Proceedings of 9th Onternational Symposium on Spatial Data Handling, Pékin, pp.1a50-63

AN ABSTRACTION-BASED MACHINE LEARNING APPROACH TO CARTOGRAPHIC GENERALISATION

Sébastien Mustière

IGN – COGIT &
Univ. Paris VI – LIP6
2-4, avenue Pasteur
94165 St-Mandé Cedex
FRANCE
(33)-1-43-98-85-45

sebastien.mustiere@ign.fr

Jean-Daniel Zucker

Univ. Paris VI – LIP6 &
CNRS
4, place Jussieu
75252 Paris
FRANCE
(33)-1-44-27-72-07

jean-daniel.zucker@lip6.fr

Lorenza Saitta

Univ. del Piemonte
Orientale
Dip. di Scienze e
Tecnologie Avanzate
Corso Borsalino 54
15100 Alessandria, ITALY

saitta@di.unito.it

ABSTRACT

This article proposes a machine learning approach to overcome the knowledge acquisition bottleneck limiting the automation of cartographic generalisation. It first explains why this automation must be guided by a differentiation of two main types of knowledge involved in this process. More precisely, it shows that cartographic generalisation is best viewed as a combination of two processes: representing (formulating, renaming knowledge) and abstracting (simplifying a given representation). The whole process of creating maps fits into an abstraction framework developed in Artificial Intelligence to account for the difference between knowledge abstraction and knowledge representation. The utility of this framework lies in its efficiency to support the automation of knowledge acquisition for cartographic generalisation as a combined learning of both abstraction and representation knowledge. The results of Machine Learning experiments show the advantages of this approach.

Keywords: Machine Learning, cartographic generalisation, abstraction, representation, knowledge acquisition.

1 INTRODUCTION

Automating cartographic generalisation is a problem of first importance in cartography. First of all, automatic cartographic tools allow to decrease **cost** and **time** necessary to produce paper maps. Then, new maps are increasingly electronic maps. Because of screen resolution these maps ask even stronger generalisation than paper maps. Map-users will also create themselves their maps. Most of the maps are so created by geographic domain experts or decision makers, who are not necessarily specialists in cartography. As a consequence today's maps are often well dedicated to specific user needs but have a poor cartographic **quality**. This quality can be re-introduced in these maps if cartographic knowledge is the basis for designing GIS tools. Furthermore, the opportunity to create and display maps quickly enhances the need for flexibility and, especially, **multi-level** space analysis. For example, João (98) emphasises the need for multi-scale analysis in environmental impact assessment studies. She also notices that, for time and cost reasons, users prefer more easily-available data than data with the right level of details. GISs should then offer efficient tools to automatically change level of details.

This paper is a contribution to the research on the automation of cartographic generalisation, and more specifically to the knowledge acquisition process that is required to develop cartographic expert systems. Part 2 presents the difficulty of automating cartographic generalisation process, especially because of the difficulty to formalise cartographic rules. Why machine learning can help to formalise these rules is then explained. Based on the lessons learnt in the field of Knowledge Acquisition (Clancey,83;Thomas,96), the necessity to differentiate, separate, and structure the different types of knowledge involved in cartography is discussed.

As an alternative to the classical distinction used in cartographic generalisation between geometric, structural and procedural knowledge (e.g. see Armstrong, 91; Weibel *et al*, 95), part 3 proposes a characterisation of the knowledge used in cartographic generalisation better fitted to its acquisition. This characterisation is based on the distinction between two dimensions: **knowledge abstraction** and **knowledge representation** (Saitta and Zucker, 98). Part 4 describes a cartographic generalisation process developed from the previous considerations and explains how to perform knowledge acquisition. Part 5 describes an experiment to acquire these kind of knowledge with machine learning techniques.

2 KNOWLEDGE BASED SYSTEM AND KNOWLEDGE ACQUISITION FOR CARTOGRAPHIC GENERALISATION

2.1 Problem specification

In the scope of this article, we focus on one of the important problem for automating cartographic generalisation: *given one geographical object (a bend, a road, a town…) how to represent it on a map, knowing the specifications of this map?* Other problems such as "the interrelations between objects" or "the way to link the representation of an object to the specification of a map" are beyond the scope of this work (see Ruas, 99, for an approach to these problems).

Several approaches have been proposed to address this problem. In particularly, many operations have been developed to specific types of object. A test on several generalisation platforms accounts for these specific operations and for there scope (OEEPE, 98). Many space analysis tools have been also developed to describe geographic objects. However, identifying the operation to apply to a given object is still an open question.

An approach to face the need for automation of cartographic generalisation is to build expert systems. Indeed, they have proved to be efficient in numerous fields where knowledge requires to be introduced. There are situations where it is difficult to acquire from experts the knowledge necessary to the system. This problem is well known as the «**knowledge acquisition bottleneck**» and has been underlined in the field of cartographic generalisation by Weibel *et al* (95).

Cartographic knowledge needs to be elicited so as to be used by expert systems for cartographic generalisation. Cartographic knowledge acquisition is problematic because cartographic rules are numerous, contradictory and not formalised. For example, in the rule "enlarge significantly non legible road bends so that they become legible but keep the planimetric accuracy and the shape as much as possible", how to formalise "significantly" "enlarge but keep accuracy" "legible" "shape" "as much as possible", etc. ? Most of the time, cartography experts cannot formalise these rules in such a way that they become understandable by a computer.

Machine learning techniques are one of the solutions developed in Artificial Intelligence area to solve this knowledge acquisition bottleneck. Their aim is to automatically build some rules from a set of examples given by an expert. The expert provides some examples in a form of, on the one hand a description of an object and, on the other hand a classification of this object. Machine learning techniques automatically build some rules from these examples to explain the classification from the object description. These rules can be then used to classify new examples provided to the system.

This paper proposes a knowledge based system to answer to our problem and we study, as suggested by Weibel *et al* (95), the opportunity to use Machine Learning to support the knowledge acquisition process.

2.2 Second generation expert systems and knowledge acquisition

The ultimate goal of our research being the development of generic cartographic expert system, this chapter briefly recalls the principles of knowledge acquisition in Artificial Intelligence.

Most traditional expert systems follow the same schema. They contain a knowledge base, a fact base and an inference engine. The knowledge base contains (usually as a set of rules) knowledge acquired from experts. Each rule is supposed to be a piece of knowledge independent from the other pieces. The fact base contains knowledge describing the initial facts about a problem to be solved. The inference engine is the program building the reasoning of the system by applying relevant rules of the knowledge base to the facts of the fact base to deduce new facts corresponding to a possible problem solution. The inference engine is supposed to be independent from the problem to be solved, for example the order of use of the rules is determined by the inference engine with generic techniques valid for all problems.

One of the key application field for Machine Learning has been and still is to automatically build rules from examples (Bergadano, Giordana and Saitta, 91). Most rule learning systems use a sequential covering algorithm where each rule is considered as an independent part of the expert knowledge to be acquired. Critics of first generation expert systems by Clancey (83) showed that these traditional systems implicitly mix a lot of different kind of knowledge. In particular they mix knowledge about basic inferences to be drawn and knowledge about how to organise these inferences. This mix of knowledge types is a limitation to the easy maintainability of the system and the comprehensibility of the reasoning done by the system.

These critics paved the way to new researches in knowledge acquisition area, and led to the development of second generation expert systems (David *et al*, 93) where control over inferences to be drawn is considered as a kind of knowledge in itself and explicitly introduced in expert systems.

It becomes necessary to understand inference steps involved in a problem solving (like generalising a map) and acquire (by learning or direct Knowledge Acquisition techniques) distinctly knowledge necessary to draw each inference. Rules are then more comprehensible (so they are easier to understand, validate, and update) and more easily acquired.

The following section first identifies which types of knowledge are involved in cartographic generalisation and how they are related each other. A knowledge based system to answer our cartographic generalisation problem is then designed from this analysis. An approach to use Machine Learning techniques to learn each identified type of knowledge that is not easy to directly acquire from experts is finally proposed.

3 CARTOGRAPHY IN A KNOWLEDGE REPRESENTATION AND ABSTRACTION MODEL

3.1 Differentiating representation and abstraction

Representing knowledge is one of the main research topic in Artificial Intelligence. The AI community has come out in the past fifty years with a large variety of languages that are more or less adapted to represent different field of humans knowledge (see Ginsberg, 93). Although a large amount of human expertise can be formulated as a set of specific procedures or inferences in one given language, the cartographic generalisation process clearly requires several knowledge representation languages to capture the different types of knowledge manipulated, ranging from the raw data to its final representation as a usable map. Saitta and Zucker (98) have proposed a model of abstraction (hereafter called the KRA model), supporting reasoning in a wide context. In this model, they distinguish two fundamental processes, namely the process of changing the nature of the language of representation and the process of abstracting it.

The KRA model originates from the observation that the conceptualisation of a domain involves at least four different levels. Underlying any source of experience there is the world (W), where *concrete* objects reside. However, the world is not really known, because we only have a mediated access to it, through our perception P(W). At this level the percepts «exist» only for the observer and only during their being perceived. Their reality consists in the “physical” stimuli produced on the observer. In order to let these stimuli become available over time, for retrieval and further reasoning, they must be memorised and organised into a *structure* S. This structure is an *extensional* representation of the perceived world, in which stimuli related one to another are stored together into tables. This set of tables constitutes a relational database. Then, in order to symbolically describe the perceived world, and to communicate with other agents, a *language* L is needed. L allows the perceived world to be described *intensionally*. Finally, a theory T might be needed to reason about the world. The theory may contain general knowledge, which does not belong to the specific domain, and allows inferences to be drawn. Let us define R = < P(W), S, L, T > as a *Reasoning Context.* The relationships among the considered levels are represented in Figure 1.

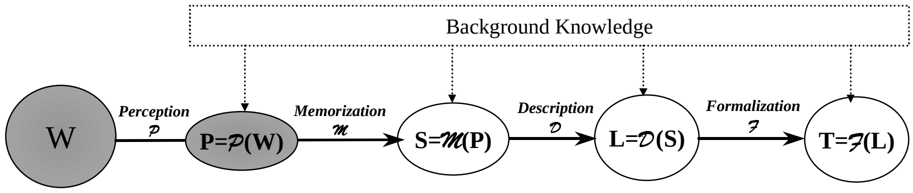

Figure 1 Knowledge representation

There is an infinity of ways in which the world can be perceived by an intelligent agent, according to the observation tools, the goal of the observation, the agent's cultural background, and so on. This variability is captured by the diversity of the world perceptions P(W). It is at this layer that is established the type and amount of information the agent will memorise, speak about, and reason about later on. The less detailed the perception, the more abstract. Sometimes the agent has control over the perception, in such a way to collect exactly the information it needs to achieve its goals. Sometimes the agent can not control the perception, so that it may receive much more information than it currently needs, or maybe it wants to perform several tasks, each one requiring different pieces of information, which, on the other hand, are easier to collect together. The preceding considerations suggest that it would be very useful to have methods to actually or virtually transform a perception into a more abstract one. The abstraction process, as defined by Saitta and Zucker (98), starts at the perception level, but propagates toward the layers of Figure 1. However, the abstraction relations between the structures, the languages and the theories are shaped from the relations defined on the perceptions. In Figure 2, the view on abstraction presented in this paper is synthetically described. The symbols ω, σ, λ and τ denote abstraction operators working between entities of the same layer.

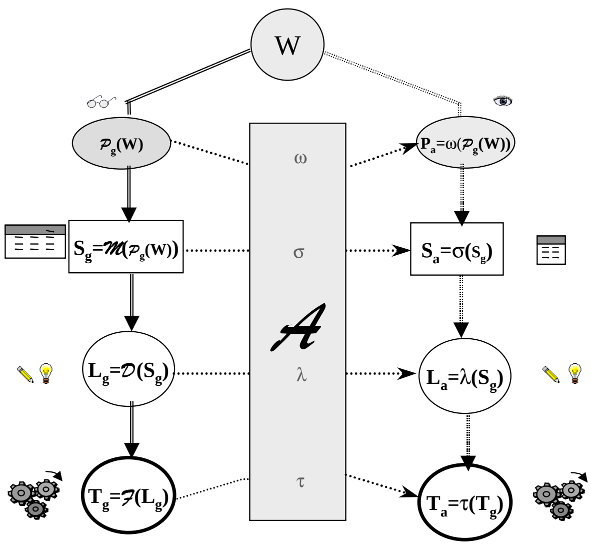

Figure 2 Knowledge abstraction and representation

3.2 Cartography in the KRA model

The topographic map production process closely parallels the KRA model, because it can be analysed according to the two dimensions, representation and abstraction. Let us first consider the scheme of Figure 1 applied to cartography. The first step of cartography is to collect data from the geographic world or part of it (W). This is usually done through aerial photographs, satellite images or field survey which are the perceived world P(W). Objects contained in these photographs are located and labelled to create a geographic database

(GDB). This GDB is the set of geographic data organised in a Structure (S). This GDB is then displayed by means of cartographic symbols applied to objects stored in it. This is the creation of a map, an iconic language (L). Finally, maps are created for specific tasks (e.g. space analysis, search for itineraries, geographic theory construction). The theory (T) is the result of map analysis and is guided by all the background facts and laws allowing one to reason about geographic configuration.

But, **Cartography is not just knowledge representation.** All steps of map creation do not only involve knowledge representation, but also knowledge abstraction, because each step retains only part of available information (e.g. a GDB does not contain all information seen on an image). In particular, map creation (which contains the generalisation process that we detail in the next chapter) is both a knowledge representation process, when objects are symbolised, and a knowledge abstraction process, when objects relevant to the theory construction are identified. So described, map creation is represented as a diagonal process in Figure 3.

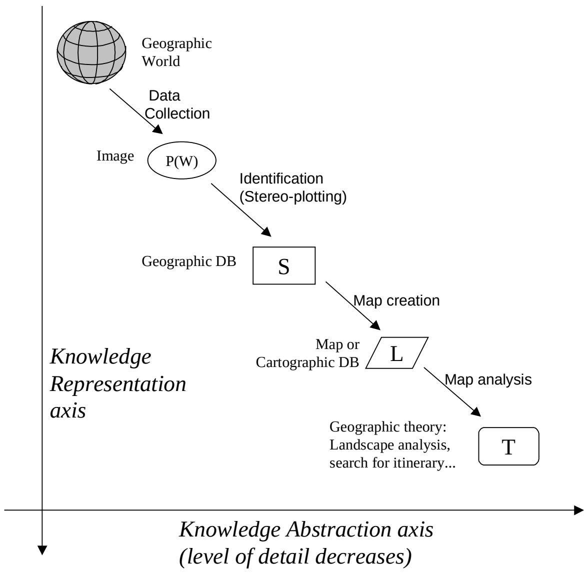

Figure 3. Cartography in the abstraction/representation space

The KRA model exhibits key properties for cartography. It allows the process of representation (change of language) to be distinguished from the process of abstraction (change of level of detail). These two processes are usually very much entangled in cartography. This distinction provides the basis for automating cartographic knowledge acquisition, as a combined acquisition of specific knowledge for abstraction and knowledge for changing representation, as we will explain in the following section.

3.3 Generalisation process in the KRA model

Knowledge abstraction in cartographic generalisation is the identification of abstracted geographic objects relevant to the theory construction that will be done from the map. By abstracted object we mean an abstracted description of the characters of the objects. For example, when we consider that the shape of a road is more important than its accurate location, we make an abstraction as we do enhance only one of the properties of the consider objects. The abstraction process goes from a detailed description of a geographic object, considering each part of the object, to a more abstract description of the object, retaining only properties of the object relevant to the map user's needs.

Knowledge representation in cartographic generalisation is the process of symbolising geographical object and is guided by the abstracted object. For example, the representation process is to determine how to represent a road (the geographical object) so that its shape (the abstracted object) is well legible. This choice is guided by the necessity to well represent the abstracted properties and is restricted by the drawing possibilities (we cannot represent all the bends of the road and keep the planimetric accuracy as we have to enlarge non legible bends).

It is important to notice that **knowledge abstraction and representation are not independent,** nor that when abstraction has been done the "ground" GDB is no more necessary. For example, considering a road as a "sinuous road" (the abstracted object) help us to change our view of the world and to decide how to represent it, but the representation process needs to look again at the ground GDB to create a simplified representation of the actual geographic object. In this way we imitate the human perception, which continuously changes the level of abstraction to analyse space. These inter-links between abstraction and representation explain why, manually, the cartographer has always performed these two steps in one time. However, this distinction between abstraction and representation is necessary for the creation of an automated process of cartographic generalisation.

4 DESIGN OF THE CARTOGRAPHIC KNOWLEDGE BASED SYSTEM

4.1 Knowledge based system description

A simple and preliminary design of a cartographic knowledge based system answering our problem (how to represent on a map a geographic object?) could be the step of figure 4.

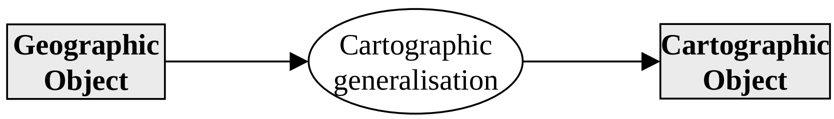

Figure 4 First design of the system

However, as explained in part 2, knowledge used by the system cannot be well understood in this process. Moreover, knowledge acquisition for this system is very difficult. This comes from the fact that a direct knowledge acquisition from experts is problematic, as shown by researches on cartographic generalisation (Weibel *et al*, 95). Machine Learning is also made difficult because the description space of inputs and outputs of the system are huge. The more complex these spaces are, the more examples are needed in order to expect an efficient learning (be it with symbolic machine learning or neural networks) and the less we can expect good quality learned rules (see theoretical works on computational learning theory, e.g. Valiant, 84, and practical experiments Saitta and Neri 98).

According to considerations of part 2 (separating different types of knowledge) and to part 3 (modelling generalisation as a combination of abstraction and representation). A more appropriate process requires two steps: knowledge abstraction and knowledge representation (figure 5). As defined in part 3, knowledge abstraction input is a geographic object, and its output is an abstracted qualitative description of it. Knowledge representation inputs are the abstracted description (to help to determine how to represent) and the geographic object itself (to perform the geometric transformation). Knowledge representation output is the cartographic object representing the geographic object. Nevertheless knowledge acquisition for this architecture is still problematic, for the same reasons than the one explained for the previous one-step architecture (figure 4).

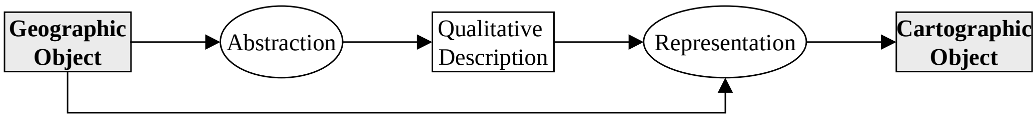

Figure 5 First refinement of the system

In order to facilitate the knowledge acquisition process we eventually designed a system architecture shown in figure 6.

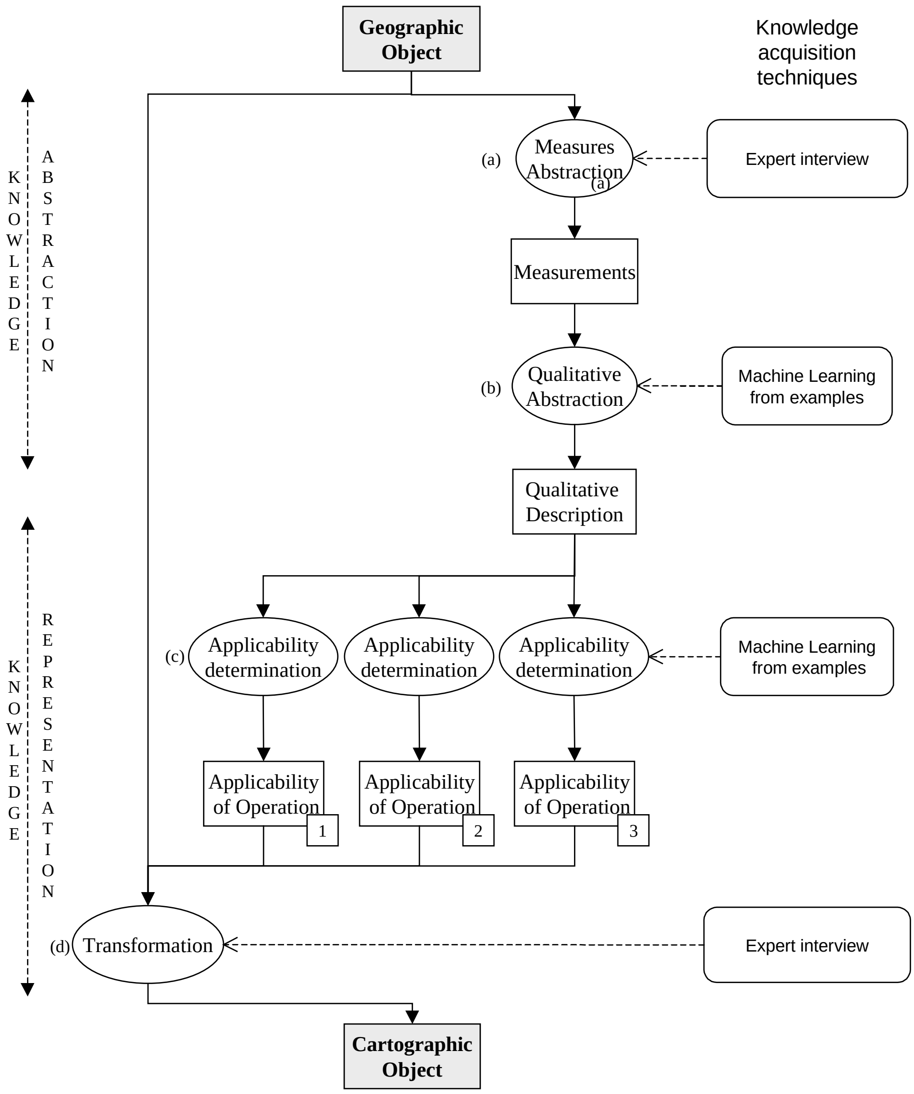

Figure 6. Inferences drawn by the system and knowledge acquisition techniques used

The knowledge abstraction is in two steps:. The first one (figure 6,a) is the description of an object as a set of measures (this abstraction is a way to supply the system with characteristics the cartographer may focus on). The measures to be used is made by an expert of the domain. The description of the considered object as a set of measures is therefore greatly simplified over the initial vector representation. Then the second abstraction step (figure 6,b) goes from this set of measures (surface, elongation…) to a qualitative description of the object. That means that a set of abstracted qualitative descriptors (size, shape…) are determined to enhance the most important characters of the object to deal with (this is the description the cartographer will use to reason about).

The knowledge representation step is decomposed in several steps. First, for a given object, and a set of possible transformation operations to apply on it, the system determined if each operation is applicable or not (figure 6,c). A distinct set of rules is necessary to determine the applicability of each operation, but this set of rules use the same abstracted description. Then from this set of applicable operations the system determine which operation to use (by mean of a strategy to be defined), and apply it on the geographic object (figure 6,d).

The reason why a one-step knowledge representation has not been chosen is that it is most of the time impossible for an expert to say which operation is the best to apply on an object. Indeed in most cases several operations are equally applicable, learning which operation is the best is then a "non-exclusive classes" problem (for one object several classes are possible), which is a rather original setting for machine learning and is a more difficult problem to handle. Designing the system the way we did allows us to work with several "exclusive classes" problem, because the applicability of an operation cannot be at the same time "applicable" and "not applicable". This design allows us to acquire necessary knowledge, as explained in the next chapter.

4.2 Knowledge acquisition for the system

Among the sources of knowledge used to fill in the knowledge based system some may be acquired through interview and others automatically learnt (right part of figure 6). This section presents different sources used to acquire this knowledge.

The first abstraction step of the system requires a selection of measures to describe an object. Knowledge necessary to this step can be directly acquired from experts on space analysis who can determine relevant measures to well describe a type of objects. This knowledge is clearly different for each type of object. For example, if we deal with *buildings* an expert will tell us to prefer measures of *surface, elongation, compactness, etc…*If we deal with roads an expert will tell us to prefer measures of *length, sinuosity, etc*

The second abstraction step goes from measurements to a qualitative abstraction of the object. Part of this knowledge can be directly acquired from experts in cartography. They can specify which characters of the object are relevant. For example, if we deal with *buildings* a cartography expert will tell us that an important character is the shape, and more precisely that rectangle-shaped and L-shaped *buildings* are both important types of *building*. But the expert in cartography is most of the time unable to link these qualitative descriptions to the previous set of measures. For example, to describe a *building* an expert will say whether it is small, medium or big, but he will not be able to say that a small *building* is one with a surface smaller than 300 m^2. He is able to show examples of small *buildings* but usually unable to provide a threshold on the surface measurement to characterise "small" *buildings*. We will so therefore learn automatically this knowledge from examples, as described in the next part.

Then, for each possible transformation operation, we need to know whether it is applicable on each given object. Experts are not able to formalise this knowledge because either they do not know the operations or operations are too complex to be easily controlled. However, if we give examples of objects transformed by a specific operation, a cartography expert is always able to say for each example if the result is acceptable or not and therefore if this operation is applicable. This “applicability” knowledge may be learned from examples, as described in the next part.

Finally, the system has to make the choice, from the set of applicable operations, of an operation to effectively apply on an object. Some complex reasoning, acquired from cartography experts, can be done to choose this operation (Ruas 99) but this aspect is out of the scope of this paper.

Knowledge acquisition for each part of the system may not come from the same expert, as some experts are specialised in space analysis (in particularly in measures) and others are more specialised in cartography (and particularly in generalisation). As the described approach separates each type of knowledge, it is possible to acquire them from different experts.

5 MACHINE LEARNING EXPERIMENT

This section presents a first experiment performed to automatically learn knowledge that the system requires to deal with the particular case of cartographic generalisation of *buildings*. As specified above both qualitative abstraction knowledge and applicability knowledge are to be learned.

5.1 Experiment design

In order to collect data necessary for our test, a "space analysis and algorithms expert"[1] was asked to define:

- a set of measures describing *buildings* (see Table 1. below). Algorithms chosen to compute these measures are defined in Regnauld (97) ,
- a set[2] of "operations" applicable to *buildings* (listed in Table 3).

Then a "cartography expert"[3] was asked to define a set of qualitative abstract descriptors for a given *building* that are somehow related to the above-mentioned measures (see Table 2. below),

Measure Name	unit	Possible values
Minimum length	m	a positive real
Surface	m^2	a positive real
Minimum width	m	a positive real
Concavity	none	a real in [0,1]
Compactness	none	a real in [0,1]
Depth of the biggest yard	m	a positive real
Elongation	none	a real in [0,1]
Squareness	radian	a real in $[0,\pi/2]$
Number of points	none	a positive integer

Table 1 Measures (Regnauld, 97)

[1] N. Regnauld from Edinburgh University
[2] Regnauld *et al.* (98) describe this combination of algorithms and show that each combination is efficient on some buildings, nevertheless they do not know when a particular combination is efficient.
[3] S. Mustière from IGN-France

Name	Possible values
Global shape	Rectangle, L-shape, other
Size	Small, Medium, Big
Number of main orientation	One, Several
Granularity	High, Small
Existence of big wings	None, Some
Existence of small necks	None, Some
Existence of special shapes	None, Triangle or Circle

Table 2 Abstracted descriptors

Name
Simple dilatation
Simplification
Squaring
Squaring / Simplification
Simplification / Squaring
Simplification / Squaring / Enlargement
Simplification / Enlargement / Squaring
Squaring / Simplification / Enlargement

Table 3 Possible operations (Regnauld *et al*, 98)

The cartography expert was then provided with a set of 80 observations of *buildings* and he was first asked to describe each *building* with the defined qualitative descriptors (this *building* has the shape "L- shaped", the size "medium", it contains "no big wings"…). The same set of *buildings*, each one generalised by each operation was presented to him and he was asked to say for each generalised *building* if the result was acceptable or not. Meanwhile the set of measures chosen by the expert (see Table 1) was computed on each *building*.

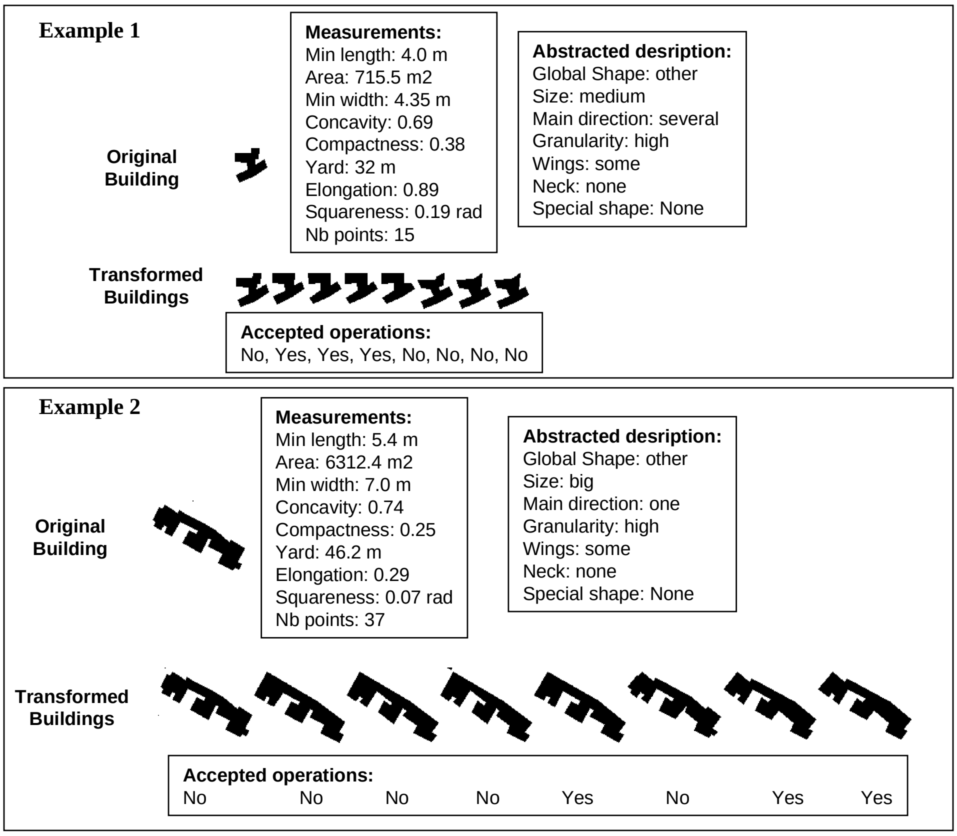

Figure 7 Two of the eighty examples of *buildings* used in the machine learning test

Two examples of *buildings* are shown in figure 7. Eighty examples were used to first learn how to link measures to abstracted descriptors, then to link abstracted descriptors and applicability to each operations.

As we intend to learn comprehensible rules, we applied on these examples symbolic machine learning tools. We used the C4.5 algorithm (Quinlan, 92) to learn the qualitative abstraction step expertise. This algorithm creates decision trees from examples. We then used ENIGME (Thomas, 96) to learn the applicability of operations expertise. This algorithm produces "if… then …" rules. To select when to use a decision tree learner or a rule learner, we did some empirical tests. We noticed that for our problem, decision trees were easier to read for the qualitative abstraction step and rules were easier to read for the applicability step.

Some results of these tests (performed on eighty *buildings*) are presented in the next section.

5.2 Machine Learning results

Due to space limitations, we are unable to show all the learned rules and did extract some of the most typical learned rules. Figure 8 shows a decision tree automatically learnt from examples to determine the abstracted character *Size* from a given set of measures. This simple tree demonstrates several advantages of using of Machine Learning and of our approach:

- the system learns the relevant measures for determining the size (Surface and Concavity).
- the system learns relevant thresholds although this is always a difficult problem in cartography. For example it learns that the expert considered *buildings* smaller than 137 m^2 as *small building* (they are actually the individual houses), and *buildings* bigger than 1168 m^2 as *big buildings* (most of the time industrial *buildings*).
- The system learns rules that the expert would not have thought useful to express. For example it learns that the concavity influences the perception of *buildings* size. This is well known in psychology that the shape of an object influences its size perception, but there are great chances that an expert does not explicitly express this rule.

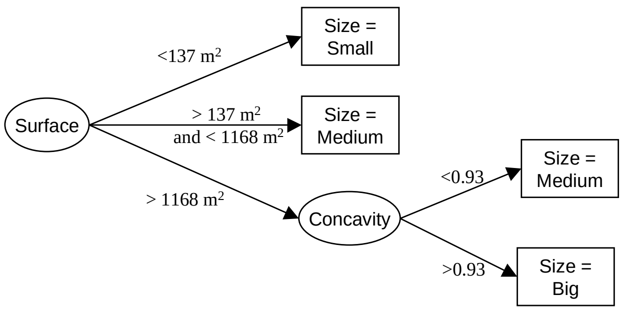

Figure 8 Learnt decision tree for Size determination

Figure 9 shows some rules automatically learnt to determine applicability of operations. In average, fourteen rules were produced for determining the applicability of each operation (seven to describe cases when the considered operation is applicable and seven cases when it is not applicable). These rules show some advantages of our approach:

- Learnt rules are simple (with a maximum of 4 abstracted characters included in the premises). They are easy to understand, which is not the case for rules learnt in one-step learning experiments.
- Each rule is independent from the others, it is as easy to remove one of them than to add a new one.

- Some rules have been created to determine the applicability of complex operations (like "Squaring, Simplification, Enlargement") which is not easy to do without machine learning because it is difficult to control what this operation does.

If Size = Big and Wings = None and Granularity = small and special_shape = none
then Squaring is applicable

If Shape = Rectangle and Granularity = small
then Squaring is applicable

If Shape = Big and Main_direction=Several
then Squaring is not applicable

If Size = Big and Special_shape = circle or triangle
then Simplification is applicable

If Global_shape = L-shape and Size = Small and Special_shape = circle or triangle
then operation « Squaring, Simplification, Enlargement »is applicable

Figure 9 Learnt rules for determining operations applicability

6 CONCLUSION

This paper support the view that knowledge acquisition is a key problem for cartographic generalisation. We showed that this knowledge acquisition bottleneck may be overcome by analysing the different types of knowledge involved in the cartography process. Such an analysis based on a model developed in Artificial Intelligence lead to consider cartographic generalisation as a combination of abstraction and representation. This analysis supports the conception of a knowledge based system developed to, on the one hand, explicitly represent knowledge involved in cartography and, on the other hand, acquire part of these knowledge through machine learning techniques. First experimental results on the knowledge acquisition for *isolated buildings* generalisation are very encouraging. Ongoing work includes an extensive evaluation of the quality of these acquired knowledge as well as the evaluation of the cartographic quality of new buildings generalised following the expert system proposed operations.

ACKNOWLEDGEMENT
We wish to thank Nicolas Regnauld, from Edinburgh University, for providing all our test data, and for being our expert who determined relevant operations and measures on buildings.

REFERENCES

Armstrong M.P. 1991 "Knowledge Classification and organisation". in *Map Generalization : Making Rules for Knowledge Representation.* McMaster & Buttendfield (eds) Longman Scientific & Technical pp 86-102

Bergadano, F., A. Giordana, Saitta L. (1991). *Machine Learning: An Integrated Framework and its Application*. Chichester : Ellis Horwood.

Clancey, W. (1983). "The Epistemology of a Rule Based Expert System – A framework for Explanation", *Artificial Intelligence,* Vol. 20, pp. 215-251

David J.-M., Krivine J.-P. and Simmons R. (eds.) (1993). *Second Generation Expert Systems*, Springer Verlag.

Ginsberg M., (1993). *Essentials of Artificial Intelligence*, San Francisco: Academic Press / Morgan Kaufmann publishers.

João E. M., (1998). "The Influence of Geographic Scale on the results of EIA when using GIS". *Proceedings of 1st GIS PlaNet conference*, Lisbon, Portugal, Sept. 1998.

Quinlan J.R. (1992). *C4.5 : Programs for Machine Learning*. Morgan Kaufmann Publishers, series on Machine Learning.

Regnauld N. (1997). *Généralisation du bâti: Structure spatiale de type graphe et représentation cartographique.* PhD Thesis. LIRM – Marseille, France

Regnauld N., Edwardes A. and Barrault M., (1998) "Strategies in Building Generalisation: Modelling the Sequence, Constraining the Choice". *ICA Workshop on Progress in Automated Map Generalization.* Ottawa

Ruas, A. (1997). "Strategies for Urban Map Generalisation". *Proceedings of ICC* Vol. 3 pp. 1387-1394

Ruas, A. (1999). *Modèles de généralisation de données géographiques à base de contraintes et d'autonomie.* PhD Thesis, University Marne-la-Vallée, France.

OEEPE (1998) "First results on the OEEPE test on generalisation" *OEEPE Newsletter* (1)

Saitta, L. Neri F. (1998). "Learning in the "Real World"." *Machine Learning*, No. 30 (2/3), pp. 133-163, February 1998.

Saitta L., Zucker J.-D. (1998). "Semantic Abstraction for Concept Representation and Learning". *Proceedings Symposium on Abstraction, Reformulation and Approximation* Pacific Grove, California

Thomas J. (1996), *Vers l'integration de l'apprentissage symbolique et l'acquisition de connaissances basée sur les modèles: le système ENIGME.* PhD Thesis, University Pierre et Marie Curie, Paris, France. December 96.

Valiant, L. (1984). "A theory of the learnable". *Communications of the ACM*, 27(11), p. 1134-1142

Weibel R., Keller S. & Reichenbacher T. 1995 "Overcoming the Knowledge Acquisition Bottleneck in Map Generalization : the Role of Interactive Systems and Computionnal Intelligence." Proceedings of COSIT'95, p. 139-156.