**Revision of cartographic generalisation rule bases funded on interactive alteration analysis**

Sylvain Bard & Sébastien Mustière
IGN – COGIT Laboratory
2-4, avenue Pasteur
94 165 ST-MANDE Cedex, FRANCE

sylvain.bard@ign.fr, sebastien.mustiere@ign.fr

**Abstract**

This paper deals with researches in the field of automatic cartographic generalisation. One difficulty to this automation resides in the need of formalising cartographic knowledge to guide the application of generalisation algorithms. This formalisation can be done by the means of rule bases. But rule bases are not perfect. Cartographic generalization results needs an alteration process to correct errors. Our aim is to take advantage of this alteration process to improve rule bases. We define thus a method which analyses the alteration phase to detect and improve deficient rules from a rule base, in order to obtain generalisation results as perfect as possible. In the field of artificial intelligence this process is called revision.

**Key words**: automatic cartographic generalisation, revision, alteration

**Gelöscht:** :

## Introduction

Generalisation process consists in representing in a simplified way a same geographic space, according to the scale and objectives of the map to realise. For this aim, less important information are deleted and the more important are emphasised. Traditionally, cartographic generalisation was manually performed. Nowadays, researches are led to automate this process. It is a well known fact that generalisation is a highly complex process to automate, but the application stakes of such process are to reduce both time and cost of production.

To achieve this aim, researchers developed first, a lot of generalisation algorithms in order to imitate the cartographer drawing, and then sequences to mime his reasoning. Nevertheless, some specific cases are not well managed, and researches in generalisation show that one main issue is to know how to apply the good algorithm or sequence of algorithms on the right place, at the right time [McMaster and Shea 92].

Knowledge bases can be used to guide the application of different generalisation algorithms on specific geographic objects. But these knowledge bases are difficult to construct, they so contain errors leading to the application of non adapted algorithms in certain cases, and then leading to badly generalised objects.

The aim of this research is to find a method to detect and improve deficient rules from a knowledge base, in order to obtain generalisation results as perfect as possible. In this paper we are discussing the use of revision techniques from artificial intelligence as a possible way to achieve this goal. The key of this approach is the determination of the enough and necessary type of information to guide the revision.

In section 1 we briefly discuss the advantage of using knowledge bases in generalisation. We show that those knowledge bases are not perfect and we investigate in section 2 some means to improve them, we focus in particular on revision techniques. In the third section we first detail our model, then the need to apply a revision process to cartographic knowledge bases. To illustrate our talk, we implement our model and discuss the results of experiments in section 4. Finally, we conclude this paper by providing some guidelines for a global framework to make a revision process.

## 1- Using knowledge bases in cartographic generalisation

In this section we present the motivation and the context of our work: why and how using knowledge bases to guide the generalisation process. In the scope of this article, we focus on one of the important problems 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).

**Gelöscht:** :

**Gelöscht:** :

### 1.1- Using knowledge bases

Many different generalisation algorithms exist. One key issue concerning the automation of cartographic generalisation is to know when and where applying these algorithms [McMaster and Shea 92; Ruas 99]. A test on several generalisation platforms accounts for the specificity and for the scope of the algorithms [Ruas 98]: one algorithm is only efficient on certain objects. Knowledge guiding the algorithms application must be formalised to face the need for automation of cartographic generalisation.

**Gelöscht:** :

Many space analysis tools have also been developed to describe geographic objects. Spatial measures can be used to analyse the objects properties, and then decide which algorithm can be applied on a given object. But, these links between measures and algorithms depend on complex and heterogeneous knowledge about the measures, the algorithms, the geographic objects and cartographic rules. An approach to formalise these knowledge is to build expert systems, using rule bases to know under which condition an algorithm is applicable. Indeed, expert systems have proved to be efficient in numerous fields where complex knowledge require to be introduced.

*1.2- Building knowledge bases*

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 [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 easily formalise these rules in such a way that they become understandable by a computer.

Several knowledge acquisition techniques exist such as text or map analysis [Nickerson 91], guided expert interviews [Kilpeläinen 2000], process tracing [Weibel 91] or machine learning from examples [Weibel et al. 95; Mustière et al. 2000].

For the experiments developed in our work (cf. part 4), we use knowledge bases built with machine learning techniques in the special case of road generalisation [Mustière 2001]. Machine learning techniques are one of the solutions developed in artificial intelligence area to solve the 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 (in our case measures computed on an object) and, on the other hand a classification of this object (in our case the chosen algorithm). Machine learning techniques automatically build some rules from these examples to explain the classification from the object description.

*1.3- Improving knowledge bases*

Whatever the used technique to build knowledge bases, these are not perfect. On the one hand, the used knowledge sources can be biased: for example in the case of machine learning some examples can be wrong and some special cases can be not illustrated in the examples. On the other hand, the knowledge elicitation techniques can provide errors: for example the machine learning algorithms can provide by chance unrealistic rules based on few examples. The application of algorithms proposed by the rule bases on a set of objects can so provide badly generalised cartographic results. In a context of production, badly generalised objects must be altered in order to improve the cartographic results. The alteration phase can be costly and should be minimised. Evaluating the quality and improving rule bases is then an important problem to improve the efficiency of automated generalisation processes.

In this article we try to explore the means to perfect such rule bases, answering questions as:
- Do any mean to perfect base of rules exists ?
- What could be those means in the particular case of generalisation ?
- Is it possible to apply them and how ?

**2- How to improve a knowledge base**

In this section, we investigate two possible techniques to improve a knowledge base, validation and revision which are artificial intelligence techniques, and we summarise them.

*2.1 Validation*

Validation consists in estimating the internal quality of a knowledge base in order to suggest improvements to the expert if necessary. In other words, giving a specific criterion, the goal of validation is to show that the knowledge base respect it. New knowledge are not introduced, it is just a reorganisation of the knowledge base. A criterion generally investigated is coherency: a knowledge base is coherent when contradictory results can never be deduce from it.

Validation do not test knowledge base exactitude, but the presence of logically incoherent rules:
- Conflicting rules, e.g. when two true rules end in contradictory results,
- not used rules (unsatisfiable, unusable, subsumed rules), e.g. when it exists two rules but one is more general
- badly organised rules, e.g. when a series of rules form a buckle.

*2.2 Revision*

Another method to improve deficient rules is called revision. The aim of revision is to detect the weak points in a knowledge base, leading to wrong conclusions. It consists in modifying a knowledge base for integrating new knowledge obtained

thanks to external knowledge.

Revision brings on a close world which can not evolve, this is only knowledge on it which change [Alchourrón, Gäredenfors et Makinson 85]. Revision allows the addition of new knowledge. It is advisable then to distinguish revision from update. Update consists in reporting a change of the world which is not closed. Revision can be extended to this case [Katsuno et Mendelson 91]. In our case, we work on the case where the world is fixed and it is our knowledge on it which evolves.

We find in artificial intelligence literature revision method close to our problematic based on three basic operations:

- modification, it consists in changing a piece of knowledge in the initial knowledge base.
- removal, it consists in removing a piece of knowledge in the initial knowledge base.
- addition, it consists in adding new knowledge to the initial knowledge base.

Those operations lay down the detection and localisation of a conflict in the knowledge base.

*2.3 Conclusion*

Validation and revision are both different operations, the first checks knowledge base consistency with internal information, whereas the second uses exact external information for improving knowledge base exactitude.

In our approach, we dispose of an external source of information: the alteration process. Consequently, we can analyse this procedure to extract relevant information, and use it to perform revision. As the interactive generalisation process could be traced to create knowledge bases [Weibel 91], the interactive alteration process could be traced to revise knowledge bases. In a more precise way alteration is the operation which consists in modifying an object, in order to return it acceptable in a generalisation point of view. For instance we can imagine an interactive correction of generalisation conflicts (alteration phase). We could extract some information which permit to detect the deficient cartographic knowledge in the knowledge base. The stake of this phase is the integration of the right information in the revision process, i.e. how to extract relevant information from the alteration process in order to guide revision.
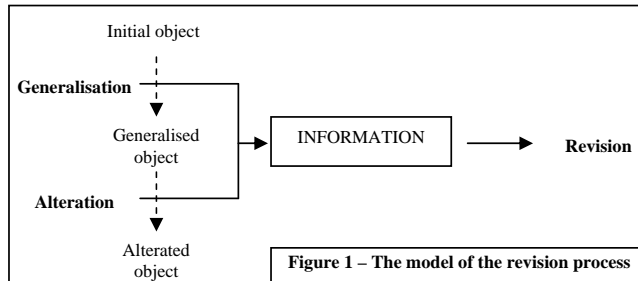
Thanks to the analyse of alteration phase, we explore in the following a possible framework to the revision of a cartographic knowledge base.

## 3- A framework using revision to cartographic knowledge base

*3.1 The model*

Revision process comes within a general process of knowledge acquisition [Bard 00]. We have decomposed it as a retroactive loop. First the knowledge base is used and then it is revised in to order to make it as exact as possible. The purpose is to improve a cartographic rule base.

In a general point of view, three steps form our approach: the application of the knowledge base, the extraction of relevant information and the revision as we can see in figure 1.

Figure 1 – The model of the revision process

The process starts with the generalisation of an initial object which becomes a generalised object by applying algorithms chosen by the cartographic knowledge base. Notice that this step is an irreversible change of state because generalisation is not a reversible operation. Later, the generalised objects are corrected thanks to an alteration process. We deduct from the previous states geographic information to perform revision. It concerns objects and the knowledge base to perform revision. The last stage consists in determining the right operation of revision to perfect the knowledge base.
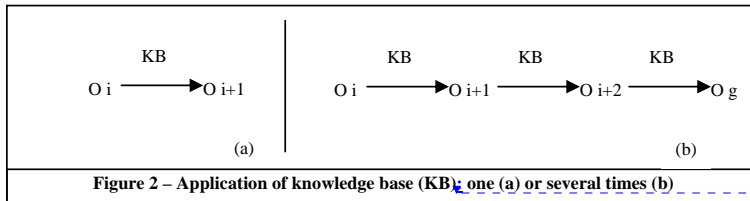
*3.2 An information dependent on the design choice*

In this section we detail the generalisation design choice to show that it influences the shape of the revision. Extracted information are different according to the type of knowledge base and alteration used.

3.2.1 The knowledge base application

The knowledge base can be applied one or several times as it illustrates in figure 2. In the first case the generalised object is directly obtained after the application of the knowledge base. On the contrary in the second a lot of intermediate objects are generalised from the knowledge base. In fact as much as object is not considered well generalised the knowledge base is applied. This impose to dispose of a stopping algorithm which decides to end the generalisation process. Those two types condition the used strategies of revision and the possible type of information to extract.



**Figure 2 – Application of knowledge base (KB) one (a) or several times (b)**

### 3.2.2 The form of the alteration phase

Alteration is the key step as it composes essential information used to guide the revision process. This can be interactively operated by an operator according to two modes:
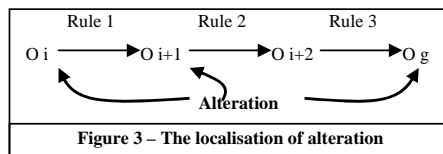- attempt / failed with generalisation algorithm. We can use either the same or another algorithm than the one advised by the knowledge base. The nature of alteration information brings on the parameter used or the new algorithm used.
- draw a better solution. In this case, we only have direct information about the fact that a alteration has been done or not on an object. If the revision process asks to know which kind of alteration has been done, it is possible to try several algorithms on the considered object until obtaining a result similar to the altered object. Anyway, computing this similarity imposes to possess efficient measures comparing the state of two objects.

If several algorithms have been applied on one object during the generalisation process, another alteration information could be the potential localisation of the alteration, on the initial, intermediate or final object (see figure 3):
- on final object, the object state of alteration (Og in fig.3) is different of the initial state of generalisation (Oi in fig.3). It is problematic for the revision because the relevant information comes from the comparison of the two generalisation result (from knowledge base and from alteration). This imposes the same initial objects to generalise. Here, it is not the case and the alteration information would be source of error to perform revision.
- on initial object, initial object states are the same. Generalisation result from knowledge base and alteration application should be compared in order to extract relevant information for revision. The main difficulty resides in the identification of this information, for instance the wrong algorithm among those applied.
- on intermediate object, it allows a backtrack process and is certainly the more precise way to generalise well the object. The main difficulty is as the first point that the information to perform revision is source of errors.



**Figure 3 – The localisation of alteration**

### 3.4 The form of the revision

Perform a right revision process requires relevant information (e.g. the alteration process operation; parameters of algorithms used, failure frequency of a rule) conditioned by both previous phases. Those two phases represent the main source of information. We show that it is necessary to remain watchful and to verify that the information of alteration is on the one hand correct and on the other hand relevant.

We determine four types of possible revision:
- no modification, when a rule is right.
- addition, when one makes defect create a rule, for example typically when a rule is estimated in half of the cases as right and in other half erroneous and retouched always in the same way
- removal, when a rule is never used, nevertheless because of a theoretical principle of revision which sets the preservation of rule base coherence, it is tricky to use.
- modification of a rule in other cases, we define five types: modification of the algorithm parameter, change of measure, rectification of measure threshold, change of algorithm, modification of rule order.

### 3.5 Extract the relevant information and use it in a decision support system

In the previous paragraphs we showed that the design choices influence the form of the collected information. We identified two types of general information that we can collect on each object generalised with the knowledge base:

- information on results of knowledge base application, it is the only source of information about the knowledge base and the used knowledge to generalised initial object.
- information on alteration process, it is the only source of information about objects badly generalised as well as about type of modification brought to correct them.

We should organise all this information to decide on the revision to begin. We decide to use a decision support system. It is based on two methods:

- one method to compare both information in order to obtain meta information. This meta information concerns the erroneous initial knowledge and the knowledge which allowed to correct it
- one method to determine the revision to operate from the previous meta information

## 4- Implementation

In this section we test on a simple application the revision process developed in the previous section. We select few criteria and limit ourselves to a simple knowledge base.

### 4.1 description

We implement this rather simple system on the platform PlaGe[1] [Lecordix, Plazannet, Lagrange 97]. PlaGe is a GIS research platform for generalisation which is composed of algorithms of generalisation and which enables the development of working modules easy to use. The used data are a road network represented in the form of a graph consisted of arcs and nodes. Hereafter we will use the term arcs for geographic objects.

For the sake of simplicity, we study a rather simple problem:
- We suppose that only one algorithm is applied to generalise each object. The used knowledge based has been learnt with machine learning technique
- We decide to feign the interactive alteration process by the mean of a new knowledge base, interactively created
- We reduced the revision phase to the determination of the operation of revision to be made, by the mean of a decision support system. It would have been too ambitious to realise a system which determines the operation of revision and makes it automatically.

The tested cartographic knowledge base is a rule base on road generalisation composed of nine rules of decision based on twelve measures. Rules are of the form "If…Then…", for instance: "If length is lower than 20 meters than use algorithm 1". The form of alteration is another rule base composed of four rules and six measures.

The main characteristic of these two rule bases is the association of an algorithm in every rule because it facilitates the determination of the erroneous rule and allows to propose a simple solution.

### 4.2 Simplified decision support system developed

The developed decision support system is used to test each rule from the initial rule base one by one. This system is based on three criteria which determines if a revision should be made or not, and fixes the modality of this revision. Those three criteria are:

- the percentage of potential error, we fix to 5 % the allowance limit of potential error for a rule, under which the rule is considered as satisfying.
- the number of corrected arcs, we calculate the number of arcs generalised by a rule and then the number of altered arcs among these generalised arcs. It gives us an idea on the representativeness of the rule. This criterion combined in the precedent enables us to define a reliability rate concerning the revision. One can note that for few corrected arcs the system can not advise revision, the information are not representative.
- the comparison of initial and alteration algorithms used to determine revision to operate. According to it, the system proposes three types of revision: removal, modification algorithm, modification algorithm parameter.

The schema of figure 4 illustrates the organisation of our decision support system. It combines the advantages of two decisional approaches: a system and a decision tree. Actually, the system is on the one hand flexible because it permits to the user to decide to revise or not and the type of revision (proposed revision) associated to a reliability rate. On the other hand the decision tree constitutes an inflexible part of the system because it permits to determine the discriminating criteria of decision.
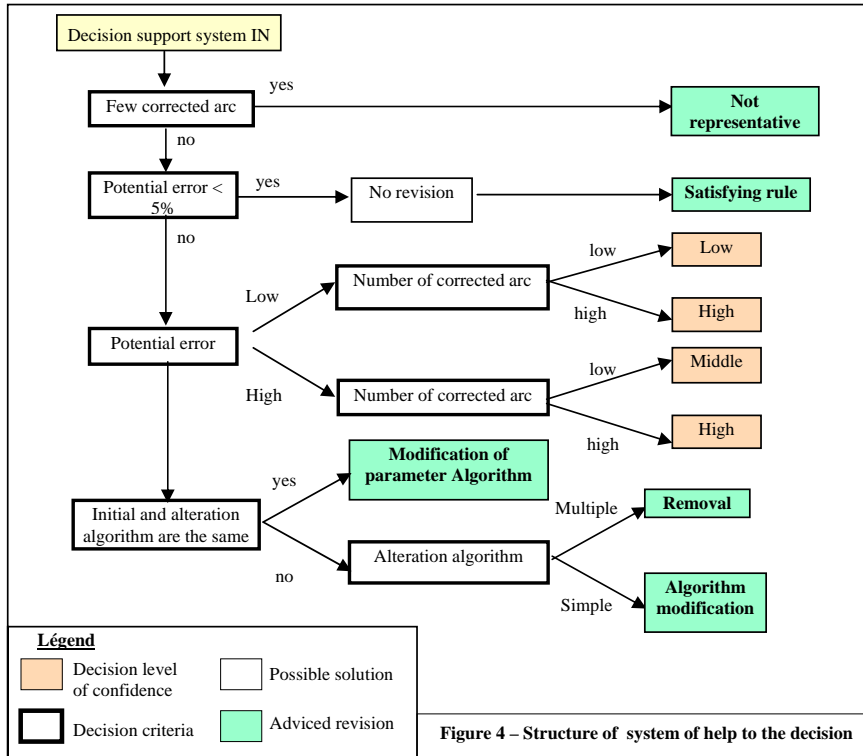
The first steps consists in testing the representativeness of the extracted information and then in testing if a cartographic rule is good or not. For this, we use number of corrected arcs, if there are few arcs the system considers that it is not representative and ends the revision process for the corresponding rule. On the contrary, it tests the percentage of potential error according to the chosen threshold (five per cent). If the treated rule is under, it is considered as satisfying. On the contrary the system continue to the next step.

---

[1] PlaGe for Platform of Generalisation

The next step computes the reliability of the decision (low, middle, high) with the combination of two criteria: percentage of potential error and number of corrected arcs. We associate a high number of corrected arc to a high reliability rate because the high number of arcs is sufficient and rather representative for the considered rule. On the opposite, we consider as less representative a low number of arcs, and in this case this is the percentage of potential error which determines the reliability rate. If it is low, the rate is low, if it is high, the rate is medium.

Finally, the system proposes a revision. If the initial algorithm and the one used for alteration are the same, the system suggests to modify the parameter of the algorithm. If need be, one looks if the replacement algorithm is often the same or not, in the first case one suggests to modify the algorithm, in the second case one suggests to remove the rule.

**Figure 4 – Structure of system of help to the decision**

## 4.3 Result and discussion

The aim of this simple test is to show that under conditions our revision model is implementable. Revision result appears like proposition of revision associated to a reliable rate on it, as in figure 5. We distinguish three cases:

- ▪ Insufficient information
- ▪ satisfying rule
- ▪ unsatisfying rule. It then proposes one of the three possible revision operations: modify algorithm, modify parameter, remove.

**Rule 6**
Altered: 82% for 654 corrected arcs
Reliable rate: high
Alteration proposed: unsatisfying rule, remove it.
Justification: Alteration algorithms are multiple and not the same as initial

**Figure 5 – example of revision proposition of the system**

We tested this method with a road network situated in a mountain area and consisted of 3500 roads. The results show 10 % badly generalised arcs, the operations of revision proposed were not tested but every situation foreseen by the system of help to the decision was met. To go farther it would be necessary to revise in a actual way the knowledge base and to test it to put in evidence an improvement of the percentage of discovered errors of generalization.

## Conclusion

The objective of this paper was to find a method for detecting and improving deficient rules in a cartographic knowledge base, in order to obtain generalisation results as perfect as possible. We first have analysed the interest of using knowledge bases in generalisation and have shown that rules were perfectible. One possible technique to improve them is revision.

This paper proposes a framework to apply revision process to cartographic knowledge base and demonstrates the requirement of an external relevant information to perform it. This imposes to fix what type of information was required and how to use it to decide the revision to operate.

For this, we developed a method to extract information based on an alteration process anlaysis and proposed to use a decision support system to determine the revision to operate. Those steps are under conditions and we have determined requirements to achieve this goal:

- Stock relevant information from initial knowledge base,
- Validate generalisation result after knowledge base application,
- Correct badly generalised features with an alteration process and stock information on it,
- Compare information to obtain meta information for guiding revision,
- Propose revision thanks to a decision support system.

In our approach we only propose a revision, for further research, to be closer to real conditions of application, we should first stock and mark out relevant information from knowledge base and alteration process, and second revise knowledge base in a real way whether manually whether automatically. Moreover, the revised knowledge base should be evaluated in order to test if it is coherent.

At last, thought should be led to reason in term of cost for revision approach. We should wonder about the revision cost in term of time and in term of what is imposed to the alteration. It is right that revision is less expensive than making another knowledge base, but it is not conceivable to persist in revising it when too much rules are wrong. That is why revision should be discussed in terms of cost to determine a point where it is not necessary to begin revision or revise more. We could imagine to validate the application results of a knowledge base and, according to a threshold, decide if revision is possible or not. For instance, it is more expensive to revise a knowledge base with fifty per cent of errors than a knowledge base with ten per cent of errors.

## References

Alchourrón C., Gäredenfors E., Makinson P., 1985. On the the logic of theory change: partial meet contraction and revision functions. *Journal of symbolic logic*. 50(2): pp. 510-530.

Bard S., 2000. *Révision d'une base de connaissances, application à la généralisation cartographique*, MsC report, university of Metz, France.

Katsuno H., Mendelson A.O., 1991. On the difference between updating a knowledge base and revising it. *Proc. of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pp. 387-394.

Kilpeläinen T., 2000. Knowledge Acquisition for Generalization Rules. *Cartography and Geographic Information Science,* vol.27, n.1, pp.41-50.

Lecordix F., Plazannet C., Lagrange J.P., 1997. A Platform for Research in Generalization: Application to Caricature. *GeoInformatica*, 1997, 1(2), pp. 161-182.

McMaster R.B., Shea K.S. 1992. *Generalization in Digital Cartography.* Association of American Geographers, Washington.

Mustière S., Zucker J.-D., Saitta L., 2000. An Abstraction-Based Machine Learning Approach to Cartographic Generalisation. *Proc. of 9th International Symposium on Spatial Data Handling*, Pekin, sec.1a, pp.50-63.

Mustière S. 2001. *Apprentissage Supervisé pour la Généralisation Cartographique*. PhD Thesis, university Pierreet Marie Curis (Paris 6), France.

Nickerson B.G., 1991. Knowledge engineering for generalization. *Map Generalization, Making Rules for Knowledge Representation*, Buttenfield et McMaster (eds), Longman Scientific & Technical: Harlow, Essex, pp.40-55.

Ruas A., 1998. First results on the OEEPE test on generalisation. *OEEPE Newsletter,* 1998-vol.1, pp.5-10.

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

Weibel R., 1991. Amplified intelligence and rule-based systems. *Map Generalization*, *Making rules for Knowledge Representation*, Buttenfield and McMaster (eds), Longman Scientific & Technical, Harlow, Essex. pp. 86-102.

Weibel R., Keller S., Reichenbacher T., 1995. Overcoming the Knowledge Acquisition Bottleneck in Map Generalization: the Role of Interactive Systems and Computionnal Intelligence. *Proc. of 2nd International Conference on Spatial Information Theory*. pp. 139-156.

**Gelöscht:** :

**Gelöscht:** :

**Gelöscht:** :