Towards Maintaining Consistency of Spatial Databases

Alia I. Abdelmoty and Chris B. Jones
Dept. of Computer Studies, University of Glamorgan,
Pontypridd, Mid Glamorgan, CF37 1DL, Wales, U.K.
email: aiabdel,cbjones@glam.ac.uk

Abstract

This paper focuses on the consistency issues related to integrating multiple sets of spatial data in spatial information systems such as Geographic Information Systems (GISs). Data sets to be integrated are assumed to hold information about the same geographic features which can be drawn from different sources at different times, which may vary in reliability and accuracy, and which may vary in the scale of presentation resulting in possible multiple spatial representations for these features. A systematic approach is proposed which relies first on breaking down the consistency issue by identifying a range of consistency classes which can be checked in isolation. These classes are a representative set of properties and relationships which can completely identify the geographic objects in the data sets. Different levels of consistency are then proposed, namely, total, partial and conditional, which can be checked for every consistency class. This provides the flexibility for two data sets to be integrated without necessarily being totally consistent in every aspect. The second step of the proposed approach is to explicitly represent the different classes and levels of consistency in the system. As an example, a simple structure which stores adjacency relationships is given which can be used for the explicit representation of topological consistency. The paper also proposes that the set of consistent knowledge in the data sets (which is mostly qualitative) be explicitly represented in the database and that uncertainty or ambiguity inherent in the knowledge be represented as well.

1 Introduction

Integrating data in spatial information systems involves the integration of diverse types of information drawn from a variety of sources requiring effective matching of similar entities in these data sets and demanding information consistency across data sets. Typically spatial information can be provided in different forms by a number of sources. For example, data sources in Geographic Information Systems (GISs) can include maps, field surveys, photogrammetry and remote sensing. Data sets may be collected at different scales or resolutions in different times. They may be collected in incompatible ways and may vary in reliability. Some details may be missing or undefined. Incompatibilities between different data sets can include incompatibilities between the spatial entities for which data are recorded, including differences in dimension, shape and positional accuracy.

For example, it may be required that a schematic representation of a certain area be stored in a GIS besides a more faithful representation (a schematic representation can be useful as an interactive tourist map). The two data sets are different. Many objects may be omitted from the schematic representation. The positional accuracy of the objects may not be maintained. However, both data sets hold the same relative position and orientation for the common subset of objects they hold.

Integrating both data sets in the geographic database involves the modelling and manipulation of multiple spatial representations for the same geographic objects. Since the two sets of knowledge are different, certain queries to the GIS may be more efficiently answered using one representation than the other.

The user of the system should have transparent use of the different data sets. Geographic information usually contains a certain amount of error which can result in uncertainty or ambiguity about the nature of some of the knowledge that can be derived from it. When integrating multiple data sets, this uncertainty may become either further or less complicated. The user of the GIS needs to be given a clear idea on the nature of the data he is using in order to evaluate the correctness of any results or analysis obtained from the system. In studying this problem research workers often refer to the need for initial processing of the data sets to make them "consistent". This term is used mostly in the sense of applying some operations on the spatial and attribute data, e.g. rectifying local geometric distortions, coordinate registration, reclassification, etc. [She91].

Integrating geographic information involves many issues [Flo91, She91, KPI87, GN90, JM86, Nye89], including, for example integrating vector and raster data [PLD90]. Work on spatial consistency between data sets has been mainly directed towards checking topological consistency [KPdB95, KPdB97, ES92, ECDF94].

In this paper, a systematic approach is proposed for studying and handling the consistency issue for integrating spatial information. The approach is based on the following steps:

- Analysing the different aspects of equivalence between the data sets. A range of spatial equivalence classes are identified which can be checked in isolation.
- Studying measures of spatial equivalence which can be applied to every class. Different levels of equivalence are proposed, namely, total, partial, conditional and inconsistent. Data sets can then be ranked as being consistent in which class to which level. This provides the flexibility for two data sets to be integrated without necessarily being totally consistent in every aspect.
- Explicit representation of the different equivalence classes and levels in the spatial database.
- Explicit representation of the set of consistent knowledge of different classes that exist in the data sets, as well as the representation of ambiguity or uncertainty inherent in this knowledge.

As an example, the representation of the topological consistency class is presented using a simple structure which stores adjacency relationships [AEG95]. This approach is aimed at providing a clear view of the nature of the data sets manipulated in the spatial information system. The user of the system could either consciously choose the type of data to manipulate or leave the system to decide which collection of data best suits his application. In both cases the user would have a clear idea about the extent of consistent knowledge that can be retrieved as well as the nature and measure of any inconsistency that may result. The issues in this paper are discussed using geographic information in GISs as example, but are also applicable to other types of spatial information systems and applications. This work is done in the context of an ongoing research project which aims at the development of methods for the modelling and manipulation of hybrid data sets in a GIS [JKL$^+$96, JKW96].

The paper is structured as follows. Different consistency classes are identified in section 2 and the different levels of consistency are proposed in section 3. Explicit representation of consistency is discussed in section 4 and a possible representation of the topological consistency is presented. In section 5, the issue of representing uncertainty is discussed and some conclusions are given in section 6.

2 Aspects of Spatial Equivalence

In integrating two sets of spatial data which relate to the same area in space, two consecutive steps are needed,

1. Object matching: where corresponding objects in both sets are identified using spatial equivalence tests. The result of this procedure is the identification of which objects in both sets can be considered to be the same, for example, matching two sets of land parcels in an old and an uptodate map or matching two road networks in maps with different scales, etc. Note that those objects could differ with regard both to positional information and geometric structure.
2. Spatial Equivalence representation: where the explicit representation of the relationship between the data sets is needed to allow the intelligent manipulation of both sets by the system and to project to the user a clear view of the nature of the data used.

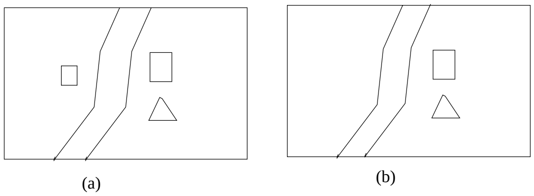


Figure 1: Existentially non-consistent data sets. Some of the objects in (a) are missing in (b).

The equivalence of two representations of a spatial object can be studied from three points of view: of an origin point, of one of the objects and of other objects in the data sets. Thus equivalence can be studied using an absolute frame of reference, an object-based frame of reference and a relation-based frame of reference. Three classes of spatial equivalence can therefore be identified as follows.

2.1 Positional Equivalence

Objects are represented by the specific coordinates describing their spatial extents. Under this reference, two objects from two different data sets match only if their representative sets of coordinates match exactly and two data sets can be considered as locationally consistent if any position (x,y,z) corresponds to the same object in both sets.

2.2 Object-Based Equivalence Classes

A spatial data set consists of the spatial properties of a set of objects in a defined space. These properties include a description of spatial extent, from which the dimension and the shape of the object can be derived. An object in the data set can be composite, i.e. consisting of or containing other objects.

Object-based consistency can be classified using the above properties. Two spatial data sets can be said to be object-based consistent of a certain class if for each object in both sets this consistency is achieved.

(I) Object Existence Equivalence
Two data sets are existentially equivalent if all the object classes and instances in one data set exist in the other data set. For example the two data sets in figure 1 are existentially non-equivalent.

(II) Object Dimension Equivalence
Two data sets are equivalent with reference to object dimension, if every object in one set has the same spatial dimension as that of the corresponding object in the other set. For example, the two spatial scenes in figure 2 are not equivalent with reference to object dimension as objects are represented using spatial representations of different dimensions (areas by points or lines).

(III) Object Shape Equivalence
Equivalence based on object shape can be as flexible as needed. On a strict level object shapes can be defined

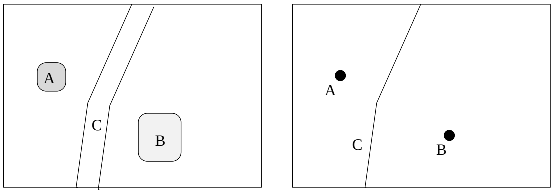


Figure 2: Non-consistent data sets with reference to object dimension.

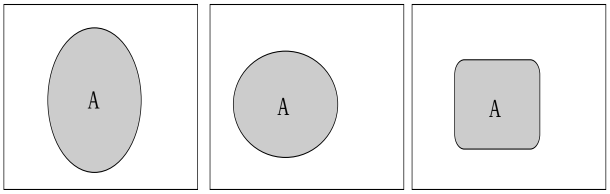


Figure 3: Object A may be considered to have equivalent shape in all three sets according to the allowed shape distortion.

using equations of the curve or set of curves defining its boundary. On a less accurate level object shapes can approximate well-known geometric shapes, for example a circle, a square, T shaped, zig-zag, etc.

Two data sets can said to be equivalent with reference to object shape if every object in the set can be described as shape equivalent to the corresponding object in the other set. For example, in figure 3 the shape of object A may be considered equivalent in all three scenes depending on the measure of shape distortion accepted in the database. The first two shapes only may be considered equivalent if different measures are used.

(IV) Object Size Equivalence
Several measures of size exist including, length of boundaries, areas and volumes of shapes. Two data sets may be considered as equivalent with reference to object size if every object in one set has a similar size to the corresponding object in the other set.

(V) Spatial Detail Equivalence
Objects in the data sets may be composite, i.e. containing other objects or made up of several connected or non-connected objects. Two data sets can be considered to be equivalent with reference to object detail if corresponding composite objects in both sets can be considered to be equivalent, as shown in figure 4.

Interdependency between Equivalence Classes
Other classes of object-based equivalence may exist. The above set of classes are possibly the most important from a general point of view. Note that the above classes may not be mutually exclusive. In particular the positional consistency implies every other type of consistency (except categorical) and is by default the strictest measure of spatial equivalence. Shape and size imply dimension and all equivalence classes imply existence equivalence. Shape equivalence may imply spatial detail consistency if the object is composed of non-connected sets, etc. Also, it is assumed that a certain degree of inaccuracy can be acceptable in the measurement of some of the properties, for example, size and shape. However, this depends on the applications intended over these data sets.

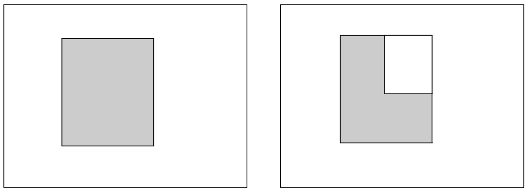
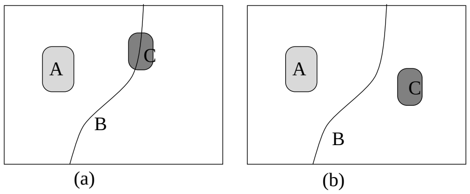
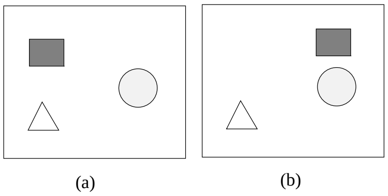


Figure 4: Non-consistent data sets with reference to object details.



(a) (b)

Figure 5: Topological inconsistency. (a) Object B crosses object C. (b) Object B is disjoint from C.



(a) (b)

Figure 6: Directional inconsistency.

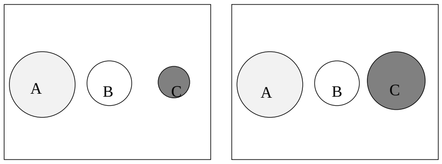
2.3 Relation-Based Equivalence Classes

The third type of consistency measures is based on the spatial relationships between objects in the data sets considered. Three classes of equivalence can be classified according to the types of spatial relationships [AW94, AEG94].

(I) Topological Equivalence
Two data sets can be regarded as topologically consistent if the set of topological relationships derived from one set are the same as those derived from the other. For example, the two sets in figure 5 are not topologically consistent.

(II) Direction or Orientation Equivalence
Two data sets can be regarded as directionally consistent if the relative direction relationship in one set is the same as the other set. For example, the two sets in figure 6 are not consistent directionally.

(III) Relative Size Equivalence
Two data sets can be regarded as consistent with reference to relative size relationships if the qualitative size relations of larger and smaller are maintained between corresponding sets of objects in the two sets. Figure 7 shows relative size inconsistency between 7(a) and (b).



(a) (b)

Figure 7: Relative size inconsistency. $A > B > C$ in (a) and $A > B < C$ in (b).

3 Different Levels of Spatial Consistency

Two geographical data sets can be consistent in more than one class of those defined above. For example, the data sets can be topologically and dimensionally equivalent, or consistent with reference to dimension, detail and category, etc. As noted earlier some consistencies do assume others. For example, topological equivalence may assume spatial detail. Up till now, the discussion is based on one level of consistency, namely, when all objects in the data sets conform to the consistency class studied. In reality this is not always the case. Ranking the level of consistency for the different classes identified is important as it would provide the user of the GDB an initial measure of the nature of the data sets in his use. Further processing of this ranking would be to identify how the data sets compare and which parts of the data sets are consistent, i.e. the nature of such consistency.

Let S_1 and S_2 represent the set of knowledge present in two data sets. This knowledge consists of all the different types of information that can be derived from every data set. It can be classified according to the object-based and relation-based classes. Let S_{1i} and S_{2i} represent the subsets of the set of knowledge S_1 and S_2 respectively, which belong to a certain class i, e.g. shape properties or directional or topological relationships, etc. Four different levels of consistency can be identified.

(A) Total Consistency
Two data sets S_1 and S_2 can be said to be totally consistent with reference to a certain consistency class i, if $S_{1i} \cap S_{2i} = S_{1i} \vee S_{2i}$, i.e. $S_{1i} = S_{2i}$. In this case a query to the GIS involving only properties of class i shall return identical results if posed to either S_1 or S_2.

(B) Partial Consistency
Two data sets S_1 and S_2 can be said to be partially consistent with reference to a certain consistency class i, if $S_{1i} \cap S_{2i} = C_i$ and $C_i \subset S_{1i} \wedge C_i \subset S_{2i}$.

In this case only part of class i knowledge is consistent in the two sets. If the two data sets are to be used together, then it is important to know which subsets of the different classes of knowledge can be manipulated interchangeably between sets.

(C) Conditional Consistency
Two data sets S_1 and S_2 are said to be conditionally consistent with reference to a certain consistency class i, if there exists a set of functions F which when applied to S_{1i} makes it totally consistent with S_{2i}, i.e. $S_{2i} = F(S_{1i})$. This can also represent the case where S_{1i} is consistent with S_{2i} but S_{2i} is not consistent with S_{1i}, i.e. $(S_{1i} \cap S_{2i} = S_{1i}) \wedge (S_{1i} \subset S_{2i})$, (an asymmetric consistency).

The set of functions F must be non-ad-hoc, i.e. predefined. For example, the set of cartographic generalisation rules used to produce maps at different scales or a set of predefined rules used to produce a schematic from a faithful representation of a map.

(D) Inconsistency Level
Two data sets S_1 and S_2 can be said to be inconsistent with reference to a certain consistency class i, if $S_{1i} \cap S_{2i} = \phi$, i.e. they do not share any piece of knowledge from that class. In this case a query to the GIS involving properties of class i shall return non-identical results if posed to S_1 and S_2.

In most cases the data sets which need to be integrated relate to a combination of classes and levels. For example, two data sets can be partially consistent in terms of shape and dimension but are totally consistent topologically, or are conditionally consistent with respect to object detail as well as partially consistent topologically.

4 Representation of Different Levels of Consistency for Different Classes

Determining the class and level of consistency between two data sets involves the extraction and comparison of the set of properties or relationships for that class. Although it is useful for the user and the system to be informed of the class and level of consistency in general, it may not be enough for certain application domains. In those cases explicit representation of the consistent set of knowledge is needed.

A closer look at the different classes of consistency reveals that they are mostly qualitative measures (apart from location, size and shape). Hence, the common set of spatial knowledge between data sets can be represented qualitatively. A structuring mechanism can be envisaged which can be applied on a geographic data set to allow the explicit representation of some of the qualitative properties and relationships and the derivation of others. Multiple spatial representations can exist for the same geographic objects, however properties and relationships are always related to objects and not to their underlying representations. Hence the structuring mechanism envisaged should be based on the geographic objects level and not on the geometrical representations. This structure can then be built for any data set irrespective of its underlying form of spatial representation.

Manipulation of such qualitative structure could make use of spatial reasoning techniques [Ege94, CRCB93, EGA96, Her94]. For example it would be possible to store only some of the topological relationships and derive others using composition tables for similar and mixed types of spatial relations.

Explicit representation of this knowledge would allow comparisons between data sets, seamless manipulation of existing sets, integration of new sets and consistent update of existing ones.

Work still needs to be done on developing the proposed structuring mechanism. Several questions need to be answered, including,

- What are the types of knowledge that can be represented explicitly and which can be derived?.
- How can the different classes of knowledge be structured?

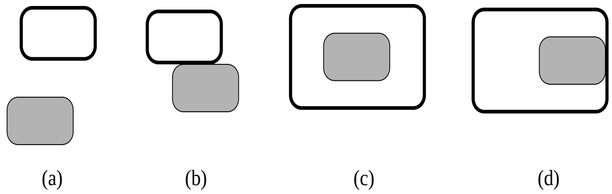


Figure 8: Set of topological relationships that can exist between two areal objects in one map theme. Overlap relationships are excluded as they only result from overlaying more than one geographic theme.

In the following section, the representation of the class and level of topological consistency is given as an example of explicit representation.

4.1 Representation of Topological Consistency with Adjacency Relationships

To represent topological consistency between two geographic data sets is to check that the same set of topological relationships between objects in one set exist for the corresponding objects in the other set. This process involves the explicit extraction and representation of topology. Several approaches to checking the topological consistency of two spatial scenes have been proposed [KPdB95, KPdB97, ES92, ECDF94]. However, these approaches didn't consider the issue of integrating both scenes and hence didn't provide ways of representing the common consistent knowledge.

In this section, a simple structure for storing the adjacency relationships between objects in the data sets is used [AEG95] from which topological relationships can be derived. The structure can then be used to represent the common set of consistent knowledge between data sets as well the ambiguity or uncertainty in the knowledge derived from both sets. The structure is based on the following assumptions.

Assumptions

- It is assumed that the data sets considered must contain non-overlapping objects, i.e. that every object in these sets occupies a unique location in space. However, it is possible for the data sets to contain objects in part-of relationships, e.g. a lake inside a forest or a city bounded by part of a motor-way, etc. As an example, figure 8 shows the different relationships that can be considered between areal objects. While the adjacency relations are enough to represent the topology of the geographic scene, an explicit part-of relationship needs to be defined when one object is part of another to distinguish the topology in the case of relationships in 8(b) and (d).

  The only limitation of this assumption is that we only compare objects from one theme and not from the overlay of more than one theme which is natural when integrating two data sets.

- The geometric representation of linear geographic objects shall assume splitting of intersecting lines. The topology of the linear geographic objects in this case is the collection of lines/arcs representing its extent, the intermediate nodes between arcs and the pair of end points as shown in figure 9. This spatial representation is typical of geometric structures of most GISs. This assumption is needed for the representation of the various topological relations involving linear geographic objects, but is not applicable for the boundary of areal objects.

- It is assumed that a geographic data set is always embedded in an infinite space, and hence the infinite complement of the extent of this set is explicitly represented in the adjacency structure proposed [EGA96].

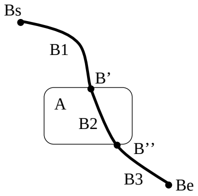
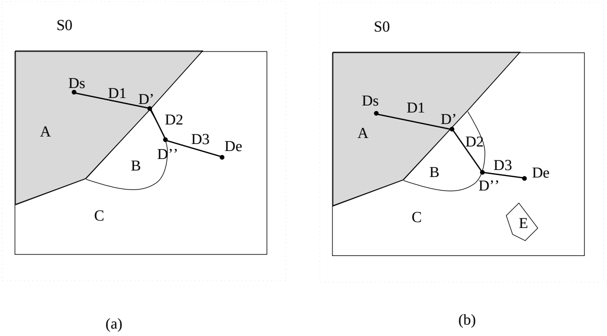


Figure 9: Representation of the topology of linear geographic objects by splitting intersecting lines.



(a) (b)

Figure 10: Two maps containing areal objects A, B and C and a linear geographic object D. The areal object E doesn't exist in (a) and the relationships between B and D are different in both cases.

We can now build a simple structure based on connectivity or adjacency relationships to represent the topology of a geographic data set. Consider the two scenes in figure 10. In 10(a), part of object D, namely D_2, lies on the boundary of B while in 10(b), object D crosses object B.

The topology of the maps in 10(a) and (b) can be represented by an adjacency structure as shown in figure 11(a) and (b) respectively. Note that adjacency is a symmetric relation and hence the structure in the figure (half a matrix) is sufficient. A (1) in the structure represents an adjacency relationship while a (0) indicates that the related objects are not adjacent. For example, in 11 (a), object A is adjacent to objects B, C, D_s, D' and D_1, and not to D_2, D'', D_3 and D_e.

There are two differences between the two scenes as can be seen from the structures. These are: in 11(a) object D_2 is connected to C while it is not in 11(b), and object E in 11(b) does not exist in 11(a).

The only relationship stored explicitly in the above structures is **adjacency** and other topological relationships can be simply derived. For example, in 11(b),

$S0$										
1	A									
0	1	B								
1	1	1	C							
0	1	0	0	D_s						
0	1	0	0	1	D_1					
0	1	1	1	0	1	D'				
0	0	1	1	0	0	1	D_2			
0	0	1	1	0	0	0	1	D''		
0	0	0	1	0	0	0	0	1	D_3	
0	0	0	1	0	0	0	0	0	1	D_e

(a)

$S0$											
1	A										
0	1	B									
1	1	1	C								
0	1	0	0	D_s							
0	1	0	0	1	D_1						
0	1	1	0	0	1	D'					
0	0	1	0	0	0	1	D_2				
0	0	1	1	0	0	0	1	D''			
0	0	0	1	0	0	0	0	1	D_3		
0	0	0	1	0	0	0	0	0	1	D_e	
0	0	0	1	0	0	0	0	0	0	0	E

(b)

Figure 11: The adjacency structures corresponding to the maps of figure 10 (a) and (b) respectively.

object E is adjacent only to C and hence it is topologically **inside** C. Also, the relationship between object D with any other object can be realised from the grouping of relationships between its constituting parts, and so on. Hence, using these structure alone we can redraw the topological equivalences of the two scenes (obviously the exact shape of each object is not meant to be represented here). The adjacency structures can be organised in a tree structure representing different levels of detail in the data sets. Also, an explicit reference to object dimension will enable a (schematic) reproduction of the topological equivalent of the data sets. However, object dimension in both data sets need not be consistent.

In [ES92] the topological consistency of a scene is checked by solving the problem as a constraint network and checking that the network is path-consistent. The sets of all possible relations between the different types of objects in a scene need to be used for the representation of the problem in a constraint network, while the adjacency structure only uses a binary relation. The adjacency structure can also be used to represent the common consistent set of knowledge in two scenes as follows.

Representing the Common Consistent Set of Knowledge

S_1 and S_2 are partially topologically consistent. The set of common knowledge in both data sets can be grouped in an adjacency structure as shown in figure 12. The structure in 12 is informative of the common consistent topological knowledge between the two data sets. In this case, the adjacency between objects D_2 and C is unknown, represented by a $(-)$, and object E doesn't exist in both data sets and hence it is deleted from this set. Using this structure one can recreate the common knowledge in both scenes with the ambiguity of the relation between D_2 and C. Figure 13 shows the integration of different sets of knowledge which are consistent in different classes and levels.

$S0$										
1	A									
0	1	B								
1	1	1	C							
0	1	0	0	D_s						
0	1	0	0	1	D_1					
0	1	1	-	0	1	D'				
0	0	1	-	0	0	1	D_2			
0	0	1	1	0	0	0	1	D''		
0	0	0	1	0	0	0	0	1	D_3	
0	0	0	1	0	0	0	0	0	1	D_e

Figure 12: The adjacency structure representing the common set of consistent knowledge in the structures of figure 11.

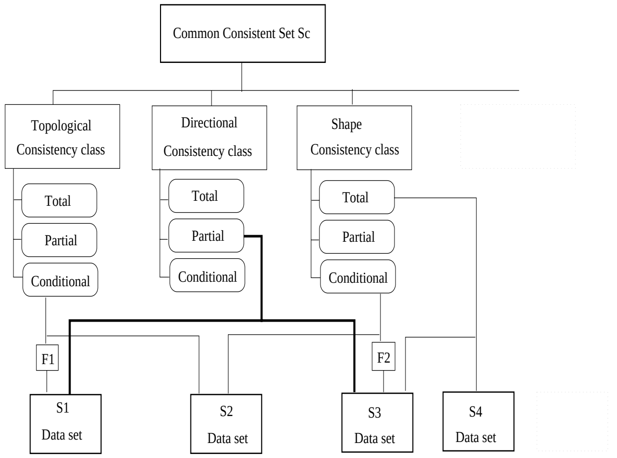


Figure 13: Integrating different data sets with different classes and levels of consistency to produce a common set of consistent knowledge. $F1$ and $F2$ represent sets of predefined functions for conditional consistency.

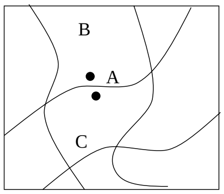


Figure 14: Small positional error can result in ambiguity of spatial relationships. Is A is in object B or C.

5 Dealing with Ambiguity or Uncertainty

Error exists in all geographic data sets. Different types of error can occur whether cartographic (error in the position of map features) or thematic (error in the values of an attribute of map features [Ver89]. In [Ope89] and in recognition of the fact that error in the GIS databases is here to stay, the need to live with error is advocated. A lot of research is ongoing for the identification of different types and causes of error, on methods of estimating the extent and importance of error and on methods for dealing with them [Goo89, MAB89, Bla84].

If error affects the manipulation and use of one geographic data set, it would have a more severe effect when integrating more than one data set in the GIS. The quality of geographic information is commonly expressed by some tolerance values within which the location of objects may vary. The problem is complicated when geographic features are approximated to point or line symbols as for example in the process of creating different map scales. Locational error can have an effect on the spatial relationships between objects in the data set. A very small error in the position of one object can be a major problem if it results in a change of relationships as shown in figure 14.

In this paper we propose the explicit representation of the consistent set of knowledge between different data sets and the explicit recognition of the class and level of consistency. Error in the data sets leads invariably to ambiguity or uncertainty. Hence, this uncertainty must be reflected in the representation of consistency of the data set. For example, tolerance values must accompany the locational knowledge (point A is at (x,y) with tolerance t). The effect of ambiguity on other classes of knowledge must be represented. For example, the relation between objects A and B is either *touch* or *disjoint*, and so on. This ambiguity must also be represented in the common set of consistent knowledge between the data sets.

Note that the ambiguity in a GIS with integrated data sets can be either from ambiguities in every data set or ambiguities resulting from the integration process itself. The adjacency structure in figure 12 is an example of representing the later type of ambiguity. The $(-)$ in the matrix implies the existence of two possible relations between objects. The same can apply on any other consistency class, for example, in studying size in two data sets where in one data set the relationship A is larger than B and in the other data set it is A is equal to B. Then it can be safe to say that the consistent set of knowledge that can be derived from those data sets includes the fact that A is either larger or equal to B.

This flexibility in representing the consistency conveys a more realistic view of the nature of the data sets and would certainly provide a more accurate base for analysis and manipulation.

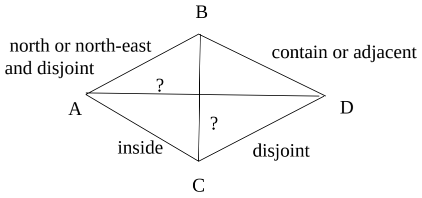


Figure 15: Constraint networks for the representation of uncertain spatial knowledge represented by a disjunction of possible relations.

Methods for the representation of uncertainty in the knowledge need to be developed. Reasoning with incomplete knowledge is a major research area in spatial reasoning where constraint networks may be used for the representation of uncertainty [SP92, Her94] as shown in the example in figure 15.

6 Conclusions

In this paper a study of the nature of consistency issues for integrating hybrid data sets is presented. In particular the paper focuses on the consistency issues related to integrating multiple sets of spatial information for the same area in space.

The proposed approach can be summarised as follows:

- The concept of consistency between data sets is broken down into two main categories: a study of the comparison of basic properties of objects and relationships between those objects. Nine consistency classes are identified under those categories which can be checked in isolation.
- For every class identified data sets can be consistent to a certain level. Four levels of consistency are proposed, namely, total, partial, conditional and inconsistent. Data sets can be ranked according to those levels, for example, totally consistent topologically but partially consistent with reference to object dimension and so on.
- The explicit representation of the different classes and levels of consistency is needed for the database to reveal a realistic view of the nature of its contents.
- The common set of consistent knowledge in the data sets needs to be explicitly expressed. A qualitative structure is proposed to hold different types of knowledge on the geographic feature or object level (as opposed to the geometric level). An explicit representation is needed of the ambiguity or uncertainty inherent in every data set and that resulting from the integration of several data sets.

As an example, the representation of the topological consistency class is presented using a simple structure which stores adjacency relationships [AEG95]. Topological relationships can be derived from the structure and any ambiguity in the relationships can also be derived.

Further work need to done for devising representation methods for the different consistency classes and for their coherent integration.

The work in this paper is done in the context of an ongoing research project which aims at the development of methods for the modelling and manipulation of hybrid data sets in a GIS [JKL$^+$96].

References

[AEG94] A.I. Abdelmoty and B.A. El-Geresy. An Intersection-based Formalism for Representing Orientation Relations in a Geographic Database. In *2nd ACM Workshop On Advances In Geographic Information Systems*, pages 44–51. ACM press, December 1994.

[AEG95] A.I. Abdelmoty and B.A. El-Geresy. A General Method for Spatial Reasoning in Spatial Databases. In *Proceedings of the Fourth International Conference on Information and Knowledge Management (CIKM'95)*, pages 312–317. ACM press, 1995.

[AW94] A.I. Abdelmoty and M.H. Williams. Approaches to the Representation of Qualitative Spatial Relationships for Geographic Databases. In *Advanced Geographic Data Modelling (AGDM'94)*, number 40 in Publications on Geodesy, pages 204–216. Netherlands Geodetic Commission, 1994.

[Bla84] M.J. Blakemore. Generalisation and Error in Spatial Data Bases. *Cartographica*, 21:131–9, 1984.

[CRCB93] A.G. Cohn, D.A. Randell, Z. Cui, and B. Bennet. Qualitative Spatial Reasoning and Representation. In P. Carrete and M.G. Singh, editors, *Qualitative Reasoning and Decision Technologies*, 1993.

[ECDF94] M.J. Egenhofer, E. Clementini, and P. Di Felice. Evaluating Inconsistencies among Multiple Representations. In *Proceedings of the 5th International Symposium on Spatial Data Handling*, volume 2, pages 901–918, Charleston, 1994. IGU Commission of GIS.

[EGA96] B.A. El-Geresy and A.I. Abdelmoty. Order in Space: A General Formalism for Spatial Reasoning. In *Proc. of the Eighth International Conference on Tools with Artificial Intelligence, TAI'96*, pages 183–193, Toulouse, France, 1996. IEEE Computer Society.

[Ege94] M.J. Egenhofer. Deriving the composition of Binary Topological Relations. *Journal of Visual Languages and Computing*, 5:133–149, 1994.

[ES92] M.J. Egenhofer and J. Sharma. Topological consistency. In P. Bresnahan, E. Corwin, and D. Cowen, editors, *Proceedings of the 5th International Symposium on Spatial Data Handling*, volume 2, pages 335–343, Charleston, 1992. IGU Commission of GIS.

[Flo91] R. Flowerdew. Spatial Data Integration. In D.J. Maguire, M.F. Goodchild, and D.W. Rhind, editors, *Geographic Information Systems*, volume 1, pages 375–87. Longman Scientific & Technical, 1991.

[GN90] M. Gargano and E. Nardelli. A Logical Data Model for Integrated Geographical Databases. In *Proc. 1ST Int. Conference on Systems Integration*, IEEE, pages 473–81, 1990.

[Goo89] In M.F. Goodchild and S. Gopal, editors, *Accuracy of Spatial DatabaseS.* Taylor & Francis, London, 1989.

[Her94] D. Hernandez. *Qualitative Representation of Spatial Knowledge*, volume 804. Springer Verlag, 1994.

[JKL$^+$96] C.B. Jones, D.B. Kidner, L.Q. Luo, G.L. Bundy, and J.M. Ware. Database Design for a Multiscale Spatial Information System. *International Journal of Geographical Information Systems*, 10(8):901–920, 1996.

[JKW96] C.B. Jones, D.B. Kidner, and J.M. Ware. Equivalence Tests for Integration of Multi-Source Spatial Data. In *Proceedings EuroCarto XIII, Ispra, Italy*, pages 13–22. Joint Research Centre, European Commission, 1996.

[JM86] M.J. Jackson and D.C. Mason. The Development of Integrated Geo-information Systems. *Int. Journal of Remote Sensing*, 7(6):723–40, 1986.

[KPdB95] B. Kuijpers, J. Paredaens, and J.V. den Bussche. Lossless Representation of Topological Spatial Data. In M.J. Egenhofer and Herring J.R., editors, *4th International Symposium on Large Spatial Databases, (SSD'95)*, pages 1–13. Springer Verlag LNCS-951, 1995.

[KPdB97] B. Kuijpers, J. Paredaens, and J.V. den Bussche. On topological elementary equivalence of spatial databases. In F. Afrati and Ph. Kolaitis, editors, *to appear in the Proceedings of 6th International Conference on Database Theory (ICDT '97)*, pages 432–446. Springer Verlag LNCS-1186, 1997.

[KPI87] T. Keating, W. Phillips, and K. Ingram. An Integrated Topologic Database Design for Geographic Information Systems. *Photogrammetric Engineering and Remote Sensing*, 53(10):1399–1402, October 1987.

[MAB89] G. Maffini, M. Arno, and W. Bitterlich. Observations and Comments on the Generation and Treatment of Error in Difital GIS Data. In *[Goo89]*, pages 55–67, 1989.

[Nye89] T. Nyerges. Schema Integration Analysis for the Development of GIS Databases. *Int. J. Geographic Information Systems*, 3(2):153–183, 1989.

[Ope89] S. Openshaw. Learning to Live with Errors in spatial Databases. In *[Goo89]*, pages 263–76, 1989.

[PLD90] J.M. Piwowar, E.F. Ledrew, and D.J. Dudycha. Integration of Spatial Data in Vector and Raster Formats in a Ge ographic Information System Environment. *Int. J. Geographic Information Systems*, 4(4):429–444, 1990.

[She91] I.D.H. Shepherd. Information Integration and GIS. In D.J. Maguire, M.F. Goodchild, and D.W. Rhind, editors, *Geographic Information Systems*, volume 1, pages 337–60. Longman Scientific & Technical, 1991.

[SP92] T.R. Smith and K.K. Park. Algebraic Approach to Spatial Reasoning. *International Journal of Geographic Information Systems*, 6(3):177–192, 1992.

[Ver89] H. Veregin. Error Modeling for the Map Overlay Operation. In *[Goo89]*, pages 3–18, 1989.