Proceedings of ACM/GIS'99: Advances in Geographic Information Systems, Kansas City, USA, November 99, pp.162-164

Cartographic Generalization as a Combination of Representing and Abstracting Knowledge

Sébastien Mustière
IGN – COGIT / 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 – CNRS – LIP6
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 shows that cartographic generalization is best viewed as *representing* (formulating, renaming) knowledge and *abstracting* (simplifying) a given representation. The general process of creating maps is shown to fit an abstraction framework developed in Artificial Intelligence to emphasize the difference between abstraction and representation. The utility of the framework lies in its efficiency to help automating knowledge acquisition for cartographic generalization as a combined acquisition of knowledge for abstraction and knowledge for changing a representation.

Keywords

Cartographic Generalization, Abstraction, Representation.

1. INTRODUCTION

In this paper we address the problem of automating cartographic generalization. This automation is needed for several reasons: first, to decrease cost and time necessary to produce maps; then, to allow geography experts who are not necessarily cartography specialists to create their own good quality maps; finally, to facilitate the crucial need of multi-level analysis of geographic data.

The lack of efficient generalization tools in GIS is due to the difficulty of the task: it must be guided by a large body of geographic and cartographic knowledge. An approach to face this need for automation is to build expert systems, which have already proved to be effective in various fields. An important problem needs then to be addressed: many authors emphasize that the main problem for the use of expert systems is the «**knowledge acquisition bottleneck**». Moreover, the analysis of first-generation expert systems [1] stressed the need to differentiate, separate, and structure different types of knowledge.

We present in this article a description of the knowledge used in cartographic generalization well fitted to its acquisition. We analyze generalization along two dimensions: **knowledge abstraction** and **knowledge representation**, as proposed in [4]. This distinction is necessary to differentiate, and therefore to acquire, the different knowledge types involved in generalization

2. DIFFERENTIATING REPRESENTATION AND ABSTRACTION

Representing knowledge is one of the main research topic in Artificial Intelligence since its birth. 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 [2]. Although a large amount of human expertise can be formulated as a set of specific procedures or inferences in one given language or paradigm, the cartographic generalization 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 have recently proposed a model of abstraction (hereafter called the KRA model), supporting reasoning in a wide context [4]. 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 conceptualization 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 memorized and organized 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. The set of these 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. At the theory level we operate through inference rules. Let us define R = < P(W), S, L, T > as a *Reasoning Context.* The relationships among the considered levels are represented in Figure 1.

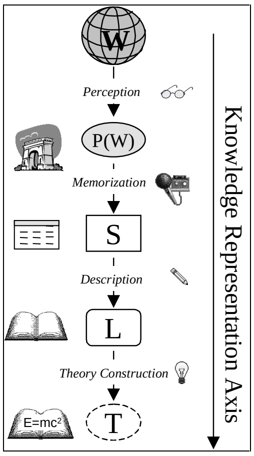

Figure 1

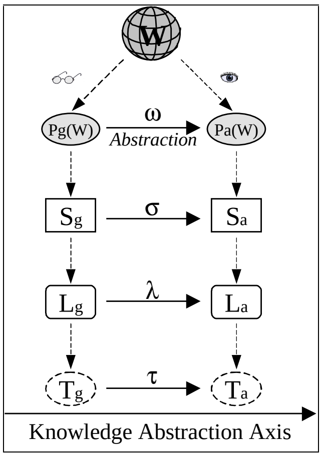

Figure 2

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 the type and amount of information the agent will memorize, speak about, and reason about later on is established. 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 following definition of abstraction tries to capture this process.

Definition – Given a world W, let $R_g = (P_g(W), S_g, L_g, T_g)$ and $R_a = (P_a(W), S_a, L_a, T_a)$ be two reasoning contexts, which we label as *ground* and *abstract*. An *Abstraction* is a functional mapping $A : P_g(W) \rightarrow P_a(W)$ between a perception $P_g(W)$ and a *simpler* perception $P_a(W)$ of the same world W.

Some comments are needed about this definition. In [4] a formal definition of "simpler" in terms of relative information gain has been given. Obviously, the process of abstracting a perception can be iterated, leading to several levels of abstraction. If no perception can be identified as a preeminent one, then any level can be selected as "ground", being the notion of "simpler" only a relative one. Another important point is that abstraction should be a reversible process; in fact, **to abstract does not mean to delete information,** but only to *hide* information, in such a way that the opposite process (concretion) becomes possible, as well. Finally, according to this definition, the abstraction process 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.

2.1 Abstraction Operators in the KRA Model

Within the KRA model framework, a set of fundamental *abstraction operators* has been defined. These operators are defined at the perceived world level. The proposed set of fundamental operators $\Omega = \{\omega_{hide}, \omega_{ind}, \omega_{set}, \omega_{ta}, \omega_{eq\text{-}val}, \omega_{red\text{-}arg}, \omega_{prop}\}$ is not exhaustive [4]. In particular contexts (such as cartography), it can be either reduced or augmented with domain-specific abstraction types.

The operator ω_{hide} is a fundamental abstraction operator that consists in hiding any kind of knowledge, be it an object, an attribute of an object or a relation between objects (e.g. hiding an isolated street). The operator ω_{ind} consists in making several objects indistinguishable (e.g. considering only one of several close isolated trees as a typical representative of them). The operator ω_{set} consists in grouping several objects that are considered not to be distinguished (e.g. grouping a set of trees into an object "forest"). The operator ω_{ta} consists in grouping a set of *different* objects to form a new *compound object* (e.g. grouping several streets and buildings to form a town). The operator $\omega_{eq\text{-}val}$ specifies what subset of the attribute or function values are considered indistinguishable and can be merged (e.g. considering at the same altitude two objects with close altitudes). Finally, the operator ω_{red_arg} specifies a relation and a subset of its arguments, which must be dropped from the relation, obtaining thus a relation with reduced arity (e.g. hiding the argument "type of crossing" in the relation between two roads). From these operators, defined at the perception level, the operators between the structures, languages and theories are shaped.

3. CARTOGRAPHY IN THE KRA MODEL

The KRA model exhibits several 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.

The topographic map production process closely parallels the KRA model, because it can be analyzed 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 or satellite images. These pictures are the perceived world P(W). Objects contained in these photographs are located and labeled to create a geographic database (GDB). This GDB is the set of geographic data organized 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 contains all the background facts and laws allowing one to reason about geographic configuration, and may be different for different tasks.

Cartography is not just knowledge representation. All The steps of map creation do not only involve knowledge representation, but also knowledge abstraction. In particular, map creation (which contains the generalization process) is both a knowledge representation process, when objects are symbolized, 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.

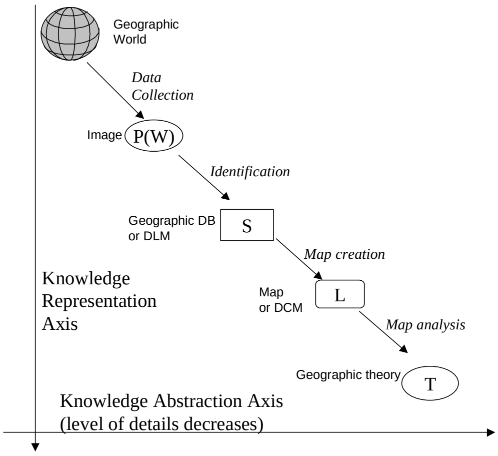

Figure 3

3.1 Generalization Process in the KRA model

Knowledge abstraction in cartographic generalization is the identification of abstracted geographic objects relevant to the theory construction that will be done from the map. "Objects" are taken in a wide sense: they may represent any basic geographic objects or any set of basic objects having a geographical meaning (e.g. a building, the set of streets and buildings of a town…).

In our model, the abstraction process is to go 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. For example, an abstraction is to go from a complete description of a set of streets in a town to the description «this is a street network».

As we explained in the KRA model presentation, abstractions at the structure level (i.e. on objects of the geographic database) are only considered as consequences of abstractions at the geographic world perception level.

Knowledge representation in cartographic generalization is the process of symbolizing abstracted objects. For example, this representation process is to determine which symbolized subset of streets is the best suited in order to well represent «a street network». This choice is guided by the necessity to well represent the abstracted object and restricted by the drawing possibilities (we cannot represent all the symbolized streets because they will overlap on the drawing).

Difference but not independence. 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, a street network is represented by a subset of the "ground" streets. The abstracted object «street network» helped us to change our view of the world, but the representation process needs to look again at the ground GDB to represent actual objects of the world. In this way we imitate the human perception, which continuously changes the level of abstraction to analyze space. This view follows Nyerges' view of cartographic generalization [3] which splits generalization in two phases: *«Geographical information abstraction mainly concerns managing geographical meaning in databases, and map generalization mainly concerns structuring map presentations»*. These inter-links between abstraction and representation explain why, manually, these two steps have always been performed in one time by the cartographer. However, this distinction between abstraction and representation is necessary for the creation of an automated process of cartographic generalization.

4. STRUCTURING KNOWLEDGE FOR ITS ACQUISITION

The emphasized distinction between abstraction and representation is necessary for efficient knowledge acquisition in cartographic generalization. Because of space limitation we only quote the different types of knowledge involved in these two processes.

On the one hand, the knowledge abstraction process is seen as the identification of geographic objects that are relevant to the user needs (and need to be drawn). It manipulates: (a) geographic knowledge (e.g. town model), which could be acquired through geographers interview; (b) perception knowledge (Gestalt), which has been studied in psychology and cartography; (c) space analysis knowledge, which should be acquired in computers through space analysis tools (e.g. Delaunay triangulation).

One the other hand, representation knowledge manipulates: (a) graphic knowledge to define when a map is legible, which is well known by cartographers; (b) drawing knowledge to define how to represent an abstracted object, and which could be acquired through cartographers interviews or manually drawn maps analysis; (c) algorithm knowledge, which could be acquired through cartographers drawing analysis by machine learning techniques.

5. CONCLUSION

One of the key problems limiting the automation of cartographic knowledge acquisition lies in the heterogeneity of the knowledge involved throughout the process. In this article, we have adapted an Artificial Intelligence model to distinguish two fundamental transformations used in cartography, namely *abstraction* and *representation*. The first contribution of this work is therefore to propose a analysis of cartographic and geographic knowledge along these dimensions in order to facilitate its acquisition.

6. REFERENCES

[1] Clancey W. (1983). "The Epistemology of a Rule-Based Expert System - A Framework for Explanation". *Artificial Intelligence, 20.*

[2] Ginsberg, (1997). "Essentials of Artificial Intelligence", Morgan Kaufmann, San Francisco (CA)

[3] Nyerges (1991). "Representing Geographic Meaning in Map Generalization" Longman Scientific & Technical pp 59-85

[4] Saitta L., Zucker J.-D. (1998). "Semantic Abstraction for Concept Representation and Learning". In *Proc. Symposium on Abstraction, Reformulation and Approximation* (Pacific Grove, CA).