Transactions in GIS v6 n2, 179-194.

A user-oriented perspective of error-sensitive GIS development

Matt Duckham
Department of Computer Science
University of Keele
Keele, Staffordshire, ST5 5BG, UK
Ph. +44 1782 724270; Fax +44 1782 713082
Email: matt@cs.keele.ac.uk
Http: tabun.cs.keele.ac.uk/~matt

Keywords: error-sensitive GIS; decision support; software

Abstract

The current lack of error-sensitive functionality found in commercial GIS is at odds with the research focus error-sensitive GIS development has enjoyed over recent years. In an attempt to address this undesirable situation, this paper explores a number of different error-sensitive GIS research themes from the perspective of GIS users. Increasingly, GIS users need to fulfil a variety of different roles, from database designers and data capture personnel through to decision makers, each of which may require an awareness of data quality issues. The user-oriented approach taken here contrasts with the predominately developer-oriented perspective in the literature. The paper concludes that from the user's perspective it is possible to identify a number of simple reformulations of current research that would help bring about the desired increase in availability and use of error-sensitive functionality within commercial GIS technology and applications.

Acknowledgements

The majority of this paper relates to work conducted during a three year PhD at the Department of Geography and Topographic Science, University of Glasgow, supported by NERC and SDS, UK, and supervised by Jane Drummond (Glasgow), David Forrest (Glasgow) and John McCreadie (SDS). The author is currently employed on an EPSRC funded project, "Managing vagueness, uncertainty and granularity in spatial information", at the Department of Computer Science, University of Keele. Gothic GIS software was provided under a development license from Laser-Scan, UK, and the assistance of Kingston Communications, UK, is gratefully acknowledged.

1. Introduction

The endemic nature of error in geographic information is widely recognised, leading to the description of error as a "function of information" (*Goodchild, 1995*) and a "fundamental dimension of data" (*Chrisman, 1991*). Correspondingly, the development of what have been termed *error-sensitive GIS* (*Unwin, 1995*) has been long accorded a high priority in GIS research (see *Goodchild and Gopal, 1989*). Despite this high priority, the translation of spatial data quality research into viable technology has not been straightforward. While considerable effort has been expended on developing spatial data quality standards and computational models of spatial data quality, there is little evidence of such work filtering through to commercial GIS applications or technology.

Buttenfield (1993) identifies three impediments to representing spatial data quality: the definition and assessment, attribution in a database and graphical depiction of data quality information. In response to these three impediments, the error-sensitive GIS development process can be characterised as comprising three distinct stages: first, deciding upon the core data quality concepts; second developing and implementing an error-sensitive data model based on these concepts; third, developing interfaces able to deliver the error-sensitive services and functionality to users. However, by focusing on the development process, the role of the user in error-sensitive GIS research is often marginalised or neglected: users' requirements are an explicit component of the third and final development stage alone. This paper argues that the primary goal of any error-sensitive GIS is to facilitate the application and understanding of uncertain geographic information by users.

As a consequence, the three development stages are more appropriately viewed from a user-biased perspective. The three user-biased themes identified and investigated in this paper can be thought of as a mirror-image of the development process outlined above: error-sensitive GIS use; error-sensitive data models; concepts of data quality. The relationship between user- and developer-biased perspectives is represented in Figure 1, where increasing depth implies increasing levels of abstraction from the user's point of view. Following a review of existing research into data quality and GIS use (Section 2), this paper explores each of these three themes in turn with reference to an example error-sensitive GIS development process. While considerable further work remains, the outlook is relatively positive. The paper concludes that the goal of off-the-shelf error-sensitive GIS suitable for general use is certainly realisable, potentially in the not-too-distant future.

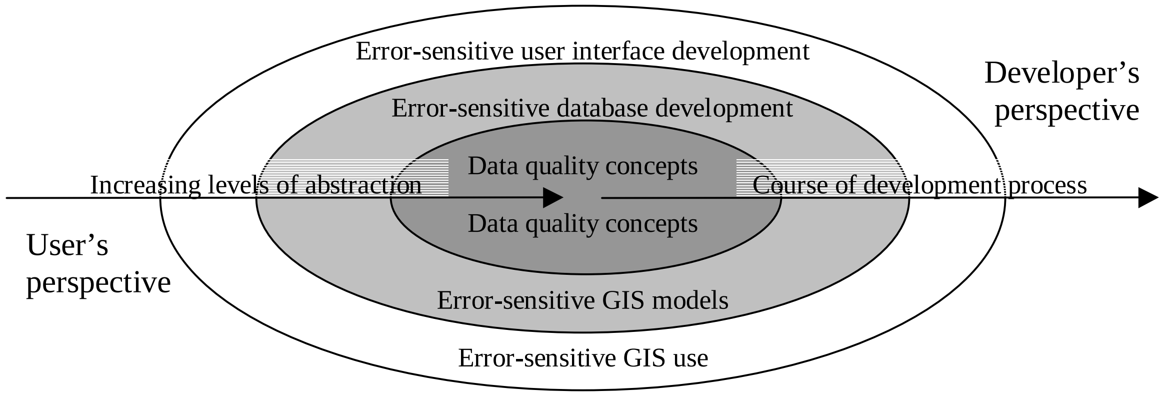

Figure 1, Different perspectives on error-sensitive GIS

2. Data quality and GIS users

An obvious criticism levelled at many spatial data quality standards and research articles is the emphasis placed on the storage, management and propagation of data quality information rather than how use such information, a point made strongly by *Hunter and Goodchild* (*1996*). As suggested above, the conflict between the developer's and user's perspectives may in part be to blame for this imbalance. One area where the developer's and user's perspectives are generally in sympathy is on the importance of the principle of *fitness for use* when using uncertain geographic information.

2.1 Fitness for use

The assessment of fitness for use requires that data producers supply enough information about the quality of a data set to enable a data user to come to a reasoned decision about the data's applicability to a particular situation (*Chrisman, 1991*). Fitness for use apportions and emphasises the responsibilities between data provider and data user. The data provider has a responsibility to provide explicit, appropriate information on uncertainty along with data. The data user has a responsibility to ensure the data is only applied to problems where such a use is apposite. Both the question of what constitutes 'appropriate' information provision and the question of whether a particular use is 'apposite' necessarily involve an element of subjectivity. Consequently, in contrast to the suggestion of some authors (eg *Agumya and Hunter, 1997*), it seems unrealistic to expect that the fitness for use of a data set can ever be assessed entirely objectively. Rather than a simple 'yes' or 'no' answer, 'fit' or 'unfit', the question of fitness for use will almost always be qualified by a degree of subjectivity.

2.2 User roles

In fact, data providers and data users are only two stereotypes of the many different roles modern GIS users are expected to adopt. *Beard and Buttenfield (1999)* note how different user types may demand different approaches to data quality issues. Database designers, decision-makers, data capture and processing personnel all have a stake in managing the quality of information in a GIS. Increasingly, a single organisation or even a single person may perform many of these different roles. The picture becomes further complicated by the various "information communities" that use GIS (*Open GIS Consortium, 2000*) and have traditionally adopted different approaches to the management of error in information. Photogrammetrists, geodesists and surveyors have highly developed techniques for increasing precision and minimising error (*Mikhail, 1978*) and for mathematically modelling error (*Chrisman, 1982*). Cartographers aim to control, represent and communicate levels of uncertainty (*Chrisman, 1991*). Computer science emphasises the enforcement of consistency within an information system (*Hunter, 1996*). Comprehensive management of spatial data quality depends on addressing data quality issues in the all the different information communities and user roles to which spatial data may be exposed during its lifecycle. Fitness for use, then, is best regarded not as a simple objective question but as a subjective process, whereby all those who have a stake in the management of geographic information discharge their responsibility to consider the quality implications of all the operations for which they use a GIS.

2.3 Visualisation of data quality

One significant area of work where the need to communicate data quality issues to the user has been tackled is in research into visualisation of spatial data quality. For example, *Beard and Mackaness* (*1993*) provide four requirements for developing spatial data quality user interfaces: the interface should be intuitive, offer scale dependent visualisations, highlight extremes of data quality and locate quality information both spatially and temporally. *Howard and MacEachren* (*1996*) offer some fundamental rules for the visual representation of data quality, while *McGranaghan* (*1993*) similarly explores visualisation of data quality from a cartographic perspective. Animation (*Fisher, 1993b*) and sound (*Fisher, 1994, Krygier, 1994*) have proved useful as more advanced tools for visualisation of spatial data quality, while *Hunter and Reinke* (*2000*) attempt to adapt familiar graphical user interface components, such as help balloons, pop-up boxes and “wizards”, for advising users of data quality issues.

While research into visualisation of spatial data quality is clearly a crucial component of error-sensitive user interfaces, visualisation techniques are primarily aimed at communicating spatial data quality to end users and decision-makers. There is, therefore, a need to provide flexible architecture able to offer assistance to all GIS users throughout the process of fitness for use, not just to end users. Seen in this broader context, visualisation is just one of a much wider range of tools than need to be supported. The following section looks in more detail at two examples of error-sensitive GIS use that cannot easily be resolved using visualisation, and explores the architecture necessary to support such use.

3. Error-sensitive GIS use

The picture of fitness for use painted in the discussion above is one of a partially subjective process in which a variety of users from different information communities must all cooperate in the management of data quality. Two requirements for error-sensitive GIS user interfaces to support the process of fitness for use follow. First, the heterogeneous approaches to spatial data quality adopted by different information communities and by different GIS user roles may demand that the range of error-sensitive GIS interfaces be similarly heterogeneous. Second, the partially subjective, context dependent nature of fitness for use may require flexible, intelligent user interfaces able to guide non-specialist users through different stages of data quality management.

The software developed during this research aimed to meet these two requirements by utilising two existing technologies. First, the need for heterogeneous user interfaces was tackled by adopting an open GIS architecture. An open GIS architecture can allow rapidly developed user interface clients to achieve very high levels of integration with core error-sensitive GIS database servers, at the same time as preserving a clear separation between user interface and database. Second, the need to support intelligent user interfaces was tackled through the use of artificial intelligence (AI). AI techniques allow domain specific user interfaces to target particular information communities' needs. The term *error-aware GIS* is used here to refer to the extension of an error-sensitive GIS using open, intelligent user interfaces. Two examples of the concept of error-aware GIS applications are given below.

3.1 Data quality capture

Based on experience gained whilst working with a telecommunications company (Kingston Communications, UK) and a data capture company (SDS, UK), the error-aware GIS interface shown in Figure 2 was developed to assist GIS users with the capture of data quality information. There exists a relatively high level of awareness of quality issues within the utility industry, especially with respect to the development of integrated multi-utility LIS.

However, for reasons of cost, lack of technology and expertise, extensive data quality information is rarely collected by utility companies. Consequently, the process of assessment of fitness for use may quickly stall. To counteract this tendency, a low cost mechanism for capturing data quality information alongside conventional data capture streams was needed. The tool in Figure 2 uses an artificial intelligence technique for learning by example, an *inductive learning algorithm* (see *Russell and Norvig, 1995*), to derive rules that describe the quality of a data set based on a small pilot quality assessment. Where particular patterns in the pilot data quality exist, for example if spatially dense features tend to be associated with lower accuracy, the algorithm is able to automatically derive rules to describe the patterns. Using these automatically learnt rules, the quality for an entire data set can be deduced from a relatively small quality assessment. In turn this allows organisations involved in data capture, such as Kingston Communications, to capture data quality information with the minimum of cost. A full discussion of the technique used is given *Duckham, Drummond and Forrest* (*forthcoming*).

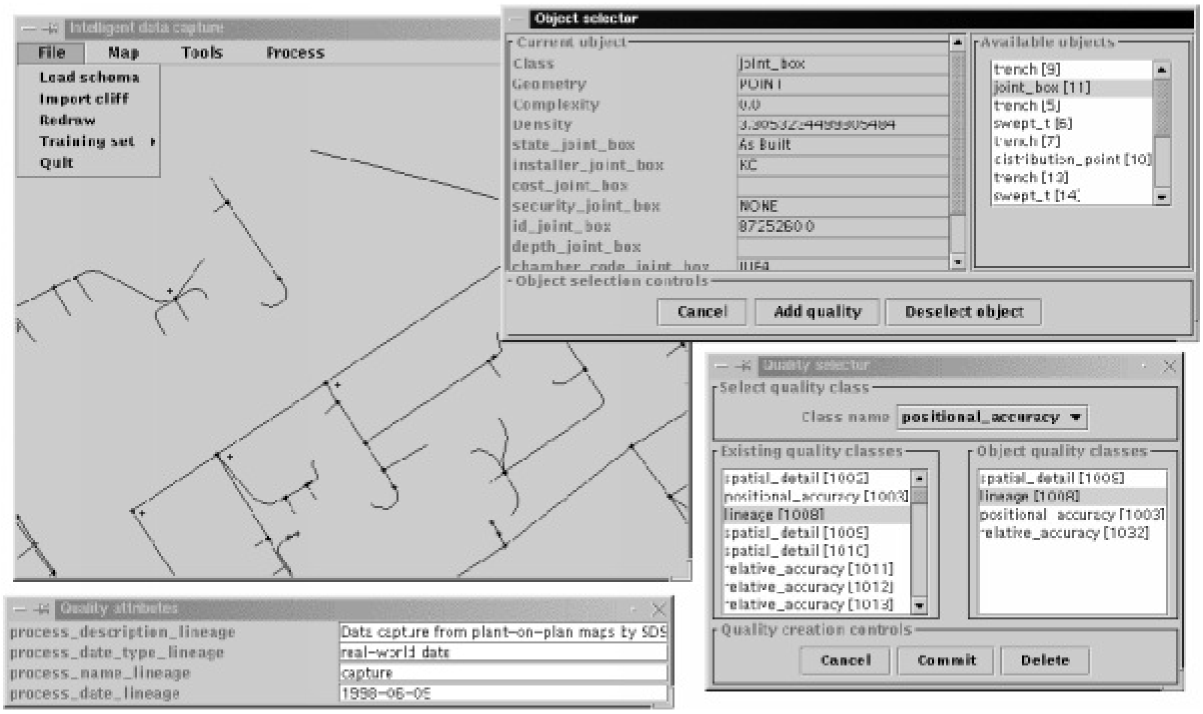

Figure 2, Error-aware inductive quality capture tool

3.2 Database design

The discussion above indicates some of the reasons why data capture organisations may need assistance with discharging their responsibilities with respect to fitness for use. Similarly, spatial database designers, unused to developing databases that incorporate both spatial information and data quality information, may also require user interfaces sensitive to their needs. Figure 3 shows the interface for a database design tool that can be used to integrate data quality classes into an existing spatial database design. The tool interface presents the database designer with a variety of questions about the type of spatial data with which the database is intended to be used. These questions are produced in hypertext markup language (HTML) by Java *servlets* (Java programs that can be integrated with a web server). The use of HTML is advantageous as it is the basis of the familiar point-and-click environment used across the Internet. The use of Java servlets allows the HTML to be dynamically generated, enabling a flexible interface that can change in response to questions already answered, further questions being triggered by the user's previous answers. The information harvested by the interface from the database designer is used to annotate the existing database design with appropriate data quality classes. The changes to the existing database design are effected using an expert system developed during the research using the Java Expert System Shell (JESS).

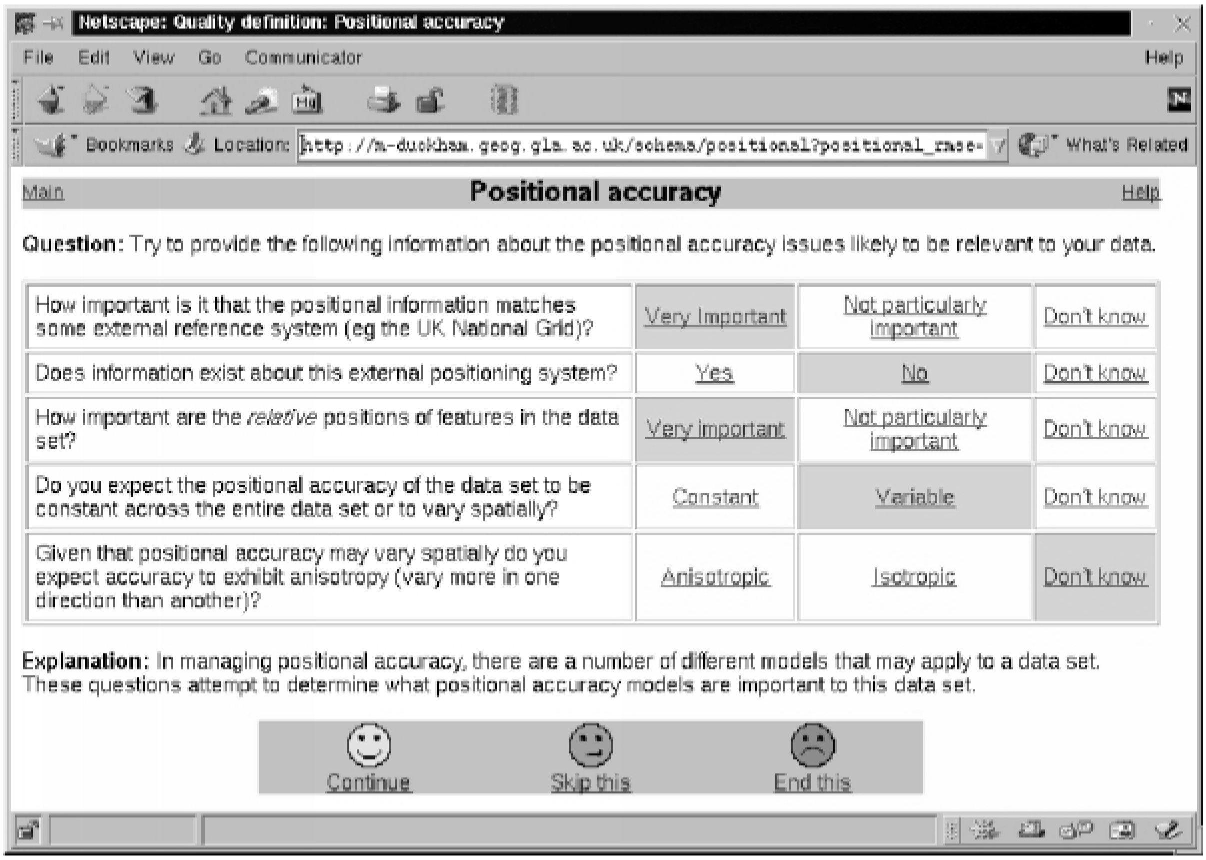

Figure 3, Error-aware database design tool

In addition to utilising artificial intelligence techniques such as inductive learning and expert systems, both the interfaces shown in Figure 2 and 3 use an open Java-based architecture to enable the user interface to achieve powerful yet flexible access to the underlying error-sensitive spatial data and functionality. The same underlying database, developed within Laser-Scan Gothic GIS, is used for both user interfaces, highlighting the efficacy of the open architecture. Communication between the user interfaces and the underlying database is achieved using the Java Remote Method Invocation (RMI) package which is a Java-only interoperable architecture similar to the Common Object Request Broker Architecture (CORBA).

The example of the two user interfaces described above indicates that the concept of an error-aware GIS is able to fulfil the key requirement of users of uncertain geographic information: support for the process of fitness for use. So far, however, very little has been said about the characteristics of the error-sensitive data models and databases needed to support such interfaces. The next section explores the demands that an error-aware GIS places upon the error-sensitive data model from the user-biased perspective.

4. Error sensitive data models

While a user's interaction with the data model employed by an error-sensitive GIS should be limited, the previous section highlighted that, even from a user-biased perspective, a consideration of the error-sensitive data model is still important (for example when users are engaged in database design). As mentioned above, the core error-sensitive functionality developed during this research was provided by a modified Laser-Scan Gothic GIS database via an open Java architecture. This section looks at some of the different design and implementation decisions surrounding error-sensitive GIS development and the importance of object-orientation (OO) to error-sensitive GIS.

4.1 Integrating data quality and geographic information

The laudable desire to maximise reuse of existing software during the development of error-sensitive GIS has often led to a dichotomous approach to data quality, where information about spatial data quality is stored and processed in a separate quality sub-system (see for example *Lanter and Veregin, 1992*; *Ramlal and Drummond, 1992*; *Wesseling and Heuvelink 1993*). While this approach may take advantage of a user's pre-existing familiarity with a GIS, it also has two important drawbacks. First it introduces additional conceptual and computational complexity. As already noted, error is an endemic, inseparable component of geographic information. Choosing a data model that artificially separates quality and spatial data entails additional conceptual and implementational structures to maintain the connections between spatial data and its quality. *Beard* (*1997*) comments that if data quality information is to remain useful in the face of increasing levels of data sharing, then it must be fully integrated with geographic information to which it refers. Second, separate quality sub-systems tend to lead to global or generalised quality information being applied to blocks of geographic information, for example on a per-layer basis (see *Lanter, 1991*). The inadequacies of such global, generalised data quality information are well documented, for example often leading to the spatial distribution of error being lost (*Fisher, 1998*). These two drawbacks of increased complexity and decreased specificity make it less likely that the dichotomous approach to spatial data quality will be able to meet the user's requirements for an error-aware GIS outlined above. Happily, this research indicated that it is possible to achieve the same high levels of software reuse through OO, without the need for a separate quality sub-system and its associated disadvantages.

4.2 Object-orientation and error-sensitive GIS

The central advantage of using OO in GIS is the superior semantic modelling capabilities when compared to other approaches, such as relational data models (*Worboys, Hearnshaw and Maguire, 1990*). Similarly with error-sensitive GIS, the complex, heterogeneous nature of uncertainty in geographic information can be more effectively modelled using OO than other techniques. From the user's perspective, OO systems should be able to decrease system complexity and better reflect users' intuitive expectations of spatial data quality as an integral component of spatial data at every level of the database.

The system developed during this research used OO to enable spatial objects to manage their own quality. Individual objects, from large aggregated geographic features down to primitive objects, like 'coordinate' or 'pixel', can be interrogated regarding their own data quality. By taking advantage of inheritance strategies available in all OO systems, pre-existing OO data models can be transformed by the addition of a handful of classes at strategic points in the inheritance hierarchy. Adding error-sensitive classes at or near the top of an existing OO inheritance hierarchy effectively transmits error-sensitive functionality to all objects in the database. In this way, data quality information can be integrated with the geographic information to which it refers enabling per-feature quality management without the need for extensive re-engineering of existing systems.

The idea is illustrated by the class diagram in Figure 4 below, which indicates how the classes in a hypothetical OO spatial database might be augmented with basic error-sensitive capabilities. In Figure 4, simple error-sensitive functionality is transmitted throughout a conventional OO GIS by inserting an "error-sensitive object" class, which is associated with classes of quality objects such as “positional accuracy” and “lineage”, at the top of the OO hierarchy. In practice, the task is somewhat more involved. A variety of different techniques were employed during the development of the Laser-Scan Gothic OO database used in this research to enable more sophisticated error handling and error propagation routines to be used (see *Duckham, submitted*).

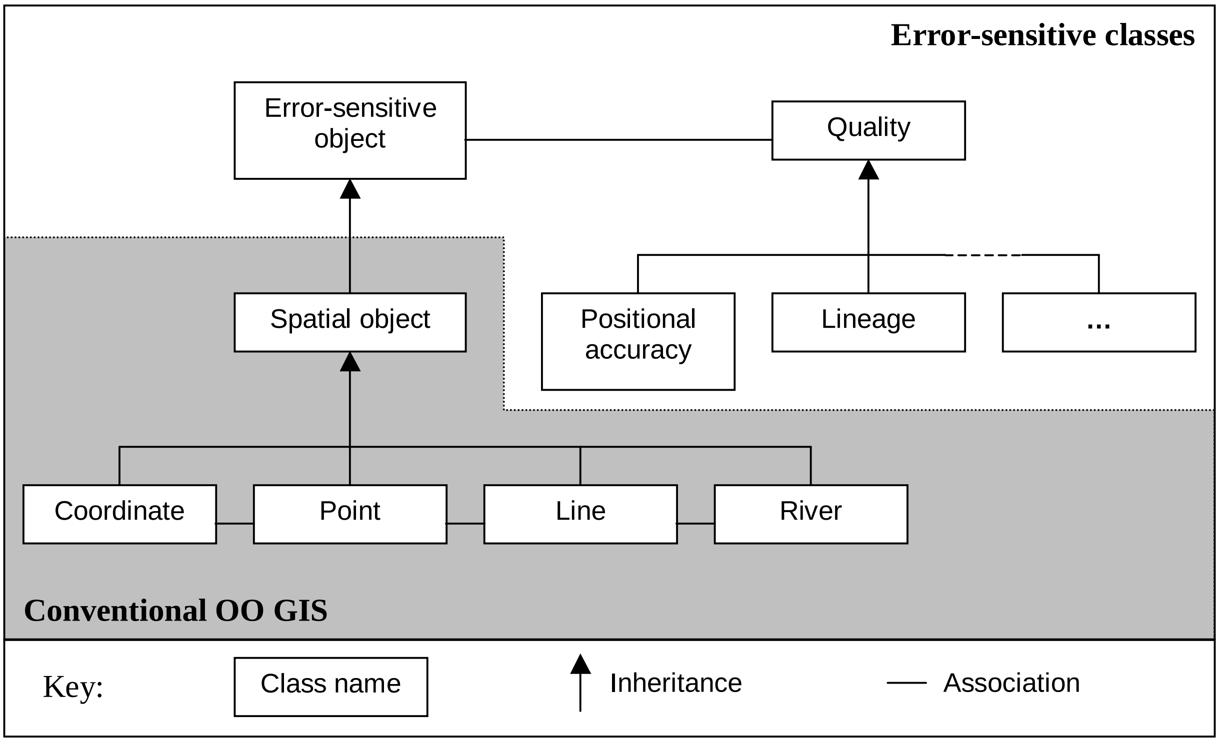

Figure 4, Example error-sensitive OO data model

By taking advantage of the generally hierarchical nature of geographic objects in OO GIS, recognised by *Fisher* (*1993a*), it is possible to achieve further efficiency gains. Rather than store quality information for every object in the database, individual objects can infer quality where appropriate from their associated or aggregate objects (*Qui and Hunter, 1999*).

4.3 Extending the OO model

An important advantage of adopting an OO error-sensitive data model is the potential for integrating data quality storage and management with data quality processing and propagation. *Fisher* (*1998*) and *Heuvelink* (*1998*) have noted the importance of retaining original observations when attempting to use and propagate error through subsequent operations upon the data. Some authors have gone further and propounded the concept of a measurement-based GIS (M-BGIS), where original survey measurements are retained to allow subsequent adjustment of the surveyed features in the light of resurveys (*Campbell,* *Carkner and Egesborg, 1994*; *Goodchild, 1999*). Such M-BGIS can be seen as a logical progression from OO error-sensitive GIS (*Duckham, submitted*). The error-sensitive GIS developed during this research was able to infer quality via association or aggregation relationships with other geographic objects. Correspondingly, a geographic object in an M-BGIS should be able to calculate its accuracy using its relationship to the underlying survey measurements upon which the object is based. Currently M-BGIS remain more of a concept than a reality (*Wright and Goodchild, 1997*), so further development of the approach outlined may find a role in the development of an OO M-BGIS. In a similar vein, a more immediate possibility is the use of encapsulated error models (*Goodchild, Shortridge and Fohl 1999*), where specialised error propagation processes are encapsulated alongside geographic objects. Arguably, encapsulated error models are also congruent with the OO approach advocated here.

5. Concepts of data quality

The discussion above highlights the importance of an integrated approach to spatial data quality in error-sensitive GIS, and champions the use of OO as an efficient route to achieve this integration. This section tackles the underlying concepts of spatial data quality needed to support the data models and user interfaces described in the previous two sections. The underlying data quality concepts form the most abstract level of the user oriented perspective of error-sensitive GIS proposed in Figure 1, and are consequently only likely to be of marginal interest to most users. However, these core concepts are important to consider as they shape both the error-sensitive data models and error-aware user interfaces. Traditionally, spatial data quality standards have formed the basis of most approaches to spatial data quality (eg *Guptill, 1989*; *Ramlal and Drummond, 1992, van der Wel et al., 1994*). Increasingly, GIS research is turning to more varied concepts of data quality as a basis for error-sensitive GIS development. The remainder of this section critically evaluates the data quality concepts appropriate for use in a user-oriented error-sensitive GIS.

5.1 Data quality standards

For many years now, research into spatial data quality has been held in thrall by spatial data quality standards. The domination of the ‘famous five’ elements of spatial data quality (lineage, completeness, consistency, positional and attribute accuracy; *NCDCDS, 1988*) over research into spatial data quality exemplifies this phenomenon. However, the close relationship between standards and research into spatial data quality has not necessarily been a healthy one. Data quality standards are undoubtedly an important area of endeavour that can help to promote and disseminate a snapshot of current best practice in the fast moving profession of GIS. Arguably, research into spatial data quality has been hindered by an unquestioning acceptance of previous standards-based work, for at least two reasons. First, research in GIS has thrived on the heterogeneity of concepts evident across the discipline. It seems unreasonable to expect standards to be flexible enough to deal with the highly heterogeneous requirements of different GIS users anticipated in Section 2.2. Evidence for this heterogeneity comes from the lack of agreement on exactly what elements of data quality should be standardised (see *Moellering, 1997*) and on the actual definition of many data quality elements across disciplines (*Drummond, 1996*).

Second, quantitative data quality elements, such as precision and accuracy, have tended to promote the stochastic model of error and data quality at the expense of other conceptualisations. The stochastic model of error assumes that observations are samples drawn from a population of possible observations. As a result of natural variation, the population of observations has predictable characteristics, such as mean and standard deviation, under central limit theorem. The stochastic model has several important advantages. It is well understood, provides detailed information about error and there exists a well developed body of work and statistical tools based on the stochastic model, even for spatial error (eg *Goodchild and Hunter, 1997; Heuvelink, 1998; Journel and Huijbregts, 1978*). However, from a user's perspective *Hunter and Goodchild* (*1996*) highlight some important disadvantages of the stochastic model. First, statistics are generally far from intuitive, and non-specialist users may have difficulty understanding statistical concepts. Further, the stochastic model demands a number of sophisticated assumptions that often limit or complicate its applicability to particular problems, for example when dealing with spatially autocorrelated data. Increasingly, alternatives to the stochastic model are being used. Fuzzy set theory is now a well established approach to certain types of uncertainty in geographic information (eg *Leung, 1987; Burrough and McDonnell, 1998*). Similarly, rough set theory is beginning to find a variety of applications to geographic information (eg *Alqvist, Keukelaar and Oukbir, 2000; Duckham, Mason, Stell and Worboys, accepted*). *Worboys and Clementini* (*accepted*) explore a variety of different logics that can be used in different contexts to integrate imperfect geographic information. The stochastic model of error is likely to continue to form the mainstay of spatial data quality handling for the foreseeable future, as it remains the most well developed and well understood model. Nevertheless, a much broader approach to uncertainty is called for than that offered by data quality standards. Consequently, error-sensitive GIS development should aim to be compatible with, rather than based upon current data quality standards.

5.2 Flexible data quality architecture

To avoid over-reliance on any particular data quality standard, the approach taken in this research was to define a set of core properties for data quality information from an OO perspective. Conventionally, geographic information is thought of as the product of a process of first abstraction and then representation of reality (eg *Maguire and Dangermond, 1991*; *Veregin, 1999*; *Worboys, 1992*). This essentially realist epistemology works well when discussing geographic information, but needs extending when dealing with spatial data quality. In contrast to geographic information there is no meaningful concept of data quality in the 'real world'; it is only as a by-product of the deficiencies of abstracting and representing reality that data quality arises as an issue at all (*David, van den Herrewegen and Salgé, 1996*). Correspondingly, an OO analysis of data quality which focussed on modelling the effects of the flawed, incomplete process of abstraction and representation was used as the basis for developing the Laser-Scan Gothic error-sensitive database outlined previously. Arguably, by focussing on the underlying processes that produce data quality, the resulting data model should be flexible enough to accommodate any data quality element. The analysis results were fed in to the OO error-sensitive data model described in Section 4.2. The results were used to define the handful of classes that provide the core error-sensitive functionality, simplistically represented in Figure 4 by the classes "error-sensitive object" and "quality" (see *Duckham and Drummond, 1999* for more information).

The approach proved successful in that it was able to support a range of different spatial data quality standards, including Spatial Data Transfer Standard (*US Geological Survey, 1999*), Spatial Archive and Interchange Format (*Geographic Data BC, 1999*) and the proposed draft European spatial data quality standard (*CEN/TC287, 1996*). At the same time, users were not obliged to adopt one particular standard, and were free to adapt existing standards for their own needs or even develop their own customised data quality elements. The ability to allow users to define their own data quality elements is a function exploited by the error-aware database design tool described in Section 3.2. Allowing data quality elements to be tailored to the specific demands of a particular spatial data set ensures the data quality requirements of heterogeneous information communities and user roles can be met. The importance of such flexible data quality concepts can be illustrated by analogy. Clearly, GIS would never have achieved the widespread usage they currently enjoy if they had only offered a standard set of geographic objects that could be modelled, say roads, trees and buildings. Similarly, error-sensitive GIS cannot be expected to be successful as long as they restrict users to a standard set of data quality objects, such as the NCDCDS five elements of spatial data quality.

6. Conclusions

This paper has reported on work that aims to promote the adoption of error-handling capabilities more widely within GIS applications. The failure to give enough weight to the *use* rather than the *development* of error-sensitive GIS technology has presented several barriers to the technology making the transition from research project to commercial functionality. However, this work indicates that it is possible to address these barriers by re-emphasising the importance of the user in error-sensitive GIS. First, the combination of an open GIS architecture and intelligent user interfaces represents an efficient mechanism for delivering error-sensitive functionality to a wide range of non-specialist GIS users. Second, the development of an integrated approach to spatial data quality, based on OO ideas and technology, is able to support the error-sensitive database functionality needed by users, without requiring extensive re-engineering of existing GIS. Finally, a widening of the debate surrounding the different conceptualisations of data quality should produce error-sensitive GIS that are better equipped to meet the requirements of the heterogeneous information communities that use GIS.

Many difficulties remain. The techniques adopted during this research were developed in response to the needs of industrial companies, such as Kingston Communications. However, despite encouraging indications the research has yet to be applied to and fully evaluated within a practical commercial setting. Further, this paper has emphasised the importance of data quality issues throughout the different stages of GIS use. Current research is biased towards data quality visualisation for decision makers. In future, as much research effort should be directed towards other, currently neglected sections of GIS users as has been directed toward decision makers in the past. Section 3 describes two tools developed to offer assistance with data quality issues to users not catered for by existing visualisation techniques, but considerable further work in this direction is necessary. The most encouraging conclusion drawn from this work is that the error-sensitive GIS functionality needed to address the highly heterogeneous and specialised requirements of potential error-sensitive GIS users can be developed within the framework of existing GIS technology. Arguably, by relying on a reformulation of existing GIS technology, rather than the development of new technology, the ultimate goal of error-sensitive GIS in practical application may become a reality within a much shorter time scale.

References

- *Agumya, A., Hunter, G., 1997,* Determining fitness for use of geographic information, ITC Journal, pp. 109–113.
- *Ahlqvist, O., Keukelaar, J., Oukbir, K., 2000*, Rough classification and accuracy assessment, International Journal of Geographical Information Science, 14, pp. 475–496.
- *Beard, K., Mackaness, W., 1993*, Visual access to data quality in geographic information systems, Cartographica, 30, pp. 37–45.
- *Beard, K., 1997,* Representations of data quality, in Geographic Information Research: Bridging the Atlantic, Craglia, M. and Couclelis, H., Eds., Taylor and Francis, pp. 280–294.
- *Beard, K.,Buttenfield, B., 1999*, Detecting and evaluating errors by graphical methods, in Geographical Information Systems, Longley, P., Goodchild, M., Maguire, D. and Rhind, D., Eds., John Wiley, pp. 219–233.
- *Burrough, P., McDonnell, R., 1998,* Principles of geographical information systems, Clarendon, 2nd Edition.
- *Buttenfield, B., 1993,* Representing data quality, Cartographica, 30, pp. 1–7.
- *Campbell, G., Carkner, L., Egesborg, P., 1994,* A GIS-based multipurpose digital cadastre for Canada lands, FIG Congress XX, Melbourne.
- *CEN/TC287, 1996*, Draft European standard; geographic information – quality, technical report prEN 287008, European Committee for Standardisation.
- *Chrisman, N., 1982,* A theory of cartographic error and its measurement in digital data bases, in Proc. Auto-Carto 5, pp. 159–168.
- *Chrisman, N., 1991*, The error component in spatial data, in Geographical information systems, Maguire, D., Goodchild, M., Rhind, D., Eds., Longman, pp. 165–174.
- *David, B., van den Herrewegen, M., Salgé, F., 1996,* Conceptual models for geometry and quality of geographic information, in Geographic objects with indeterminate boundaries, Burrough, P., Frank, A., Eds., Taylor and Francis, pp. 193–206.
- *Duckham, M., submitted*, Object calculus and the object-oriented analysis and design of an error-sensitive GIS, GeoInformatica.
- *Duckham, M., Drummond, J., 1999*, Implementing an object-oriented approach to data quality, in Innovations in GIS 6, Gittings, B., Ed., Taylor and Francis, pp. 53–64.
- *Duckham, M., Drummond, J., Forrest, D., forthcoming,* Spatial data quality capture through inductive learning.
- *Duckham, M., Mason, K., Stell, J., Worboys, M., accepted*, A formal approach to imperfection in geographic information, Computers, Environment and Urban Systems.
- *Drummond, J., 1996,* GIS: the quality factor, Surveying World, 4(6), 26–27.
- *Fisher, P., 1993a*, Conveying object-based meta-information, in Proc. Auto-Carto 11, pp. 113–122.
- *Fisher, P., 1993b*, Visualising uncertainty in soil maps by animation, Cartographica, 30, pp. 20–27.
- *Fisher, P., 1994*, Hearing the reliability in classified remotely sensed images, Cartography and Geographic Information Systems, 21, pp. 31–36.
- *Fisher, P., 1998*, Improved modeling of elevation error with geostatistics, GeoInformatica, 2, pp. 215–233.
- *Geographic Data BC, 1999,* Spatial archive and interchange format release 3.2 formal definition, Accessed 21 December, http://www.env.gov.bc.ca/gdbc/saif32/.

- *Goodchild, M., 1995*, Sharing imperfect data, in Sharing geographic information, Onsrud, H., Rushton, G., Eds., Rutgers, pp. 413–425.
- *Goodchild, M., 1999*, Measurement-based GIS, in Proc. International Symposium on Spatial Data Quality, Shi, W., Goodchild, M., Fisher, P., Eds., pp. 1–9.
- *Goodchild, M., Gopal, S., 1989*, Eds., Accuracy of spatial databases, Taylor and Francis.
- *Goodchild, M., Hunter, G., 1997*, A simple positional accuracy measure for linear features, International Journal of Geographical Information Science, 11, pp.299–306.
- *Goodchild, M., Shortridge, A.M., Fohl, P., 1999*, Encapsulating simulation models with geospatial data sets, in Spatial accuracy assessment, Lowell, K., Jaton, A., Eds., Ann Arbor, pp. 123–130.
- *Guptill, S., 1989*, Inclusion of accuracy data in a feature based object oriented data model, in Accuracy of spatial databases, Goodchild, M. and Gopal, S., Eds., Taylor and Francis, pp. 91–98.
- *Heuvelink, G., 1998*, Error propagation in environmental modelling with GIS, Taylor and Francis.
- *Howard, D., MacEachran, A., 1996*, Interface design for geographic visualisation: tools for representing reality, Cartography and Geographic Information Systems, 23, pp.59–77.
- *Hunter, A., 1996*, Uncertainty in information systems, McGraw Hill.
- *Hunter, G., Goodchild, M., 1996*, Communicating uncertainty in spatial databases, Transactions in GIS, 1, pp.13–24.
- *Hunter, G., Reinke, K., 2000*, Adapting spatial databases to reduce information misuse through illogical operations, in Proc. 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Heuvelink, G., Lemmens, M., Eds., pp. 313–319.
- *Journel, A., Huijbregts, Ch., 1978*, Mining geostatistics, Academic Press.
- *Krygier, J., 1996*, Sound and geographic visualisation, in Visualisation in Modern Cartography, MacEachren, A. and Taylor, D., Eds., Elsevier, 149–166.
- *Lanter, D., 1991*, Design of a lineage-based meta-data database for GIS, Cartography and Geographic Information Systems, 18, pp. 255–261.
- *Lanter, D., Veregin, H., 1992*, A research paradigm for propagating error in layer-based GIS, Photogrammetric Engineering and Remote Sensing, 58, pp. 825–833.
- *Leung, Y., 1987*, On the imprecision of boundaries, Geographical Analysis, 19, pp. 125–151.
- *McGranaghan, M., 1993*, A cartographic view of spatial data quality, Cartographica, 30, pp. 8–19.
- *Maguire, D., Dangermond, J., 1991*, The functionality of GIS, in Geographical information systems, Maguire, D., Goodchild, M., Rhind, D., Eds., Longman, pp. 319–335.
- *Mikhail, E., 1978*, Observations and least squares. IEP Dun Donnely.
- *Moellering, H. (Ed.), (1997)*, Spatial database transfer standards 2: characteristics for assessing standards and full descriptions of the national and international standards of the World, Elsevier.
- *NCDCDS, 1988*, The proposed standard for digital cartographic data, American Cartographer, 1, pp. 11–142.
- *Open GIS Consortium, 2000*, OpenGIS guide, Accessed 13 March, http://www.opengis.org/techno/guide.htm.
- *Qui, J., Hunter, G., 1999*, Managing data quality information, in Proc. International Symposium on Spatial Data Quality, Shi, W., Goodchild, M., Fisher, P., Eds., pp. 384–395.

- *Ramlal, B., Drummond, J., 1992*, A GIS uncertainty subsystem, in Archives ISPRS Congress XVII, 29.B3, pp. 356–362.
- *Russell, S., Norvig, P., 1995*, Artificial intelligence: a modern approach. Prentice Hall.
- *Veregin, H. 1999*, Data quality parameters, in Geographical information systems, Longley, P., Goodchild, M., Maguire, D., Rhind, D., Eds., Longman, pp. 177–189.
- *Unwin, D., 1995*, Geographical information systems and the problem of error and uncertainty, Progress in Human Geography, 19, pp. 549-558.
- *US Geological Survey, 1999*, SDTS home page, Accessed 19 December, http://mcmcweb.er.usgs.gov/sdts/.
- *van der Wel, F., Hootsmans, R., Ormeling, F., 1994*, Visualization of data quality, in Visualization in modern cartography, MacEachren, A., Taylor, D., Eds., Pergamon, pp. 313–331.
- *Wesseling, C., Heuvelink, G., 1993*, Manipulating qualitative attribute accuracy in vector GIS, in Proc. Fourth European Conference on Geographical Information Systems, pp. 675–684.
- *Worboys, M., 1992*, A generic model for planar geographical objects, International Journal of Geographical Information Systems, 6, pp.353-372.
- *Worboys, M., Clementini, E., accepted*, Integration of imperfect spatial information, Journal of Visual Languages and Computation.
- *Worboys, M., Hearnshaw, H., Maguire, D., 1990*, Object oriented data modelling for spatial databases, International Journal of Geographical Information Systems, 4, pp. 353–372.
- *Wright, D., Goodchild, M. 1997*, Data from the deep: implications for the GIS community, International Journal of Geographical Information Science, 11, pp. 523–528.