# Cartography, Database and GIS: Not Enemies, but Allies!

**Paul Hardy & Thierry Kressmann**
ESRI, Redlands, California, USA
phardy@esri.com,tkressmann@esri.com

**ABSTRACT**

Commercial GIS software such as ESRI ArcGIS has historic strengths in geography, spatial data modeling, and data analysis, but has traditionally been perceived as less strong in cartographic representation, artistic freedom and map publishing. However, a set of major software advances in cartographic functionality has recently become available, which together with further developments under way, will greatly automate high quality cartographic production, while empowering the human cartographer.

This paper overviews a related set of technology advances arising from research and development at ESRI. The aim is to provide the optimal tools and environment for the production cartographer, centered on the rigor of the master geodatabase but allowing artistic freedom where needed. It will release cartographers from the drudgery of repetitive actions and free them to concentrate on applying their unique human visual abilities for interpretation and design.

## 1 INTRODUCTION

### 1.1 Context

The majority of commercial cartographic publishing today relies on using file-based graphics software to create the visual product and to carry out cartographic edits. This is typically either a desktop graphics software package (such as Macromedia Freehand or Adobe Illustrator) or a specialist graphic map finishing system (such as Star Mercator or Lorik Dry). Almost always, the source line, point, and area geographic data for this cartographic process comes from a Geographic Information System (GIS), such as ArcGIS [ESRI 2004B], which maintains the data in a spatial relational database (such as Oracle or SQL Server).

While this split workflow has the possible advantage of using best-of-breed tools, there are a number of drawbacks. Having to export then import data to transfer it between packages is time consuming, and duplicating changes and updates in both environments is inefficient and expensive. This approach is further limited by not providing a WYSIWYG connection throughout the process, by making feature attributes unavailable during finishing, and by the need to maintain separate databases to make multiple products at different scales.

### 1.2 Requirements

Many national mapping agencies and commercial mapping companies have a strategic goal of using a common database and common environment for all map publishing, and a 'capture once and use many times' ethos. This requires that it be:

- A single software environment from capture to finishing
- Centered on  an enterprise database
- Supporting multiple representations for multiple products
- Capable of generating high-quality cartographic outputs
- Extensible to handle generalization and incremental update as these advanced facilities mature

### 1.3 The challenge of database versus freedom

The introduction of a GIS and database at the heart of the map production process has many advantages that are covered in more detail below. In particular it facilitates data sharing, process automation, and the handling of updates. Most cartographic production agencies have adopted a GIS database environment because of these strengths. However, in the past, the GIS database approach to cartography was criticized from two fronts:


ISBN: 0-958-46093-0

- **Not being comprehensive enough**: storing only geographic features in the database, with all other associated information such as symbology and map marginalia being kept in files outside the database.
- **Not being flexible enough**: being bound by simplistic symbolization rules based on feature class properties or geographic attributes.

### 1.4 Concepts

The solution to the above requirements and challenges necessitates the introduction of new concepts and capabilities into the GIS/cartographic software:

- cartographic representations
- overrides and exceptions
- cartographic editing tools
- cartographic data models
- generalization and product derivation
- production automation and optimization
- automation of map series and atlas page layout
- graphic control and visual output

For some years ESRI has been researching the marriage of geography with cartography, the extension of the GIS geodatabase [ESRI 2004C] to handle map presentation and cartographic details, and the fusion of database with freedom [ESRI 2004]. These concepts are described in detail in later sections.

## 2 REPRESENTATIONS

### 2.1 Cartographic representation

The conceptual basis for cartographic representation has been the subject of extensive academic analysis [MacEachren 1995], [Fairbairn et al 2001], but there has been continued difficulty in resolving the conflicting pressures of automation (rule-driven visualization) with those of cartographic clarity (freedom of expression). Giedre Beconyte in a recent paper on "Conceptual Models for Cartographic Representation" states "*Other than in the simplest cases, it is impossible to limit cartographic design to a single set of rules at all; hence thematic mapping can hardly be subject to automated processing functions*" [Beconyte 2004]. The Representations and Overrides system which is summarized in this paper and described more fully in [Hardy, Eicher, Briat, Kressmann 2005] unifies automation and freedom capabilities, and hence contradicts aspects of the above analysis. Its fundamental advance is to add minimal information to geographic feature classes in a GIS database to store representation rules and graphical overrides to individual features.

### 2.2 Line example

Representing linear features unambiguously with patterned or dashed lines has traditionally been a challenging problem for cartographers. A database-centered mapping system provides a new solution to this and similar problems, by supporting high-quality automatic representation of GIS features, while at the same time providing the flexibility to override the automated rules. Clear and attractive maps can be efficiently produced.

Fig. 1 below shows five stages of symbolization for linear road or 'track' features from a vector topographic GIS dataset (data copyright swisstopo). Traditional GIS-based mapping systems support only the first two stages.

In stage 1, default GIS symbology (a red line) is assigned to the linear features, and in stage 2 a dashed line symbol is applied, as this is the typical line pattern used in cartographic products for tracks. This GIS symbolization falls short of many cartographic requirements because the poorly symbolized line intersections and bends (highlighted by the red circles), lead to ambiguity as to where the tracks start and end.

Stage 3 shows how the new representation capabilities can automatically produce better symbology at line intersections by adjusting the line dash pattern to ensure intentional (half-dash) connections at the ends of all such line features. Stage 4 shows how using this improved representation as a starting point, the cartographer can further perfect it by forcing the centre of a dash to be placed at the sharp bend in the track in the northeast corner. This modification is stored in the database as an override to the representation geometry.

Stage 5 shows the ultimate graphical freedom and escape from the rules, where the cartographer has decided to change the color of some of the dashes, and to delete one dash from the other track. However, this 'Free Representation' is still closely associated with the original feature.

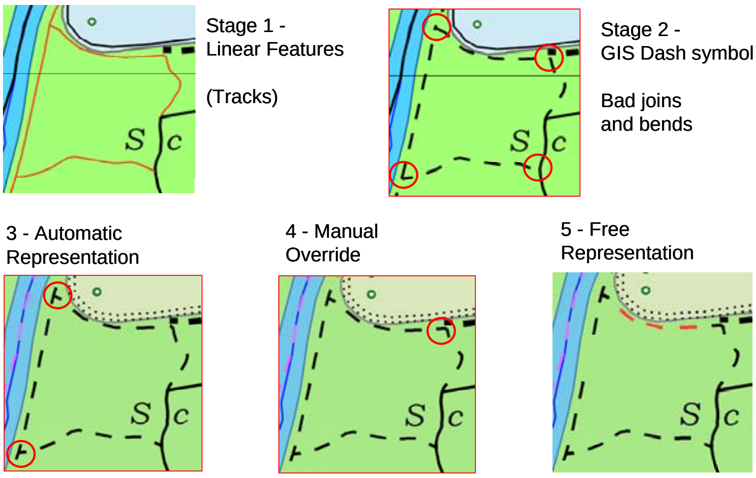


*Fig. 1 - Five stages in symbolization.*

## 2.3 Representation Storage

Physically, a cartographic representation adds additional columns to a standard ArcGIS feature class table within the spatial database. As in any vector GIS, the source feature class stores point, line, or polygon geometries, as well as a set of additional attribute columns used for mapping, analysis, and data management. The added representation columns store data that defines the representation rule used to symbolize a feature. They also store cartographic overrides, which are exceptions to the representation rule for a given feature.

A representation system design premise was to avoid unnecessary duplication of data. Therefore, the extra columns that hold cartographic representations and overrides are minimal in size, and wherever possible the representation information is derived dynamically from the existing GIS feature as it is needed. The structures used to hold overrides are flexible enough not to require separate columns for individual overrides.

## 2.4 Representation Rules

Each cartographic representation added to a feature class can refer to different rules for subsets of features within the feature class. For example a roads feature class will typically have different rules for streets, first, second and third class highways, and for freeways. It may also have variant rules for highways on bridges, in tunnels, or for unique circumstances not normally part of the standard data model (such as a highway temporarily interrupted by a fair).

Rules are made up of one or more visual layers, each of which start from the feature shape geometry and have an optional chain of 'geometric effects' and 'placement styles' that are applied dynamically prior to rendering with a basic symbol (marker, stroke, or fill).

Fig. 2 shows the data flows during symbolization of a cartographic representation added to a GIS line feature. The shape field of the feature has a representation rule applied, which generates two visual layers, the first of which goes through two geometric effects (an offset to one side, then a dash pattern) before having a basic symbol (stroke) applied. The second layer has one effect applied (a marker placement pattern) before the basic symbol (marker) is applied. Fig. 3 shows a typical visual result of such a rule.

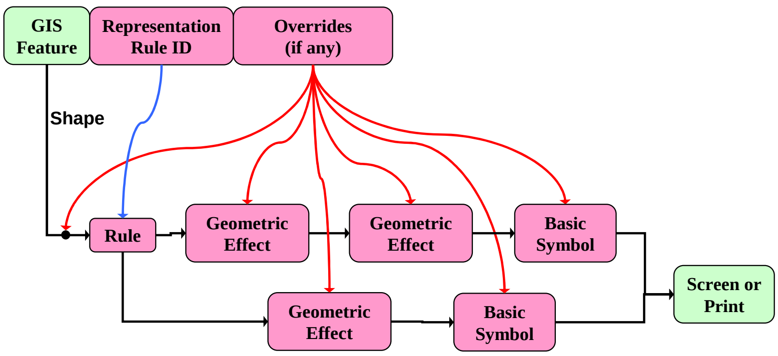


*Fig. 2 - Drawing pipeline for representations with overrides.*

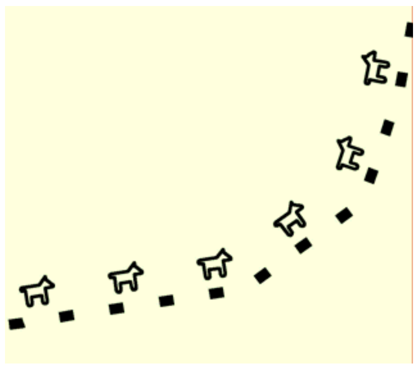

*Fig. 3 - Result of Representation Rule.*

Rules can also be set up to use any existing field in the database as an ‘Explicit Representation Field’ to control the feature representation appearance. Such field values can be set by geoprocessing processes, which can use the full power of the GIS toolkit to determine the need and calculate the required result. A typical and powerful example is the use of the topology engine from within a geoprocessing tool to find all the cul-de-sac roads, and set a database field which is then used to control their line end style to be square rather than round ended (Fig. 4).

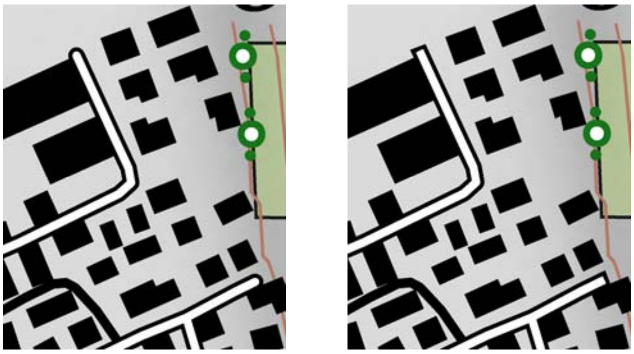

*Fig. 4 - Before and after Cul-de-sac processing.*

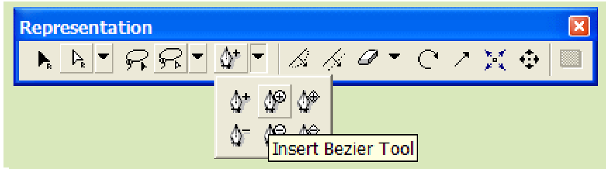


*Fig. 5 - Cartographic Editing Tools.*

## 2.5 Exceptions to Rules

### 2.5.1 Overrides

Overrides allow the user to make exceptions to the rules, while remaining within the data model. In Fig. 2, the override field can modify the input shape, the properties of the geometric effects, or any of the graphic properties of the symbols. A set of intuitive geometric and representation property editing tools are provided for defining and modifying overrides, based on tools and palettes familiar to a user of desktop graphics packages such as Adobe Illustrator or Macromedia Freehand (Fig. 5).

### 2.5.2 Free Representations

A further level of exception is provided by the ability to convert any representation into a ‘Free Representation’. This makes an in-line copy of the rules affecting the particular representation, so that the rules can be changed for this one feature. This can include change of geometry type (area to point), adding new rules or symbol layers, or introduction of arbitrary new graphics. Being able to liberate a particular representation from the data model in this way gives freedom to successfully represent features with appearance too rich to model otherwise, such as a railway siding area where the representation should just show a typified subset of lines indicating ‘there are lots of railway lines here’. It also allows repositioning or suppression of individual graphic elements of the symbolization, such as individual dashes of a road tunnel to avoid important features at ground level, as in Fig. 6 & 7.

*Fig. 6 - Rule-based Representation of Tunnel.*

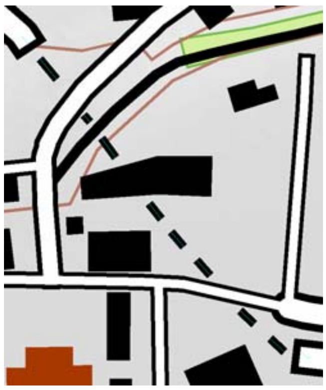

*Fig. 7 - Free Representation of Tunnel, with edited dashes.*

# 3 CARTOGRAPHIC EDITING TOOLS

## 3.1 Cartographic Editing Tools

The introduction of focused cartographic representation editing tools (Fig. 5) allows cartographers who currently prefer to graphic software to work even more efficiently in a GIS-based system. Representation editing tools that work similarly to tools found in graphic packages such as Illustrator or Freehand are introduced to ArcGIS. Furthermore, many of the newly introduced tools are particularly efficient because they are designed with specific cartographic tasks in mind. Editing of representations takes place within the same versioned editing environment supported by ArcGIS for editing vector feature classes. See [Eicher & Briat 2005] "Supporting Interactive, Manual Editing of Cartographic Representations in GIS Software" for a detailed exploration of the cartographic editing tools and their applicability.

# 4 CARTOGRAPHIC DATA MODELS AND WORKFLOWS

## 4.1 Simple workflow case

In the simplest case, an organization has existing GIS data in a feature class, and wants to use it to produce a high-quality cartographic product. In this case, they add a cartographic representation to the feature class (see Fig 8).

## 4.2 Multi-product Case

In the next case, an organization has existing GIS data in a feature class, and wants to use it to produce more than one high-quality cartographic product at similar scales. In this case, they add a cartographic representation for each product to the one feature class (see Fig. 9).

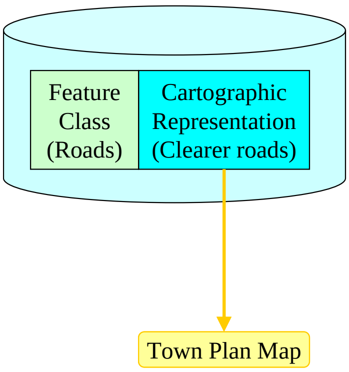


*Fig. 8 - Simple case – Existing Feature Class.*

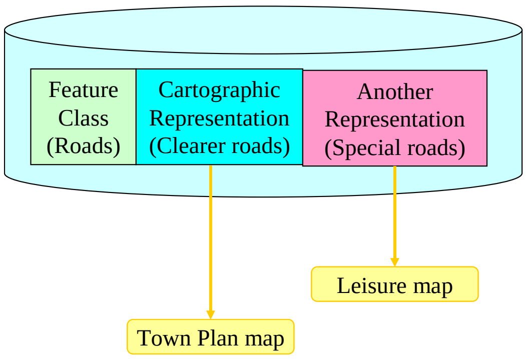


*Fig. 9 - Multiple Cartographic Representations.*

## 4.3 DLM/DCM Case

In the third case, an organization has a master landscape model data (DLM), and wants to use it to produce more than one high-quality cartographic product at different scales, as well as non-cartographic products (such as navigation routes for an in-car voice guidance system). In this case, they require the extraction of requisite data from the DLM by selection and generalization into a Digital Cartographic Model (DCM), which can then be enhanced with multiple representation capabilities as in the multi-product case (see Fig. 10). See also [Buckley & Frye 2005] "An Information Model for Maps: Towards Cartographic Production from GIS Databases" for background on DLM/DCM models and the model transformations involved in their use.

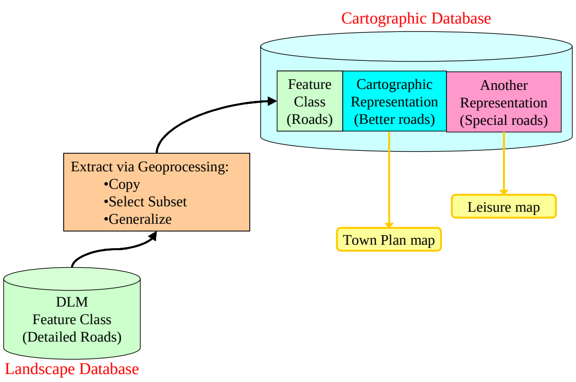


*Fig. 10 - Enterprise case with DLM/DCM.*

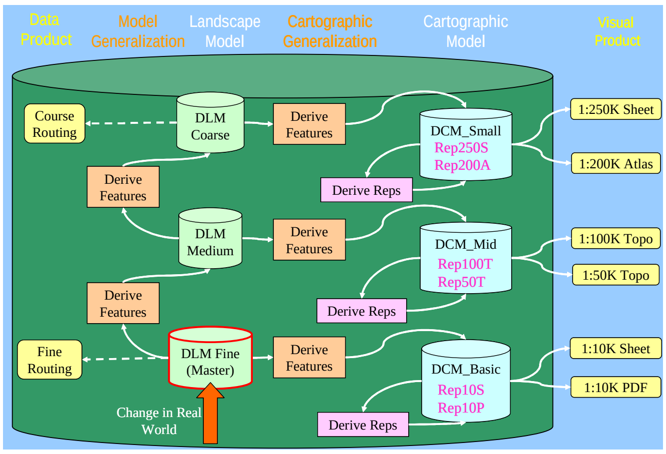


*Fig. 11 - Multi-scale DLM/DCM Data Flows.*

## 5 GENERALIZATION AND DERIVED PRODUCTS

### 5.1 Enterprise Case

Fig. 11 extends the enterprise DLM/DCM case to where multiple products at different scales are to be produced. Here, the data flow encompasses both ‘model generalization’ (deriving landscape model features at coarser resolution by selection, aggregation and simplification), and ‘cartographic generalization’ (deriving visually appropriate features by applying displacement, exaggeration and typification), taking into account the symbolization widths and sizes.

Model generalization is applied first, to derive a set of reduced-resolution landscape models. Starting from each DLM, cartographic generalization is applied, to produce a digital cartographic model for a particular ‘scale band’. Once cartographic data appropriate to the scale band has been derived, then the multiple representation and override capabilities can be applied. These handle the symbology requirements and geometric differences for the product specifications and sheet/page limits needed to produce the various different products.

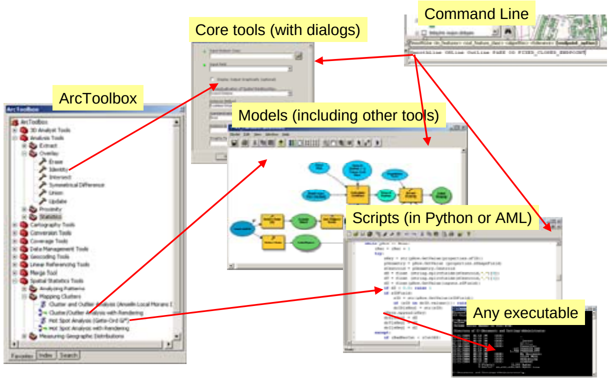


*Fig. 12 - Geoprocessing framework, models and tools.*

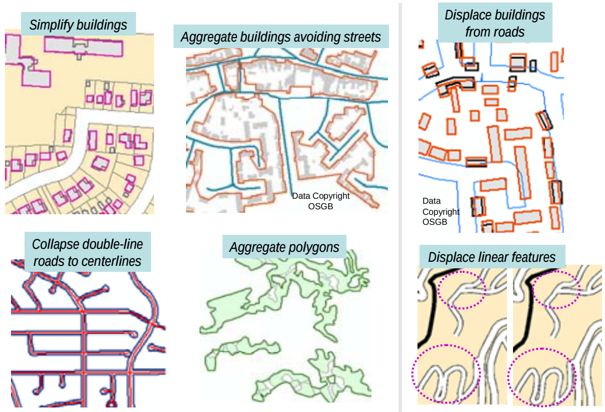


*Fig. 13 - Generalization Algorithms.*

### 5.2 Generalization Framework

Both model and cartographic generalization are executed using the geoprocessing environment of the GIS. This environment has a large and increasing range of algorithms for manipulation of geospatial data, including ones specifically aimed at generalization for product derivation [Fig. 13]. Algorithms can be combined into sequences, either by scripting languages such as Python, or using a visual ModelBuilder [Fig. 12]. The system is extensible, and user-written tools can easily be added as scripts or executables. See [Lee & Hardy 2005] for more details on database-driven generalization.

### 5.3 Contextual Generalization

Good generalization is not an easy task, whether done by a computer or a human. It requires understanding and charactering the geography of the mapped area, and involves finding patterns in the data, and abstracting them. In particular:

- It can’t be done one feature at a time – It involves relationships with neighbors
- It can’t be done one layer at a time – Relationships are with features of many classes
- It can’t be done uniformly – Even features of one class have different surroundings
- It can’t be done across whole map at once – Must localize or partition
- It can’t be done just on geometry – It needs understanding of topology and of attributes

An earlier paper [Lee 2004] analyzed and illustrated in detail what geographic and cartographic contexts should be considered in generalization. The multiple representation and generalization project is moving on to extend the framework to include contextual generalization. The project team will gain from the knowledge of other ESRI development team members who participated in the European AGENT project [Lamy et al 1999], and in other leading cartographic functionality, such as the Maplex text placement system.

The resultant design and architecture will exercise concepts of partitioning areas into geographic zones, recognizing patterns and distributions, setting rules and priorities to guide the generalization analysis and decisions, controlling and assessing the generalization status and quality, and supporting post processes and representation refinement.

### 5.4 Update Data Flows

Once the DLM/DCM work flow has been established and the products generated in their initial editions, then the data flows for update should follow the same paths, starting with capture into the master DLM, flowing through the reduced resolution DLMs, the cartographic model and the representations to the products. See [Briat 2005] on "Incremental Update of Cartographic Data in a Versioned Environment" for more information on strategies for handling update.

For maximum efficiency, just the delta changes should be propagated through the framework, which requires incremental generalization – a concept which cannot be fully implemented until a second stage of development, learning from experience of initial generalization. A prerequisite for incremental generalization is the creation and maintenance of 'parentage' attribution for all derived features.

In the interim, the practical work flows will assess the accumulated amount of change, and determine for each product whether it is better to reapply generalization starting from the current state of the DLM, or to apply minor edits at the DCM stage. In either case, the versioning capabilities of the geodatabase will be used to control the time dependencies often required for cartographic publishing.

Note that rule-based representations (other than the very small minority where overrides have been made) automatically apply any changes made to the source feature. This minimizes the effort in propagating change through to products.

## 6 PRODUCTION AUTOMATION

### 6.1 Automation and Optimization

In a production cartographic organization, such as a national mapping agency or a commercial atlas producer, every button pressed by an operator costs money. The Production Line Tool Set (PLTS) has been developed by ESRI as a software layer on top of ArcGIS to automate and optimize the cartographic work flow and the daily tasks of operators. It provides a set of interactive toolbars and automated processes targeted directly at cartographic production.

Crucially for bulk cartographic production, PLTS provides database driven capabilities for validating attributes and using attribute combinations to select representation styles. It optionally provides standard data models and tailoring for defense and nautical data preparation.

### 6.2 Map Series and Atlas Page Generation

PLTS, using the underlying geodatabase, provides facilities for defining a map series, then generating the map sheets automatically with appropriate marginalia and titles. It can also create an atlas made up of multiple pages with overlaps and page grids. Styles, layouts, and marginalia for common civilian and defense mapping styles are provided. The approach of using the database not just for the feature data, but also for layouts and product metadata is a key aspect of the database-centered development strategy.

### 6.3 Text and Label Placement

ArcGIS has clean and comprehensive capabilities for storing text in the database as annotations, which can be freestanding or feature linked. Feature linked annotations are tied back to the GIS features from which they were created so that they can be automatically updated as change happens in the real world. Annotation support now includes such cartographic elements as callouts and leader lines.

The market-leading Maplex text placement application has been re-engineered and built into ArcGIS as a label placement engine (see Fig. 14). This can massively reduce the previously labor-intensive task of generating and positioning text for cartographic clarity [Murad 2005B]. Development continues on improving labeling and leader placement, and will continue on specialist labeling strategies such as contour laddering.

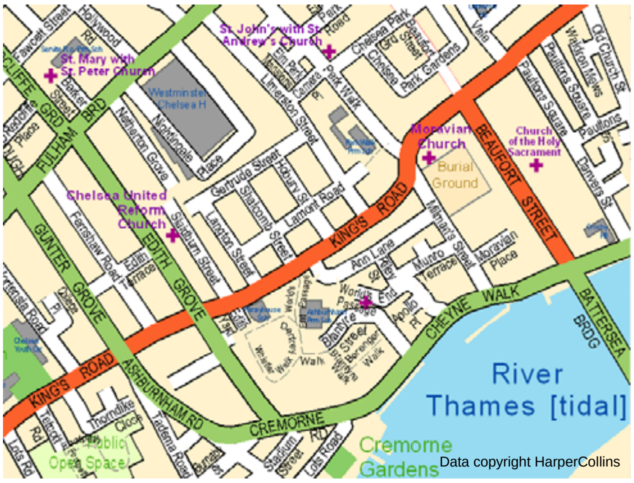


Fig. 14 - Text placed by Maplex Engine.

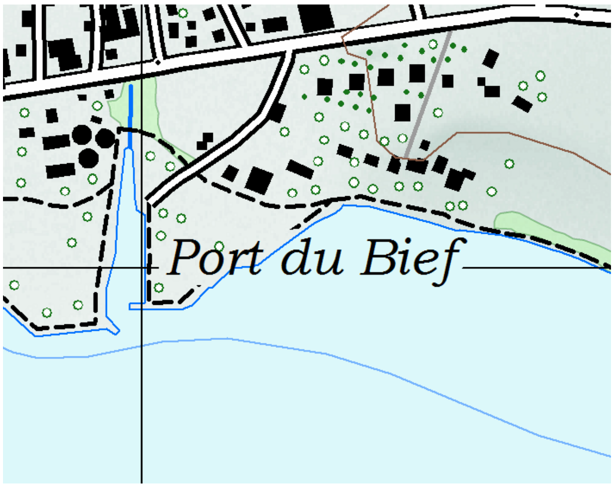


Fig. 15 - Selective Masking of Black behind Text.

## 7 GRAPHIC CONTROL AND VISUAL OUTPUT

Underpinning the visualization capabilities of ArcGIS is a powerful graphics output pipeline that handles stroking, filling, patterning, color, priority, transparency, visibility, masking, and much more. It drives screen output, plotting, and graphic export formats.

The graphics pipeline supports powerful visualization capabilities that were previously only available in specialized graphics applications, such as selective masking, which allows the user to specify a set of target layers to be masked or erased for each source polygon layer (see Fig. 15). This can greatly increase cartographic clarity, for example, by suppressing road casings where crossed by text without suppressing the road infill.

Work continues to revise the graphics pipeline for greater performance, new functions, and platform portability. It will improve handling of complex graphics primitives, such as the clipping and merging of patterned fill areas, and hence reduce plot file sizes, particularly on output to intelligent plotting devices. Support is also planned for new document color models to provide a framework for CMYK, spot, and Pantone® colors, and to facilitate the use of gradients and textures.

## 8 CONCLUSIONS

Advanced cartographic representation mechanisms and cartographic editing tools will soon become available in a commodity GIS. In conjunction with production automation and graphics capabilities described above, they will dramatically change the scope of automated cartography, facilitating the generation of multiple products from a central database. The human freedom enabled by the override system, combined with the symbolization rule pipeline, will permit high-quality attractive cartography within a database-centered environment.

These mechanisms and tools are part of an overall vision to provide a single, consistent, intuitive, efficient, and liberating mapping environment, used throughout the cartographic communication process. This environment is geodatabase-centered, holding master data, plus specifications, processes, and results of derived products. It releases cartographers from the drudgery of repetitive actions, but provides the freedom and tools to amplify our human creative and expressive skills.

## 9 NOTES

1. The sample data used in Fig. 1, 4, 6, 7 & 15 is swisstopo VECTOR25, copyright Swiss Federal Office of Topography. Samples in Fig. 13 are courtesy of OSGB, Crown Copyright. Data for Fig. 14 is copyright HarperCollins.

2. Some contextual matter in this paper is revised and updated from other documents jointly authored by members of the ESRI team working on representations and DLM/DCM work flows. In particular Cory Eicher, Marc-Olivier Briat, Dan Lee and Edie Punt are thanked for their contributions.

3. This paper is a forward-looking document, and the capabilities it describes are still under development. As such, it is intended to give guidance as to likely future direction and should not be interpreted as a commitment by ESRI to provide precise capabilities in specific releases.

4. This is an overview paper, and more detail on the capabilities summarized are provided in other papers being presented at the same conference – see [Briat et al 2005], [Eicher & Briat 2005], [Lee & Hardy 2005], [Buckley & Frye 2005] and [Murad 2005].

## 10 REFERENCES


- Beconyte G, 2004, "Conceptual Models for Cartographic Representation", Geovizualization 2004, Kolymbari, Crete http://www.kc.gf.vu.lt/Publications/ConcptModelling2004.doc
- Briat M, Monnot J, Kressmann T, 2005, "Incremental Update of Cartographic Data in a Versioned Environment" , 22nd ICA Conference Proceedings, A Coruña, Spain
- Buckley A & Frye C, 2005 "An Information Model for Maps: Towards Cartographic Production from GIS Databases", 22nd ICA Conference Proceedings, A Coruña, Spain
- Eicher C & Briat M, 2005, "Supporting Interactive, Manual Editing of Cartographic Representations in GIS Software", 22nd ICA Conference Proceedings, A Coruña, Spain
- ESRI 2004, ESRI White Paper, "Cartographic Design Process – Artistic Interpretation With The Geodatabase", http://www.esri.com/library/whitepapers/pdfs/cartographic-design.pdf
- ESRI 2004B, "What is ArcGIS", ESRI Press, ISBN 1-58948-090-2, 2004, http://store.esri.com/esri/showdetl.cfm?SID=2&Product_ID=109&Category_ID=28
- ESRI 2004C, "Building a Geodatabase", ESRI Press, 2004, http://store.esri.com/esri/showdetl.cfm?SID=2&Product_ID=90&Category_ID=28
- Fairbairn D, G. Andrienko, N. Andrienko, G. Buziek, J. Dykes, 2001, "Representation and its relationship with cartographic visualization: a research agenda" - Cartography and GIS, Vol. 28, Nr. 1
- Frye C & Eicher C, 2003, "Modeling Active Database-Driven Cartography within GIS Databases", 21st ICA Conference Proceedings, Durban, S. Africa
- Hardy P, Eicher, Briat, Kressmann, 2005 "Database-stored Representations and Overrides, Supporting Automated Cartography with Human Creativity", Auto-Carto 2005, Las Vegas, USA
- Lamy et al 1999, "AGENT Project: Automated Generalisation New Technology", 5th EC-GIS Workshop, Stresa, Italy, June 1999. http://agent.ign.fr/public/stresa.pdf
- Lee D, 2004, "Geographic and Cartographic Contexts in Generalization", ICA Workshop on Generalisation and Multiple Representation, Leicester, UK, August 2004 - http://ica.ign.fr/Leicester/paper/Lee-v2-ICAWorkshop.pdf
- Lee D & Hardy P, 2005 "Automating Generalization – Tools and Models", 22nd ICA Conference Proceedings, A Coruña, Spain
- MacEachren A, 1995, "How Maps Work. Representation, Visualization and Design". New York-London: The Guilford Press.
- Murad-al-shaikh M, 2005, "Professional Labeling and Annotation Techniques with ArcMap" 22nd ICA Conference Proceedings, A Coruña, Spain


*[29-Apr-2005, Revised 4-May-2005]*

## BIOGRAPHY

Paul Hardy has been active and influential in the digital mapping and GIS industries for 30 years. He gained his degree in 1975 from Cambridge University in England, studying Natural Sciences and Computer Science (B.A. + M.A. Cantab).

He then joined Laser-Scan (Cambridge, UK) which was a pioneer and market leader in computer graphics and digital mapping, and remained there for 28 years, leading the creation of hardware (digitizers, displays and plotters) and software (map production systems and spatial databases). After many years leading software development teams, and as architect of the LAMPS mapping software, he moved to roles in Business Development and Product Management.

In 2003, he moved to ESRI in Redlands, California, as Product Manager for Cartography. He combines this with an ongoing role as a software project leader, and as consultant and presenter, building on his particular experience with geospatial data and data models, production cartography, automated generalization, and data re-engineering techniques.

He is a Member of the British Computer Society (MBCS), a Chartered Engineer (C.Eng), and a Fellow of the British Cartographic Society (FBCartS)