it uses map, controls, overlays, services (such as geocoding, transit), map types, layers, street view, events, base, and libraries (such as geometry, place, drawing, weather, and visualization) to classify its functions. The categories are not clearly defined and tend to overlap. For example, it is hard to cognitively differentiate the functions in *google.maps.Polygon/Polyline class* and in *google.maps.geometry.poly namespace*, because

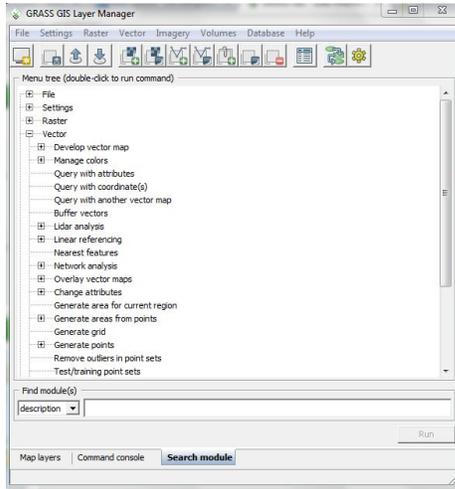


Fig. 1. The interface of GRASS GIS.

the latter also contains functions for computations involving polygons and polylines. It is not a problem of having multiple classifications, depending on the expertise of the user and the nature of the application. But again, as the functionality in web-map service increases, we need a clearer principle to organize them. In general, we do not think there might be perfect classification system, but instead the user interface could be improved with interactive suggestions and the best-practice of question examples.

III. TAXONOMIES OF GIS FUNCTIONS

Simple questions should be answered directly by using the analysis tools, whereas complex questions may need several consecutive operations to be connected together for problem-solving. Golledge (1995) stated that both simple and complex spatial concepts can be derived by using primitives of spatial knowledge (identity, location, magnitude and time)[8]. However, the user does not know which tools should be adopted unless he or she can link the intention from a problem domain to a set of specified spatial analysis tools. This process is most likely to be depending on the user's own memory. Unfortunately, it is hard for non-experts or new users who are not familiar with GIS software to establish these connections because of the convoluted classification of functionalities. Some progress has been made to classify GIS functionalities; for example, Dangermond (1983) described a classification of commonly used GIS components which include: map automation, analytic manipulation, database manipulation and graphic manipulation techniques[9]. Berry (1987) presented four classes of primitive GIS operations by organizing a set of analytic methods into a mathematical structure[10]. Goodchild (1988) grounded his taxonomy of functional operations in GISs database based on formalizing the object representation of spatial data[11]. Rhind and Green (1988) classified GIS functions based on data operations (data input and encoding, data manipulation, data retrieval, data analysis, data display, database management)[12]. Albrecht (1998) proposed a task-orientated systematization (search, location, terrain, distribution/neighborhood, spatial analysis, measurements)[13]. Good-

child (2001)³ introduced a six-way conceptual classification of spatial analysis (query and reasoning, measurement, transformation, descriptive summaries, optimization, hypothesis testing). However, there is no single taxonomy of functionality which is ideal in all situations, thus we still need to make efforts on capturing the user's spatial questions being asked. Kuhn (2012)[14] proposed a set of ten core concepts of spatial information (location, neighborhood, field, object, network, event, granularity, accuracy, meaning, value), aiming at giving meaningful demonstration to scientists and researchers who are otherwise not specialists with regard to spatial information.

After five decades of development in the GIS field, the user interface of GIS software remains cognitively complex and lacks standard terms for functionalities. Therefore, we need to build a semantic bridge between spatial thinking and computer implement. The solution we propose here is to design a new GIS user interface based on spatial questions. The framework contains mainly two parts: (1) a list of spatial questions; and (2) a semantic interpreter incorporating different levels of ontologies to identify which analysis tools might be helpful.

In addition, the analysis tools in GIS software are data-format sensitive, most functions are designed for specified data models, e.g. vector (V), raster (R), TIN (T), and network (N). We could take data types as filters when the user selects these questions. Because of the length limit on this paper, we will focus on the design of the conceptual framework as below.

IV. THE DESIGN OF QUESTION-BASED USER INTERFACE FOR GIS FUNCTIONALITY

A. A semantic framework of the user interface

The semantic framework of question-based GIS user interface at the vertical level can be divided into three layers (see Fig.2): the question layer, the semantic layer, and the implementation layer. The framework uses different levels of ontologies to guide the process of extracting the core spatial concepts from questions and translating them into a set of computational or operative tasks to invoke GIS functions. A detailed description for these components on this framework is as follows:

1) *Question layer*: The interface allows users to ask spatial questions based on their thoughts or to select some question examples listed on the user interface.

2) *Semantic layer*: It uses natural language processing (NLP) methods to parse questions to the core spatial concepts by referring to the *spatial concept ontology*. It also employs specific domain ontology to understand the descriptive structure of questions for capturing the user's purpose. The output of the semantic layer is a single task or a workflow of multi-step tasks that need to be implemented to answer users' questions.

3) *Implementation layer*: It resorts to possible GIS tools or functions to accomplish the task on desktop-GIS software. Also, it could provide a set of uniform and simple RESTful style geo-processing web services in the age of cloud computing[15].

³http://www.csiss.org/learning_resources/content/good_sa/

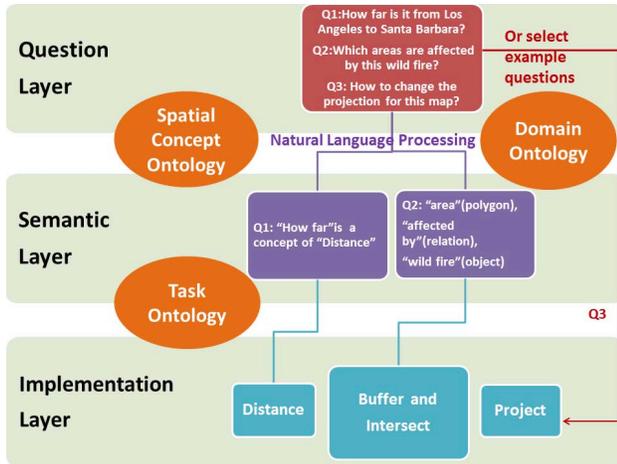


Fig. 2. A semantic framework of question-based user interface for GIS.

4) *Spatial concept ontology*: It describes the ontology design pattern for spatial concepts based on Kuhn's spatial information theory[14] and Albrecht's universal elementary GIS functions[13]. It presents a semantic net of GIS functionality, represents common properties and connections between spatial concepts. A collection of the core spatial concepts includes location, neighborhood, field, object, network, event, distribution, connectivity, proximity, adjacency, pattern, direction/orientation, hot-spot, relation, overlay, buffer, and join etc.

5) *Domain ontology*: This part could be linked to domain ontologies existing elsewhere in other collections and standard formats, such as the ontology of USGS land-use and land-cover classification which support a remote sensing image processing in a GIS environment.

6) *Task ontology*: It describes the combined actions to complete a specific task and presents relationship between different functions, e.g. in hydrology GIS application, the function to generate watershed only can be executed if the user has finished the preliminary function named as *Flow Direction* based on the digital elevation model (DEM) dataset.

Fig.2 also shows three examples of asking questions in this new user interface and how this framework support to answer them.

Question 1: How far is it from Los Angeles to Santa Barbara? This is a simple question. The semantic layer recognizes that the word "how far" is a concept of distance and thus provides the distance function for the user.

Question 2: Which areas are affected by this wild fire? This is a cognitive question at a higher level that requires some knowledge about spatial relations and the environment (like the spread speed, direction of the fire). A possible solution for this question is firstly to generate a buffer polygon that represents a potential affected area based on the domain knowledge. Then it operates another GIS function "intersect" to the target layer (such as U.S. county map layer) to get the influenced areas.

Question 3: How to change the projection for this map? This is a simple question and could be listed on the drop-down menu of question examples that can be automatically performed with

specified function. Such questions can be answered directly at the implementation layer without semantic interpretation.

Modern GIS software is a multidisciplinary tool that should allow both expert and non-expert users to utilize it for discovering spatial patterns and for sharing geographic knowledge. This principle introduced above could be applied into not only desktop GIS software but also web map service. The question-based user interface bridges the gap between the user's spatial thinking and the computer's implementation or programming language in describing GIS functionality.

B. List of examples of spatial questions

Although we cannot enumerate all the spatial questions in various domains, we could try to identify the typical spatial questions that would be interesting to most of the users. There are numerous ways of classifying them, e.g. Nyerges (1991) suggested that geographic questions can be categorized into those dealing with: location and extent, distribution and pattern or shape, spatial association, spatial interaction; and spatial change[16]. By extending this classification, we group some typical spatial questions that might be posed frequently and identify their corresponding analysis tools in ArcMap which is taken as the mainstream commercial desktop-GIS software. This could be helpful for developers to design the question-based user interface. We would also list corresponding functions for open-source GIS tools and for web map services in future work.

- Search/Location and Extent

1) *Where is the location for a place name, or for a ZIP code:* -Query table attributes

2) *Which features on the map have specified attribute values:* -Query table attributes

3) *Which features lie in a specified area:* -Select by Location

4) *How many households are there in the selected area:* -Frequency or Summary Statistics

5) *Which street intersection is closest to a specified building:* -Locate Features Along Routes

6) *What is the shortest path from an origin to a destination:* -Route, OD Cost Matrix (Network Analyst Tools)

- Data Basics/Processing/Conversion

7) *How to convert census tract data into raster or grids:* -Polygon to Raster

8) *How to create points from the geotagged photos in my iphone:* -GeoTagged Photos To Points

9) *How can we convert GIS-format data to CAD format for urban planning:* -To CAD

10) *How to visualize map layer in Google Earth:* -Map to KML, Layer to KML

11) *How to symbolize this layer by using the same symbology from another map:* -Apply Symbology From Layer

12) *How to extract a remote sensing image or land-use raster in this study area:* -Extract by Mask

- 13) *How to reclassify a land-use raster:* -Reclassify
- 14) *How to assign different values to each pixel in a raster:*
-Raster Math
- Distributions/Patterns/Neighborhood
- 15) *How to summarize the locations of a set of features (e.g. places of interest):* -Central Feature
- 16) *Where are the clusters of an epidemic:* -Density
- 17) *Is the distribution of points dispersed or clustered:*
-Average Nearest Neighbor
- 18) *What kind of distribution does it look like:* -High/Low Clustering
- 19) *Which city is the closest to Santa Barbara:* -Near
- 20) *Where are the nearest storage facilities:* -Cost Distance and Closest Facilities
- Relations/Associations
- 21) *What areas are affected by freeway noise:* -Line Buffer and Intersect
- 22) *What soil types occur in this area:* -Spatial Join
- 23) *What schools or supermarkets are there within 1km of my home:* -Buffer and Overlay
- 24) *How to combine different polygons into one layer:*
-Union or Merge
- Terrain/Surface
- 25) *What is the slope of this terrain:* -Slope
- 26) *What is the aspect of this terrain:* -Aspect
- 27) *Where are the contours in this area:* -Contour
- 28) *Where is the skyline from this point:* -Skyline
- 29) *Which area is visible from a specific location or from a set of observer features:* -Viewshed
- 30) *How to interpolate a surface from the sampling points:*
-IDW, Kriging, Natural Neighbor, Spline, and Trend
- 31) *Where are the drainage basins in this area:* -Flow Direction and Basin
- Time
- 32) *How did Hurricane Sandy move across the Northeast:*
-Make Tracking Layer
- 33) *How do GPS-enabled taxis move around the city:* -Make Tracking Layer and Track Intervals To Line

V. CONCLUSION

The contribution of this research is two-fold. First, it presents a semantic framework for designing a new GIS user-interface driven by spatial questions. Second, it lists some typical spatial questions that might be posed frequently and the corresponding analysis functions in ArcMap. Both of them offer insights on the design of the next generation of semantic user interface for accessing GIS functionalities.

In future work, we will employ artificial intelligence in the development of new user interface to answer users spatio-temporal questions. Furthermore, a catalog of spatial concepts⁴ developed by spatial@ucsb will also be integrated into our semantic framework. This semantic framework with different level of ontologies would be useful to enhance the ability of spatial reasoning in web search engines (e.g. Google semantic search) and answering questions in location-based services as well (e.g. iPhone Siri assistant).

ACKNOWLEDGMENT

The authors would like to thank the colleagues in the Department of Geography at UC Santa Barbara and the reviewers for their insightful comments.

REFERENCES

- [1] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, *Geographic Information Systems and Science (Third Edition)*. Hoboken, NY: Wiley, 2011.
- [2] M. F. Goodchild, "Spatial thinking and the gis user interface," *Procedia-Social and Behavioral Sciences*, vol. 21, pp. 3–9, 2011.
- [3] F. T. Fonseca, M. J. Egenhofer, P. Agouris, and G. Cmara, "Using ontologies for integrated geographic information systems," *Transactions in GIS*, vol. 6, pp. 231–257, 2002.
- [4] D. Mark, M. Egenhofer, S. Hirtle, and B. Smith, "Ontological foundations for geographic information science," *A research agenda for geographic information science*, pp. 335–350, 2004.
- [5] Y. Gao, S. Gao, R. Li, and Y. Liu, "A semantic geographical knowledge wiki system mashed up with google maps," *Science China Technological Sciences*, vol. 53, no. 1, pp. 52–60, 2010.
- [6] K. Janowicz, "Observation-driven geo-ontology engineering," *Transactions in GIS*, vol. 16, no. 3, pp. 351–374, 2012.
- [7] N. Guarino, "Semantic matching: Formal ontological distinctions for information organization, extraction, and integration," in *Information Extraction A Multidisciplinary Approach to an Emerging Information Technology*, M. Pazienza, Ed., 1997, pp. 139–170.
- [8] R. G. Golledge, "Primitives of spatial knowledge," in *Cognitive aspects of human-computer interaction for geographic information systems*. Springer, 1995, pp. 29–44.
- [9] J. Dangermond, "A classification of software components commonly used in geographic information systems," in *Design and Implementation of Computer-Based Geographic Information Systems*, D. J. Pequet and J. O'Callaghan, Eds., 1983, pp. 70–91.
- [10] J. K. Berry, "Fundamental operations in computer-assisted map analysis," *International Journal of Geographical Information System*, vol. 1, no. 2, pp. 119–136, 1987.
- [11] M. F. Goodchild, "Towards an enumeration and classification of gis functions," in *International Geographic Information Systems Symposium: The Research Agenda II*, 1988, pp. 67–77.
- [12] D. Rhind and N. Green, "Design of a geographical information system for a heterogeneous scientific community," *International Journal of Geographical Information System*, vol. 2, no. 2, pp. 171–189, 1988.
- [13] J. Albrecht, "Universal analytical gis operations: A task-oriented systematization of data structure-independent gis functionality," *Geographic information research: transatlantic perspectives*, pp. 577–591, 1998.
- [14] W. Kuhn, "Core concepts of spatial information for transdisciplinary research," *International Journal of Geographical Information Science*, vol. 26, no. 12, pp. 2267–2276, 2012.
- [15] S. Gao, H. Yu, Y. Gao, and Y. Sun, "A design of restful style digital gazetteer service in cloud computing environment," in *Proc. 18th International Conference on Geoinformatics*. IEEE, 2010, pp. 1–6.
- [16] T. L. Nyerges, "Analytical map use," *Cartography and Geographic Information Science*, vol. 18, no. 1, pp. 11–22, 1991.

⁴<http://teachspatial.org>