22

ID-MARGARY: an Inference Database for the MApping Recognition and Generation of Ancient Roads and trackwaYs

Stephen G. Bullas

Strategic Decisions Ltd., The Sycamores, Leatherhead, Surrey KT22 9EX UK

Email: 100111.2665@compuserve.com

22.1 Overview

This paper outlines the concepts, ideas and methods behind a prototype modelling system suitable for a GIS research project. The principal, short-term and technical aim of the project is to investigate the feasibility of using a standard Geographical Information System in combination with Artificial Intelligence techniques to:

- recognise known Roman Roads
- train the system in their typical characteristics
- automatically detect ones hitherto unknown
- map and store them on the GIS
- provide statistical information as to their likelihood

If successful, then the long-term objective would be to provide a central repository of nation-wide GIS information to enable:

- Data retrieval locally, at the site of the repository and, via modem, remotely to registered users
- Remote access (by registered users) using standard Windows software
- Remote update of the central repository (under its control) where subsequent field work confirms, for example, the existence of roads conjectured by the system or deviations in the course of known roads from that previously understood.

22.2 System requirements

In order to fulfil the above objectives and to service all of their required attributes, the system will need to possess the essential components shown in Table 22.1. Of these six components, Artificial Intelligence may be further broken down into the following functional items as shown in Figure 22.1:

- Knowledge Base
- Rule Base
- Inference Engine
- Repository Interface

The *Knowledge Base* is that part of the system which stores information about the geographical, topographical, environmental and other feature-specific details of the terrain under investigation. The principal difference between this and a simple database is that in the former, there is an inferred 'intelligence' that may predictively affect the mere data contained in the latter.

The Knowledge Base receives its information from two main sources:

1. a richly populated GIS which itself may be extended and updated with application-specific data
2. a training process in which a Neural Network or other

	GIS	*AI*	*Repository*	*Catalogue*	*GUI*	*Export*
Recognition	✓	✓	✓			
Training	✓	✓				
Detection		✓	✓			
OS-compliance	✓					
Mapping	✓					
Data storage	✓		✓	✓		
Knowledge /Rules		✓	✓			
Statistics		✓	✓			
Control			✓			
BLOB support				✓	✓	✓
Remote access					✓	✓

Table 22.1: Main Components and Attributes of the proposed ID-MARGARY system

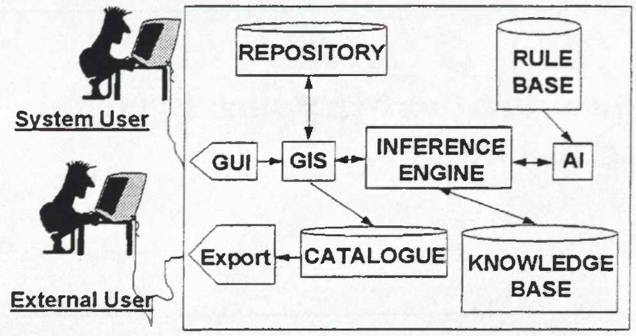


Figure 22.1: System Components: Overview.

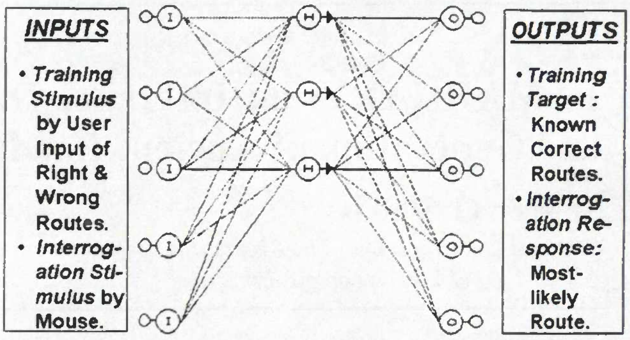


Figure 22.2: Neural Network Inference Engine.

suitable technique is used to enhance and control the behavioural aspects of the system, the results of which may then update the Knowledge Base as in Figure 22.2.

The Rules that may lead to such updates in the Knowledge Base are contained in the *Rule Base.* These rules form a set of logical statements about the nature, attributes and possible consequences of the subject under review. Because they express our own human understanding in a logical form, the systems embracing them are often known as expert systems. Dependent upon the programming language (or, alternatively, the expert system 'shell') being used, rules may be qualified by a particular premise – known as an *antecedent* – in order to reach the correct conclusion. Rules that are not qualified in this way are known as *Doctrinal Rules.* Some trivial examples of those which may be applicable to our system are:

- Roman roads run in straight sections
- They lie on well-drained high ground with clear lines of sight
- They follow footpaths and field/parish boundaries
- Surrounding field names/areas are called 'Stony', 'Black', 'Cold Harbour' etc.
- They pass through towns with known etymologies (—caster, —chester, caer—, strat—, etc.)
- The towns lie 15 miles apart.

Just as a human expert is not infallible, the statements comprising the Rule Base and the premises under which they apply may be prone to uncertainty. It is one of the purposes of the *Inference Engine* to be able to reason with such uncertainties and to draw the correct conclusions from them. This may be achieved by either applying a form of Bayesian Theory or using so-called fuzzy logic techniques within the Inference Engine itself. Because of these uncertainties, an expression of the degree of faith in the probability that the resulting conclusions are correct needs to be stored along with the conclusions themselves.

All of the resulting information is kept in a Central Repository. In general, the purpose of the Repository is to ensure that:

1. Referential Integrity is maintained.
2. Only one copy of on-line data is stored at a single place at any point in time.

The *Repository Interface* ensures that the Repository is kept properly up-to-date; it is effected through a single route to and from the GIS. In practice, the Repository will need to support the following data formats:

- Text data (for reports and general information);
- Multidimensional spreadsheet data (for the contextual information);
- Graphical data (for line drawings, maps and photographs);

This may be accomplished by using either BLOBs (Binary Large Objects) in Borland format or standard Microsoft Windows application file formats.

A summary of the requirements and deliverables for such as system is therefore:

- Able to draw, zoom and update maps
- Ordnance Survey compliant
- GIS BLOB support for
    - Text processing
    - Multidimensional spreadsheet functionality
    - Line drawings, maps and photographs
- Export to third-party formats
- Secure, intuitive and user-friendly.

22.3 Hardware and software requirements

Storing information in one or other of the above ways has important consequences for the remote user. Because these formats are independent of the GIS being used to create the data, information may be retrieved by the user without the need to purchase an additional copy of the GIS in question. Furthermore, an Intel-based 80486 (or higher) compatible personal computer running Microsoft Windows together with a reasonable amount of unused disk space, a modem and, optionally, Borland's Paradox or Quattro Pro or Microsoft Office, is all that is needed to retrieve information from the system once the user has been registered.

Hardware	*Software*
Pentium, RISC or equivalent	32-bit Windows or Windows NT
16MB RAM	32-bit C++ and Prolog compilers
Fast 500MB hard disk	32-bit C++ libraries
20" S-VGA screen; graphics accelerator card with 2MB memory	KPWinPro++; NeuralDesk
Mouse	GIS with programmable user exits

Table 22.2: Anticipated hardware and software for the developer and repository

The anticipated requirements for the System Developer/Central Repository are somewhat less modest but nevertheless still well within the reach of most academic budgets. A large part of the unseen requirements would be that surrounding the acquisition by and entry onto the system of the source data which form the basis of its data extraction and presentation capabilities.

The anticipated requirements for both the developer of the central repository and the remote user are summarised in Table 22.2

22.4 Simple example of capabilities

In order to demonstrate the power and potential of the proposed system, we shall use a simple (albeit somewhat crude) example which nevertheless demonstrates very clearly the principles that may be applied.

Let us suppose that we wish to reference the course of the Roman road between Chichester *(Regnum Noviomagus)* and Silchester *(Calleva Atrebatum)* in the south of England.

The GIS would firstly present us with an outline map of Great Britain upon which we could choose the towns of interest by either entering their names in response to a prompt from the place locator or by clicking at the appropriate position on the map (see Figure 22.3).

The system would then automatically consult the Repository for known information on the area. This information could be nominal or large and varied. Examples might be abstracts from site reports, references to a number of external publications, copies of tithe maps complete with field names, Ordnance Survey detail, statistical data, surveys, aerial or site photographs and so on. The most important item, however, would be the on-line retrieval of the immense amount of information originally collated by Ivan Margary (1955) amended and extended where appropriate by the subsequent field-work of his successors.

Next, the area would be zoomed in on to show detail of, say, known Roman roads together with their Margary classification. From this, it would be seen that such a road exists (Margary classification number 155) heading in a north-north-westerly direction from Chichester. Its course is known only for part of its length, our present knowledge ending just north of the modern village of Milland at a small Roman posting station, some 15 miles from Chichester.

Simply being able to access this sort of on-line information would be of significant use to archaeologists and research students in its own right. However, the inference capabilities of the system would allow us to better this by continuing on where present human knowledge leaves off. Looking for the existence of hedgerows, parish boundaries, topographical features such as marsh, streams and unscaleable high ground, modern place- and field-names and comparing these to the desired goal using appropriate antecedents in the Rule Base enables the system to generate and plot the possible missing course together with estimated probabilities of its existence at any point along the conjectured route. For simplicity, such a route has been shown as a single dashed line in Figure 22.4; in reality, this route would almost certainly comprise of multiple straight sections joining the start and end points in question.

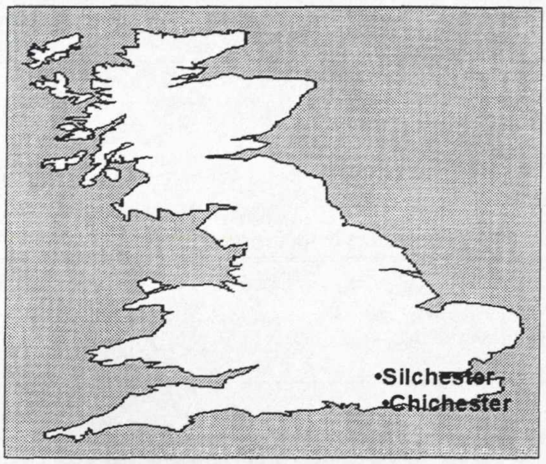
22.5 Summary and conclusions

The proposed system provides the following capabilities:

- Maps and Recognises existing roads.
- Suggests routes and possible courses for new ones.
- Determines their probability.
- Creates a General Reference Catalogue for nation-wide use.
- Updateable and accessible remotely using non-specialist software.

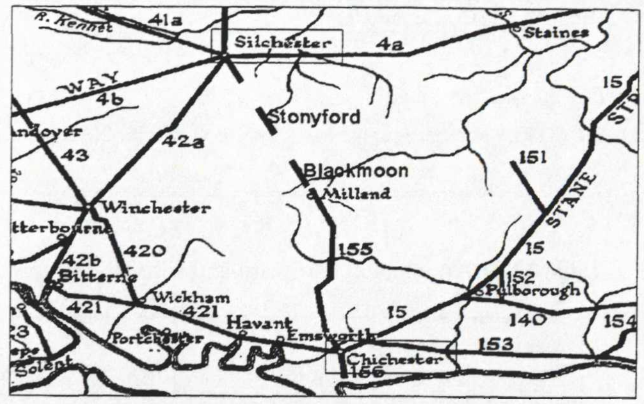
A General Catalogue of ancient roads and trackways providing a central repository of information - eventually on a national basis - has a number of attractions:

- On-line access to the works of Ivan Margary and others
- Independent updates by disparate field units whilst still maintaining central control and referential integrity
- Co-ordination of related field work on a national database
- Results (in BLOB and/or general Windows formats) available to all



SELECT START AND END POINTS

Figure 22.3: Simple example of capabilities (1)



ZOOM AND APPLY RULES & KNOWLEDGE BASE

Figure 22.4: Simple example of capabilities (2).

- Enormous cross-reference capability
- Large potential for educational use

Against these advantages must be weighed the very substantial amount of work necessary to implement such a scheme. Although modern hardware and software are capable of adequately supporting the technology involved, it must first be determined whether:

1. Nation-wide coverage is desirable, manageable and financially justifiable
2. The Inference and Artificial Intelligence capabilities are not just theoretically possible but also capable of being programmed
3. It would be possible to legally transfer the data sources to such a system
4. An adequate amount of source data could be entered in a reasonable time frame
5. A suitable body would be prepared to undertake or underwrite such a project.

If these difficulties can be overcome, a truly useful system implementing concepts at the very forefront of modern technology could provide significant kudos to the institution progressive enough (and brave enough) to attempt it. A prototype designed to address the viability of the above points would provide an interesting and worthwhile research project.

Acknowledgement

Portions of the map shown in Figure 22.4 were scanned from Margary 1955, pp.76–77

Reference

MARGARY, I. D. 1955. *Roman roads in Britain*, Phoenix House Ltd., London.