# Issues on Modeling Spatial Granularities [*][†]

Elena Camossi[1] Michela Bertolotto[2] Elisa Bertino[1] Giovanna Guerrini[3]

[1]Dipartimento di Informatica e Comunicazione
Università degli Studi di Milano
via Comelico 39/41 - 20135 Milano, Italy
{camossi, bertino}@dico.unimi.it
Tel. +39 010 353 6717, +39 02 503 16227

[2]Department of Computer Science
University College Dublin
Belfield, Dublin 4, Ireland
michela.bertolotto@ucd.ie
Tel. +353 (0)1 716 2913

[3]Dipartimento di Informatica
Università di Pisa
via F. Buonarroti 2 - 56127 Pisa, Italy
guerrini@di.unipi.it
Tel. +39 010 353 6635

**Abstract**

The formalization of the spatial granularity concept is a topic that has recently received a growing interest, but none of the proposals presented so far has been acknowledged by the spatial community as a reference definition. This is because developing such a formalization entails addressing a large number of issues, due to the intrinsic characteristics of spatial information. In this paper we discuss some of these issues, and we propose a formal definition of spatial granularity that addresses most of them. Our definition supports the representation of spatial entities at multiple granularities by applying model oriented map generalization principles. In particular, we consider a set of generalization operators that guarantee topological consistency and their inverse functions, for performing, respectively, the conversions to coarser and to finer granularities.

# 1 Introduction

Among the topics related to spatial information representation, one of the most relevant ones is the formalization of *spatial granularity*. This notion, desirable for representing spatial data at the preferred level of detail, is crucial when integrating spatial information from heterogeneous sources, since it is often the case that information at different sources are stored at different levels of detail. Such a scenario is becoming quite common, since a large number of organizations and companies geographically distributed over large areas need to use information systems that take into account spatial components of data. For the same reasons, INSPIRE project specifications [INSPIRE] suggest the development of integration techniques for spatial data represented at different levels of detail. Unfortunately, despite the huge amount of work carried out during the last few years in this direction, a reference formalization of spatial granularity does not exist yet. Recent work on the formalization of spatial granularity has focused on issues related to the concepts of vagueness, imperfection and imprecision [DMSW01] of spatial information, mainly for use in qualitative reasoning [Ben02]. In particular, Bittner and Stell discuss qualitative aspects of spatiotemporal information and propose a granularity notion for this kind of approach [Bit02, BS01, Ste03]. Katri et al. [KRSO02] have presented an annotation-model (extension of the Unifying Semantic Model formalism) allowing for the specification of spatiotemporal data at multiple granularities, where the spatial granularity system relies on the concept of spatial *imprecision* [DMSW01].

In this paper we discuss some of the issues related to the definition of a spatial granularity system for application to spatial databases, and we present a formal model addressing such issues. In our work we refer to the guidelines given by Stell and Worboys [SW98], who have developed a theoretical framework for the specification of a spatial granularity lattice that can be integrated in a spatial database model. With respect to the specification in [SW98], the main steps involved by the design of a spatial granularity framework can be identified as:

---

[*] Contact Author: Elena Camossi. Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano, via Comelico 39/41 - 20135 Milano (Italy). Email: `camossi@dico.unimi.it`. Tel: +39 010 353 6717.

[†] **Acknowledgments:** The work of Elena Camossi was partially supported by the University of Genova, Italy. The support of the International Collaboration Initiative of Enterprise Ireland is also gratefully acknowledged.

1. *the identification of a meaningful and connected set of granularities for representing data*;
2. *the specification of the functions for converting different granularities in the set.*

Referring to the second step, we can equivalently define a set of conversion functions for converting spatial values instead of granularities.

The issues related to those two steps, that we analyze in this paper, are mainly due to the complex semantics of spatial data. In particular, as we will see, a large number of different sets of granularities can be devised for different spatial domains, but is difficult to define a set that is suited for every database system. Moreover, different relationships can be devised among granularities in the same set, resulting in different connecting structures. It's crucial to obtain a structure of relationships allowing for comparing any pair of data expressed at different granularities. Finally, a lot of conversion functions for spatial data has been defined, respect to the data representation considered and to semantics of data itself, with conflicting characteristics and properties. The set of functions supported must result in consistent data conversions, allowing the largest set of conversions as possible respect to a given set of requirement. We analyze these and other issues for three different conceptual categories of granularities, defined by considering different aspect of spatial data: scale changes, semantics of data, and specifically data conversion functions. Finally, we discuss relationships between granularities and map overlay operation.

We address the issues we discuss also by proposing a spatial granularity definition that is suitable to be applied to vector data and does not deal with indeterminacy related to spatial data representation, although we take into account domain semantic dependency in the general design of spatial granularities. Our notion of spatial granularity supports hierarchical partitions of the space, by applying granularities designed on the specific application domain considered. Therefore, even though this notion is application-domain dependent, it has a very high expressive power. Examples of spatial granularity are those representing different levels of detail for an administrative boundaries map, such as *municipalities*, *provinces*, *states* and *countries*, or those related to standard units of length, such as *ms*, *Dms*, and *kms* (respectively meters, decameters, i.e., ten meters, and kilometers, that in a two-dimensional space represent squares with 1 meter, 10 meters, and 1 kilometer long sides). Moreover, our approach allows for relating a database schema to multiple sets of spatial granularities, by simply introducing equivalence relationships between granularities in different sets.

For converting spatial data expressed at different granularities, we refer to map generalization operators [MS92, Mul91, WD99], specifically those used by *model-oriented generalization* [MLR95]. In particular, the operators we consider for converting data to a coarser granularity are those defined in [Ber98], that guarantee topological consistency through composition, while through the application of the operators that perform the inverse functions we obtain conversion to finer granularities.

The paper is organized as follows. In Section 2 we review work related to our proposal of spatial granularity. In Section 3 we discuss issues in the spatial granularity design. In Section 4 we present a proposal of formalization for spatial granularity notion addressing the issues discussed. We formalize the set of spatial granularities we consider and we report on how the model handles multi-representation of spatiotemporal data. Finally, Section 5 concludes the paper, outlining possible extensions to our proposal.

## 2 Related Work

The spatial granularity framework proposed in this paper relies on spatial *generalization* [Mul91] process principles. The term generalization was first used in cartographic context [MS92], and usually refers to the choice concerning the representation of features that must be still represented in a less detailed map, as a consequence of scale reduction, for purposes of graphical representation and visualization. The choice of which information to highlight in a map, and what representation to use for them, is strongly dependent on applications. Several generalization operators have been devised for describing and formalizing this human driven process [WD99, DP96], in order to automatize it. *Model-oriented generalization* uses cartographic techniques for representing spatial data at different levels of abstraction, by taking into account also semantics aspects of data and some notion of consistency, as, for example, the preservation of topological relationships at the generalized level [PD95]. The generalization operators we integrate in our proposal are those defined by Bertolotto [Ber98]. Bertolotto [Ber98] has defined a set of minimal generalization operators and has proved that they guarantee topological consistency. In [BE99, BE01], where a model for progressive transmission of vector data over the world wide web is presented, these operators are used for creating multiple representations of vector data. A set of spatial data, basically a map, is generalized at different levels of abstraction, and each logical level is computed by generalizing a more detailed one by using the generalization operators defined in [Ber98]. Each logical level is represented by a set of vector data, topological relationships and related vertical links. Vertical links bind a geometric entity represented in a map on which the generalization process has been applied to its finer representation in the immediately finer level, and each vertical link is an ordered sequence of functions (inverse of minimal operators defined in [Ber98]) to apply to an entity in a level in order to (re)obtain the precomputed representation of the entity at finer level. In this paper we formalize also these inverse operators, to be consistent with the guidelines given by Stell and Worboys in [SW98] for integrating the support of spatial granularities in a spatial data model. Moreover, with respect to [SW98], we propose a more concrete notion of spatial granularity and we specify the operators to use for converting a spatial data at different spatial granularity.

Other related approaches propose structures for relating different spatial granularities, most of which term these structures as "hierarchies". In particular, Timpf [Tim99] uses an approach opposite as ours, for deducing levels of detail organized in hierarchies. Such approach is based on the analysis of a "map series", that is, a set of spatial data representing the same geographical area at different resolutions, with respect to a given collection of human driven abstraction processes: aggregation, filtering and generalization. Vangenot [Van01] proposes a timestamped model for multiple representation of data, where each representation is related to a view point and a resolution (spatial and semantic resolution), but relationships supported between different representations are semantics driven. In [CJ00] nested hierarchies are considered for spatial and temporal dimensions in order to make quantitative evaluations about the relationships that can be obtained. With respect to those approaches, we propose a framework for spatial multirepresentation by considering a significative set of multi-resolution operators. Scaling operations, that is, granularity conversions, can be performed on every set of spatial data in vector format. The spatial conversions we support are performed explicitly, and allow for obtaining generic sets of representation that are hierarchically organized with respect to the user needs.

# 3 Design Issues for Spatial Granularities

This Section discusses some of the issues entailed by the formalization of the spatial granularity notion. We first analyze issues related to the specification of the relationships that must hold among spatial granularities referred by a database schema. Indeed, to achieve comparison among data expressed at different levels of detail, all database objects must refer to a common set of granularities. The granularities in each set must be related by some relationship among them, with respect to which the conversion of data at different granularities is specified. This conversion capability is crucial for comparing data expressed at different granularities. Moreover, we discuss the specific issues entailed by different conceptual kinds of granularity sets. We devise three different categories of spatial granularities, based on the geometry of spatial information; on the spatial data semantics; on a given set of spatial data conversion operations. Finally, we analyze the map overlay operation by discussing its relationships with granularity concept.

## 3.1 Specification of Relationships Among Granularities in the Same Set

The theoretical framework proposed by Stell and Worboys [SW98] for the specification of spatial granularities, requires that granularities used for achieving spatial multi-representation form a connected set, represented, in the most general case, by a lattice structure: granularities are the nodes of the lattice, whereas edges represent relationships among granularities, generically referred as *finer-than* and *coarser-than* relationships. Generic scaling functions are introduced for converting a certain granularity to a coarser or a finer one. Equivalently, we can define scaling functions for converting spatial data to a different granularity. Scaling functions are directly related to existing relationships between different granularities. In particular, by considering scaling functions defined for converting data we obtain a lattice structure among spatial data represented at different granularities that is equivalent to the one representing the relationships between granularities.

Actually, in several real cases, we can consider also different structures. For example, granularities can be related by a total order, that results in a hierarchical structure (e.g., among the set of granularities $mms$, $ms$, $kms$), or by a partial order that allows for a tree-structure representation (e.g., as it happens by considering the set of granularities $\mu ms$, $mms$, $ms$, $inches$, $feet$).

Note that, in the second case, some conversions are not possible. For example, we can not convert directly a data expressed at granularity $ms$ to granularity $feet$, since such conversion would produce data referring non-entire granules of granularity $feet$[1]. Fortunately, we can compare such data by converting both of them to granularity $\mu ms$. Then, both structures allow, as the lattice one, for queries involving data expressed at different granularities.

Once a meaningful set of spatial granularities for a database schema have been defined (a problem discussed in detail in Section 3.2), we must specify the semantics of scaling functions for performing conversion of data at different granularities. Referring to real conversions performed on vector spatial data, these scaling functions can be easily represented by generalization and specialization operations, respectively for obtaining a coarser and a finer views of data. Generalization, in particular, is widely addressed by scientific literature on spatial data, and a large number of operators have been defined in order to formalize this operation. Unfortunately, their use as scaling functions to coarser granularities data is not straightforward, despite the appearance. First of all, several conceptual abstractions on spatial data can be devised, and some of them may conflict. Indeed, each abstraction reflects a different way of performing generalization and requires the application of a different set of operators. For example, *zooming out* a map for increasing its scale is, conceptually, a different operation with respect to reducing the map detail for highlighting some specific map features, although the second generalization is usually performed as a consequence of a scale increment, as the first one. Moreover, some abstractions exist that, even if referring to specific generalization operators, cannot be carried out in a fully automatic way. This is the case of *line simplification*, an operation that, when automatically applied, usually introduces data inconsistencies that have to be corrected by hand.

The last issue we consider is the specification of *specialization* of spatial data, i.e., how to convert a spatial data to a finer granularity. Specializing a spatial data implies introducing details in the data representation. This can be obtained only if a more detailed level of representation exists, from which to retrieve these details. Unfortunately, this detailed representation could not always be obtained. This problem is widely discussed by literature dealing with *spatial imprecision* [DMSW01]. A similar issue arises when considering multiple-granularity temporal data, and is called *temporal indeterminacy* [DELS98]. However, we can note that, from a multigranular database perspective, such a problem is limited. Indeed, since the granularity specified for an entity in the definition of a database schema is that used for acquiring data, it represents an integrity constraint on data, specifying the finest level of detail achievable for acquiring and then for representing them. Then, a query involving data at a granularity finer than that specified for acquiring them can be correctly managed as a violation of data integrity constraints.

Moreover, when a more detailed level can be obtained, links between different representation levels must be established, even if the granularity transformation considered is local, that is, it has been applied to a single feature. Those links can be easily represented in terms of the applied conversion functions, by specifying origin and destination of each function.

## 3.2 Which Kind of Granularity?

In order to avoid referring only to simplistic spatial granularities as those that can be obtained according to standard length or area measures and providing regular subdivisions of the space not very significative for most application domains, we can take into account granularities that can be devised by considering,

---

[1] 1 meter corresponds to 3.28084 feet.

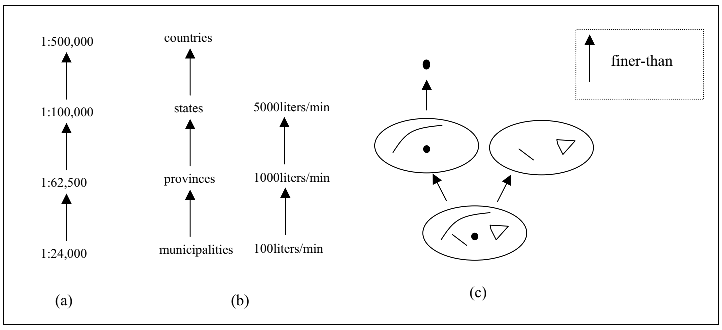

Figure 1: Granularity hierarchies examples

for example, the operations ones wants to perform on spatial data, or by highlighting particular aspects of data itself. In the following we discuss three different categories of spatial granularities: granularities related to scale changes of maps, granularities that reflect semantic aspects of spatial information, and granularities defined with respect to conversion functions applied on data.

### 3.2.1 Scale Changes

The most simple set of spatial granularities one can devise is represented by different *scales* used for data representation. The application of these granularities requires the use of a geometric data representation involving numeric characteristics. Scales commonly used for maps by government agencies, such as USEPA and USGS, belong to such kind of granularity. An example is shown in Figure 1(a), where $1 : 24,000$, $1 : 62,500$, $1 : 100,000$, $1 : 500,000$ are reported. Despite their intuitive application, granularities based on scale changes complicate the generalization operation, that is, the increase of the scale. The reason is that the choice of which details must be discarded in the more general map is difficult to automatize, since it is based only on geometric characteristics of data and does not take into account their semantics. For example, when increasing the scale of a Europe map, we would like to maintain the representation of the most important cities, as country capitals, although their spatial extension (or population, or each other numeric characteristic taken into account for performing the scale increase) would imply deleting some of them from the map. Those types of correction are usually made by hand by cartographers.

### 3.2.2 Semantic Granularities

Given a set of spatial data, a connected granularity set can be devised for these data by taking into account the semantics of the information they represent. Consider, for example, a map reporting administrative boundaries of Europe. Such a map can be easily represented at different decreasing granularities by considering all municipalities boundaries, or only the ones corresponding to provinces, or regions, or states or countries, since each boundary type provides a different abstraction on data. The corresponding set of granularities is shown in Figure 1(b). Such kind of classifications usually produces a proper inclusion among granularities, thus resulting in a strong relationship between granules of different granularities.

Note that the terms used for naming each granularity are usually dependent on the particular area to which data refers; as an example, French departments are more or less equivalent to Italian provinces. This characteristic results in *heterogeneous schemas* [HM01] for the spatial dimension, where granularity sets are connected as complex structures. For simplifying the approach, we can require, as an alternative, to specify the interested area with respect to which give the correct interpretation of the used terms.

Moreover, different granularity sets can be devised with respect to different spatial data involved in the same application domain, that can have different semantics. For instance, a database can manage several kinds of maps, such as different thematic maps, each representing the same geographic area. Since they have different semantics, different sets of granularities can be considered for representing those maps. For example, the application of administrative boundary granularities to an hydrographic map is not intuitive. This is more evident when considering, in the same database, different kinds of spatial objects, like maps

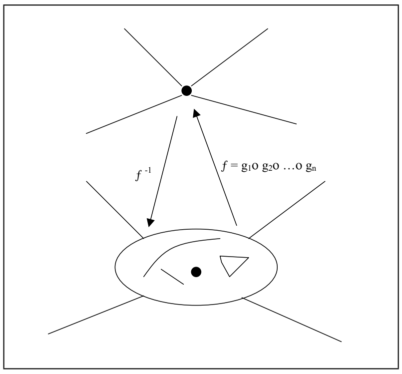

Figure 2: Topological generalization

and moving objects [GBE$^+$00], that have very different purposes. Furthermore, the choice of granularity names and semantics depends not only from the application domain but also on specific personal criteria of the users. For example, rivers in an hydrographic map can be classified with respect to their flow, using the granularities $100 liters/min$, $1000 liters/min$, $5000 liters/min$, as shown in Figure 1(b), or equivalently with respect to their length, or simply as *small*, *medium* and *huge*. As we pointed out, for developing a query language, that is, for relating/comparing spatial data representing different real entities, a reference common set of granularities must be defined.

Another related issue is related to aggregation operations on non-spatial data related to the geographic area represented, like population, and meteorological information. Referring to the administrative boundaries set of Figure 1(b), in order to obtain a consistent representation we can, for example, redefine the concept of population of a country to be the sum of the amounts of population of the cities represented in a map.

#### 3.2.3 Granularities Based on a Given Set of Scaling Operators

Given a vector map and a set of generalization operators, we can define different granularities for the map representation by considering the set of maps that can be obtained by generalization of the given map. Different map representations are related according to a tree organization, as shown in Figure 1(c). An advantage of this approach is that it involves generalization only at spatial level, whereas non-spatial information related to the map, like population and area, remain consistent.

Several generalization operators have been proposed in the past [Mul91, WD99], each one having different characteristics and satisfying some properties, that we must take into account when choosing which ones for implementing granularity scaling functions. In particular, for maintaining data integrity the preservation of topological consistency is crucial. In Figure 2, function $f$ represents a composition of operators that generalize a set of map features converting them to a point representation. Because of the preservation of topological consistency, the input map features, depicted as straight lines, still exist after the generalization.

The support of specialization can be obtained by considering inverse operations, requiring the existence of a detailed level from which to retrieve details to be added to a representation. This can result in a duplicate representation of topological relationships, if they are maintained explicitly.

### 3.3 Considerations on Map Overlay Operations

The overlay operation as usually applied in GIS systems among thematic maps, for example land use and rainfall maps, is apparently an approach alternative to the granularity definition discussed above. Indeed, given a set of thematic maps of the same geographic area, we can achieve representations at different levels of detail by overlaying the maps. We can give a tree representation of the area at different details, where the tree root is represented by the map of the area itself obtained by considering all the information

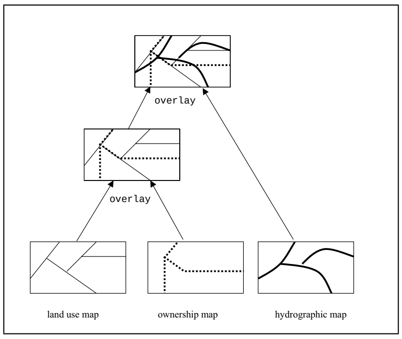

Figure 3: Map Overlay Example

represented by the maps, and each tree leafs is a thematic map. Inner nodes of the tree are obtained by overlaying a set of maps. This representation is shown in Figure 3. The leafs of the tree represent rainfall, land use and hydrographic maps of the same area. Although we obtain a tree-hierarchical structure, the overlay operation actually does not produce the desired result, since each level we obtain conveys different information.

However, the use of overlay operation among maps is orthogonal to the other granularity solutions proposed, and can integrate a multigranular model for retrieving interesting information from data by applying the algebra proposed in [DHT94]. For example, we can define spatial units based on the application domains, and then join a set of regions, for specifying some interesting aggregation information, as to highlight which are the regions where a given cultivation is tilled.

# 4 A Design Example for Spatial Granularities

In this Section we present a framework for handling spatial granularities that we have integrated in a spatiotemporal multigranular object data model [CBBG03]. Such framework addresses the issues discussed in the previous Section, by formalizing the notion of spatial granularity and the conversion of spatial values between different granularities.

## 4.1 Definition of Spatial Granularity

For representing the Earth surface we consider a two-dimensional Euclidean space. Spatial entities, that is, entities that are related to the reference space by having a spatial extension or position, are represented by means of usual vector features: *points*, *lines* and *regions*, respectively of *dimension* 0, 1, 2. A subset of vector features composing an entity can be referred as a *sub-entity.* Entities can be represented at different granularities by considering hierarchical representations as the ones that can be obtained by considering subdivisions of the reference space into regular grids, or the ones that can be obtained from semantics characteristics of data, like administrative boundaries, roads categories, lands use classifications, as discussed in Section 3.2.

The set of spatial granularities referred by a database schema is denoted with $\mathcal{G}_{\mathcal{S}}$. Granularities in $\mathcal{G}_{\mathcal{S}}$ are related by the *finer-than* and the *equivalent-to* relationships.

A granularity $G$ is said to be *finer-than* a granularity $H$, denoted by $G \preceq H$, if the representation of a spatial entity $E$ at granularity $G$ contains more features than those at granularity $H$, or, for each sub-entity $e$ present in both representation, $e$ is represented at granularity $G$ by using the same set of features or a set of features of greater dimension. We also say that $H$ is *coarser-than* $G$. We can say, for instance, that *municipalities* is finer-than *states*, and that *Dms* (i.e., Decameters, ten meters) is coarser-than *cms* (i.e., centimeters). In what follows, the symbol "$\prec$" denotes the anti-reflexive finer-than relationship.

**Example 1** *Consider a spatial database managing data related to a particular city crossed by a river. Assume that such data include the geographical extension of those entities. By using the set of spatial*

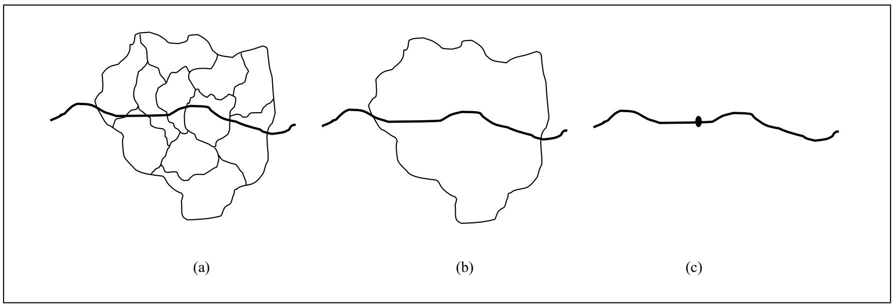

Figure 4: Representation of the extension of a city at different granularities *hms* (a), *kms* (b), *hkms* (c)

*granularities hms, kms, and hkms (i.e., hectometers, kilometers, hundred of kilometers, that in a two-dimensional space represent squares with 100 meters, 1 kilometer, 100 kilometers long sides*[2]*), we can obtain multiple representation of the extension of the city, as shown in Figure 4.*

The *equivalent-to* relationship allows for the specification of equivalences among the interpretations of granularities referring to different semantic information. The semantics of the granularities used for representing data stored in the database is related to the semantics of the spatial objects represented in the database, but can also be given by explicitly specifying which features must be represented at each granularity.

**Example 2** *Consider again the database introduced earlier. Suppose to represent, in the same database, an hydrographic map and a boundary map both representing the same geographic area. The set of boundaries granularities municipalities, provinces, states and countries of Section 3.2 is designed to be applied to the boundaries map. Moreover, we can give a (meaningful) semantics that allows for their application to the hydrographic map. Operationally, this semantics specifies which features must be represented in the hydrographic map when it is specified at different levels of details.*

*For instance, when representing the hydrographic map at countries granularity, only the big flow rivers can be represented, or only the those that are longer than a certain amount of kilometers. Note that such specifications are application-domain dependent and also subject to personal interpretation by users. In our model [CBBG03], these specifications are represented as conversions of spatial and spatiotemporal values at different granularities.* ◇

In order to simplify the approach, we allow for considering different semantic sets of granularities that can be specified for characterizing specific data, but we require the database designer to introduce relationships specifying how to relate granularities in such sets when comparing data. This is achieved by using the following syntax:

$$< G_1 > \texttt{ is equivalent to } < G_2 >$$

where $G_1$ and $G_2$ are two granularity names. After this specification, $G_1$ is said to be *equivalent-to* $G_2$ (and viceversa). This results in a unique connected set of granularities, $\mathcal{G}_\mathcal{S}$, that allows for the comparison of spatial data stored at different granularities in a non-ambiguous way. The specification of such equivalences is performed by the database designer when designing the database schema as it happens for integrity constraints, since they are based mainly on specific user choices, that, in turn, can rely on geometric considerations.

**Example 3** *Suppose, for example, that a database schema refers three sets of granularities: one representing different levels of detail for an administrative boundaries map, municipalities, provinces, states and countries; one for an hydrographic map, that classify rivers respect to their flow* $1000 liters/min$,

---

[2] A possible straightforward interpretation for these granularities can be, for instance, that for a spatial entity represented at *kms* granularity, all its subcomponents that are less than 1 kilometer long are not represented.

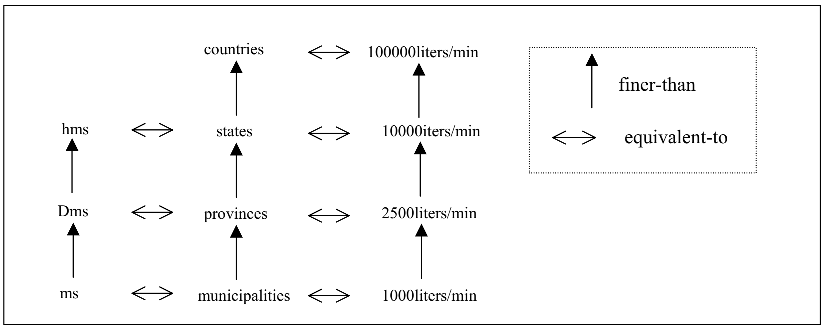

Figure 5: Extended set of granularities

$2500 liters/min$, $10000 liters/min$, $100000 liters/min$; *one used for representing different levels of detail of a truck crossing the geographic area represented by the maps, as the set ms, Dms, hms.*

*When representing the boundaries map at states granularity, for instance, we can establish that the truck must be referred at at hms granularity, by specifying the clause: hms* `is equivalent to` *states, and that the rivers reported by the the hydrographic map must be referred at granularity* $10000 liters/min$ *(that only those with flows greater than 10000liters/min have to be considered), by specifying the clause:* $10000 liters/min$ `is equivalent to` *states. The complete connected set of granularities is shown in Figure 5*

## 4.2 Conversion of Spatial Value at Different Granularities

We handle the conversion of data at different spatial granularities by referring to model generalization theory [MLR95]. In particular, the conversion of a spatial value from a given granularity to a coarser one is obtained by applying a composition of the operators defined in [Ber98], whereas by the application of the composition of their inverse functions we perform specialization (in a similar way as proposed in [BE99, BE01]).

The operators defined in [Ber98], that are shown in Figure 6, allow for contracting of lines and regions to points, thinning of regions to lines, merge of regions and lines sharing edges and endpoints, respectively, and for abstracting isolated points and lines in regions. It has been proved that such operators preserve topological consistency through composition, an essential property for usability of spatial data [Ber98]. Moreover, it has been proved that this set is minimal and sufficient for providing, through composition, every generalization that preserve topological consistency, that is, they can be composed to obtain macro-operators with the same characteristics of basic operators [Ber98]. Actually, such operators are only a subset of those that can be used to generalize spatial data, since they do not support some of the traditional generalization operations, namely those that do not guarantee preservation of topological consistency, like aggregation, and line simplification.

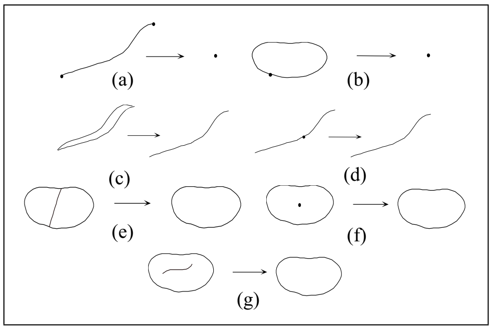

Figure 6: Generalization operators

In our framework, the set of these generalization operators is denoted by $\mathcal{O}p_g$ and its elements are:

(a) `l_contr`, that contracts an open line, endpoints included, to a point;

(b) `r_contr`, that contracts a simple connected region and its boundary to a point;

(c) `r_thinning`, that reduces a region and its bounding lines to a line;

(d) `l_merge`, that merges two lines sharing an endpoint into a single line;

(e) `r_merge`, that merges two regions sharing a boundary line into a single region;

(f) `p_abs`, that eliminates (abstracts) an isolated point inside a region;

(g) `l_abs`, that eliminates (abstracts) a line inside a region.

Their inverse functions, expanding points to lines and regions, lines to region, splitting regions and lines and adding isolated features to regions, supports specialization through function composition. The set of specialization operators is denoted by $\mathcal{O}p_s$ and its elements are:

(a) `p_exp_to_line`, that expands a point into an open line;

(b) `p_exp_to_region`, that expands a point into a simple connected region;

(c) `l_exp`, that expands a line into a region;

(d) `l_split`, that splits a line into two lines sharing an endpoint;

(e) `r_split`, that splits a region into two regions sharing a boundary line;

(f) `add_p`, that adds an isolated point inside a region;

(g) `add_l`, that adds a line inside a region.

*Expansion* operators are the inverse of contraction ones, (e.g., `p_exp_to_line` performs the inverse operation of `l_contr`), *split* operators are the inverses of merge ones (e.g., `l_split` is the inverse of `l_merge`), and *add* operators are the inverses of abstraction ones (e.g., `add_p` performs the inverse operation of `p_abs`).

The specification of the conversion of a spatial entity at granularity $G \in \mathcal{G}_\mathcal{S}$ to a different granularity $G' \in \mathcal{G}_\mathcal{S}$ is obtained by by applying a composition $f$ of the operators just described:

$$f = \{f_1 \circ f_2 \circ \ldots \circ f_n\} \text{ where, } \forall i = 1, \ldots, n \begin{cases} \text{if } f_i \in \mathcal{O}p_g & \text{then } G \preceq G' \\ \text{if } f_i \in \mathcal{O}p_s & \text{then } G \succeq G' \end{cases}$$

The choice of which operators have to be combined for converting a spatial entity to a different granularity is domain dependent. Moreover, the same conversion can be obtained by composing operators according to a different order. This observation motivates the requirement that the specification of spatial granularity conversions be expressed as a constraint in the database schema. Our framework supports conversion specification of the form:

`convert` <`entity`> `from` $< G >$ `to` $< G' >$ `by applying` < *op_composition* >
[ `where` < *conditions_seq* > ]

where *entity* is a spatial entity[3], $G$ and $G'$ are spatial granularities s.t. $G \preceq G'$ or $G \succeq G'$, and *op_composition* is an operator composition as described above.

The optional clause `where` allows for the specification of a list of additional conditions that make the conversion process automatic. These additional conditions can be based on geometric and dimensional characteristics of features, like length, width, area, or on the semantics of represented data. For example,

---

[3] In [CBBG03] we model spatial entities as object spatial attributes.

when converting a map of Europe form granularity *states* to granularity *countries*, we can specify that all the cities, that are represented as regions at granularity *states*, must be represented as points at granularity *countries* if they are less large than 20kms, by applying `r_contr` operator, or that all cities that are not states capitals have not to appear any longer after the conversion, that is performed by applying the operator compositions `p_abs` $\circ$ `r_contr`.

If no conversion condition is specified, the composition of operators is applied to all the features of the spatial entity that are suitable for the application of the operators, as illustrated by the following Example.

**Example 4** *The conversion specification of the extension of the city as shown in Figure 4 is as follows, where* `extension` *represents the extension of the city.*

`convert extension from` *hms* `to` *kms* `by applying r_merge;`
`convert extension from` *kms* `to` *hkms* `by applying r_contr;`

*The first specification implies that, for converting the extension of the city from granularity hms to granularity kms, the* `r_merge` *operator must be applied. This results in the city extension representation as a single region as shown in Figure 4(b). The second specification implies that, for converting* `extension` *attribute from granularity kms to granularity hkms (i.e., 100 kilometers), the application of* `r_contr` *operator to the region representing the city is required, and the result of the conversion is shown in Figure 4(b). Note that, as shown in Figure 4, the conversion process as specified above does not involve other spatial objects that are in some relationship with the city we generalize, such a river crossing the city.* ◇

## 5 Conclusions and Future Work

In this paper we have discussed some of the issues involved into the definition of the concept of spatial granularity, and we have proposed a framework for handling spatial granularities that addresses these issues based on model generalization principles. We are currently working for extending this framework in several directions. First We are extending the set of generalization operators we consider in this framework with other operators that are commonly used in model-oriented generalization [WD99, DP96] and that cannot be represented through the operators defined in [Ber98]. In particular, we are interested in considering also *aggregation*, elsewhere called amalgamation [SW99] or granulation [Ste99], and *line simplification* [IM88].

## References

[BE99] M. Bertolotto and M. Egenhofer. Progressive Vector Transmission. In *Proceedings of 7th ACM Symposium on Advances in Geographic Information Systems*, pages 152–157. C. Bauzer Medeiros, 1999.

[BE01] M. Bertolotto and M. Egenhofer. Progressive Transmission of Vector Map Data over the World Wide Web. *Geoinformatica*, 5(4):345–373, 2001.

[Ben02] B. Bennett. Physical Objects, Identity and Vagueness. In D. Fensel, Deborah McGuinness, and Mary-Anne Williams, editors, *Proc. of the Eighth International Conference on Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann, 2002.

[Ber98] M. Bertolotto. *Geometric Modeling of Spatial Entieties at Multiple Levels of Resolution*. PhD thesis, Università degli Studi di Genova, 1998.

[Bit02] T. Bittner. Reasoning about qualitative spatio-temporal relations at multiple levels of granularity. In *Proceedings of the 15th European Conference on Artificial Intelligence*. F. van Harmelen IOS Press, 2002.

[BS01] T. Bittner and B. Smith. A Unified Theory of Granularity, Vagueness and approximation. In *Proceedings of COSIT Workshop on Spatial Vagueness, Uncertainty, and Granularity*, 2001.

[CBBG03] E. Camossi, M. Bertolotto, E. Bertino, and G. Guerrini. *ST*_ODMG: A Multigranular Spatiotemporal Extension of ODMG Model. Technical Report DISI-TR-03-09, DISI, Università di Genova, 2003.

[CJ00] C. Claramunt and B. Jiang. Hierarchical Reasoning in Time e Space. In *Proc. of 9th International Symposium on Spatial Data Handling*, pages 41–51, 2000.

[DELS98] C.E. Dyreson, W.S. Evans, H. Lin, and R.T. Snodgrass. Supporting Valid-time Indeterminacy. *IEEE Transactions on Knowledge and Data Engineering*, 23(1):1–57, 1998.

[DHT94] V. Delis, Th. Hadzilacos, and N. Tryfona. An Introduction to Layered Algebra. In *Proc. of 6th International Symposium on Spatial Data Handling*, 1994.

[DMSW01] M. Dunkham, K. Mason, J.G. Stell, and M.F. Worboys. A formal approach to imperfection in geographic information. *Computer, Environment and Urban Systems*, 25:89–103, 2001.

[DP96] G. Dettori and E. Puppo. How Generalization Interacts with the Topological and Geometric Structure of maps. In *Proceedings of 6th ACM Symposium on Advances in Geographic Information Systems*, 1996.

[GBE$^+$00] R.H. Güting, M.H. Bhölen, M. Erwig, C.S. Jensen, N.A. Lorentzos, M. Shneider, and M. Vazirgiannis. A Foundation for Representing and Querying Moving Objects. *ACM Transaction On Database Systems*, 25:1–42, 2000.

[HM01] C.A. Hurtado and A.O. Mendelzon. Reasoning about summarizability in Heterogeneous Multidimensional Schemas. In *Proc. ICDT*, 2001.

[IM88] H. Imai and MIri. Polygonal Approximation of a curve - formulations and algorithms. In G.T. Toussaint, editor, *Computational Morphology, Machine Intelligence and Pattern Recognition*, volume 6. North-Holland, 1988.

[INSPIRE] The INfrastructure for SPatial Information in Europe. WWW page. http://www.ec-gis.org/inspire/.

[KRSO02] V. Katri, S. Ram, R.T. Snodgrass, and G. O'Brien. Supporting User Defined Granularities and Indeterminacy in a Spatiotemporal Conceptual Model. *Special Issue of Annals of Mathematics and Artificial Intelligence on Spatial and Temporal Granularity*, 36(1-2):195–232, 2002.

[MLR95] J-C. Muller, J.P. Lagrange, and R.Weibel, editors. *GIS and Generalization: methodology and practice.* London:Taylor and Francis, 1995.

[MS92] R.B. McMaster and K.S. Shea. *Generalization in Digital Cartography.* Washington, DC: Association of American Geographer, 1992.

[Mul91] J-C. Muller. Generalization of Spatial Databases. In D.j Maguire, M.F. Goodchild, and D.W. Rhind, editors, *Geographical Information Systems: Principles and applications*, chapter 30, pages 457–475. Harlow UK: Longman, 1991.

[PD95] E. Puppo and G. Dettori. Towards a Formal Model for Multiresolution Spatial Maps. In M.J. Eghenofer and J.R. Herrings, editors, *Proc. of 3th International Symposium Advances in Spatial Databases*, volume 951 of *Lecture Notes in Computer Science*, pages 152–169. Springer-Verlag, 1995.

[Ste99] J.G. Stell. Granulation for Graphs. In *Proc. of COSIT*, 1999.

[Ste03] J.G. Stell. Qualitative Extents for Spatio-Temporal Granularity. *Spatial Cognition and Computation*, 3(2-3):119–136, 2003.

[SW98] J.G. Stell and M. Worboys. Stratified Map Spaces: A Fomal Basis for Multi-Resolution Spatial Databases. In T.K. Poiker and N. Chrisman, editors, *Proc. of 8th International Symposium on Spatial Data Handling (International Geographical Union)*, pages 180–189, 1998.

[SW99] J.G. Stell and M. Worboys. Generalizing Graphs using Amalgamations and Selections. In R.H. Güting, D. Papadial, and F. Lochovsky, editors, *Proc. of 6th International Symposium on Advances in Spatial Databases*, volume 1651 of *Lecture Notes in Computer Science*, pages 19–32. Springer-Verlag, 1999.

[Tim99] S. Timpf. Abstraction, levels of detail, and hierarchies in map series. In *Proc. of COSIT*, 1999.

[Van01] C. Vangenot. Supporting Decision-Making with Alternative Data Representations. *Journal of Geographic Information and Decision Analisys*, 5(2):66–82, 2001.

[WD99] R. Weibel and G. Dutton. Generalizing Spatial Data and Dealing with Multiple Representations. In P.A. Longley, D.j Maguire, M.F. Goodchild, and D.W. Rhind, editors, *Geographical Information Systems: Principles, techniques, management and applications*, chapter 10. New York: John Wiley, 1999.