VECTOR BASED GENERALIZATION APPLICATION IN THE PRODUCTION OF SMALL SCALE DATABASES OF THE NLS OF FINLAND

Veijo Pätynen
veijo.patynen@nls.fi

Jouni Ristioja
jouni.ristioja@nls.fi

National Land Survey of Finland
Development Centre
Opastinsilta 12 C, P.O.Box 84, FIN-00521 Helsinki, Finland

Abstract

Production of small scale databases (1:100k, 1:250k, 1:500k, 1:1M and 1:4,5M) of the National Land Survey of Finland was started in 1995. The production line of each database was separate from the others. Although the result was topologically incompatible, the data sets, however, were comprehensive. Validity of these databases for present-day needs has become inadequate, and also the production applications, which were based on ArcInfo and aml-scripts, have become old. International co-operation projects require the existence of high quality map databases.

The development project to produce a new production system for small scale databases was started in 2003. The new system is based on ESRI ArcGis (Windows XP) application. ESRI's ArcObjects technology provides a wide component library for developing own generalization tools. These tools are developed using Visual Basic and C# scripts and they are integrated into ESRI's standard ArcMap application. Its database format is ESRI ArcGis Geodatabase. Topological relationships between features are maintained in Geodatabase with ArcGis topology tools. The basic idea of production is to avoid interactive generalization work and use automation whenever it is possible.

Whole generalization process is done in a vector format. The process can be seen as a model-oriented generalization because it is not intended for any specific map representation. The parameters of generalization are based on the level of detail and minimum size, width or on distance between objects depending on the desired scale of data. All parameters for controlling the generalization process are pre-defined and are stored into separate control database tables. This ensures that all data is generalized in the same way and users can't use wrong parameter values by mistake. For example, there are 630 control parameter values for generalization 1:10k to 1:100k.

Source data for generalization is derived from the Topographic Database, which incorporates the most accurate positional data about Finnish topography, and in this respect is comparable to maps on scale 1:5k - 1:10k. The first phase is to produce a 1:100k map database and after that a 1:250k map database using 1:100k data. The production application contains all necessary generalization tools for selection, simplification, combination, smoothing and enhancement of points, lines and polygons. Tools for polygon aggregation, especially, are very versatile. All different classes of land cover (agricultural areas, rock areas, wetlands, etc.) can be generalized in the same process as the application controls the priority and processing rules of different area classes. For example, fusion and merge of different class of land cover polygons can be carried out automatically in one run. Generalization tools can be used for generalization of different scales after the parameters of a desired generalization step have been defined.

Production with the new generalization application started in 2006. The production team consists currently of about 10 persons. Efficiency of the new process is remarkably better than that of the old production system. The new 1:100k map database covered two-thirds of the area of Finland in the end of 2008. Full coverage will be achieved during this year. Production of the new 1:250k map database is still in early stages but the functionality of generalization tools has already been assured.

1. INTRODUCTION

Production of small scale databases (1:100k, 1:250k, 1:500k, 1:1M and 1:4,5M) of the National Land Survey of Finland was started in 1995. The production line of each database was separate from the others. Map database 1:100k was generalized from the Topographic Database 1:10k but rest were scanned and vectorized from existing printed maps. Although the result was topologically incompatible, the data sets, however, were comprehensive. The needs of the customers of that time could be fulfilled.

Validity of these databases for present-day needs has become inadequate. International co-operation projects, such as EuroRegionalMap (ERM) and EuroGlobalMap (EGM) of EuroGeographics, require the existence of high quality map databases

The production applications, which were used in the production of small scale databases, were based on ArcInfo and AML-scripts. ESRI underwent a major change in its GIS product family when it released ArcGIS version 8.0 in 1999. So, technical architecture of the production application has gradually become old.

2. DEVELOPMENT PROJECT

The development project to produce a new production system for small scale databases was started in 2003. The new system is based on ESRI ArcGIS Desktop architecture. Whole generalization process is done in a vector format. The basic idea of production process is to avoid interactive generalization work and use automation whenever it is possible.

The goal of the original project plan was very ambitious including master and cartographic databases in different scales and version management of objects using unique identifiers. During the realization part of these objectives had to be given up because they were too complicated to be managed in production environment. Still, the process can be seen as a model-oriented generalization because it is not intended for any specific map representation.

In the project there have worked mainly three application developers and three definers. Production with the new generalization application, called Piekka, started in 2006.

3. TECHNICAL REWIEV OF PIEKKA APPLICATION

3.1 Introduction

Piekka application is build by customizing ESRI ArcMap application, and by extending ArcObjects architecture with custom components. The custom components provide automated processes for generalizing topographic data based on user-supplied specification. A set of generalization operators is implemented to compress data, to reduce the level of details in all feature types. Generalization tasks can be done using a single generalization command, tool or edit task, or commands can be made up as procedures to solve more complicated generalization tasks. This chapter provides a brief description of the development environment and technology decisions, and takes a look at developed generalization functions and procedures derived for specific generalization tasks in Piekka application.

3.2 Developing Environment

VBA development environment is used for interface customization. Code for custom objects is written in Visual Basic 6 and Visual Studio .NET development environments, because they both adhere to the COM specification. Extending ArcMap is implemented using both COM and .NET API. When using .NET API object proxies are used to communicate with the native ArcObjects COM components. In .NET, the object proxies can be created using wrapper classes, contained with interop assemblies. ESRI supports this technology through the COM interop. Technical Documentation is produced by using InnovaSys DocumentX!, which is integrated into Visual Studio. The whole Developer Help (VS, ArcGIS, and Piekka Developer Help) is also integrated into the IDE.

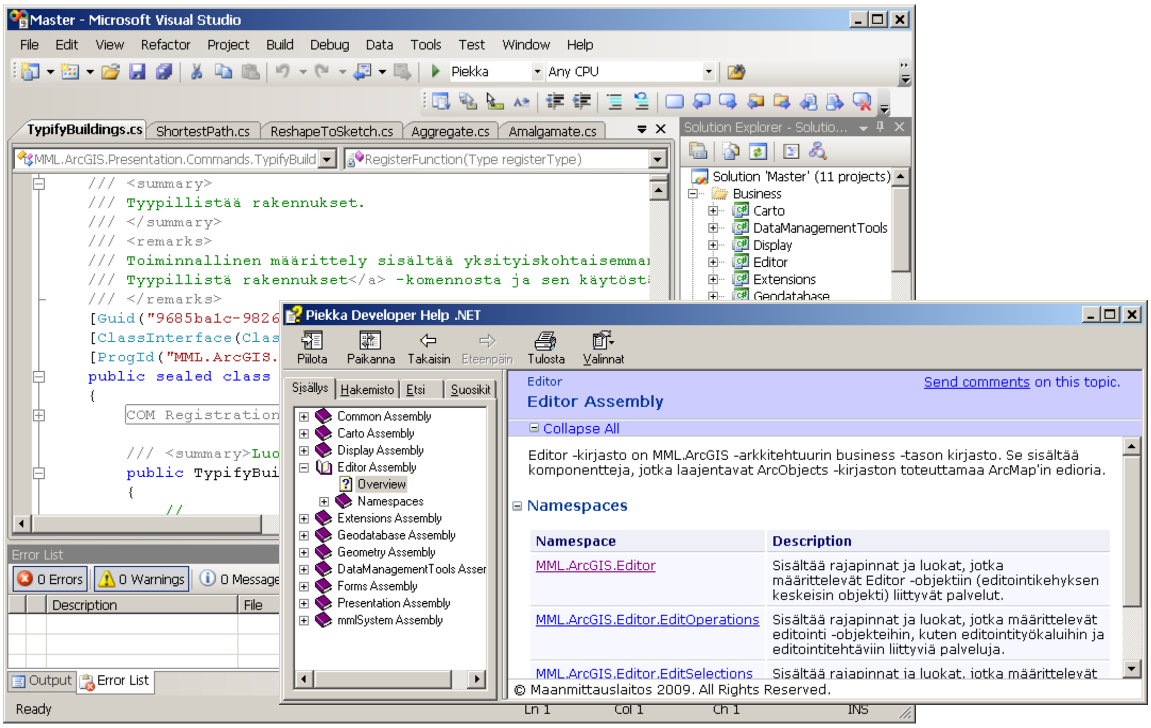


Figure 1. Integrated developing environment.

3.3 User interface customizations

User interface customizations are mainly performed via the Customize dialog box, which is part of the VBA development environment. Using the Customize dialog box does not involve programming, and adding controls is as simple as drag-and-drop operations. Map templates are used to disseminate the UI customization throughout the production organization. All the UI customizations are stored in the Normal templete, which stores the default or original state of the application. When Piekka application is started it first reads the customizations from the Normal template, and they will get propagated to all map documents when they are next opened. This is similar to how templates are used in Microsoft Office applications.

3.4 Extending the functionality of ArcMap

ArcMap provides a full range of tools for extensive data manipulation and high-end cartography. Because ESRI used Microsoft COM to create the ArcObjects platform, upon which ArcMap is built, the entire application is open to customization down to a low level. After adding a custom component to ArcMap, it can be used as any built-in command or tool. This is how the custom components for generalization tasks are plugged in to the ArcMap framework as controls, such as commands, tools, menus, and toolbars – and thus composing Piekka application with rich functionality to satisfy the requirements for both database and cartographic generalization.

Adding a new command is quite straight-forward: one has to create a COM object and plug it into application. The main steps are as follows:

1) Create a COM/.NET project,
2) Create a COM class,
3) Reference the ArcObjects libraries,
4) Implement ArcObjects interfaces,
5) Compile project, and
6) Register in an ArcGIS component category.

3.5 Core libraries and the main functionality

Piekka application is split into separate layers that have distinct roles and functionalities. Each layer contains a number of discrete component types grouped into sublayers, with each sublayer performing a specific type of task. This helps to maximize maintainability of the code, and provide a clear delineation between locations where certain technology or design decisions must be made.

At the most abstract level, the Piekka application architecture can be considered to be a set of cooperating services grouped into the following layers (libraries):

1) Presentation services (Forms, Presentation). Responsible for user interaction; provides a common bridge into the core business logic encapsulated in the business services.
2) Business services (mmlSystem, Carto, Display, Editor, Extensions, Geometry, DataManagementTools). Implements the core functionality of the system, and encapsulate the relevant business (generalization) logic.
3) Data services (Geodatabase). Exposes data to the business layer through generic interfaces designed to be convenient for use by business services.
4) Cross-cutting services (Common). Implement specific types of functionality that can be accessed from components in any layer.

The core generalization functionality is implemented as edit commands. Each edit command orchestrates one or more generalization operators providing functionality for both attribute and spatial transformation. Some of the operators change only the geometry of individual objects, some transforms a group of objects, and some of them operate with both classes. In all cases each command takes all related or nearby features into account to avoid conflicts and maintain spatial relationships.

Most of the generalization edit commands require some parameters, which can be set either by using interactive user dialogs, or by predefining them into a product specific database. In general, these parameters can be set without understanding all the details about the techniques or algorithms used. Because many of these parameters are directly or closely related to spatial constraints, such as minimum distance or minimum size of the features, the effects of changing these values are usually obvious.

Developed generalization functions and procedures provide functionality for:

1) Simplification (e.g. Simplify > Weed)
2) Collapsing (e.g. Collapse > To Point Feature, To Line Feature)
3) Enhancement (e.g. Simplify > Simplify Bends, Straighten Twists)
4) Selection (e.g. File > Load data),
5) Elimination (e.g. Eliminate > Smaller than Regional Minimum)
6) Displacement (e.g. Displace > Move from Nearest, Move to Nearest)
7) Aggregation (e.g. Classify > Reclassify; Collapse > To Line Feature and Fusion; Amalgamate; Fusion > Short, Narrow, With Alike, With Similar, With Adjacent; Typify Buildings etc.)

The main toolbar provides computer-assisted and automated edit commands for all these database and cartographic generalization tasks from the Generalization menu, as shown in Figure 2.

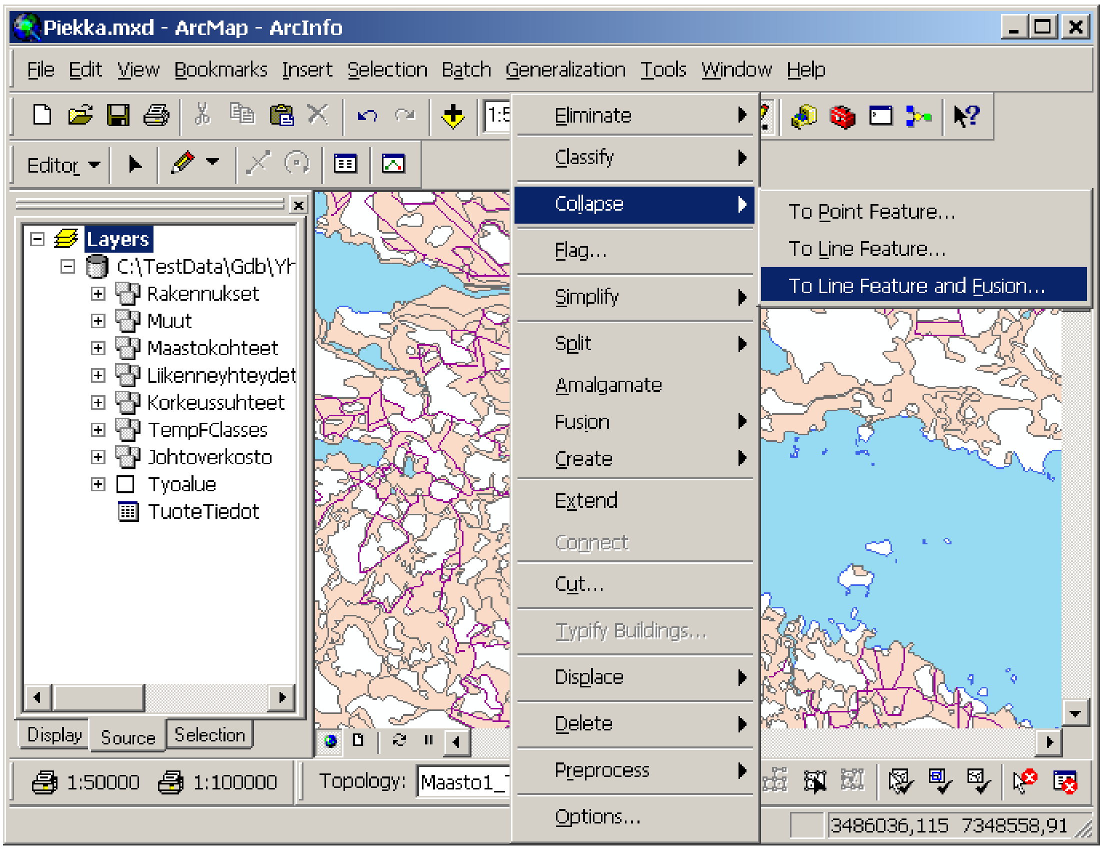


Figure 2. The Generalization menu provides the core edit commands for database and cartographic generalization.

4. SPECIAL FEATURES OF PIEKKA APPLICATION

Source data for generalization is derived from the Topographic Database (TDB), which incorporates the most accurate positional data about Finnish topography, and in this respect is comparable to maps on scale 1:5k - 1:10k. The first phase is to produce a 1:100k map database and after that a 1:250k map database using 1:100k data.

Piekka application contains all necessary generalization tools for selection, simplification, combination, smoothing and enhancement of points, lines and polygons. The parameters of generalization are based on the level of detail and minimum size, width or on distance between objects depending on the desired scale of data. All parameters for controlling the generalization process are pre-defined and are stored into separate control database tables. This ensures that all data is generalized in the same way and users can't use wrong parameter values by mistake. For example, there are 630 control parameter values for generalization 1:10k to 1:100k.

In the following chapters some of the special features of Piekka application are presented.

4.1 Model based river network selection

In Finland there are a lot of marshlands, which are drained with very dense ditch network. The TDB contains nearly 15 million ditch lines. Ditches of the old 1:100k map database have been generalized interactively in the foundation process of that time.

These old ditches are used as a model for selection of ditches of the TDB. Great work contribution done in previous process is not wasted but can be utilized.

Figure 3a. Start: Original ditches of the Topographic database.

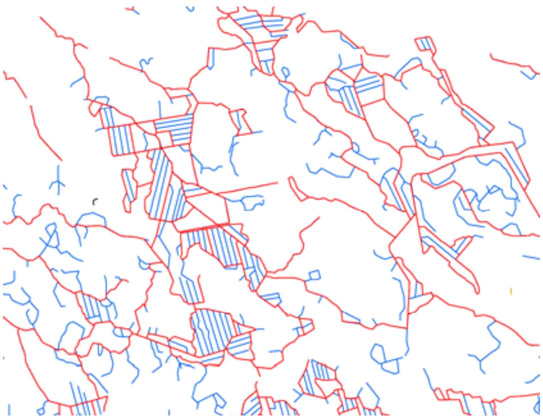


Figure 3b. Ditches of the old 1:100k map database (red lines) control the automated selection process.

Figure 3c. Result: Generalized ditches.

4.2 Collapse of narrow polygon river

Many functions of Piekka application are based on finding and exploiting centerlines of polygons. There is also a special tool for collapsing narrow polygon type rivers to line type rivers. This tool finds narrow sections of polygon rivers, creates automatically the centerline and connecting lines to tributaries and deletes redundant river polygons.

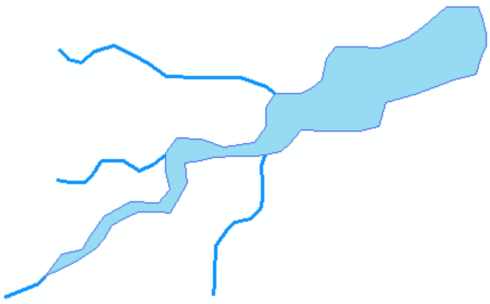


Figure 4a. Start: Polygon type and line type rivers.

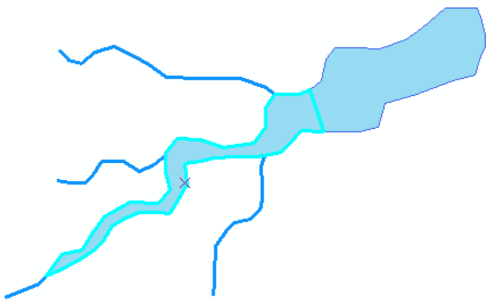
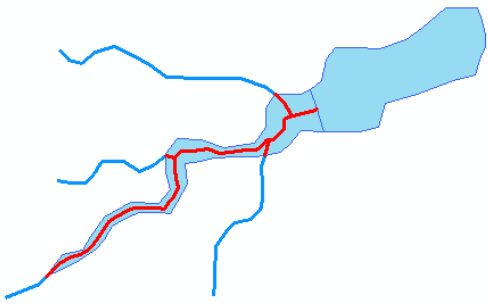


Figure 4b. Narrow section of the river found.



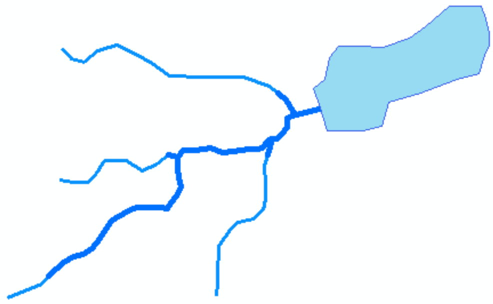


Figure 4c. Centerline and connecting lines created automatically.

Figure 4d. Result: Generalized topologically compatible river network.

4.3 Aggregation of polygon features

Tools for polygon aggregation, especially, are very versatile. All different classes of land cover (agricultural areas, rock areas, wetlands, etc.) can be generalized in the same process as the application controls the priority and processing rules of different area classes. For example, fusion and merge of different class of land cover polygons can be carried out automatically in one run.

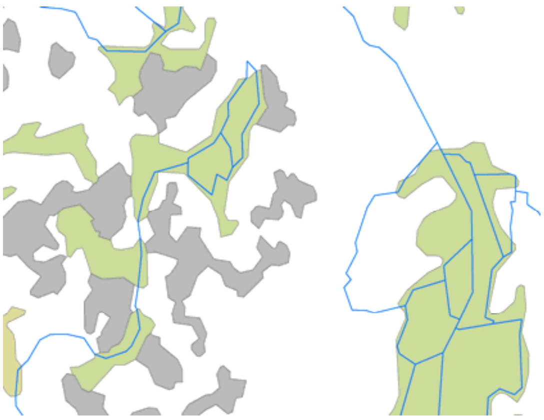


Figure 5a. Start: Original land cover polygons of the Topographic database.

Figure 5b. Result: Generalized land cover polygons.

5. PRODUCTION OF SMALL SCALE DATABASES

Production with the new Piekka application started in 2006. The production team consists currently of about 10 persons. Efficiency of the new process is remarkably better than that of the old production system.

Size of a normal working area is 40 km x 40 km. Processing time of one area is an average one month and it is divided between different themes as follows:

- Terrain features and Hydrography — 2 weeks
- Transportation (roads and railroads) — 1 week
- Elevation (height contours) — 2 days
- Buildings — 3 days
- Others — 2 days

The new 1:100k map database covers currently over 80 % of the area of Finland. Full coverage will be achieved during this year. Production of the new 1:250k map database is still in early stages but the functionality of generalization tools has already been assured.

6. CONCLUSIONS

Piekka application allows establishing of homogeneous, high-quality small scale databases to cover whole Finland. Topological relationships between features are maintained in Geodatabase with ArcGis topology tools. Generalization tools can be used for generalization of different scales after the parameters of a desired generalization step have been defined.

Piekka application is currently intended for the creation of databases. Development work for appropriate maintenance system of data sets is still to be done.