Isaac Emrick
Geog 752: Advanced GIS
12/9/2008

Functional Limitations of Current GIS

The threshold of functionality of a Geographic Information System is continually being pushed farther and farther to include ever increasing numbers of functions yet there remain some major areas of limitation within the software available. Computer technology has inherent problems dealing with reality stemming from the reliance on binary coding at the most basic levels. As a level of abstraction computers have problems approximating nuance and human logic such as the ability of the human brain to discern patterns visually. These functional limitations fall into three basic categories: Multi-dimensionality, Analysis, and Data capture. These correlate closely to Michael Goodchild's "Grand Challenges." Each of these fields pose significant dilemmas within the software framework of the GIS that are currently being explored to find better solutions.

I. Multi-dimensionality

GIS most often represent reality as a flat two-dimensional digital surface resembling paper maps that is visualized with the computer monitor. This does not capture or represent physical reality in a natural way that a user might recognize. One of the initial original reasons for this abstraction was a limitation on processing power. Despite the exponential increase in computer processing power has increased exponentially the ability to deal with multiple dimensions remains elusive. Beyond the standard x and y coordinates, multidimensionality suggests z, i.e. elevation. The level of abstraction is limited by visualizing the physical world in three-dimensions, as it more naturally depicts the world as users experience it. The depiction of a three-dimensional environment though is impeded by the technology used to view it. A computer monitor can produce two-dimensional flat images or two-and-a-half dimensional renderings of the three- dimensional model produced by the computer, this is due to the flatness of the monitor. True three-dimensionality can be achieved through the use of VR goggles or CAVES. These technologies though require a great deal of expense and thus are not available to most users.

Tools for handling and representing elevation have increasingly been incorporated into GIS, but it is not the only dimension that has been difficult to accommodate. Time has provided some of the most sophisticated challenges for GIS designers. Time is a consideration in all questions applied to the GIS. All data in a GIS is time specific, whether past, present, or in predicting the future, yet there is currently no intuitive way to deal with spatial data through time that maintains relationships through time. Things change continually through time. This not only increases the amount of data stored but causes problems of how to structure databases. Most commonly time is sampled at discrete moments, creating snapshots. This is a form of generalization that captures moments but does little to show the relationships between them. The key would be to create a topology of time but the complexity two-dimensional topology suggests that increasing it chronologically would exponentially increase the development and analytical workload.

The ability to interact with data temporally also becomes problematic. Even if a researcher has digitally formatted historically accurate data, which are extremely rare, there are issues with formatting the data such as how to order it. Are we starting from point A and moving forward or does the focus of the project start at the present and move backward. Gail Langran discussed such structural complexities when dealing with issues of spatio-temporal dimensional dominance. Even beyond this, Peuquet suggests there are three problems with temporal information which a computer has trouble managing. As mentioned above, the more “snapshots” of data the more the data volume increases. Time-consuming comparisons of each snapshots is the only way to gather information about the changes that occurred through time. If more than one change occurred between T1 and T2, this will be generalized in the analysis. Lastly, the location in time of any change can not be determined with this method. This produces fuzzy points in time that correlate to discrete location data. It is difficult to query time when events are ambiguously located. To cope with these issues, two methods of visualization have been used. The first take the snapshots and produces a time-series of maps. A more coherent method is the space-time cube whereby x and y are geographic locations and the z is time. Neither of these visualizations come much closer to true analytical tools. ESRI’s Tracking Analyst deals with time artificially as discrete moments is rather clunky, it does little to help with the explanations of the relationships between moments on time.

Much like the problems posed by the space-time cube, the visualization of GIS data have proven limiting. This area though has seen greater advancement than the issues of multidimensionality in the data structure. Visualization, though, has become much more than just a view technique. It has become a mode of exploration in itself. Paper maps are a very basic form of visualization that depict the world in abstract symbols and generalizations. Recent visualizations have capitalized on increasing computer graphics capabilities in creating ever more rich multi-dimensional images. The user becomes able to more naturally interact with the landscape in a GIS when viewing it in three-dimensions. This though returns the developer to issues of representation. The ability to represent the three dimensional world in a way that users understand requires art and skill. Another issue with rendering in three-dimensions is the issue of equipment limitations. In order to experience true three-dimensional environments, the user has to be immersed in a virtual reality. CAVES and VR/AR goggles place the individual into the environment but remain cumbersome devices to control at the user and programming level. There are severe limitations on user interaction with the data especially in a CAVE. The user cannot directly control or brush the visualization thus limiting their ability to find meaning in the images.

Though researchers always bring the real world into their work at a level of logical abstraction, multi-dimensionality poses both the challenge and potential of representing reality in a more natural manner within the GIS. This will require ever increasing computer processing power but also new wider development of ways to visualize the multiple dimensions of human perception and inquiry.

II. Exploratory Analyses

There are almost as many ways to draw meaning from the data placed within a GIS as there are users of the technology. Some of the traditional methods of statistical analysis though are flawed. There are issues of efficiency and utility with complicated techniques such as multiple regressions. Three methods have been suggested to deal with the deficiencies of the traditional methods. GIS users have begun to utilize the exploratory data analysis methods suggested by John Tukey in the 1970s. New forms of computer pattern analysis in the form of Neural Nets and Expert Systems have begun to influence the modeling within GIS technology. Lastly, a new geographic awareness and access to technologies like Google Earth have brought GIS-like analysis to a wider public audience. All have contributed to the development of more exploratory analysis of data sets outside of the standard framework of statistics.

Though the technique is about the same age as GIS, Tukey's Exploratory Data Analysis (EDA) has a vigor that traditional statistics lacks. This largely due to the reliance on the user's natural mental computer coupled with collaborative exploration of data to produce questions and lines of research without transforming the data. This has produced the offshoot ESDA, Exploratory Spatial Data Analysis. This is closely tied to the effective use of visualization mentioned above. Though EDA/ESDA techniques have begun to be applied to GIS data, the tools for displaying non-spatial attribute data are weak. Likewise, EDA/ESDA becomes immeasurable more complicated when dealing with the multi-dimensional reality of geography. The representation and exploration of multivariate data sets is relatively new and poorly supported by GIS software in general. Though spatial data is what GIS technology was designed to handle, the statistical analyses it performs are very basic.

Like EDA/ESDA, there are two similar program concepts that are increasingly being suggested for dealing with large amounts of complicated spatial data: Expert Systems and Neural Nets. These are both powerful pattern analysis programming packages that can manage complex GIS data but have some major functionality issues. Primary among these issues is that they are expensive and require a great deal of programming ability. This makes them inaccessible to all but the most experienced and well-funded GIS users. The power of these packages, especially Expert Systems, to create analysis processes and forecasting models is not integrated easily into GIS software like ArcGIS. One specific problem with Neural Networks comes from the inability of the user to manipulate the processing of data. Neural Nets are "black boxes" that turn inserted data sets into a product with rules, but the user does not know how this was done. Expert systems show the rule building process and circumvent this limitation.

Neogeography shares much with the underlying philosophy of EDA in that it suggests that you can with limited manipulation geo-reference data for a much wider audience than traditional GIS. With the advent of programs like Google Earth, non-GIS users were able access some of the most basic of the tools available in a GIS. It is even suggested that a user can produce 80% of the product of a GIS with 20% of the labor. This “new” geography has the benefit of increasing geographic awareness. The addition of application programming interfaces (APIs) further allowed the incorporation of GIS-like tools in a wider variety ways. This also aided in the development of user-generated data. Neogeography has some severe limitations in the larger realm of GIS. Slowly, ESRI is incorporating the use of neo-geo techniques but integration of tools like Google Earth is still poor. Beyond user-generated data and the use of Google Earth, there is a push towards indexing of geographic data in a manner that would allow for greater access. Until the vast amounts of data maintained at all levels of society are identified and accessible to GIS users, the full-functionality of the GIS will be unattainable. Neogeography poses a problem also from its relationship to the GIS community, despite its utility, the more analytical uses of a GIS are still absent from applications like Google Earth. This limits exploratory value of neogeography.

As Tukey suggested with EDA, these three methods attempt to tease out information from data sets with minimal amounts of transformation. The exploratory nature of each though is limited by the integration of those technique sin the larger framework of GIS software. This is most apparent with the complexity and steep learning curve of Expert Systems and Neural Networks. Coupled with the already difficult problems of visualization and multidimensionality, exploratory functionality in a GIS is limited by the vast amounts of data pouring in from the real world and how to appropriately manipulate that data for analysis.

III. Data capture and maintenance

Much like the rise of neo-geography for the handling of spatial information, the collection of data has also been brought to a wider audience. This has proven to be a double-edged sword of limitations of the GIS. The types of data available are still heavily biased by the operating rules of the three main organizations producing GIS data, namely the government, the military, and businesses. Getting new forms of data from users is problematic in both content formatting and maintenance. Individual users may represent and store data in very different ways that may not be compatible. Combining data indifferent formats is highly problematic. With addition of increasing amounts of data, researchers come into issues about storage and accessibility of these data sets. This can mean that different agencies may have multiple formats of basically the same data all being stored in different places not easily accessible to users.

Global Positioning System (GPS) devices have become a major part of modern society, from handheld devices to chips embedded into cell-phones. Increasingly, the public is using GIS-like tools and are largely unaware of it. The growth of GPS devices has opened up for more collaborative data gathering. The limitation here is in transferring this information to a GIS in a systematic and meaningful way. A user can get information into Google Earth from an iPhone but transferring and manipulating that information in a GIS is still complicated at best. GPS also pose a concern about the accuracy of the device. Off-the-shelf GPS devices are appropriate for locating houses, but are not suited for marking the corners of the house. The issue of scale is one that needs to be addressed when acquiring data for a GIS database.

To take this issue one step further, data is created for a purpose at a specific time and place, details that need to be available to users of GIS data. Though highly mundane, metadata is imperative for the proper use of data sets. Currently, metadata is rarely complete and this largely due to the difficulty of creating the metadata files. Formatting is equally difficult for metadata as different organizations handle the metadata differently, if they create it all. By standardizing the process of metadata and by limiting the types of information required, we can increase the creation of this data. Without the metadata, indexing of data sets is nearly impossible. The ability to search for data efficiently and effectively remains elusive. This limits the utility of the GIS by forcing individuals to often recreate data that could otherwise be located. Data management and creation seem almost rudimentary but they strike at the core of the GIS. Without good data sets, the GIS is useless.

IV. Overcoming the Limitations

As computer technology advances and processing power becomes available, many of the complicated technical issues of GIS analysis and data management may be solved. Increased processing power though is not going to overcome all of the limitations of GIS. Developers and researchers need to be creative to come up with solutions. Data management, analysis, and interoperability are three areas that the GIS can advance to overcome some of the limitations mentioned above.

The internet is a valuable tool for the management of data. The process has begun for the development of internet based applications for the finding and distribution of GIS data. Warehouses and portals allow users to connect to data yet there is no larger framework for searching for datasets. This could be achieved by instituting standards but enforcement becomes an immediate issue. This of course suggests that there might be a codified set of ethics addressing such matters. Closely connected to this are database management issues. Applying more logical database structures to data, though complicated help in the querying of data in meaningful ways. The sophistication of databases must match the level of abstraction a user is willing to accept. Finding ways to combine the power of Neo-geo applications and the GIS might also remove some of the limitations of both. Combining these two will also promote a community involvement and collaborative effort in the creation and maintenance of data sets.

The issue of analysis is a great deal more difficult overcome, especially in a multi-dimensional GIS. The avenue of research for tackling these issues appears to move away from temporal objects to features. Features have mutable attributes and locations and thus can change over time. Instead of taking time and viewing it as a series of snapshots, analysis needs to construct it as continuous. The efforts of dealing with temporality will be aided by the development and maintenance of a library of time-sensitive data sets, thereby creating a history of GIS data to work from. Another issue with time and multi-dimensionality is the problem of fuzziness. Events occur over a period of time and have lingering impacts, how these connections get inputted into a database need to be explored in greater detail. Events are not discreet they have blurred or fuzzy beginnings and ends as individual enter and leave. How the GIS deals with uncertainty is a frontier that computer developers have only recently begun to grapple with meaningfully.

Along the same lines of uncertainty and multi-dimensional analysis is the potential for developing semantic applications. To take a passage of text and tease out spatial information requires a complex understanding of the meaning of language. Semantic logic may be a way for computer technology to grapple with the uncertainty and mixed meanings of data. By determining an application that can understand the fuzziness of “behind” or “around” as much as more specific references to location, the GIS might be able to come up with a geometry and topology from text. This would be immeasurably powerful for humanities researchers dealing with text based information, i.e. historical documents. Semantic applications have also been suggested for use in the indexing of data sets.

GIS technology is still growing, albeit slightly behind the computer technology on which it is predicated. By utilizing both technical and social advancements for dealing with the limitations of the GIS, researchers may move the frontier of GIS towards an even more integrated and interdisciplinary field.

Bibliography

Anselin, L. "Interactive techniques and exploratory spatial data analysis." in Paul A. Longley, Michael F. Goodchild, David J. Maguire, and David W. Rhind, eds. *Geographical Information Systems: Principles, Techniques, Management and Applications, Vol. 1*. 2nd Ed. New York: John Wiley & Sons, 1999: 253-266.

Armstrong, Marc P. Temporality in Spatial Databases. 880-889.

Batty, Michael. "Virtual Reality in Geographic Information Systems." in S. Fotheringham and J. Wilson, eds. *Handbook of Geographic Information Science*, Blackwell, Oxford, UK: 317-334.

Belcher, R. C. and A. Paradis. A Mapping Approach to Three-Dimensional Modeling. 107-122.

Berners-Lee, Tim; James Hendler, Ora Lassila. The Semantic Web. 1-9.

Brimicombe, Allan J. Location-based Services and Geographic Information Systems. 581-595.

Butler, Declan. Virtual globes: the web-wide world. 1-7.

Charlton, Martin E. "Quantitative Methods and Geographic Information Systems." in John P. Wilson, and A Stewart Fotheringham, eds. *The Handbook of Geographic Information Science*. Blackwell Publishing, 2008: 379-394.

Chatfield, Chris. "Exploratory data analysis." *European Journal of Operational Research*, 1986, 23(1): 5-13.

Cox, N. J., and K. Jones. "Exploratory Data Analysis." in Neil Wrigley, Robert John Bennett, eds. *Quantitative Geography: A British View*. Routledge, 1981: 135-143.

Fischer, M. M. "Spatial analysis: Retrospect and Prospect." in Paul A. Longley, Michael F. Goodchild, David J. Maguire, and David W. Rhind, eds. *Geographical Information Systems: Principles, Techniques, Management and Applications, Vol. 1*. 2nd Ed. New York: John Wiley & Sons, 1999: 283-292.

Fisher, T. R. and R. Q. Wales. 3-D Solid Modeling of Sandstone Reservoirs Using NURBS. 39-41.

Fisher, T. R. and R. Q. Wales. Three Dimnesional Solid Modeling of Geo-objects using non-uniform rational B-splines (NURBS). 85-105.

Fonseca, Frederico. The Geospatial Semantic Web. 367-376.

Gahegan, Mark. Multivariate Geovisualization. 292-316.

Getis, A. "Spatial Statistics." in Paul A. Longley, Michael F. Goodchild, David J. Maguire, and David W. Rhind, eds. *Geographical Information Systems: Principles, Techniques, Management and Applications, Vol. 1*. 2nd Ed. New York: John Wiley & Sons, 1999: 239-251.

Goodchild, Michael F. Citizens as Sensors: The World of Volunteered Geography. 1-15.

Harris, Trevor, and John Boardman. Alternative Approaches to Soil Erosion Prediction and Conservation Using Expert Systems and Neural Networks. 461-477.

Hedley, Nicholas R., Mark Billinghurst, Lori Postner, Richard May, and Hirokazu Kato. "Explorations in the Use of Augmented Reality for Geographic Visualization." Presence, *Vol. 11, No. 2, April 2002,* 119 –133.

Houlding, Simon W. 3-D Component Modeling of Thin-Seam Deposits. 28-31.

Houlding, Simon W. 3-D Computer Modeling. 45-47.

Hutchinson, Michael F. Adding the Z Dimension. 144-168.

Jones, Christopher B. Data structures for three-dimensional spatial information systems in geology. 15-31.

Kirk, Keith G. Use of the OSM technical information processing system for three-dimensional analysis of overburden quality. 225-233.

Lange, A. F., and C. Gilbert. Using GPS for GIS data capture. 467-479.

Langram, Gail. A Conceptual Model of Cartographic Time. 27-54.

Langram, Gail. Temporal GIS Design Tradeoffs. 890-899.

Leipnik, Mark R. GIS and 3-D Modeling Fight Subsurface Contamination at Federal Site. 38-41.

Lima, En-Mi; and Tsuyoshi Honjo. Three-dimensional visualization forest of landscapes by VRML. *Landscape and Urban Planning* 63 (2003) 175–186.

Mahoney, Diana P. Mapping Toxic Spills.

Merrill, Duane. Mashups: the new breed of Web app. 1-8.

Nöllenburg, Martin. "Geographic Visualization." in A. Kerren et al, eds. *Human-Centered Visualization Environments 2006*. LNCS 4417: 257-294.

Openshaw, S. "Developing appropriate spatial analysis methods for GIS." in D. Maguire, M. F. Goodchild, D. Rhind, eds. *GIS Principles and Applications Volume 1*. London: Longman, 1991: 389-402.

Pack, Skip, and Gene Bressler. Modeling & Mapping: Three-Dimensional Property Data. 26-28.

Parsaye, Kamran, and Mark Chignell. Expert Systems for Experts. 29-117.

Parsons, Ed. The Role of Web Services for Spatial Data Delivery. 1-10.

Peuquet, Donna J. It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. 441-461.

Peuquet, Donna J. Time in GIS and geographical databases. 91-103.

Raper, Jonathan. "Multidimensional geo-representations for exploration," *Multidimensional Geographic Information Science*. CRC, 2000: 173-193.

Raper, Jonathan. Multidimensional representations of Space and Time. 85-119.

Raper. Jonathan. Multidimensional geo-representations for exploration. 173-193.

Raper. Jonathan. Multidimensional geo-representations for modeling. 121-171.

Sibley, David. *Spatial Applications of Exploratory Data Analysis*. Concepts and Techniques in Modern Geography No. 49. Norwich: Geo Books, 1988.

Skapura, David M. Building Neural Networks. 1-64.

Skupin, Andre, and Pragya Agarwal. Introduction: What is a Self-Organizing Map? 1-44.

Turner, A. Keith, and Kenneth E. Kolm. Three-Dimensional Geoscientific Information Systems for Ground-Water Modeling. 217-225.

Turner, A. Keith. The Role of Three-Dimensional Geographic Information Systems in Subsurface Characterization for Hydrogeological Applications. 115-124.

Turner, Andrew J. Introduction to Neogeography. 1-54.

Vasilopoulos, Audrey. Mapping Mine Fires: 3D graphics system helps OSMRE locate and track potential mine fire disasters quickly.

Yuan, May. Adding Time into Geographic Information System Databases. 169-184.