UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN GEOGRAFÍA

THE DEFINITION OF A MINIMUM SET OF SPATIAL RELATIONS

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN GEOGRAFÍA

PRESENTA:
LUIS MIGUEL MORALES MANILLA

TUTOR
DR. JOSÉ LUIS PALACIO PRIETO
INSTITUTO DE GEOGRAFÍA, UNAM

MORELIA, MICHOACÁN, JUNIO 2014

a MIS PADRES

a MAYELA

a PABLO

THE DEFINITION OF A MINIMUM SET OF SPATIAL RELATIONS

CONTENTS

INTRODUCTION 11

CHAPTER 1 14

THE NEED FOR A COMMON APPROACH TO SPATIAL PROBLEM SOLVING 14

- 1.1 ON GEOGRAPHY 15
  - The integration of time and space 15
  - The fragmentation of Geography 16
- 1.2 SPATIAL ANALYSIS 17
  - The limits of Spatial Analysis 18
- 1.3 GIS AND GISCIENCE 19
  - GIS 19
  - GIScience 19
- 1.4 A SPATIAL PROBLEM-SOLVING FRAMEWORK FOR GEOGRAPHY 20
  - From values to patterns 21
  - The goals of spatial problem-solving 22

CHAPTER 2 25

THE THEORY OF SPACE-EVENT INTERACTION 25

- 2.1 A BRIEF HISTORY OF THE CONCEPT OF SPACE 26
  - The perception of space 26
  - The representation of space 26
  - The conception of space 27
  - The relativization of the concept of space 27
  - A couple of ‘final’ ideas about space 28
- 2.2 SPACE AND EVENTS 29
  - Properties of space 29
  - Properties of events 30

2.3 THE THEORY OF SPACE-EVENT INTERACTION (TSEI) 32

Rationale 32

Foundations 33

Spatial Relations 34

Interactions between properties of space and properties of events 36

CHAPTER 3 38

A MINIMUM SET OF SPATIAL RELATIONS 38

3.1 SPATIAL RELATIONS 39

The concept of spatial relations 39

The meaning of spatial relations 40

The importance of a minimum set of spatial relations 41

Relevancy of spatial relations 41

Topological qualities of spatial relations 42

Non-topologic relations 43

Relative-topologic relations 44

Absolute-topologic relations 45

3.2 FROM STRUCTURE TO ORGANIZATION 46

The structuring of events and the organization of space 46

Structure – behavior - organization 47

3.3 SPATIAL RELATIONS IN NATURAL LANGUAGES 48

Space in natural language 48

Spatial relations in English and Spanish 49

Syntactics and semantics of spatial terms 51

3.4 RELATIONSHIPS BETWEEN SPATIAL RELATIONS 52

CHAPTER 4 60

STRUCTURAL RELATIONS 60

4.1. PROXIMITY ... 61
  Definition ... 61
  Proximity and distance ... 61
  The concept of Proximity ... 62
  Logic properties ... 63
  Topologic properties ... 64
  Qualitative expressions ... 64
  Quantitative measurements ... 66
  Crisp measurements ... 67
  Fuzzy measurements ... 69
4.2 ORIENTATION ... 71
  Definition ... 71
  Orientation and direction ... 71
  The concept of Orientation ... 71
  Logic properties ... 73
  Topologic properties ... 74
  Qualitative expressions ... 74
  Qualitative systems of orientation ... 75
  Quantitative measurements ... 77
4.3 EXPOSURE ... 78
  Definition ... 78
  The concept of Exposure ... 78
  Barriers ... 79
  Logic Properties ... 79
  Topologic Properties ... 79
  Qualitative Expressions ... 80

Quantitative Measurements 80
CHAPTER 5 81
NEUTRAL RELATIONS 81
5.1 ADJACENCY 82
Definition 82
Adjacency and the concept of boundary 82
The concept of adjacency 83
Logic properties 83
Topologic properties 84
Qualitative expressions 85
Quantitative measurements 85
5.2 CONTAINMENT 87
Definition 87
Containment and explicit order 87
The concept of containment 87
Logic properties 88
Topologic properties 89
Qualitative expressions 89
Quantitative measurements 90
5.3 COINCIDENCE 91
Definition 91
Coexistence in time and space 91
The concept of coincidence 91
Logic properties 92
Topologic properties 92
Qualitative expressions 93

    - Quantitative measurements 93
- CHAPTER 6 94
- ORGANIZATIONAL RELATIONS 94
  - 6.1. CONNECTIVITY 95
    - Definition 95
    - Links, flows and connections 95
    - The concept of connectivity 96
    - Logic properties 96
    - Topologic properties 97
    - Qualitative expressions 97
    - Quantitative measurements 97
  - 6.2 AGGREGATION 98
    - Definition 98
    - Emergence without (apparent) purpose 98
    - The concept of aggregation 99
    - Properties 100
    - Qualitative expressions 101
    - Quantitative measurements 102
  - 6.3 ASSOCIATION 103
    - Definition 103
    - Awareness and purpose 103
    - The concept of association 104
    - Properties 105
    - Qualitative expressions 106
    - Quantitative measurements 107
  - 6.4 FINAL NOTES ON SPATIAL ORGANIZATION 108

CHAPTER 7 109

A GEOGRAPHIC APPROACH TO VULNERABILITY ASSESSMENT 109

Abstract 110

7.1 INTRODUCTION 110

7.2 THE NEED FOR A COMPREHENSIVE APPROACH TO THE ASSESSMENT OF VULNERABILITY . 113

7.3 THEORETICAL FRAMEWORK 116

Place 116

Spatial Relations 117

7.4 PROBLEM-SOLVING FRAMEWORK 120

Phase 1. Search of Vulnerability Patterns 125

Phase 2. Description of Vulnerability Patterns 128

Phase 3. Explanation of Vulnerability Patterns 128

Phase 4. Prediction of Vulnerability Patterns 129

Phase 5. Design of Vulnerability Patterns 130

7.5 CONCLUDING REMARKS 131

CHAPTER 8 133

LAND ATTRACTIVENESS AS A SUITABILITY MEASURE FOR AGRICULTURAL PURPOSES: THE AGRILOCAL MODEL 133

8.1. INTRODUCTION 134

8.2 THEORY 136

Ideal practices versus actual practices in agriculture: technical knowledge versus local knowledge 136

The container of farmers' expert knowledge 136

Spatial relations and agriculture 137

8.3. THE AGRILOCAL MODEL 138

Key characteristics of the model 138

Step one: finding relevant spatial relations 139

Step two: specifying preference functions.. 144

Step three: establishing the degree of influence 145

Synthetic mode of the AGRILOCAL model 146

Parametric mode of the AGRILOCAL model... 147

8.4 DISCUSSION 150

Validation of the AGRILOCAL results.. 150

Limitations of the model 152

Utility of the AGRILOCAL model 155

Concluding remarks and future research... 156

CONCLUSIONS 158

REFERENCES 161

INTRODUCTION

When exploring and solving spatial problems in astronomic, geographic, architectural or subatomic spaces, one is tempted to point out that there are fundamental aspects clearly differentiating these fields, although all of them use of the concept of space.

The primary difference appears to be one of scale, which intervenes in many aspects of the spatiality of problems, from the possibility to perceive events to the organization of events into other events. A second difference, somewhat related to the first, is that spatial events appear to be of different kind: celestial bodies in astronomic space, cities or rivers in geographic space, buildings and rooms in architectural space, elementary particles in subatomic space. Thirdly, processes seem to be different as well: intergalactic collisions in astronomic space, traffic and erosion in geographic space, wall weathering and renovation in architectural space, fission and beta decay in the subatomic space. Lastly, forces driving processes appear to be also different: gravitational and electromagnetic in astronomic space, economic development and barometric pressure in geographic space, construction costs and building permits in architectural space, strong forces and weak forces in subatomic space, to mention some examples.

Even within a single space, geographic for instance, there appear to be different events, forces, processes and patterns, depending on the discipline we practice. Geophysicists, economists, ecologists or geographers, work on spatial problems in geographic space, but do it so with different approaches, presumably because subjects of study are all different. Wouldn't be ideal if, no matter the scale or the discipline, all spatial problems could be approached essentially in the same way? That would be a reality if we all had attended the 'Unified Space-Time Science 101' course in High School, which unfortunately is not taught. Yet.

Space is a unique notion, meaning that differences in scales, events, forces or processes are only apparent and should not translate into differences of spatial approaches. Knowledge of this uniqueness would allow us to approach spatial problem-solving in a common way. Uniqueness of the concept of space can be understood at some degree if we look for a common denominator among events, forces, processes and patterns in spatially-aware disciplines. This common denominator is 'interaction'. Everything interacts with everything, directly or indirectly. We may observe that events at one scale interact with other events in similar or different scales, with interaction driven by different forces that, with time, create processes, which in their turn create patterns. At the end we find that these patterns are themselves events at another scale of space-time, in a cyclic fashion.

But 'interaction' is a too general concept. For practical reasons we need specific concepts to describe specific forms of interaction, particularly spatial interactions. How could we describe the interactions between us and our parents, between us and the place we live in, between us and the planet, between us and the Universe? As an example, the following set of questions could serve to guide the description of the interaction between us and our parents: Are you *close* to your parents? Are you in *contact* with them? Do you *see* or *hear* each other often? Do you *share* your time and space occasionally with them? How well or bad is the *integration* as a family? Are they *part of* your life? Are you *member of* a family?

Pay attention to the italics. What are all these terms about? Relations, of course; relations are the specific concepts we use to describe interactions. In fact, they are *spatial relations*, although they may not seem too spatial. That might be because the above questions allude to relations in a social, affective, space, but we can replace the subject and object if we wish to use the same italicized terms in different spaces.

The importance of the concept of spatial relations to Science lies in that they provide the cognitive structure to understand interactions in all kinds of spatial problems. To give an idea of their relevance, a simple search on the internet (Google, September 2012) with the exact words “spatial relations”, “spatial relation”, spatial relationships” and “spatial relationship”, resulted in a total of 2,889,000 references (and that only in English), which of course might seem not as relevant as fundamental concepts such as “love” (8,370,000,000 references) or “life” (1,480,000,000 references), but start to appear relevant when the context is limited to less fundamental notions such as “spatial” (82,700,000 references), and “geography” (45,000,000 references).

Spatial relations are of common use in everyday situations and they are highly relevant in Science. Yet, there is no single comprehensive account of their origin and structure. Moreover, these concepts are at the very core of our discipline. No matter what particular perspective, orientation or field we adopt in our work as geographers, they are essential part of any geographic description, explanation or prediction task, for spatial interactions are the only way to understand geographic events of any kind: facts, phenomena, process, etc. However, as in the case of other sciences, no standard or unique perspective on their origin or structure has been ever given. Not to mention a minimum number of them. Even with the advent of GIS technology, no serious effort was made to fill this gap, in spite that these systems are no more than toolboxes for calculating, estimating, measuring and representing spatial relations.

A general theory of spatial relations is missing in the structure of science. Since theory is the basis of methodology, this implies that such a theory could contribute to improve spatial methodologies that are based on spatial interaction. In this last implication I found the motivation for inquiring into the possibility of building a general theory of spatial relations that could serve as the foundation of a geographic methodology (I would say a spatial methodology). My specific contribution, described in this work, consist in the systematization of these concepts so that they can be understood to use them properly.

This dissertation has been structured in two parts. The first illustrates in six chapters the theory on which my concept of spatial relations is built, providing definitions and descriptions of each category of spatial relations. The second includes two seemingly unrelated chapters, presented with the purpose of demonstrating the usefulness and usage of the theoretical concepts, given in the first part, in two different types of geographic problems: vulnerability and land suitability for agriculture.

This work has been written with the intention of being primary beneficial to practitioners of geography, however, I could only wish that this benefit would extent to practitioners of any spatially-aware discipline.

CHAPTER 1

THE NEED FOR A COMMON APPROACH TO SPATIAL PROBLEM SOLVING

1.1 ON GEOGRAPHY

Perhaps no other perception is as relevant in general as that of space (and implicitly, of time). The existence of animals, including us, and of plants, depends strongly on their perception of what we call space. Perceptions, however, do change with the perceiver, and in the case of human beings, both as individuals and as groups, these perceptions change often. So, perception seems to be a weak basis to construct a more stable concept of space, at least one useful to science. Moreover, we need a unified concept of space.

The integration of time and space

Space, or more properly space-time has been a subject of study since humans are capable of reasoning. I will not give here an account on how ideas of this concept have evolved (see Chapter 2). Ever since Einstein unified previous efforts to show that space and time were different views of the same thing, it is now recognized that, at least for physical space, any approach to spatial problem solving requires paying attention to both components. Interestingly, the word "space" derives from Latin *spatium*, which besides the meaning of "distance" it also signifies "stretch of time" (did Romans know about relativity?). I would like here to recuperate this integral meaning, so that whenever the word "space" is used in this work, the duality of the concept should be recalled.

As usually, physics has had the lead in understanding the concept and its unity, while geography and other earth and social sciences have been rather slow in integrating the study of time and space, mainly because we still stick to the old idea that time is only a framework to locate and measure change in space. But we must recall that everything is movement and time and space are the measures of movement, therefore, to understand space, time should be considered more often and explicitly.

This delay in integrating time and space in spatial problem-solving can be exemplified in the fact that most spatial analysis techniques do not take time into consideration. Some do it implicitly, such as when soil erosion rates are estimated using the USLE, or when we need to find a shortest path in a street network, but fewer do it explicitly, as in climate / whether forecasting, or land use / cover change analysis.

But the lack of integration pertains not only to the analysis of information, it is also present in information itself. The problem is twofold: on one side, having spatial information for different periods of time means larger data sets and the design of data structures suitable to store and retrieve them; on the other side, there are very few implementations of temporal data processing models in information systems capable of properly handling spatiotemporal data. One could say that although the study of time as an integrated component of space has experienced some delay, the study of geographic space alone has reached new heights, especially when considering the enormous advantage that having sophisticated information systems represent to handle geographic data. But in spite of the advances, it is necessary to be aware that such systems cannot solve spatiotemporal problems without a previously defined conceptual solution.

The fragmentation of Geography

This denial or ignorance of the unity of time-space, in its turn, has lead to a long-lasting identity crisis in geography, which in my view reveals itself in a lack of a unified geographic problem-solving approach. If we are unable to see the unity of space-time we are also unable to see the unity of everything, and hence the falsity of the idea of different spaces.

Along the last century, the idea of the different spaces, has led to the fragmentation of what we stubbornly call geography, here broadly defined as the study of the origin, evolution and outcome of human and environmental spatial interactions. Facing a lack of a common spatial approach, first the physical and human viewpoints of geography broke apart. Then, within each of these branches new disciplines emerged: geomorphology, climatology, industrial geography, urban geography, radical geography, behavioral geography, etc., each one having their own approaches to geographic problem solving. This disintegration was reinforced in many cases by borrowing from other disciplines such as geology, ecology, physics, economy, anthropology or sociology, where space is seen also from fragmented perspectives.

With a few exceptions, methodological approaches in geography lack a common theoretical ground, in spite of the common use of fundamental disciplinary concepts such as place or scale. This way of approaching the problem is typical of a "geography-is-what-geographers-do" vision of the discipline. A main reason for this approach to geographic problem-solving lies partly in what de Blij said almost twenty years ago: "only a small fraction (of geographers) are much concerned over the roots and lineages of their discipline, or where their work fits in its greater design" (de Blij, 1987). He was talking about Ph. D. students, but, as is evident in geographic literature, the same can be said of many professional geographers.

Specialization within a discipline is not something that one should worry about, provided that the specializations preserve the focus of the main discipline. What happened to specializations in geography is that they lost the focus on the spatiality of geographic phenomena to place it on their non-spatial features. Thus, for instance, the study of migration in geography is often more centered on its social and economic causes and less on the spatial interactions of social and economic processes as causes of migration. Similarly, the study of risk in geography is centered either on the physical characteristics of natural or anthropogenic hazards or on the social or economic vulnerability of human groups, but less on the spatial interactions within and between hazards and vulnerability.

The diversity of existing approaches can undoubtedly be seen as richness of concepts, but on the other hand, some of them reflect a fragmented thought that, even though, not totally unwelcome (and at times necessary), adds to the ambiguity of geographic contributions to science. Again, we find support to this assertion in de Blij (1987) words: “such diversity illustrates both the strengths and weaknesses of professional geography... The strength lies in the versatility and adaptability of the discipline's practitioners ... But the weakness is one of coherence, of commonality -- and ultimately, of identity and image”. These words were said two decades ago but unfortunately remain as current as then.

Although many geographers would not agree, I share the view that geography defines itself by its approach (a geographic approach) not by its subject (USGS, 2005a). I have to agree with this assertion, because otherwise I could not explain the variety of subjects that have been the center of my scientific inquiries during the last thirty years. I would add that such a geographic approach should follow the general steps of the scientific method: observation, description, explanation and prediction. Starting by finding and describing the ways phenomena occur on geographic space, i.e. their patterns, the approach then, in order to explain the found/observed spatial patterns, would proceed to establish the spatial causes and mechanisms of such patterns. The next step of the approach would be a procedure to forecast or predict the phenomena’s future spatial outcome, given past and current states. Finally, if decisions need to be taken, an attempt to design spatial patterns satisfying desirable, acceptable, or legal conditions should be the approach’s last step. This rarely happens in practice.

As said above, this departure from spatial focus in geography has much of its origin on the absence of a unified theory of space-time, which has us lead to think that different spaces exist (physical, social, economic, etc.). The most severe consequence has been the lack of a common methodology for spatial problem-solving. Imagine that you could approach the study of migration and the study of erosion using the same basic methodology to understand the spatial behavior of these phenomena. You could argue (probably smiling or frowning) that patterns in both cases result from very different processes (socioeconomic and physical respectively), and therefore such an approach is infeasible. But then, it is necessary to recall that in both cases you are, as a geographer, studying the *spatial* behavior of the phenomena, not the physical and the social or economic ones alone, although knowledge of them is required.

No wonder geography has been disappearing as an integrated discipline. It is my conviction that to reverse the situation (if that thing is yet possible) we should look at the unification of the methodological basis of the geographic sub-disciplines, that is, to pay attention to the approach. For that we need to understand the unity of space-time and see the unreality of different spaces. To have an idea of how to accomplish such a feat let us first briefly recall the conventional, fragmented, ways of approaching spatial problems.

1.2 SPATIAL ANALYSIS

Spatial analysis embodies all existing methods and techniques in use by all disciplines requiring analysis of spatial data to solve spatial problems. At first sight it could serve as the basis of the long-sought unified approach, but in its current state it is more a name to refer to a 'bunch' (or at best to a set) of concepts and procedures rather than to a well-structured field of knowledge.

Over time, any new technique or methodology that used, or could be applied on, spatial data (not only geographic) became automatically part of the field, including mathematical, logical, statistical, cartographical or computational procedures mostly.

Spatial analysis procedures are more or less organized in subfields. Thus, we have multivariate analysis, interpolation techniques, terrain analysis, map algebra, land use / land cover change analysis, spatial regression and autocorrelation, network analysis, geostatistics, fractal analysis, etc., just to name a few of these subfields. The common characteristic is that all of them work with spatial data, although in some cases their outcomes are not clearly spatial.

The limits of Spatial Analysis

Spatial analysis procedures have been used traditionally to solve some kinds of spatial problems. While the success in doing so cannot be denied, it is also undeniable that most of them work well for very specific problems, but not for complex spatial problems. Most spatial analysis procedures are data oriented not problem-solving oriented. The spatiality in their application is reduced to the simple association (and manipulation) of some type of coordinates to observed or estimated data. Spatial analysis techniques are limited as tools for spatial problem solving because few of them go beyond coordinates as the basis for their spatiality, or take spatial relations into account. Geostatistics is one exception, but it considers only basic forms of proximity and orientation relations.

My argument is that the inability of most spatial analysis techniques to analyze spatial interactions by means of spatial relations occurring between spatial events, result in their failure to cope with complex problems and in the absence of integral solutions. This deficiency cannot be compensated only by linking isolated measures of properties of spatial events to sets of coordinates.

I maintain here that the inclusion of true spatiality in spatial analysis, by taking into account spatial relations and not only coordinates, would improve our ability to solve spatial problems. Moreover, this consideration could help to face the challenge of devising a common basic methodology for geographic sub-disciplines, given that spatial relations are common to all types of spatial problems. It is evident that in the study of migration or erosion processes it is equally important to consider proximity, coincidence, and other spatial relations between events relevant to each type of phenomena.

However, accounting for spatial relations in spatial analysis procedures is a necessary, but not a sufficient condition, to provide the solution for a common spatial problem solving methodology. There are, currently, spatial analysis procedures that can measure, calculate or estimate different spatial relations, but they are still used in isolation. What is needed is a general framework to study the spatiality of geographic phenomena.

What we need then is a methodology (not a method) comprising the spatial concepts relevant to *any* spatial problem-solving situation. Such methodology should incorporate the spatial analysis techniques or procedures to measure, calculate or estimate the values of spatial relations, but also, it should indicate the way events interact (relate) to form spatial processes that lead to existing or future spatial patterns. From this common departing point you could add any other, non-spatial, methodological consideration depending on the type of phenomena you are studying.

Given the complexity of the required common methodological framework, it is desirable to include in it some tools that can help us to handle that complexity, especially in the operationalization of the approach.

1.3 GIS AND GISCIENCE

GIS

Geographic Information Systems (GIS) developed as a useful technology to handle geographic data. In addition to its spatial data handling capabilities, a GIS provides tools for spatial analysis and procedures to generate new geographic information, usually in the form of maps and tables, which might tell us something relevant about the spatiality of our subject of study.

Powerful and sophisticated as they are nowadays, these systems are still toolboxes: if you do not know how and which tools to apply, they cannot provide by themselves the solution to a particular spatial problem. Even worse, they could give you a wrong or false solution, misleading your interpretation of results. The key to the correct use of the technology lies in knowing the relevant pieces of geographic knowledge you should look for, and not only in knowing how to use a particular GIS operation or function.

GIScience

Geographic Information Science (GIScience) emerged as a relatively recent field to provide help in that direction. This concept implies a shift from merely developing a technology (GIS) to help us to handle geographic data, into developing a framework to help us to understand and organize spatial information as spatial concepts, so that we can make a better use of geographic technologies. The term was first used in 1992 (Goodchild, 1992), and has a somewhat similar meaning to those of Geomatics or Geoinformatics, although these have a more technological orientation than GIScience, and are much more limited in scope.

Although there is a well-established research agenda for GIScience (www.ucgis.org), there are many issues which have not yet received all the attention they deserve. One of this is related to the development of a general methodology for solving spatial problems (or the spatial component of a problem) using a GIS. This issue is closely connected to the one of developing a common methodology for solving geographic problems (without a GIS). In fact, the GIS problem-solving approach can be looked upon as a specific case of the geographic problem-solving approach.

In my opinion, the major contribution of GIScience so far has been the raising of awareness on the importance of understanding spatial concepts, as used in many disciplines and not only in geography. In this sense, GIScience is the so much needed first step into a Space-Time Science. However, its ties to technology have hindered its advance in that direction.

1.4 A SPATIAL PROBLEM-SOLVING FRAMEWORK FOR GEOGRAPHY

The framework I here introduce was born out of the disappointment with the conventional, and multiple (fragmented), problem-solving approaches in geography and from three decades of practice as geographer. In this framework, the concept of spatial relations has a central place because, in my view, they provide the bridge between spatial information and spatial knowledge.

My findings can be synthesized in a general schema (Figure 1.1) illustrating how to proceed from a value to the solution of a spatial problem. The framework has been used already to approach the study of complex geographic phenomena, as shown in the last two chapters. Also, I have introduced it to many geography students, as a guide to organize their work with GIS, having received an enthusiastic welcome and confirmed its thought-structuring power, although it was not conceived as a GIS problem-solving approach but as a geographic problem solving framework instead.

At the bottom of the diagram in figure 1.1 there are five concepts. These indicate the possibilities for spatial problem-solving. Any spatial problem you may think of can be classified as belonging to one of these possibilities, or to a combination of them. Spatial patterns (geographic distributions) are, according to this framework, the main subjects of any spatial problem-solving situation. What the diagram tells you in general is that whether you want to find, describe, explain, predict or design a pattern, you must consider all knowledge elements shown in the central portion of the diagram, using the tools provided by different geospatial technologies and applying the relevant spatial analysis techniques. Let me explain how these elements are important in any spatial problem solving situation.

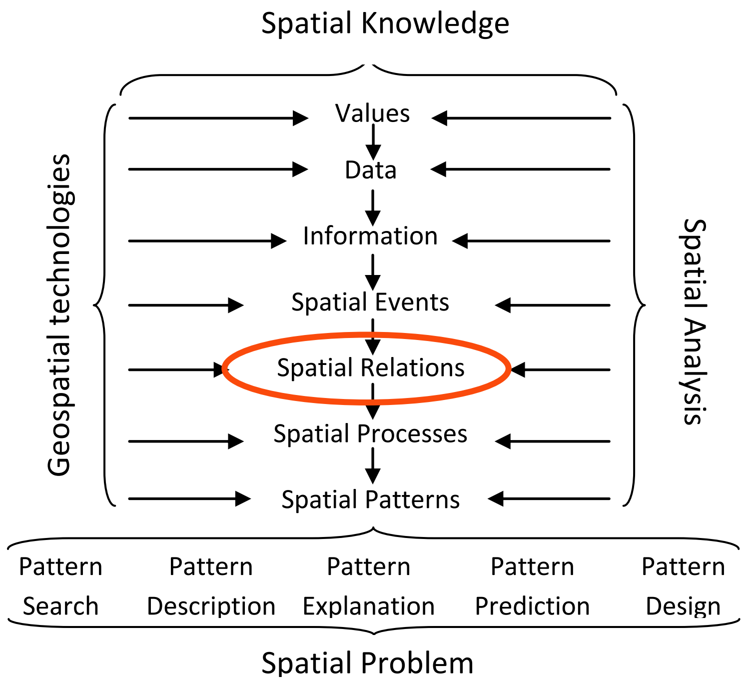


Figure 1.1 Main concepts in spatial problem-solving. The concept of spatial relations, enabling us to go from the simple (values) to the complex (patterns), is encircled in red to highlight the research topic of this dissertation.

From values to patterns

It all starts with values. A value might be a number, or a word, even a symbol representing a class of values. Values implicitly convey knowledge about magnitudes, intensities, or pertinence. For instance, you may have values such as “A”, “low”, or “10”. They tell you that much.

When you add a reference to the units of measurement in which these values are expressed, then you have data. This is more explicit than values. For instance, “10 degrees”, “low risk”, or “class A” tell you a bit more than the simple values.

Still, “10 degrees” may refer to slope or temperature; “low risk” may refer to low risk of fire or low risk of flooding; “class A” may refer to soils or vegetation cover. It is only when you add some thematic meaning to data that you start having information. Thus, for instance, “10 degrees of slope” goes beyond data to become information.

But to complete the picture you must link this information to a spatial event. For example, expressions such as “10 degrees of slope at kilometer 56 of Federal Highway 101”or “low risk of fire in the eastern slope of Mount Rainier during the dry season", tell you about the particular places (and times) which those values belong to.

To proceed in our search for a solution of a spatial problem according to the proposed framework, we now must look into the interactions existing between spatial events, i.e. the spatial relations. As this is the main research topic of this work, let me say for now that these concepts describe the spatial interactions taking place among events. Indeed, the understanding of these concepts is the key to the solution of most spatial problems. It is not by chance that this concept is at the very center of the diagram (figure 1.1). It is the 'bridge' concept in this structure, both between knowledge and problems and between tools and techniques.

But one important characteristic of spatial events, and that is why they are called that way, is that they exist in space as well as in time. Hence, their interactions (spatial relations) change with time too, and this change in the interactions between events is called a process henceforth. Spatial processes, then, can be conceptualized as sequences of spatial relations occurring in between spatial events, causing a change of state in the events holding the relations and a change of values in the relations.

In the end, the outcome of any spatial process is a pattern, that is, the spatial distribution of events. Without pretending that patterns are all that geography is about, I can reasonably argue that much of geographic research is committed to some activity related to patterns. In this context, the goals of spatial problem-solving can be considered as any, or a combination, of five types of pattern-oriented activities as shown at the bottom of the diagram and described next.

The goals of spatial problem-solving

In most cases, when solving a spatial problem, the goal is to find some kind of order or disorder in a spatial situation, an arrangement, i.e. a pattern. It is necessary then to apply tools and techniques on values, data, and information to find the ways spatial events relate to each other to form processes that result in a particular pattern. This is usually the first goal in any geographic endeavor. For instance, if the task is to map the risk of flooding in an area, all elements in the central part of the diagram must be used to search and reveal a pattern that, although might not visible, exists nonetheless, such that one can tell if a specific place is in a high, moderate, or low risk zone.

Some other times the need arises to describe a pattern that can be seen on the ground, on an image, or has been already revealed in a map. Then, again all concepts can be used to characterize every aspect of the observed pattern. For example, if a complete description of the distribution of supermarkets in a city is required, then the description should provide measurements of relations between supermarkets and areas where consumers live and work, or in between supermarkets themselves to describe competition, etc. This goal is usually accomplished in conjunction with the first.

A more demanding problem-solving goal is that of explaining a pattern. Here, although all concepts in the diagram may be used, the emphasis has to be put on the understanding (of causes and mechanisms) of the spatial processes whose outcome is the pattern we are trying to explain. This is so because processes imply sequences of interacting events. This knowledge allows us to find how a pattern has reached the observed or current state.

A somewhat related goal is that of predicting a pattern, for if one is already capable of explaining how a pattern occurs, it is easier to predict what could be its specific state at a specific time-place. But it is important to keep in mind that this prediction task not only must be done in time, as in conventional forecasting, but also in space, that is, we should be able to tell how space will change to accommodate a new pattern, or an event or set of events in an existing pattern.

The most complex of all five types of goals depicted at the bottom of figure 1.1 is that of designing a pattern in such a way that its implementation is optimal in some sense. This usually entails the search for some other patterns that might be related to the one we want to design. It also requires the description of the existing related patterns and of the pattern to be designed. Moreover, to be sure that the designed pattern is optimal, it is necessary to explain how it might work once implemented in space-time, and also needs a prediction of how the new pattern will interact in the future with existing ones.

Pattern design is common to planning activities, where different possible scenarios of a specific activity must be created in order to take the best possible decision. For instance, the location of a sanitary landfill in a region requires the insertion of a new single spatial event (the landfill) into a variety of existing patterns to create a new pattern. This must be done in such a way that, some patterns, for instance the pollution pattern in the region, are not changed negatively (increasing pollution). It also may be needed to see if the road pattern has to be changed in order to optimize waste transportation costs, etc. It can become a very complex activity because at some point you must consider patterns as events by themselves, albeit complex events, and look to their interactions as well, for it is very common that an optimal pattern for a specific set of events disrupts the optimality of another one. For example, an alteration in the pattern of roads, in order to minimize transportation costs to the new landfill, may result in an increase of traffic and pollution or in land use change, in such a way that these modified patterns might not be optimal with respect to the current pattern of population distribution, nor to the pattern of water supply facilities, or to the pattern of quality of life, and so forth.

Besides knowledge of central elements described before, to accomplish any of the goals in solving a spatial problem, we need tools and techniques. Together, spatial analysis techniques and geospatial technologies form a powerful instrument to analyze and synthesize spatial knowledge.

The problem-solving framework here succinctly described can be used to guide the study of any geographic phenomena. To complete the framework depicted in the diagram, one could place the name of any geographic phenomena at the very bottom, to serve as the name of the spatial problem to study. Although more work is needed in order to fully develop it, its structure can serve as the basis for the long-sought common geographic approach.

The keystone of the framework, and hence of the proposed approach, is the concept of spatial relations. Given their importance to the creation of a unified methodology in geography, the next chapters are devoted to the study of their nature, structure and use in spatial problem-solving.

CHAPTER 2

THE THEORY OF SPACE-EVENT INTERACTION

2.1 A BRIEF HISTORY OF THE CONCEPT OF SPACE

Understanding spatial concepts entails first the understanding of the concept of space. The concept has undergone several changes throughout history. There are excellent reviews on the subject (Grunbaum, 1963; Swinburne, 1968; Lefebvre, 1991; Jammer, 1993), so here I will only refer to those concepts that represented a revolution in spatial thought, and from them, advance the one used in this dissertation.

The perception of space

Perhaps the first concept of space in human history was that of place. It was more a percept (the object of perception) than a concept (the abstraction behind a class of percepts), but however primitive it represented an integrated vision: the place was perceived in a space-time context, it contained different events, and spatial relations between them were implicitly accounted for. No abstraction of the place was needed because every place was like a 1:1 scale map.

As people moved to remote places, it was necessary to keep an account of the different places along a route or in a region, and of what each one afforded (shelter, food, water, dangers), this probably took the form of mental maps of the most relevant characteristics of each place. It was also necessary to have a mental record of some of the most fundamental relations between and within places, basically proximity, orientation and exposure (in the sense of visual exposure) relations.

The concept of place was then an unbounded one, or at most its boundaries were very fuzzy and ambiguous. When the idea of property was developed, it became necessary to set limits and frontiers, and then abstractions of geographic space enclosed by those limits had to be represented, thus giving birth to physical maps.

The representation of space

From the first maps we can see that the concept of space was one based on relations between the most representative events within the territory. From those relations the ones more useful were adjacency, connectivity and inclusion (or containment), whereas orientation and proximity relations were expressed in a very approximate way. Such concept of space was more topologic than metric.

Civilization implanted more firmly the idea of property. Wars and trade made necessary the use of a more formal concept of space. Geometry was born. The first philosophical questions that we know of about space, matter, and time were posed: Had space an independent existence from matter? Was matter what gave existence to space? Was time a framework to measure movement of matter in space? Was space equivalent to the void?

The conception of space

Aristotle’s answers to those questions formed a concept of space that dominated for centuries. In his view, space was conceived as the sum total of all locations occupied by bodies, hence negating the possibility of an independent existence of space from matter. Indeed, space was considered a quality of matter.

This view prevailed till the sixteenth and seventeenth centuries when European natural philosophers (Descartes among them) postulated a new concept of space in which space (and time) existed independently of matter, and, unlike it, had a homogeneous structure. Moreover, they advanced that the study of space should come before the study of matter because in that way it was possible to understand natural phenomena. Although this notion was challenged by Galileo around 1632, who introduced a primitive concept of relative space, his idea did not have the acceptance it deserved, probably because, a year before, the Inquisition had undermined Galileo’s credibility.

These ideas prepared the road to one of the most influential concepts of space in modern science, that of absolute space, where space, immovable and homogeneous, had a real existence (although imperceptible) which acted on bodies but could not be acted upon. This concept was derived from Newton’s system of mechanics. Newton elaborated this concept as a logical and ontological necessity to support the laws of motion he developed. As a complement to his concept of absolute space he admitted the existence of relative spaces, being the former the ‘true’ space and the later nothing but measures of the former.

Due to the success of Newton’s system of mechanics, his view of space prevailed over that developed simultaneously by Leibniz. Leibniz’s theory of space postulated that space was but a system of relations in between bodies. This idea, however interesting (I would say, outstanding!), did not attract attention at that time. Even Kant, who at first supported Leibniz’s ideas about relative space, later abandoned them in favor of Newton’s absolute space and time (and we should recall the influence that the Kantian view had on the development of the concept of geographic space).

The relativization of the concept of space

The ideas of absolute space and absolute time were to be challenged during the nineteenth century, due principally to the fact that they had no practical use in physics. A renowned physicist of the time, James Clerk Maxwell is quoted (Jammer, 1993) as having remarked: “We cannot describe the time of an event except by reference to some other event, or the place of a body except by reference to some other body. All our knowledge, both of time and space is essentially relative”. By the end of that century the idea of absolute space was declared contrary to scientific reasoning. Relative space became the accepted concept.

Through all these changes in the concept of space only one characteristic had survived: the structure of space was Euclidean. That was going to change soon also with the development of non-Euclidean geometry. This started with the independent work of mathematicians Lobachevski and Gauss in the nineteenth century, but it was until another mathematician, Riemann, who mathematically proved the existence of ‘physical curved space’, that the conception of the Euclidean geometry as the basic structure of physical space began to disappear. Non-Euclidean geometries also had the effect of reconcile mathematical and physical spaces, which until that moment have been considered two different concepts. I must emphasize this last achievement because it was the first step in the unification of the ideas of space. It is also my view that there is no fundamental difference between any type of space, no matter how concrete or abstract they might be. As we shall see, spatial relations do exist in any space, which also supports the idea of a unified concept of space.

The discovery on non-Euclidean geometries opened the door to the concept of Relativity, in which physical space was a continuum of space-time whose structure was curved by the distribution of matter, specifically through gravitational forces. In the relativist conception, the dimensionality of space also changed from the Newtonian three-dimensional absolute space to a four-dimensional space – time continuum. Hermann Minkowski, in a now famous paper in physics wrote: “…space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality “ (Minkowski, 1909).

A couple of ‘final’ ideas about space

In spite of everything said about the current apparent validity of the concept of relative space over that of absolute space, we should not be too fast in discarding the later conception. After all, it seems to me that the discussion of the concept has been experiencing a malady also affecting some other concepts in science from time to time: that of going from one opposite to the other.

A reason to give absolute space a second thought is that there are other related concepts whose existence is also not entirely proved but that nevertheless are somehow accepted in science and have a strong link to the idea of the absolute. In particular, one of these is the concept of infinity, briefly defined as something with no beginning and no end. The idea of an infinite universe (and therefore, of infinite space-time) has been one of the most intriguing and unresolved questions in astronomy, however, the existence of infinite space is absolutely certain for some spaces. Take for instance the infinity of the surface of a sphere in a 2D space: unless we arbitrarily choose an origin, we cannot say that there is a beginning and an end in such a surface, although it has limits. Hence, we accept the current notion that the universe is infinite but bounded. Until we establish an arbitrary origin, the space defined by that surface, besides being infinite, can be considered as absolute, that is, its existence does not depend on anything existing on it. But the fact that we can actually relativize the same space by choosing an origin makes one wonder if both, the relative and the absolute, can coexist.

It is my view that, in order to advance knowledge in science, we should avoid being trapped in one of the two opposites of a dichotomy. Although dichotomies are useful as analytic tools, truth is not to be found in one extreme or the other, but in both simultaneously.

2.2 SPACE AND EVENTS

Physical space (spacetime, herein after, except when explicitly told otherwise), as we can conclude from the previous section, has an existence equal to that of matter, but unlike the Newtonian concept, it can be acted upon by matter, being this effect more apparent in the vicinity of concentrated matter. Its structure (its ‘geometry’) can be described as apparently curved by forces derived from the concentration of matter, but since a great concentration of matter is required for its curvature to be noticed, it is reasonable (but not quite right) to assume that, for all practical purposes, spaces such as the physical part of geographic space (where concentration of matter is relatively low compared to astronomic spaces), pose a structure that can be approximated very well by Euclidean geometry. Surprisingly, the conceptual or abstract side of geographical space (the 'social' spaces) is often curved, being that the reason why Euclidean geometry is badly suited to describe and analyze this space. Likewise gravity for physical space, organization for a conceptual or abstract space has the same effect: it distorts the 'original' flat fabric of space.

On the other hand, any object, subject, thing, entity, concept, percept, etc., existing in any space can be called an event. Why an event? Event is a better name since it reminds us that they, and their properties, also have a temporal quality. According to physicists, 'an event is a stretch of time, a continuous variable' (Pinker, 2008). Indeed, this definition reminds us of the *spatium* of the Romans!

Properties of space

Setting aside the absolute vs. relative dichotomy, I establish, for the purpose of this dissertation, that space is not a quality of matter as in the Aristotelian view but its perception or detection is determined by matter, as in Einstein’s view. However, unlike Einstein's view, where space is curved due to the proximity to matter, my view is that space *is* matter and concepts, with matter being the concentrated physical manifestation of concepts, and where space is always curved without regards to proximity to matter or concepts, with curvature detection depending on the concentration of matter and concepts and on the capacity of our physical and intellectual instruments.

This is the idea of space that I want to promote here in order to give sustenance to what I intend to say about spatial relations. Thus, in some ways, but only apparently, space is a product of matter, and it also can be seen as a product of concepts or ideas, as advocated by the social constructivism theory (Lefevbre, 1991), but functionally there is no real distinction between both views. That is the reason why the same number and types of spatial relations exist in any space, even though their expressions and representations may appear quite different.

The structure of this concept of space can be detected only through the field created by forces deriving either from the concentration / dispersion of matter, or from the organization / disorganization of concepts and ideas. Therefore, I define *spatial structure as the material / conceptual field by which we detect space*. This field has some properties, even though, strictly speaking, they are only affordances. In the same way that, for instance, a chair affords sitting, any type of space affords the measurement of distances, affords the measurement of directions, and also affords the measurement of different levels of concentration or of organization.

To be precise, these are not affordances of space itself, but affordances of the field by which we detect space. For the rest of this dissertation, the following two-component affordances are taken as the three fundamental properties of the space-time continuum:

- Distance, with a spatial component expressed as extension and a temporal component expressed as duration. Extension is the spatial stretch of an event; duration is the temporal stretch of an event.
- Direction, with a spatial component expressed as path and a temporal component expressed as age (past, present, future). Path is the spatial course of an event; age is the temporal course of an event.
- Concentration, with density as the spatial component and periodicity as the temporal component. Density is the spatial frequency of an event; periodicity is the temporal frequency of an event.

These properties define the structure of space in such a way that everything that exists in it experiences the effect of this structure. In other words, space structures events. Specifically, spatial events can either have a physical structure given by matter (matter-energy) or can have an abstract structure determined by concepts (abstractions–ideas). This dual structure suggests that there exist material and conceptual spatial events although in reality all events have this dual structure. For instance, a nation is a concept whose abstract structure is determined by other concepts (ethnicity, culture, economy, etc.), but it is also a physically structured portion of land (territory, vegetation, water bodies, etc.).

Properties of events

Spatial events have properties (affordances) too, but the quantity and the qualities of these may vary depending on the amount and organization of matter and concepts. Here we need to distinguish between instantiated and fundamental events. Fundamental events are classes or categories of events, instantiated events are particular cases of fundamental events: the Nile River is an instance of the category 'river'.

Some of the properties of instantiated events are explicitly spatial, like size and shape, while others may be explicitly temporal, such as age or rate of change from a certain state into another. Most properties of instantiated events are apparently neither spatial nor temporal, but physical or chemical such as sound, color, odor, flavor, weight. However, these are produced by spatiotemporal interactions of small events that create bigger events, from which it can be concluded that all physical and chemical properties of instantiated events are in fact spatiotemporal properties too. A last set of properties may be called nominal or abstract, like name, function, use, and others, they are also spatiotemporal.

In contrast, fundamental or generic events do not have such properties; instead their properties are those that arise from organization of matter or concepts. Most importantly, these fundamental properties can be inherited by particular instances of fundamental events.

Instantiated events, through their inherited fundamental properties, can have influence on the structure of space in two ways: increasing the amount and/or increasing the organization of matter and/or concepts. We know that in geographic space there are physical limits for the concentration of matter (because of gravity constraints), and because of these limits there cannot be any concentration that changes (curves) the physical structure of geographic space in a significant way. However, there exist various levels of organization, of matter or of concepts, which substantially change spatial structure.

Organization is the basis of the fundamental properties of events. I can now define *spatial organization as a genetic process that relates events, and results in the creation of high-order events, the so-called organizations*. Organization, as a process, is an affordance, not a property of fundamental events: they can organize by themselves or can be organized by other events.

Also, I can now formally define a *spatial event as any organized assemblage of matter and concepts in space*, which can be differentiated from other assemblages, including the parts of an event, which are also events themselves, and the parts of the parts of events, and so forth, up and down the scale of the space-time. As we are mainly concerned here with geographic space, we will be dealing chiefly with those properties of spatial events which may have an impact on the spatial organization of geographic space, namely:

- Capacity for connecting (capacity for creating links and flows).
- Capacity for combining (capacity for producing aggregations).
- Capacity for grouping (capacity for forming groups).

Interestingly, and unlike the properties of space, these properties create three fundamental types of high-order spatial events: connections, aggregations, and associations. These types of events are the basic components of any simple or complex spatial organization.

Simple spatial organizations are based on two low-order fundamental types of events: links and flows. These are generic events that only come into existence when organization starts to develop, in matter or concepts, thus forming the emergent high-order spatial events known as connections. More on connections is said in Chapter 6.

From that parting point, a more advanced type of organization is built using connections to create an emergent high-order event that may take either the form of a network or a hierarchy of events, or a combination of both: a spatial aggregation. Here, the qualities and spatial interactions of many events are combined, by means of connections, to produce an assemblage (the aggregation) that has its own high-order properties, among them the capacities of feedback and auto-regulation (self-organization). From a functional point of view, the key feature of this type of organizations is the strong interdependence among the low-order events that form them: its interdependence is either mechanistic (as in most of natural aggregations) or stochastic (as in most of social aggregations). More on spatial aggregations is discussed in Chapter 6.

A different type of organization results when events *decide* to get together or the events are grouped by the decision of an external agent. The fact that a decision to group or associate is required in order to create a higher-order event implies the existence of a level of intelligence only found in some animals and humans, which does not necessarily means that associations are the perfect form of organization (actually, I have seen some associations that left one wondering whether there was any intelligence at all!). More on spatial associations is told in Chapter 6.

2.3 THE THEORY OF SPACE-EVENT INTERACTION (TSEI)

Having defined my own concepts of space and events I am now in the position of building a theory that links both, and whose main outcome is the concept of spatial relations.

Rationale

The late Reginald Golledge, in his Presidential Address to the AAG (Golledge, 2002), said that geographic knowledge evolved from the phenomenal to the intellectual, being defined the former as the knowledge *of* the space and the later as the knowledge *about* space (Eliot, 2000). He also said that knowledge of this last type is concerned with the recognition and elaboration of the *relations* (the italics are mine) among geographic primitives and the advanced concepts derived from these primitives, adding that the development of such knowledge has been hindered mainly because of the following reasons (Golledge, 2002):

- Geography has not developed a widely accepted vocabulary, we have dictionaries and lists of terms but these are merely collections of concepts, they are not organized as low and high order concepts.
- The language of geography is ill-structured, under-researched, and poorly taught and learned.

- Biases occur in the assimilation of the fundamental concepts that are the essence of geographic knowledge, mainly because of improper thinking and reasoning, perceptual differences among geographers, and the commission of errors in the processes of encoding, manipulation, and decoding of knowledge.

To these problems I could add the lack of a formal general theory of space (including geographic space). Such a theory could serve to organize geographic concepts in categories according to the level of cognition to which they belong, thus helping us to handle complexity, avoiding confusion and facilitating communication of knowledge; in Golledge's (2002) words: "Without such a base, our knowledge structure is speculative and hard to justify and defend". It is because of that need that I present here a theory that may contribute to the partial fulfillment of this long-pursued goal.

The TSEI is developed in order to serve as the foundation for a unified theory of geographic space and, in particular, to give formal support to the notion of spatial relations. In this last context, the theory has a double purpose:

- To help in the definition of a minimum set of spatial relations.
- To help in the definition of the nature of spatial relations.

Inquiring about the basic notions involved in the term 'spatial relations' we can observe that two basic notions are already made explicit in this expression: "spatial", which refers to space-time, and "relation", the condition of entities, objects, subjects, things or facts of being related, which refers to an interaction among events. So, what we are looking for are those interactions held by events in space-time.

There have been many attempts to come across a minimal set of spatial relations (cf. Nystuen, 1963; Freeman, 1975; Robinove, 1977; Youngman, 1978; Burton, 1979; Peuquet, 1984a, 1986; Smith & Peuquet, 1985; Peuquet & Zhan, 1987; Abler, 1987; Pullar, 1987, 1988; Feuchtwanger, 1989; Molenaar, 1989; White, 1991; Egenhofer & Herring, 1991, OCG, 1999). The one proposed here could just appear as another one except for the fact that, unlike the others, it has a strong theoretical base. The other efforts have been rather empirical and tied to particular sub-disciplines (topology, GIS, geometry).

Foundations

For the following discussion, we will assume that we can distinguish space and events through their properties (as defined in the last section) and consider them as two different concepts (but keep in mind that they are only two different views of the same concept).

The Theory of Space-Event Interaction (TSEI) is based on two axioms, seven postulates and three theorems:

Axiom 1: there exist a minimum set of basic spatial concepts, the so called spatial relations, which are shared by all scientific disciplines.
Axiom 2: spatial relations derive from the interaction of the properties of space and the properties of events.

Postulate 1. Space and Events are two categories of different fundamental entities.
Postulate 2. Interaction may arise if two or more fundamental entities coexist.
Postulate 3. Fundamental entities have fundamental properties.
Postulate 4. Interaction is achieved through the fundamental properties of the interacting entities, such that each property of an entity interacts with one, and only one, property of another entity.
Postulate 5. In any interaction either there can be a dominant property and a subordinate property, or else some kind of neutrality can be reached between interacting properties.
Postulate 6. Fundamental properties of Space are called structural properties.
Postulate 7. Fundamental properties of Events are called organizational properties.

Theorem 1. By postulates 1 and 2, whenever space and events coexist there is the possibility of interaction.
Theorem 2. By postulates 3, 4 and 5, if space and events interact there are three possible cases in which this interaction can be realized:

a) The interaction is dominated by a property of space, such that this property minimizes the effect of one property of the events.
b) The interaction is not dominated neither by the properties of space nor of those of the events, such that the effects of both types of properties are neutralized.
c) The interaction is dominated by a property of events, such that this property minimizes the effect of one property of space.

Theorem 3. By postulates 6 and 7, the concept of space without the concept of events supports the idea of structure but not that of organization; conversely, the concept of events without the concept of space supports the idea of organization but not that of structure.

Neither the postulates nor the theorems, can be proved to be true (they cannot be proved to be untrue also), but can only be assumed to be true, as far as they rest on the assumptions made in the axioms.

Spatial Relations

Assuming that interactions of space and events result in relations, as proposed by the TSEI, Theorem 2 tells us the number of possible spatial relations and their origin. As I established that both space and events have three properties each, this theorem implicitly states that the maximum possible number of spatial relations is nine, which, on the basis of the dominance in the interaction, can be arranged in three groups of three relations each (Table 2.1).

The first three relations result from the dominance of the three properties of space; the second three relations derive from the equal interaction of the properties of space and events; and the last three reflect the dominance of the three properties of events. The arrangement in the table also reflects the condition established by Postulate 4, in the sense that each property of space has a counterpart in the properties of the events. This arrangement is explained in the next section.

Table 2.1 Spatial relations resulting from the interaction of space and events.

PROPERTIES OF SPACE	**SPATIAL RELATIONS**	PROPERTIES OF EVENTS
DOMINANCE OF THE PROPERTIES OF SPACE		
Distance	PROXIMITY	Connecting Capacity
Direction	ORIENTATION	Combining Capacity
Concentration	EXPOSURE	Grouping Capacity
EQUILIBRIUM OF BOTH SETS OF PROPERTIES		
Distance	ADJACENCY	Connecting Capacity
Direction	CONTAINMENT	Combining Capacity
Concentration	COINCIDENCE	Grouping Capacity
DOMINANCE OF THE PROPERTIES OF EVENTS		
Distance	CONNECTIVITY	Connecting Capacity
Direction	AGGREGATION	Combining Capacity
Concentration	ASSOCIATION	Grouping Capacity

The spatial relations, as depicted in table 2.1, are not specific relations but categories or types of relations, with each type comprising a number of specific relations. Instances of these categories are, for example, *near, close*, both belonging to the category of proximity relations; *within, outside*, both belonging to the category of containment relations, etc.

In agreement with Peuquet (1988), who advocated the use of theory to derive spatial relations, I derive them using the TSEI, not categorizations such as whether they are implicit/explicit in a particular data structure (Peucker & Chrisman, 1974, Peuquet, 1984b; White, 1991), or on the basis of mathematical properties such as transitivity, order, reflexivity, symmetry, etc. (Pullar, 1988; Egenhofer & Herring, 1991), or how they arise from the interaction of geometrical or topological entities (Youngman, 1978; Molenaar, 1989; Gatrell, 1991; White, 1991; de Hoop & van Oosterom, 1992). The common characteristic of these efforts is that they are based on too specific concepts: data structures, geometry, topology, etc. I do not regard such previous conceptualizations as unimportant, but I believe that they are not broad enough to provide the fundamental concepts on which a general theory of space-event interaction could be based, and therefore the spatial relations derived from them are not general enough to apply to all types of spaces (physical or conceptual).

Interactions between properties of space and properties of events

It is important to remark that, as declared in Postulate 4, each space property has its counterpart in the properties of events, as seen in Table 2.1. This is not fortuitous. Let me explain this peculiar arrangement of properties.

The first type of relation in the first group is Proximity, indicating the dominance of the space property of distance. But although this property dominates and defines the nature of the relation, the proximity between two events can only be measured along a physical or virtual line (straight or curved) that connects two or more events, no matter how long, indirect, convoluted, temporal, or artificial the connection may be, therefore implying a minimum participation of the connecting capacity of the events.

The second type of relation in this group, Orientation, is mainly determined by the space property of direction. However, to establish the value of the relation between two events it is necessary to create a physical or a virtual trajectory by combining the geometries of events or locations into a frame of reference (see chapter 4), so that the trajectory's direction can be measured. This then implies a minimum but necessary participation of the combining capacity.

The last relation of the first group is Exposure, which is primarily determined by the amount of events concentrated in a place. But while the simple density or frequency of events can be used to express the degree of exposure, it can only be determined by the way events group together (or are grouped by an external event), forming or not barriers in specific ways, according to physical laws or conceptual criteria. This requires a minimum of capacity to group.

In the second group of relations, the pair of properties formed by the distance property of space and the connecting capacity property of events gives raise to the relation of Adjacency. This relation unites characteristics of both properties: on one side, for adjacency to exist among two events it requires the distance in between them to be zero, that is, that they are in contact with each other; now, if they are in contact, it is very likely that they can connect, that is, that some kind of flow (heat, energy, money, ideas, matter, etc.) may exists in between them.

The same happens to the pair comprised by the property of space called direction and the combining capacity property of events, which give raise to the Containment relation. For this relation to take place it needs an event to be surrounded in all directions by another event, that is, that one contains the other. This also implies that both events combine their shapes so that a third, two-part (or more-part) higher order event, is formed, where one of the events acts as an "island" or "hole" of the other event by combining their boundaries: the internal boundary of one is the external boundary of the other.

The interaction of the last pair of properties in this second group, the concentration of events in space and the grouping capacity of events, creates the relation of Coincidence. For events to be in a coincidence relation it is required that there is a concentration of at least two events, so that they are able to share the same space in some specific number of dimensions, forming at the same time a (concentrated) group of events.

The first relation of the last group is Connectivity. This relation arises primarily from the capacity of events to create connections, but although its effects may be minimized, the property of space known as distance still plays a role in determining some properties of links and their flows, such as the strength of the link or the reach of a flow.

The second relation in this group is Aggregation. This is created mainly by the capacity of low-order events to combine their properties and form a higher-order event. However, because the ways in which such a higher-order event is structured, as a hierarchy, as a network, or as a combination of both, it is essential to define the directions, within the aggregation, in which the effects of the combined properties of the lower-order events propagate, and the directions in which the effects of the aggregation properties are inherited to the lower-order events, that is the directions in which interdependence exists..

The last relation is that of Association. While the grouping capacity of events determines mostly the occurrence of this relation, this necessarily, and evidently, implies as well the existence of a physical or virtual concentration of low-order spatial events that participate of the association, and certain amount of exposure among them or of their properties.

CHAPTER 3

A MINIMUM SET OF SPATIAL RELATIONS

3.1 SPATIAL RELATIONS

The concept of spatial relations

Simply defined, spatial relations are interactions held by events.

Specific spatial patterns arise from the relative locations of two or more events in space, or two or more parts of a single event (events are made of or are part of other events). These patterns are used to investigate actual or potential spatial interactions among events. As derived in the last chapter, the relations in the first group, resulting from the dominance of space, indicate passive patterns, and therefore are used to investigate the *potential* of an interaction. Relations in the last group, arising from the dominance of the events, represent active patterns, and therefore are used to investigate the *actual* characteristics of an interaction. The relations in the intermediate group, of no dominance, may be used to explore both, *potential and actual*, interactions. This means that, for instance, a proximity relation between two events do not necessarily implies that there is an interaction among them, but only that interactions may take place depending on the degree of separation. In the case of a connectivity relation between two events, the interaction necessarily takes place, since otherwise a connection could not be identified.

Spatial relations are considered as something that takes place between events, but spatial relations can also be regarded as spatial events. At this respect Vickers (1981) wrote: “We are accustomed to regard ‘relating’ as something which entities do, rather than as something they are”. This last meaning should also be always taken into consideration. This is especially true of connectivity, aggregation, and association relations, because they create emergent, high-order events, as we shall see in Chapter 6. Spatial relations should be regarded as representations of interactions between events but also as events themselves.

Spatial relations are concepts that have been in use since humans were able to think. Nowadays, people still use these concepts intuitively, as in everyday life, or use a more formal idea of them, as in science. The intuitive notions are more or less linked to the perception of space by an individual or a by a cultural group. In science, these notions are usually related to a conceptualization within the discipline where they are used.

In cognitive map theory (O’Keefe and Nadel, 1976), spatial events are located using spatial relations, where these relations are specified by means of three concepts: places, directions and distances. While place is not a property of either space or events but an event in itself (although sometimes place is used as an attribute of an event to specify their location or position), direction and distance correspond to orientation and proximity relations respectively. O’Keefe and Nadel (1976), argue that animals (including us) use these three elements to create cognitive maps, but it seems to me that this old conception is somewhat limited. In my view not only proximity and orientation relations are used in constructing these types of maps, but also exposure, in the sense of sensorial exposure (visual, auditory, olfactory, tactile, and gustative), since we can locate places, albeit imprecisely, even if the direction or distance cannot be estimated or measured. For instance, when hearing sounds (of a shout, or of a waterfall, for example) in a forest, we cannot sometimes tell how near we are to or in which direction is the source of the sound, however we can sense its presence in our environment simply because its auditory exposure.

I could even argue that cognitive maps can be created using the other six spatial relations as well, since we can use them to relatively locate events in space, in the same way that we use the first three. But beyond their use in locating events in space, I believe that the most useful and advanced function of these concepts is to describe the interactions between spatial events. To fully realize this function it is essential to standardize their meanings.

The meaning of spatial relations

While hardly anything can be done to standardize meanings pertaining to individual or cultural group notions (why should we want to do that in the first place?), there are good reasons in science to accomplish that goal. The most important is to avoid confusion; a complementary reason is to facilitate cross-disciplinary work on spatial problems. Often, a single term denoting a spatial relation is defined differently in different disciplines, presumably because the spaces each one studies are different, which, as we show here, is not true, only subjects of study or approaches might be or are different. Even worse, researchers tend to use these terms attaching arbitrary or *de facto* meanings without bothering with definitions, least to say a formal definition of them.

The diversity of meanings of spatial relations is not only revealed in the variety of terms used but also in the definition of what a spatial relation is. For example, the proximity relations expressed by the pairs *near – close* and *far – remote* might appear to have clear, although undefined, meanings for the common public, but in science they need to be defined in some precise, or at least fuzzy, way, before they can be used. A problem with definitions is that many of them expose a lack of spatial knowledge. In other cases, the concepts are more or less subjectively defined, based mostly on the researcher’s experience and intuition than on a theory of space. Furthermore, the spatial relations used are no minimal sets, but instead some of the relations can be defined as combination of others, or are specific cases of more generic relations.

The need for the implementation of spatial relations as tools in GIS has led to their definition as concepts tied to geometric or topologic representations of spatial situations. But we should remember that representations are just views, and therefore, incomplete. At this respect, Cassirer (1953, quoted in Couclelis & Gale, 1986) said: “To represent a thing is not enough to be able to manipulate it in the right way… We must have a general conception of the object, and regard it from different angles in order to find its *relations* to other objects”. To conceptualize spatial relations only as a set of geometric or topologic concepts is an oversimplification. For this reasons, it is a bad practice to use GIS-derived definitions of spatial relations to denote interactions between real-world spatial events.

Confusion about the meanings of spatial relations could be ended if the essence of the concepts were to be understood. This work can contribute to that aim by showing that, disregarding the space of study, there is only a minimum set of spatial relations, each one with a well differentiated meaning, hence preventing each discipline and each researcher of having to devise their own sets and attach non-standard meanings.

The importance of a minimum set of spatial relations

Marvin Minsky, one of the founders of the knowledge field known as Artificial Intelligence, once said that “the most useful set of properties are those whose members do not interact too much” because then we “can put them together in any combination whatsoever” (Minsky, 1988). This is true of spatial relations. A minimum set must consist of very well differentiated relations because they are to serve as the basis to construct or describe complex spatial relations or spatial situations (Golledge, 2002). The minimum set of spatial relations, as derived from the TSEI (Chapter 2), has the characteristic described by Minsky, with every relation having a very distinctive quality. Only the properties of events and space interact to create the relations, but relations do not interact too strongly themselves. Thus, we can combine basic relations to explain or describe a complex spatial situation or spatial process with a minimum of redundancy.

Complex concepts, such as ‘location’ and ‘position’, can be described using primitives represented by spatial relations. Hence, it is possible, for instance, to describe the location or position of an event in a way such as: “the place you look for is 50 km away (proximity) to the south (orientation), behind that mountain range (exposure), it is located by river X (adjacency), within county Y (containment), can be reached through highway 101 (connectivity), and is part of the city of Z (aggregation)”.

Relevancy of spatial relations

Spatial relations exist independently of scale (though their values might not be scale independent), and can be found in all physical spaces, from the astronomic to the subatomic, but also in abstract or conceptual spaces like mathematical or social spaces. However, not all relations between two or more events need to exist at a given time, nor are all of them relevant to solve a specific problem or to understand a particular spatial situation.

A good example of the importance of determining which spatial relations are relevant to solve a specific spatial problem is given by the situation of navigating through a subway network. Proximity relations exist, and can be measured, among all stations belonging to all subway lines, but those proximity relations measured between stations *along* the subway lines are more relevant than the others. However, in this problem, proximity relations are usually not as relevant as connectivity relations, that is, those relations maintained between two stations in two different lines, which tell us how many and which connections we should go through to travel from one station to another. Orientation might be also relevant, since to go from one station to another is necessary to know which direction (from two) to follow. Containment is important in specific cases where it is necessary to know whether a station is in between another two. Relations such as exposure, adjacency, and aggregation might not be as relevant. Association relations are an exception because an important piece of spatial knowledge is whether stations belong to the same line or route, or how are all stations associated, and which lines they form.

In general, I would say that solving the subway navigation problem consists primarily in paying attention first to topological relations (connectivity, containment, and association) and then use non-topological relations (proximity and orientation) to refine the solution: 'Topology Matters, Metric Refines' (Egenhofer, 1995).

Scale plays also a role in determining the relevancy of spatial relations for the solution of a given problem. In ecology it is frequent to find that relations determining the functions of an ecosystem, and indeed its very existence, vary depending on scale. While on a macro scale one may find that those relationships considered as structural or neutral are of more importance to study (specific coincidence relations between water, vegetation and faunistic groups over a portion of a region or ecosystem must exist for the ecosystem to exist), at a more detailed level, the organizational relations between plants, animals, and their connections with the abiotic resources, are more important to understand the functioning of the ecosystem.

In a conceptual context, spatial relations are relevant as indicators of potential and actual capacity for interacting. Structural relations provide indicators of potential for interactions. Organizational relations provide indicators of actual interactions. Neutral relations, given their dual nature, can serve as indicators of potential or of actual interactions. In other words, *the potential for interaction is determined by the structure of space, while the organization of space determines the actual interactions*. The transformation of a potential into an actual interaction takes place the moment an event becomes part of or member of some kind of organization, by creating connections with other events. This potential-actual-indicator quality of relations is very relevant to understand spatial processes.

Topological qualities of spatial relations

The grouping of relations in three distinct sets as established by the TSEI is also related to their topological qualities, as indicated in table 3.1. Knowledge of the topological qualities of spatial relations is decisive to determine the right form of measurement and the invariance of relations in a given space.

Table 3.1 Grouping of spatial relations according to their topologic qualities.

RELATIONS	TOPOLOGICAL QUALITY
PROXIMITY	**NON-TOPOLOGIC RELATIONS**
ORIENTATION	
EXPOSURE	
ADJACENCY	**RELATIVE-TOPOLOGIC RELATIONS**
CONTAINMENT	
COINCIDENCE	
CONNECTIVITY	**ABSOLUTE-TOPOLOGIC RELATIONS**
AGGREGATION	
ASSOCIATION	

Topology deals with the changes in the values that spatial relations adopt when being subject to spatial transformations such as scaling, rotation, translation, distortion and dimensional shifting. According to the scale of analysis or observation, values of proximity, orientation, exposure, adjacency, containment and coincidence relationships, may change, that is, they are not topologically invariant. As for the role of the nature of events in precise measurement of relations, it is also well known that many events in geographic space have undetermined boundaries (Burrough and Frank, 1996), therefore only absolute topologic relations may be expressed accurately.

In all cases, the topological qualities of spatial relations are related to the topological qualities of the properties of space and events they derive from: properties of space are non-topologic; properties of events are topologic.

Non-topologic relations

Since the properties of space are all of them non-topologic, the relations deriving from the distance, direction and concentration properties also share this quality.

Consider for example the case of proximity relations between two events. If there is a change in the scale of space, the value of a relation may increase or decrease its magnitude. A most striking example of this is that of measuring the distance between two points along a coast line; if we use maps at different scales we will find that this distance seems to be greater the larger the scale of the map is. Also, as an example of how re-dimensioning of space affects proximity consider the distance of two stars as measured on the sky: two stars that appear very close on the plane of the sky may be very far in three-dimensional space; despite what astrology can says, stars in constellations are usually unrelated in 3D space (the space of astrologists is 2D). Rotation of space may also affect measurements of proximity, as when measuring the distance in between the position of events projected on a plane using a non-equidistant projection: the distances measured on the plane are not the true distances between the events. Scaling, rotation and re-dimensioning of space also affect the measurement of orientation and exposure relations.

Relative-topologic relations

The relations in the second group are considered relatively topologic because they are invariant to most transformations of space, with the exception of changes in scale and/or dimensionality, and hence their topologic meaning is relative to the number of dimensions in which they are measured. This duality is partly explained if we recall that these relations participate equally of the qualities of space and events properties.

During at least the past two decades these relations were thought to be pure topologic relations, invariant to scaling, rotation, translation, or re-dimensioning of space. The pioneer work (in GIScience) of Egenhofer and others (Egenhofer & Herring, 1991; Egenhofer & Franzosa, 1991) contributed a lot in this respect. This belief obeyed to three reasons:

- most of the work was done assuming a 2 dimensional space;
- the events subject of study were usually abstract, geometric objects;
- objects of study were considered mostly as having crisp boundaries, which may apply only to some conceptual and physical objects.

Regarding a change of scale of space, to say that two events that touch each other are adjacent, could be a relative truth because from physics we know that no matter how close two physical events might be, there always will be some space (subatomic if you want) between them, and the particles forming them will never be touching or adjacent, in a strict sense. The same happens with the other two relations of this group speaking of physical events. Hence, their meaning is relative to the scale of the space. In contrast, conceptual events may hold true adjacency, true containment or true coincidence relations.

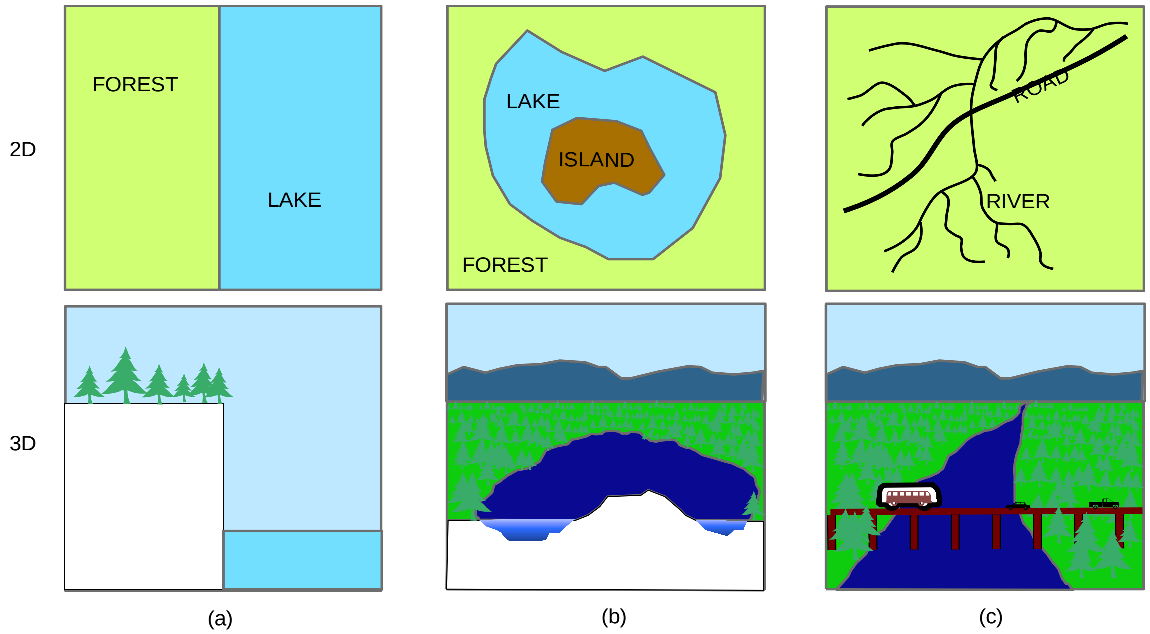


Figure 3.1 Relative topologic relations may not hold when space is shifted from 2D to 3D: (a) adjacency; (b) containment; (c) coincidence. See text for explanation.

In the case of change of dimensionality, if we shift the representation of relations between two events, say from two to three dimensions, the relations may not hold. For instance (figure 3.1), in 2D space, two events that are adjacent (a forest and a lake) may not be so if we consider them in 3D (the forest may be on top of a 90-degree cliff which, in its inferior end, is the one event adjacent to the lake according to figure 3.1a). An island may not be enclosed by a lake in 3D (figure 3.1b) since it is possible to access the island by digging a tunnel from mainland. A road and a river may not intersect in 3D (figure 3.1c), as the map in 2D could mislead us to think.

Absolute-topologic relations

The relations in the last group are true topologic relations. Their values remain the same even if the space where the related events occur is transformed, because the properties of events control how space is structured.

So, for instance, when we say that two roads are connected, it does not matter if we rotate or change the scale of the space where the roads exist, they still will be connected. A change in the number of dimensions or a shift of the space will have the same lack of effect on the relations.

Also, transformations of space do not affect relations of aggregation or association. Think of a mall as an example of spatial association of businesses: it does not matter if you represent the mall in space as a point in a 2D plane, or as a building with certain extent or as a set of buildings in 3D, the fact remains that all businesses in the mall are associated to form it.

3.2 FROM STRUCTURE TO ORGANIZATION

The structuring of events and the organization of space

The TSEI establishes two cases for the interaction of events and space in terms of structure and organization: the space without events has structure but no organization; conversely, the events without space may have organization but no structure (Theorem 3).

Because of those conditions, in the first case, the structure of the field by which we detect space influences the location and qualities of events (Jones, 1984, Mazur & Urbanek, 1983). This is the "scheme of the things in the field" of Körner (1966), but I rather call it 'spatial structure', in agreement with Abler et al (1971). For example, an irrigated crop exists primarily because its proximity to a water source (a channel or a well, in 2D or in 3D). That is, the existence or quality of an event is determined by some relation, with another event, concerning to the structure of space.

In the second case, location and qualities of events can be the outcome of the qualities and interactions of events (Jones, 1984; Mint & Preobrozhensky, 1970), and the original structure of space is modified to reflect the nature of events (Mazur & Urbanek, 1984). This is "scheme of the fields" of Körner (1966), but I rather call it 'spatial organization'. The organization of space has the primary objective of making the space more efficient for the realization of specific interactions (although sometimes the result is the opposite). For instance, a road network transforms the space in such a way that ground transportation activities can be accomplished more efficiently on the space occupied by the network (the organized part of space) than on the rest of the space.

The first three types of relations (Proximity, Orientation and Exposure) are termed structural relations because they result from the dominance of the properties of space, or more properly, from the structure of the field that allows us to detect space. The next three relations (Adjacency, Containment and Coincidence) form the group of neutral relations because there is no dominance of space or event properties. The last three relations (Connectivity, Aggregation and Association) are the organizational relations because they result from the dominance of the properties of the events, which are also properties of organizations.

Actually, there is a gradient of increasing presence of spatial organization starting at its minimum with proximity and ending at its maximum with association. Conversely, there is a gradient of increasing presence of spatial structure with its maximum in proximity relations and its minimum in association relations. This means that, in practice, organization cannot totally override the influence of structure of space, only reduce it to a minimum, and also it means that structure cannot make disappear the influence of organization of events, but only reduce it. Both, structure and organization coexist in any spatial situation. It is perfectly possible to measure or characterize all nine relations (although some of them might not be relevant), whether we are investigating the structure of events or the organization of space. It is all a matter of choosing to pay attention to one or the other, or both, accordingly to the purpose of the study or to the goal of the problem-solving situation.

In our example of the stations and the subway lines, the structure of space makes possible to measure all distances and all directions between all stations, and even see if they are exposed (visually or in other sense), but we know that the network of subway lines and stations *is* an organized space, in which the only relevant part of space is that occupied by the tracks and the stations, therefore we should restrict our measurements to that space. Although, within that space, we can measure all nine relations, more important is to measure only those relations which allow us to understand how to use the subway (how to benefit from spatial organization), i.e. the relevant relations. Only then we can link our findings to measures of relations in other spaces, such as the proximity of the place where we are going to the nearest station on the subway network, etc.

Structure – behavior - organization

Helen Couclelis (1986) asserted that spatial structure governs spatial behavior. While this assertion is true, it is also incomplete. Whereas, initially, the structure of space may determine the behavior of events, behavior tends toward organization, and with organization comes the possibility of modifying structure. I would complete Couclelis' assertion in the following way: structure governs behavior, behavior governs organization and organization governs structure.

To exemplify this cycle, let us consider the process of human settlement over a territory. If we assume that the process starts with a single settlement from which other settlements derive, we can see that the locating of subsequent settlements may be initially dominated by the need of the later of being near to the former. The behavior of settlement process, in its beginning, is thus governed by the structure of space, in particular by the characteristics of the territory that determine: which of its portions are near to the first settlement (proximity); and within this "nearness", in which directions (orientation) the territory is still available, and which portions of the territory have suitable characteristics such as gentle slope, fertile soils, good climate, contiguity to a water body or to the source of raw materials, etc. (i.e. adjacency and coincidence).

In particular, we may find that this behavior is the same for locating subsequent settlements while organization is relatively absent: new settlements tend to be relatively near to former ones and are established were suitable or desirable characteristics of the territory coexist, or settlements develop in the directions where land is still available.

This behavior creates a set of settlements with particular forms of interaction. Let us assume that one of these forms is trade. Trade promotes the construction of a road network (connectivity relations) to make it more efficient, thus beginning an organization (based on links and flows). The organization of trade leads to the creation of markets. Markets, as emergent events, have their own properties, among them the levels of offer and demand for goods, which regulate prices.

Prices control the production of goods, thus leading to the specialization of settlements where the production of specific goods is profitable. This specialization has influence on the type and amount of labor force required to produce and sell goods. The level of job offer of a place has a direct influence on the potential for population growth in a settlement, through migration or births, which has also influence on the level of consumption of goods and therefore on the demand. Thus, we have already a set of dependences (aggregation relations), albeit simplistic, between markets, population and production places, and the emergence of a system of places and a system of trade (markets) where each place performs a certain amount of functions as demanded and regulated by the markets.

Since dependencies may jeopardize the existence of a system, the organization may tend to improve its resilience by purposely diversifying its interactions with other systems, forming groups of spatial organizations (association relations) which decide how to modify the territory (its structure), in theory to make better use of it (more effectively, more efficiently, and more equitably), which in practice depends primarily on the amount of intelligence (in its widest sense) concentrated by the group of spatial organizations. On this modified structure new settlements will be created, starting a new cycle. These new settlements may have interactions with existing settlements without necessarily being near because they are part of or a member of an organization, thus proving that organization governs structure.

3.3 SPATIAL RELATIONS IN NATURAL LANGUAGES

So far, I have been using some terms as names of spatial relations assuming the reader is familiar with their meaning. I will defer for the next chapters the formal definitions of these terms and the discussion on their main characteristics as relations. However, to show that these concepts are not only a product of a theory but also exist in natural language, I next offer the results of a comprehensive (but not exhaustive) analysis of spatial terms in English and Spanish languages, and their linkage with the definition of a minimum set of spatial relations.

Space in natural language

Since space plays a fundamental role in life, we can expect that all possible spatial concepts have some expression in terms of natural language, the so-called 'spatial terms'. All languages have spatial terms, although, from language to language, their form may vary drastically.

Sometimes the spatial terms share common roots, some even may be absent from one language and present on others, but one thing is common: meanings are much the same. Take for instance the English word “*near*” and the Spanish word “*cerca*”; their meanings are the same although their forms (roots) are different.

More important, for the subject matter of this work, is that spatial terms in natural languages function as indicators of spatial relations. The various spatial terms we use in natural language describe different perceived aspects of the nine fundamental categories of spatial relations.

Investigations into the spatial content of natural language have concluded that neither language syntax nor language morphology reflect spatial information (Bierwisch, 1996), instead, it is supposed that meaning carries the spatial content. I disagree with that view. I argue that syntax does carry spatial content as well, as in the following sentences involving the spatial term *end*:

(a) *At the end* (of a situation), *she signed the letter*
(b) *She signed the letter at the end* (of the letter)

Bierwisch (1996) concludes from these types of examples that the use of relations such as *at*, *in*, *enter*, *leave*, and others, are not restricted to space, that there is no clear distinction between spatial and non-spatial terms. For instance, he interprets the first sentence above as having a non-spatial content, but a temporal one. However, his notion corresponds to a conception of space where time is not part of it. As I established in Chapter 2, time and space are components of the same thing (the '*spatium*' of Romans), hence, my interpretation in this context is that both sentences have spatial content and the modification of meaning is related to the type of space in which both events occur (a "situational space" or "social space", and a "letter space", respectively), not to a non-spatial meaning as Bierwisch believes. He gives other examples to demonstrate that a word with spatial significance may have a non-spatial interpretation, for instance the sentence:

he entered the church

may have, according to Bierwisch, a non-spatial interpretation with the meaning '*he became a priest'*, instead of being interpreted as entering a certain type of building, which is clearly a spatial situation. Again, I disagree with his view that this first interpretation is not spatial. According to the TSEI, spatial relations occur in any kind of space, from the concrete to the abstract, therefore, while the later interpretation refers to a physical space (to enter a building), the former implies a social space, specifically it indicates the existence of an association relation in which, the person becomes a member of (an associate of) a religious congregation. Thus, I conclude that a term that has spatial meaning preserves it, disregarding the type of situation it is used in; only the type of space where the situation takes place may change.

Spatial relations in English and Spanish

I have chosen English and Spanish because of their different linguistic origin (and also because these are the only two languages that I know well enough), although there is a bias towards western cultures since both have Indo-European roots. Thus, if we found the same or similar spatial categories in both languages, we can reasonable assume that those are fundamental concepts to humans (at least for those speaking these languages), and therefore deserve a place of their own on any concept of space. That is, they are true spatial primitives. The analysis first identified "all" spatial terms in use in both languages, using two 50,000-entry dictionaries, and then, based on their meanings, assigned them to one of the TSEI's nine categories of spatial relations.

Table 3.2 Number of terms in English and Spanish languages assigned to each spatial relation

SPATIAL RELATIONS	NUMBER OF TERMS IN ENGLISH	NUMBER OF TERMS IN SPANISH
Proximity	128	149
Orientation	145	161
Exposure	44	46
Adjacency	36	40
Containment	47	47
Coincidence	47	57
Connectivity	59	67
Aggregation	19	20
Association	32	32
TOTAL	**557**	**619**

The first part of the analysis resulted in the identification of 521 terms in the English language and 610 terms in the Spanish language, which have a distinct spatial meaning (those that may be categorized as "strictly spatial" according to Bierwisch, 1996). The discrepancy between these figures and the total number of terms, as the sum of the terms assigned to all relations (557 and 619, in Table 3.2), is explained because some terms are used indistinctly in two types of relations. The pattern of the amount of terms per spatial relation is very similar for both languages, indicating, perhaps, that cultures speaking these languages have a very similar concept of space.

Very likely, the number of spatial terms in both languages is slightly bigger, but those identified here are the most frequently used. I must mention that this list does not include derivatives of the terms identified as spatial, for instance, terms such as *narrowly* or *narrowness*, derived from *narrow*, were not counted. In any case, it is a surprising low number considering that the notion of space is a fundamental concept for humans, and the number of words in both languages is high (more than 50,000 each). Hence, one could expect more spatial terms in natural language. Two very likely reasons why the number of spatial terms is relatively small are:

- The number of basic categories of spatial concepts is small, with each category requiring only a few spatial terms to completely differentiate from others.
- It is possible to use a reduced set of spatial terms to describe any possible spatial situation in any possible space.

These two reasons conform very well with the TSEI's notion that there is a minimum set of spatial relations (the basic categories of space) and that this minimum set is common to all scientific disciplines.

Assigning the identified spatial terms to the basic categories of spatial relations, as postulated by the TSEI, was easy in most cases because each relation has a very distinctive nature. Somehow expected, but very interesting to notice, is that Orientation in first place, and Proximity in second (I was expecting it to be the other way around), together concentrate most of the terms (52% of the English terms, 51% of the Spanish terms).

Syntactics and semantics of spatial terms

Table 3.3 Findings related to the grammatical use, meaning and form of spatial terms in English and Spanish languages, in the context of spatial relations.

Referring to the **grammar** of the terms:

- Nouns indicate generic events or parts of events, attributes, and generic processes through which spatial relations take place.
- Adjectives, adverbs, and prepositions indicate specific forms of occurrence of a spatial relation, or the specific qualitative values it may take.
- Verbs indicate different ways to initiate, modify, or terminate a spatial relation.

With regards to their **meaning** :

- Some terms may be assigned to more than one category of spatial relations. This indicates that those categories hold some kind of relationship.
- A few terms like “space” or “spatial” have fundamental meanings and therefore could not be assigned to any category of spatial relations.
- Many spatial terms have more than one meaning.
- Sometimes a term in one language has more meanings than the equivalent term in the other language.
- Some terms have an explicit spatial and temporal meaning simultaneously; some others only have spatial or temporal meanings but not both.
- Some terms have very specific meanings while others have very ambiguous meanings, with these last terms being used in many contexts

Concerning the **form** of the terms:

- Some terms have similar forms in both languages because of common roots (usually Latin or Greek).
- Some terms are composed of more than one term, used as single or separate words

Further analysis of the lists of spatial terms in both languages produced some interesting findings, listed in Table 3.3, being the most relevant those describing the role of the grammatical use of terms with regards to spatial relations. They suggest that spatial relations are implicit in natural language and there is no need for them to be made explicit, since spatial terms give at the same time information of the generic and the specific relations, as in the term “near”, whose meaning implicitly tell us that it is a proximity relation (because is related to the space-time property of distance), and also informs about the particular form of the relation (it tells us that there is a *short* distance implied). Grammatical use of spatial terms reflects a structure of space that, unlike language syntax or language morphology, is very clear. As I indicate it in table 3.3, up to three groups of grammatical uses can be distinguished, representing different ways of expressing spatial relations in language.

Nouns admit one of three different meanings in the context of spatial relations (parenthesis indicate the relations they are used in):

- Events or parts of events, as in *top* (orientation), *section* (aggregation and association), and *entry* (connectivity).
- Attributes of events, as in *inclination* (orientation), *visibility* (exposure), *region* (coincidence).
- Processes, as in *turn* (orientation), *approximation* (proximity), *crossing* (coincidence).

Prepositions, adjectives, and adverbs play two passive roles:

- Mode of a relation, as in *square* (orientation), *flowing* (connectivity), and *separately* (adjacency).
- Value of a relation, as in *close* (proximity), *within* (containment), and *hidden* (exposure).

Verbs convey dynamic meanings of spatial relations, so they can be used to:

- Initiate a relation, as in the verbs *to meet* (coincidence), *to open* (connectivity), and *to enter* (containment).
- Modify a relation, as in the verbs *to concentrate* (aggregation), *to close* (proximity), and to *hide* (exposure).
- Terminate a relation, as in the verbs *to disconnect* (connectivity), *to stop* (proximity), and *to divide* (aggregation).

By comparing meanings of the spatial terms with those of the properties of space and events, as discussed in Chapter 2, I was able to form groups with similar meanings. The final step was to give names to the groups of terms. Since there are no “official” names for the generic concepts of spatial relations, their names had to be chosen, in each group, from the terms that were less specific, more encompassing and that more closely resembled the meaning of the space or event property which they derive from.

3.4 RELATIONSHIPS BETWEEN SPATIAL RELATIONS

From the analysis of spatial terms in natural languages it was observed that some terms could be assigned to two or more spatial relations, meaning that those two relations might be closely related. For instance the term “contact” can be used within a context of adjacency (as in *the two* *objects were in contact*) but also within the domain of connectivity (as in *I made contact with the dealer*). Similar situations are indicated by terms such as “piece” or “section”, which can be used either in aggregation or association relations. Other examples include “crossing” (used in coincidence and connectivity relations), “underneath” (used in exposure, coincidence and orientation relations), or “there” (used in proximity and orientation relations), just to mention a few. This suggests that spatial relations hold relationships among themselves. These relationships indicate that the existence of a specific relation depends at some degree on the simultaneous occurrence of other relations. I classified these relationships on the degree of influence a relation exerts on other relations, resulting the following categories (see figure 3.2):

- *Necessarily implies* (generic), means that the occurrence of one relation always implies the occurrence of another.
- *Necessarily implies* (specific), means that the occurrence of one relation implies the occurrence of another but only in specific cases.
- *May imply* (generic), means that the occurrence of one relation sometimes implies the occurrence of another.
- *May imply* (specific), means that the occurrence of one relation sometimes implies the occurrence of another but for specific cases.
- *Mutually exclude* indicates that the two relations cannot occur simultaneously, but only one of them.

While some relationships are rather obvious, others were unsuspected. Their importance rests in that they may help us to infer the existence of certain spatial relations in the absence of evidence. For example, knowing that adjacency might imply connectivity, if in the course of a research we find that two types of events are always (or at least most of the time) adjacent or coincident, we can infer that there is a connection, that is, they also hold a connectivity relation, which can then be investigated to look for evidence.

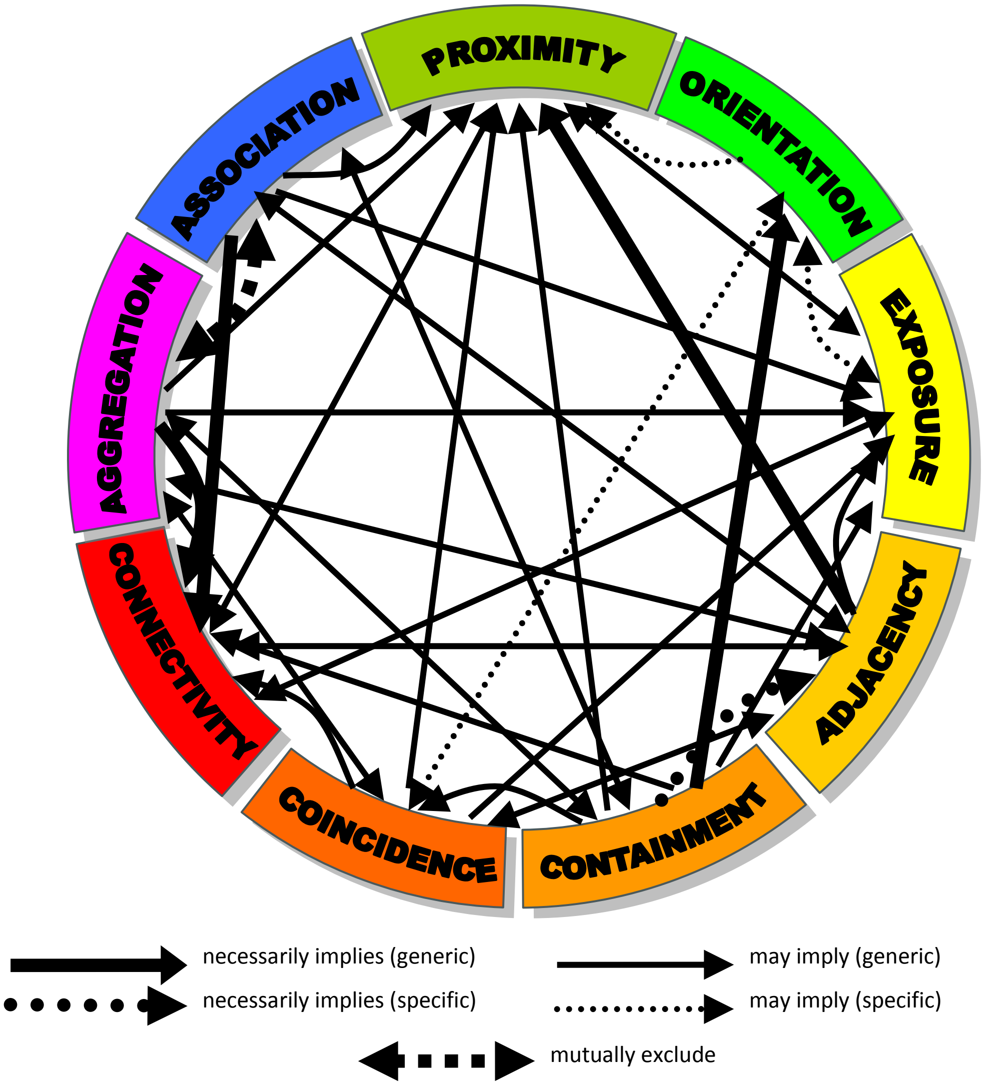


Figure 3.2 Relationships among spatial relations.

From the diagram in figure 3.2, some observations can be made (further research is needed to ascertain what their meaning might be for spatial structure and spatial organization):

- Some of the relationships are unidirectional, some other are bidirectional.
- Most relationships are generic, that is, they apply to all cases; some other (five relationships) only apply to specific cases of relations.
- There are only two relations that are mutually exclusive (aggregation and association).
- Relationships among spatial relations are transitive, that is, if relation A implies relation B, and B implies relation C, then A implies C.
- Proximity, exposure, containment and coincidence have eight relationships each (the maximum); the other relations have less than eight relationships.
- Orientation has only four relationships (the minimum), which might partly explain why is the relation with the most of spatial terms.
- Any relation implies or may imply proximity and exposure.

Since the character of each relationship might be somewhat unclear, we give examples of them in Table 3.4.

Table 3.4 Examples of relationships between spatial relations (Q = question; A = answer)

RELATIONSHIP	EXPLANATION	EXAMPLE
Orientation may imply Proximity (specific)	If an event is described as oriented in the direction indicated by the term "there", it might be also relatively near, since the use of that term implies pointing to a location that can be seen, and if it can be seen it is probably not remote.	Q. Where is the bus stop? A. There (pointing at it)
Proximity may imply Exposure	If two events are relatively near they also might be exposed to each other, visually or otherwise.	If a sanitary landfill is located relatively near the limit of a city, population living near the limit might be exposed to, bad odors or gases from the landfill
Exposure may imply Proximity	If two events are exposed to each other, it is very likely they are near.	If population is exposed to traffic noise it may be because they live near to a highway or some other major type of road.
Adjacency necessarily implies Proximity	If events are adjacent, they are in contact, therefore necessarily they are near (for practical purposes the value of the proximity relations is zero).	Two buildings may be adjacent; therefore it can be assumed that no space is in between their adjacent walls.
Containment may imply Proximity	If an event contains a set of events it is likely that the contained events are closer to each other than to the majority of non-contained events.	Most settlements within a municipality tend to be nearer to each other than to the (non-contained) settlements in adjacent municipalities.
Coincidence may imply Proximity	If two events coincide it is almost certain they are near.	A flood prone area coinciding with a river plain also indicate they are near to each other.
Proximity may imply Connectivity	If two events are near each other they might have more chances to become connected than if they are remote.	A bank branch may establish more connections with potential customers living near the branch.
Connectivity may	If two events are directly connected it	If two cities are connected by

RELATIONSHIP	EXPLANATION	EXAMPLE
imply proximity	may be because they are near each other.	a direct road they might be nearer than other cities which are indirectly connected.
Aggregation may imply Proximity	If several events are aggregated into a bigger event, it is more probable they are near each other.	In an ecosystem all or most of its components must be near each other to combine their functions.
Association may imply Proximity	If several events decide to associate, it is likely they are near.	Countries associated in the European Union are relatively near to each other.
Orientation may imply Exposure (specific)	This may apply to the case where three or more events are oriented in the same direction, and therefore it is more feasible for at least one of them to be exposed to the other two, than if the events were all oriented in different directions	A city, oriented downwind from a pollutant factory, may be exposed to air pollutants if dominant winds are also oriented in the direction defined by the city and the factory.
Exposure may imply Orientation (specific)	If two events are exposed it might well be because they are oriented in a specific direction.	The eastern slopes of the Sierra Madre Oriental, in Mexico, are exposed to rain because they are oriented to the East, where the humidity comes from.
Containment necessarily implies Orientation	If one event contains another then the former necessarily surrounds the later in all directions.	In two dimensions, a lake surrounds in all directions any island within the lake.
Coincidence may imply Orientation (specific)	If two events coincide, one may be located on top of or underneath the other.	High-yield crops coinciding with fertile soils are on top of them (soils are underneath the crop, mostly).
Orientation may imply Coincidence (specific)	If two events are oriented in a specific direction respect to each other, they may coincide in a plane normal to that direction.	If an aquifer is beneath a sedimentary deposit of clay, both coincide in the vertical direction.
Adjacency may imply Exposure	If events are adjacent they may be exposed to each other.	If a city is adjacent to a river it may be exposed to floods.
Containment may imply Exposure	If an event contains another they might be exposed to each other.	If a region contains an active volcano it may be exposed to volcanic hazards.
Coincidence may	If two events coincide they might be	If farmlands coincide with an

RELATIONSHIP	EXPLANATION	EXAMPLE
imply Exposure	exposed to each other.	area of low rainfall, they may be exposed to droughts.
Connectivity may imply Exposure	If events are connected they might be exposed to each other.	If several cities are connected by a transportation network, they might be exposed to an outbreak of a disease in one of them.
Aggregation may imply Exposure	If several events are aggregated they might be exposed to each other in the aggregation.	If state economies are aggregated into a national economy they might be exposed to the failure of the economy of some states.
Association may imply Exposure	If several events are associated they might be exposed to each other.	If farms are associated into a cooperative to produce dairy products they might be exposed to an outbreak of a disease in the cattle of one farm.
Adjacency may imply Coincidence	Two adjacent events may also be coincident when represented in a 2D map.	A geologic formation may be adjacent to a subjacent one and therefore they will appear to coincide when looked upon from above.
Coincidence may imply Adjacency	Two coincident events may be in contact with each other.	Vegetation is adjacent to soils.
Adjacency may imply Connectivity	If two events are in contact then there is the possibility of interchanging something (energy, matter, money, ideas, etc.).	Two adjacent cities may connect through their respective street networks.
Connectivity may imply Adjacency	Many events that are connected are also adjacent because is easier to interact being in contact.	If an irrigation channel is connected to a river, they are adjacent at some point.
Adjacency may imply Aggregation	If several events are adjacent to each other they might be aggregated.	Houses in cities are usually adjacent, thus forming city blocks.
Aggregation may imply Adjacency	In an aggregation, some of the aggregated events may be in contact with each other.	City blocks and streets aggregated into a city usually are adjacent.
Adjacency may imply Association	If several events are adjacent to each other there is the possibility that they	Adjacent industries may associate into an industrial

RELATIONSHIP	EXPLANATION	EXAMPLE
	decide to associate.	park.
Association may imply Adjacency	In an association, some of the associated events may be in contact with each other.	Businesses associated into a mall are usually adjacent.
Containment may imply Coincidence	If an event contains another it is likely that the later shares its space with the former.	If a natural protected area contains a forest, both events are also coincident where the forest is located.
Containment necessarily implies Adjacency (specific)	If an event directly contains another, then both are adjacent.	A lake containing an island is also adjacent to it.
Containment may imply Connectivity	If contained events are more or less isolated by the limits of the container event, then they may seek a connection among themselves.	People living (contained) in a city may be connected to provide services to each other.
Containment may imply Aggregation	If an event contains many other events then it may derive its properties from the combined properties of the events it contain.	Supermarkets, markets and retail stores within a city aggregate into the food supply service of the city.
Containment may imply Association	If events are contained they may organize to modify its container event.	People living (contained) in a specific neighborhood may associate to improve their quality of life: increasing children safety, reducing noise and pollution, etc.
Aggregation may imply Containment	Events which are combined to form a higher order event are usually included within this higher order event.	Trees in a forest are aggregated to form the forest which includes the trees.
Coincidence may imply Connectivity	Two coincident events might also be connected.	Vegetation connects to the soil to extract water and nutrients.
Coincidence may imply Aggregation	Two or more coincident events may combine their properties to form an aggregated event.	Coincident soils, climate, water bodies, animal life and vegetation may aggregate to form an ecosystem.
Aggregation necessarily implies Connectivity	In an aggregation all aggregated events have to be connected directly or indirectly with each other.	Components of an ecosystem are interconnected.
Association	In an association all participant	To be efficient, different

RELATIONSHIP	EXPLANATION	EXAMPLE
necessarily implies Connectivity	events can connect or disconnect depending on specific needs.	modalities of transportation services, in a transportation system, need to be interconnected.
Aggregation excludes Association	Since aggregated events do not decide how to combine, they cannot be associated.	Components of an ecosystem do not decide to associate they can only be aggregated.
Association excludes Aggregation	Associated events decide if they want to be part of the association or not, therefore they are dependent like aggregated events.	Businesses in a mall usually do not depend on each other for their existence; therefore do not form aggregations but associations.

CHAPTER 4

STRUCTURAL RELATIONS

This set of relations encompasses those resulting from the dominance of the properties of space**.** Relations in this group afford measuring the potential for interaction depending on distance, direction, and the existence of barriers between events. The key feature in these relations is that their measurements provide us with insights about the structure of the field by which we detect space.

4.1. PROXIMITY

Definition

I define a proximity relation as:

The degree of separation between events.

This category of relations derives from the property of space called distance. Indeed, the above definition could also be considered as that of distance; however, distance is only the concept we use to measure or express a proximity relation.

The occurrence of the relation is not dependent on the type of space; all spaces, physical, mathematical, social, etc., admit the existence of this concept. It is also nondependent on the dimensionality of a particular space: it is possible to measure the amount of separation in between two or more events in one or more dimensions.

Proximity and distance

The relation of proximity is more than the single concept of distance because proximity, besides the measure of the amount of time-space, requires a qualifier for that measure. Consider for instance the proximity relation between two points, as depicted in Figure 4.1. If the straight distance between points A and B is measured on the plane, we have a first expression of proximity which can be qualified only with respect to another relation using the criteria described in the next pages. If we measure again the straight distance between both points, but this time taking into account terrain inclination, we can have a second expression of proximity that, initially, can be qualified as less proximal than the first because the distance is longer. A third measurement of the same distance can be done considering the inclination as before but made along the road going from one point to the other. The proximity relation in this third case would be qualified as even less proximal than the previous two because the distance is greater. A fourth case of proximity relation would arise if the measurement of distance were done using travel time instead of conventional distance units. In this last case, in spite that the physical distance between the two points is the same disregarding the direction of measurement, the proximity relation, when it is measured from B to A, would be qualified as nearer than from A to B, simply because the time-distance is longer in this last situation (it usually takes longer to travel uphill than downhill if the same amount of effort or energy is used in either situation).

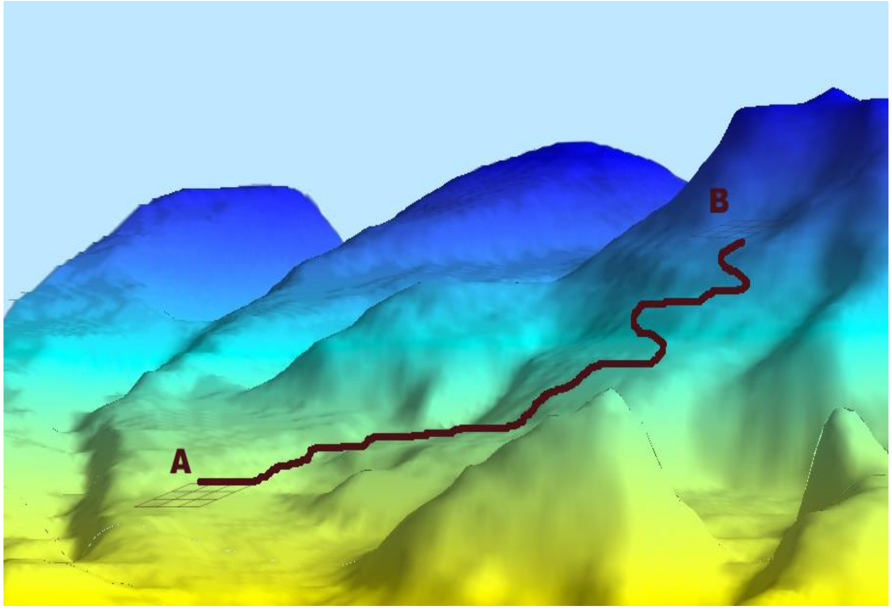


Figure 4.1 Proximity between places A and B is not always symmetric in time-space or in cost-space. Time or costs of traveling from A to B can be longer or costlier than from B to A, although the physical distance is the same.

The concept of Proximity

Proximity relations are perhaps the most common of all spatial relations, maybe because, with regards to physical space, they are inseparably linked to movement. In science, and particularly in geography, proximity plays a key role in describing, explaining and predicting the spatial behavior of many phenomena. Tobler (1970) stated one of the fundamental aspects of proximity in what it is now known as the First Law of Geography: 'everything is related to everything else, but near things are more related than those far apart'. Of course, this is a general rule and not a true law, because there are cases where proximity is not the most important relation, and events which are not near are more related than those that are near.

Proximity is one of the most studied characteristics of spatial events. The idea of proximity as the amount of separation between events is closely related to two basic aspects of spatial events: extension and duration. Although in the relativistic conception of space one implies the other, in the Euclidean view of geographic space both aspects can be artificially treated as somewhat independent magnitudes.

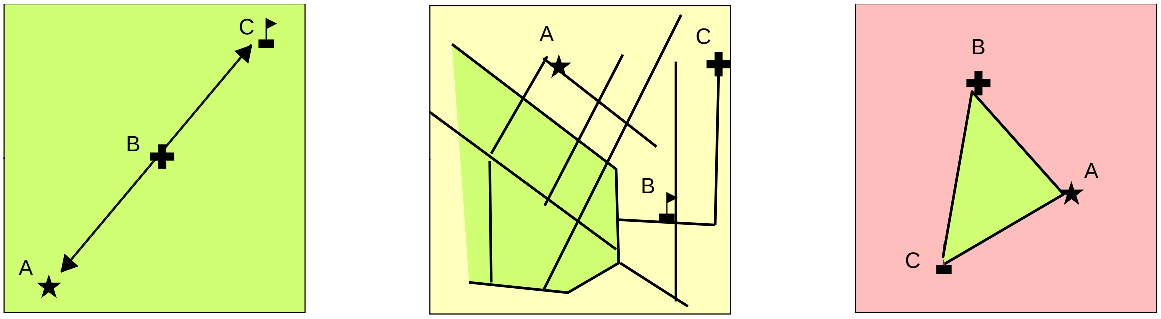
Consider for instance a spatial event such as a flood. We can measure the extension of a flood, both, as the amount of separation existing in between any two opposite points of the flooded zone (the two more distant, for example) or as the flooded area. These two measures give an idea of the proximity relations in between parts of a single event. The former gives an exact value of proximity but only for two points of the event; the later gives an approximate idea of the value of a proximity relation, in this case, a minimum possible value for the extension of the flood, but for the entire event. If we know that the flood covers an area of 10 km2, then because Area of Circle = π* r2, and r=√ Area of Circle /π we know that one of its dimensions (length or width) has to be at least 3.5 km or more, i.e. the minimum possible extension of the flood in between any two points on its boundary should be at least that value if the flooded area has the form of a circle, or greater if the flood has another shape.

Conversely, we could measure the duration of a river flood. In this case, our only chance of finding the amount of space covered by the flood would lie in having a measure or an estimate of the velocity of the process as well.

Logic properties

The relation has the following logic properties:

- It is not transitive; the value of the relation between two events is not passed on to other related events. If event A is near event B, and event B is near event C, event A is not necessarily near event C.
- It is additive for some situations; the value of the relation does not accumulate in between three or more related events. If distance from A to B is *x* and distance from B to C is *y*, distance from A to C is not necessarily x + y (Figure 4.2c). This could be only possible if the three events were collinear points (Figure 4.2a), or non-collinear points were on a network and proximity was measured as the shortest network distance between two points (Figure 4.2 b).
- It is not always symmetric; the value of the relation between two events may not be the same when measured from one or the other. This may be particularly true of proximity relations measured in time units or other units different from conventional distance units. Consider for instance two places connected by a road as in Figure 4.1., if we are traveling by car, then it is possible that the amount of time can be shorter when traveling from B to A than from A to B because higher speeds can be achieved easier when the vehicle is going down. Also, in cost-distance space, fuel consumption can be noticeable less when going from B to A than for traveling from A to B.
- It is not reflexive; the value of the relation is not the opposite when measured from two different events. If event A is the nearest to event B, this does not necessarily implies that the nearest event to A is B.



(a) (b) (c)

Figure 4.2 Proximity can be additive only if events are collinear as in (a)**,** or are located on a unique path on a network as in (b), where AC = AB + BC. The relation in (c) is clearly non additive with AC ≠ AB + BC.

Topologic properties

The relation has no topologic properties. Transformations of space such as scaling, rotation, translation and re-dimensioning affect the value of the relation.

Scaling has a strong effect on the value of proximity relations. For instance, if we measure the distance between two places represented as points in a 1:1,000,000 scale map and then measure the distance between the same places (points) on a 1:250,000 scale map, it is very likely that there is a difference (the distance measured in the 1:250,000 scale map will be greater), however small the increase, between both measurements.

Also, if we translate, without rotating, the position of two points on the equator to, say parallel 60° (a translation on the y axis), the decrease in distance will be noticeable: one degree of longitude at the equator equals 111.321 km, while at 60° latitude it is only 55.802 km (that is, half of the distance!).

A shift in the number of dimensions affects the relation value as well. If we measure the separation between two points on a topographic map, as a straight distance on the plane (2D), and then measure the same separation taking into account terrain slope (3D), we may find that both distances differ. This difference will be greater the steeper is the slope in between both points.

Qualitative expressions

Very likely, the first qualitative expressions of proximity relations were not verbal but made of signs. Pointing to a specific spot or place in the visual landscape is perhaps the most primitive, but effective, way of giving an indication of proximity (and of orientation too). As languages evolved, words were invented to convey the meanings given to signs. Moreover, language allowed creating proximity expressions for which no signs existed.

Some languages differ in the way they handle proximity. Levelt (1996) says that while English or Dutch use a bipartite system (proximal-distal), Spanish and Japanese use a tripartite system (proximal-medial-distal). Actually, as a native speaker of Spanish, I can say that it uses a four-element system (aquí-ahí-allí-allá) that markedly contrasts with the two-element system of English (here-there). These differences reveal different needs for constraining the exactness of qualitative values of proximity, so that approximate quantitative values could be estimated more precisely without using quantitative measures.

In all natural languages there is the possibility of using words as qualitative measures to describe proximity relations. For example, in English, some proximity terms are used when an indication of an event's position is required: r*emote, far, close, near, nearby, distant, deep, high, extreme, immediate*, etc. These words indicate the relative proximity of a single point or part of an event, or of a place.

Some other terms, the "dimensional adjectives" of Bierwisch (1996), without being explicitly qualitative measures of the relation, give an idea of proximity through an indication of size: *short, small, reduced, long, big, compact, thick, tiny, extensive, wide, infinite, enormous, immense,* etc. In these cases, size refers, directly or indirectly, to a measure of proximity between two or more given points or places located on the boundary of an event or to its extension. Additionally, these terms appear in antonymous pairs in language (Bierwisch, 1996), so that they are used to represent opposite proximity values.

Still, other terms may only imply the existence of different degrees of proximity between two or more events. For example: *smooth, flat, narrow, bottomless, parallel, square, oval, here, there*. Notice how, in some terms, such as "smooth" (meaning that the difference in vertical distance of the various parts of a surface is negligible) or "narrow" (meaning that the distance between two parts of an event or two events is relatively short), the approximate degree of proximity is directly specified, while in others, such as "parallel", the term only implies that the degree of proximity is the same for all parts of the events that maintain the relation.

In any case, a qualitative expression of proximity has a context dependent meaning, it "…takes its meaning from prototypical distances or interactions among the kinds of objects in the statement" (Mark & Frank, 1989). To increase the exactness of its meaning we use some sort of rule which can be based on one or on a combination of four possible criteria (figure 4.3):

- The necessary condition (Jackendoff, 1983)
- The centrality condition (Jackendoff, 1983)
- The typicality condition (Jackendoff, 1983)
- The graceful degradation condition (McClelland, 1986)

Figure 4.3 illustrates the cases where these criteria can be applied to determine if a proximity relation can be qualified as 'near' or 'far'.

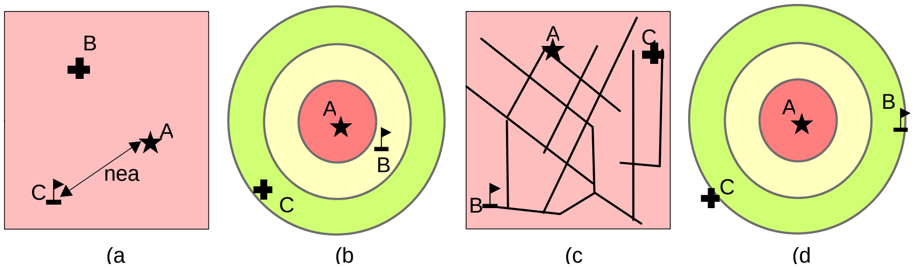


Figure 4.3 The four criteria to determine the meaning of qualitative expressions of proximity, applied here to the near – far relations: (a) necessary condition; (b) centrality; (c) typicality; (d) graceful degradation.

The necessary condition establish that, according to figure 4.3a, for event B to be qualified as near to event A there must exist another event, C, presenting a degree of proximity to A similar to that of event B, and that has already been defined as ‘near’.

The centrality condition is exemplified in figure 4.3b. Since proximity can be seen as a circular concept (in isotropic or near-isotropic spaces), event B is nearer to event A than C, because it lies in an inner circle, which is closer to the center (A).

The typicality condition is represented in figure 4.3c. If the walking time from A to B is five minutes and from A to C is twenty minutes, then event B is typically near to A because we typically associate the concept ‘near’ to a distance we can traverse in five minutes or less. For the same reason event C is typically far, although its straight distance to A is shorter than to B.

McClelland’s “graceful degradation” criterion is illustrated in figure 4.3d. If event B is 15 minutes’ walk from A and event C is 16 minutes’ walk from A, neither of them could qualify as typically near, but since B is the nearest of both events, it could qualify as ‘near’ in absence of a better instance.

But while qualitative expressions of proximity may be useful and, in many situations, of sufficient precision to solve a problem, there are other circumstances in which higher levels of precision are needed, thus requiring quantitative measurements.

Quantitative measurements

When measurements of distance are used as an expression of proximity, the relation has been always considered as binary, in the sense that it requires at least two events to take place. To this regard, White (1991) says that distance is an attribute of a pair of events. However, it is essential to point out that although measuring can be done in a binary style, the resulting value does not necessarily express a binary relation. The binary situation occurs if the two events are represented as points in space. If more than two events (or non-point events, or parts of an event, or parts of different events) are involved, it is possible to provide an overall or a specific measure, such as the average or the maximum distance from all calculated distances, or have a multivalent expression of the relation.

When quantifying proximity relations two aspects of the quantification must be observed: precision and accuracy; the first being the ability to measure with sufficient detail and the second the ability to measure with sufficient agreement with reality (sometimes called exactness).

Precise measurement of distance in physical space is relative to scale. There is no absolute precise measuring in this type of space. Paraphrasing French artist Noelie Altito: “The shortest distance between two points is under construction”. Precision of a proximity measure depends on the scale, and implicitly, on the spacing in between units on a rule, or on the resolution of the measuring instrument. This impossibility of having absolute precise measurements in physical space also means also that there are no absolute accurate measures of proximity. However, once a scale of measurement has been set up as sufficient, all measures can be considered as precise or accurate for practical purposes (provided there are no measurement errors imputable to the instruments or to the measurers).

In mathematical or abstract spaces it is possible to have accurate measures of the relation. Take for instance the distances between elements in the space of the integer numbers: any measure of proximity between two given numbers is necessarily precise, and, at the same time, accurate. However, when a physical space is mapped onto a mathematical space (such as when the Earth's surface is projected onto a plane), only the precise character of the mathematical space is retained, while the accuracy quality depends on how well the mapping was done.

Crisp measurements

When we measure proximity, we usually wish our measures to be as exact as possible. As we said before, this implies to pay attention to precision and accuracy. We already know that if we are taking measures of the real world, either physically or by means of a model (usually a map), precision has to be subordinated to scale, while accuracy depends on how well our model matches reality at that particular scale.

Once the acceptable precision level has been established for a particular scale, we can regard any measure complying with that level, as crisp and precise enough, although we must know that there is always the possibility of measuring the same event with greater precision at another scale, but perhaps such improvement in precision would lack significance for any practical purpose.

Consider the situation where we are traveling from a city to another. We may want to know what the distance we have to travel is, and this could be accomplished by measuring that distance in a map. For example, if we use the map in figure 4.4, we could say that the (straight) distance between city A and city B is 52.6 Km (between the centroids of the cities, as indicated by line b). In any case, the value of our measurement will depend on the scale of the map we are using.

But if the required precision can be easily achieved by selecting the appropriate scale of measurement, accuracy needs careful deliberation of the following: the value we are looking for is to be measured as a straight distance or as a road distance? If we are traveling by land we should measure the road distance, especially if the road is winding. It would be far more exact.

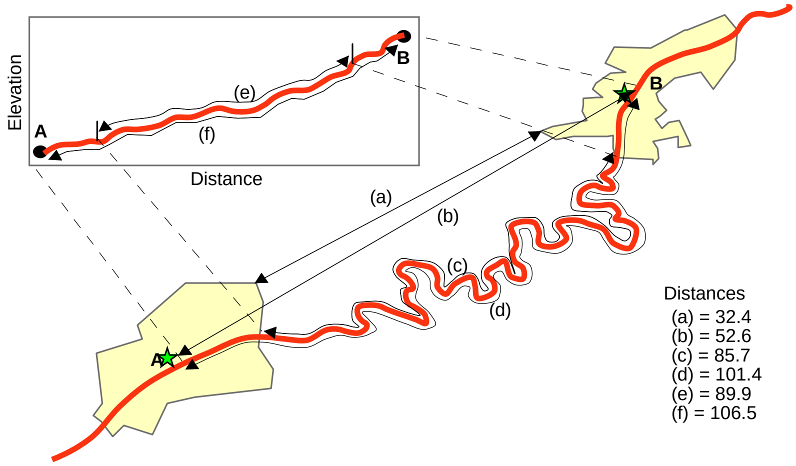


Figure 4.4 Different ways (black lines with arrows) of expressing proximity between two cities connected by a road (in red) result in six different distance values.

Furthermore, do we want to measure the distance between the centers or centroids of both cities or only between the cities' limits? If we are using a small scale map such as 1:500,000, it is very likely that the cities are represented as points and we can reasonably expect that such points correspond to the centers of the cities. If we have larger scale maps we could decide to take the measure between the limits of both cities, but then we may wonder if those limits are current or are 10 years old, and therefore outdated.

Moreover, should we consider the territory as flat or should we take into account terrain steepness? It would be more exact to account for road steepness if the terrain is very rough, but then this would require data that is not usually available in a road map, and very probably you would also need a GIS to carry out calculations. Unless it is very important to estimate fuel consumption, you could stay with a flat distance.

So, in response to a question such as "how near (or far) is city A from city B?" we could answer: it is between 36 km (if we fly from city limit to city limit) and 106 km (if we go from city center to city center by road and take into account road winding). This answer dramatically shows that proximity relations in the real world are more complex than maps could make us believe. It also emphasizes the importance of accuracy in measuring proximity relations: should we have used the first figure we gave of 52.6 km (which might be quite precise) to travel by road from one city to the other, we could have received an unpleasant surprise discovering that the real distance is twice as more, with the implications this has on time and cost of traveling.

All these considerations would result in a set of possible proximity measures as those indicated in Figure 4.4, all crisp and as precise as our map and measuring instrument would allow but all with different degrees of accuracy. This set of values can be regarded as the fuzzy set (Zadeh, 1965) of proximity measures between the two cities.

Fuzzy measurements

Fuzzy proximity measures consist of a set of crisp values, which provide a range of acceptable values for the relation between two events. These values can come from different ways of taking a measure (as in our above example) or can be derived statistically (the minimum, the mean and the maximum measures of a set of measures, for example). In addition, a fuzzy set could also be described by a function which provides information on the degree of membership of values with respect to a specific proximity relation. This membership is in practice the degree of certainty (or possibility) that a given value belongs to a specific class, and it ranges between 0.0, indicating non-membership, to 1.0, signifying complete membership.

In the context of a spatial relation, a fuzzy set can be interpreted as a class of relation without crisp boundaries. This suggests that a relation between two events or parts of a single event can be measured and expressed as a set of values, every one of them indicating different degrees of magnitude or intensity of the relation.

As a way to illustrate the possibilities for fuzzy logic regarding proximity relations, I have depicted in Figure 4.5 a set of fuzzy functions to describe the proximity values for relations ranging from "very near" to "very far".

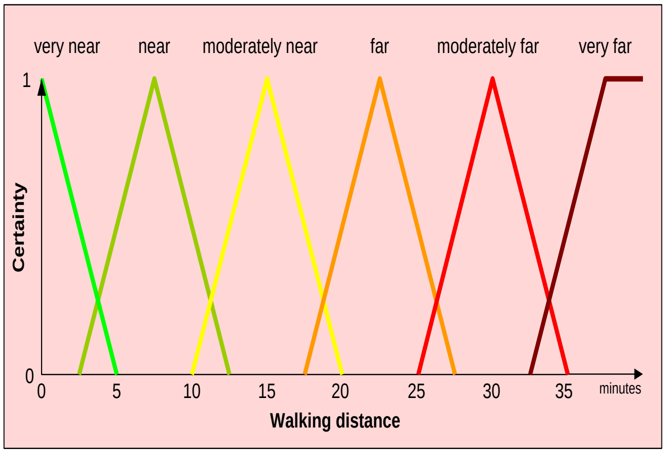


Figure 4.5 Fuzzy functions for different proximity relations measured as walking distance in minutes with their corresponding certainty or membership values.

For each relation we are using two functions: a monotonically increasing function (MIF, absent for the "very near" relation) and a monotonically decreasing function (MDF, absent for the "very far" relation). Given these functions we can calculate the degree of membership to any of the relations using the following formulas:

for a MIF, DM = (x – L) / i

for a MDF, DM = (L – x) / i

where,

DM = degree of membership to a specific case of a proximity relation

x = value of a location for which its DM is to be asserted

L = x value of the lower limit of the function (the one for which the DM=0.0)

i = interval of x units spanned by the function (on the x axis) between its lower and upper limits.

Since the functions overlap at some parts we can have the case where a single event can be qualified as holding two relations simultaneously (Table 4.1). In practical terms, these values mean that, for instance, a place located at 4 minutes of walking distance, still could qualify as “very near”, although, “near” would be a better qualification (since the DM value is higher for “near” than for “very near”).

In the example, I use linear functions to describe relations but other type of functions can also be used, depending on the situation.

Concerning the measurement and expression of spatial relations, fuzzy logic is in fact closer than crisp logic to the way we perceive the real world, especially when qualitative expressions need to be translated to quantitative values, or vice versa.

Table 4.1.1 Degree of membership for the “very near” relation in Figure 4.5

Walking distance (x) in minutes	Degree of membership (DM) for the “very near” relation
0	1
2.5	0.5
4	0.2 (0.3 for the “near” relation)
5	0 (0.5 for the “near” relation)

4.2 ORIENTATION

Definition

I define an orientation relation as:

The degree of alignment of an event with respect to other events.

The relation arises from the property of space which affords the measurement of direction among events. This category of relations is present in everyday life and every space because, like proximity, it is also linked to motion. Motion not only implies a certain distance but also occurs in specific directions.

Although all spaces include the relation, in some of them the number of directions may be limited. For instance in space-time there are only two possible directions: inwards and outwards / past and future; same applies to any number-space: increase or decrease. Some other spaces afford measuring in multiple directions like 2D geographic space (North, South, East, West, etc.), 3D egocentric space (in front, behind, up, down, left, right), nD political space (left, right, center, democrat conservative, liberal, environmentalist, socialist, republican, communist, etc.).

Orientation and direction

The concept of direction, used to express orientation relations is not equivalent to the relation. If I measure the direction of an event with respect to another event I can obtain a value, either nominal or numeric. This value, although useful, might not be sufficient to establish the orientation relation between both events; this is because orientation, besides the specification of a directional value, requires the specification of an orientation framework.

Suppose that the value of the direction between two events is 23 degrees. I'd still not be able to establish the orientation between them since I do not know where is the origin of the orientation framework. I need to know if the framework is azimuthal, zenithal, or other.

Even in those cases were we would assume that an orientation framework is already implicit in a directional value, we might find that this is not the case. For instance if I say that a place is North of another place, I still might be left with the question of whether the reference is to the magnetic or the geographic North.

The concept of Orientation

Although the word literally makes reference to the ancient custom of drawing maps with East (the Orient, the place where the sun raises) at the top, the concept of orientation has nowadays a broader meaning: to align (oneself or an event) with respect to an event or system of reference.

The ways this alignment takes place vary depending on the needs for orientation or on cultural issues. In the old days maps were oriented towards the East (figure 4.6) and nowadays they are oriented towards the North, but some other maps are oriented in different ways: Edo maps in

Japan are oriented towards the imperial palace; Portolan charts are oriented towards the shores they depict; Polar maps are oriented towards the map center.

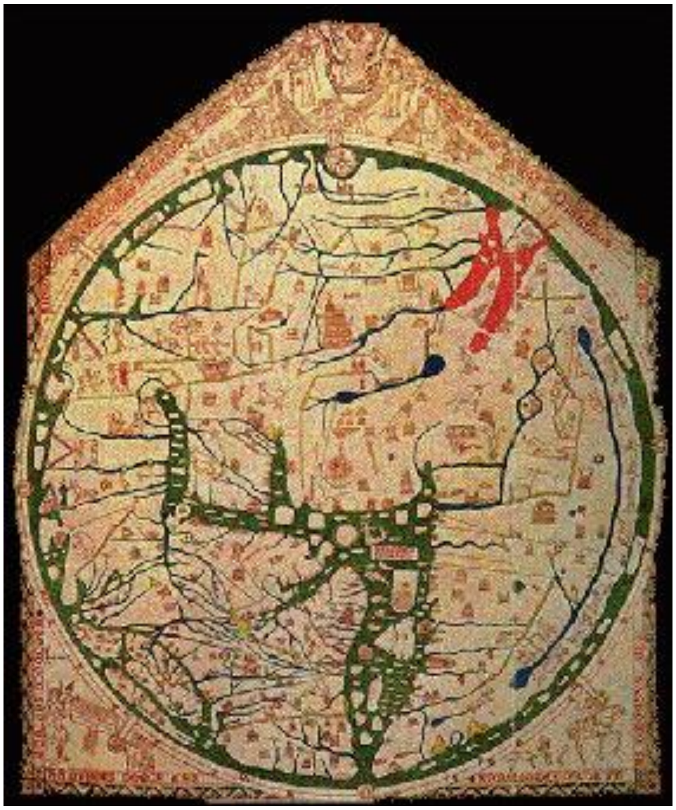


Figure 4.6 The Hereford Mappamundi, Richard de Bello, c. 1290. The map shows Asia at the top, Europe at the lower left and Africa at the lower right.

Orientation complements proximity when describing the location of events. Location descriptions would remain very vague without an indication of direction, no matter how precise and accurate distance could be specified. Orientation helps to navigate in any space by specifying a set of consecutive directions in the form of a path.

In science, besides its *de facto* use to correctly locate things in maps, the analysis of directions between events serves to discern patterns and consequently to establish hypothesis about the processes leading to the arrangement of events into specific directions. It also complements prediction by pointing the direction in which observed or forecasted magnitudes are expected to move. For instance, the slope of a mathematical function contains valuable orientation information: while the sign tells the direction in which the values of an event move (decrease or increase) in response to the values of other events, the slope value itself gives an indication of the rate of change of this direction.

Orientation is conceptualized and expressed differently in physical-mathematical and social-psychological spaces. In the former, the relativistic notion of space-time dictates that the number of possible directions between events can be either 2 or infinite within the limits of the number of dimensions or axes of the space, but the kind of directions is restricted to one of the following categories equal, opposite, orthogonal, convergent and divergent, where 1D-spaces can only have the first two kinds, while higher dimension spaces may have all of them. In comparison, social (cultural and economic) and psychological (mental and emotional) spaces, possess only a very limited range of possible directions, and their kind can only be categorized as convergent or divergent to or from in relation to a third event or concept: the origin of the measuring framework.

Logic properties

The relation has the following properties:

- It is mostly asymmetric; the value of the relation is not the same when measured from events with opposite locations. In figure 4.7a, event A is north (up) of event B, but event B is not north (up) of event A, it is south. In an azimuthal framework it would also be asymmetric: A is located at 355° of B, while B is located at 175° of A. However, in figure 4.7c, A (the umbrella) is in front of B (the sea), but B is also in front of A.
- It is sometimes reflexive; the value of the relation is sometimes the inverse when measured from an opposite event. In figure 4.7a, if event A is north (up) of event B, then event B is south (down) of event A. However, in figure 4. 7c, if A is in front of C, C is not necessarily at the rear of A.
- It is sometimes transitive; the value of the relation passes on in between three or more related events. In figure 4.7b, if event A is west of event B and event B is west of event C, then event A is west of event C. However, in figure 4.7c, if event A is in front of event B and event B is in front of event C, it does not necessarily follows that A is in front of C.
- It is additive only for a single case; in a triangle, the sum of two angles is equal to a third angle only if the third angle has the value of 90° (figure 4.7d).

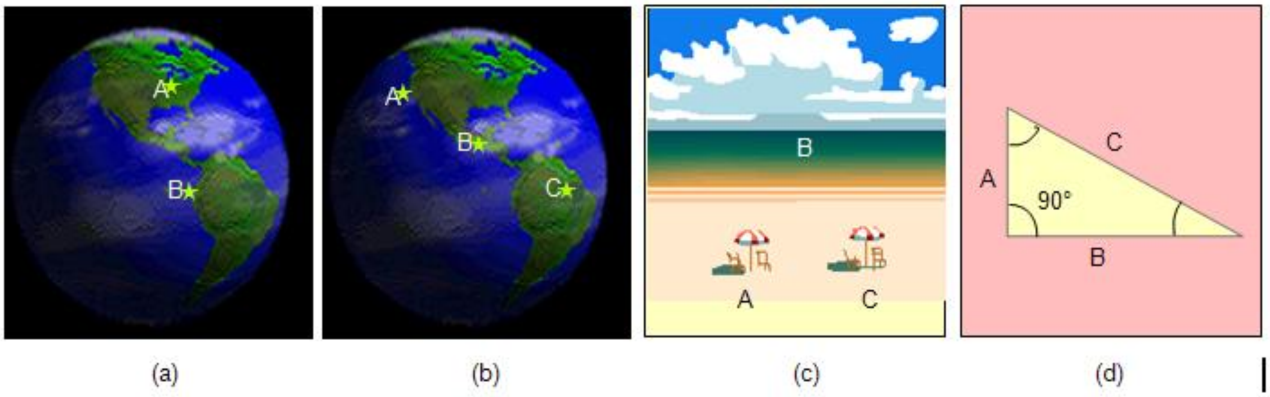


Figure 4.7 Logic properties of orientation relations (a) asymmetry and reflexivity; (b) transitivity; (c) symmetry, non-reflexivity and intransitivity; (d) additivity.

Topologic properties

The relation possesses topological properties, but only when a qualitative intrinsic system is used, since changes in scale or rotation, as well as changes in the number of dimensions and translation affect the value of the relation in the other qualitative systems, and in any quantitative system. Moreover, values can be further affected by the extent of events, in such a way that the only case where it is possible to have a precise, single valued, expression for the relation, is when the location of events correspond to points.

Qualitative expressions

Different concepts of orientation are embedded in natural languages and therefore it would be convenient to study such expressions to complement our understanding of this relation.

It appears that the majority of spatial terms in natural languages (at least in English and Spanish) can be categorized under this relation. The main reason for the abundance of orientation terms appears to be the fact that orientation relations are used to describe things in terms of shapes, mainly through terms that indicate the basic shapes of events. Terms such as ‘pentagonal' or ‘triangular’ convey how angles occur in the shape, that is, the orientation relations between parts of an event or between events, without necessarily indicate them in some quantitative form.

Other terms, that I will call horizontal prepositions, are used to relatively locate events with respect to other events in a horizontal direction; to this group belong terms such as ‘side’, ‘face’, and ‘in between’. These require the specification of a reference object and sometimes other directional qualifiers to eliminate ambiguity in their meaning. Thus for instance, it is usually necessary to indicate which side (*right, left*, etc), or to specify in between which objects another object is found.

Another group of qualitative terms corresponds to what is called vertical prepositions (O’Keefe, 1996). This group requires that there is a framework allowing the recognition of a vertical axis. In a gravity-free environment these propositions would lack meaning unless a reference object is chosen (a spaceship, the Moon, etc.). These propositions include *below, above, beneath, down, up, under* and *over*.

The expression of orientation relations in a temporal context requires the specification of terms related to time (*past, present, future*). Propositions like *before* and *after* are used to indicate the position of events along a temporal line, but in reference to another event. Other terms, like *past, present*, and *future*, admit several grammatical uses. For example the term *past*:

- as an adjective, indicates an early position of an event in time, but unlike *before*, it is not always needed to make reference to a specific event;
- as a noun, it is used to indicate a part of time, that is, a portion of the temporal line, the time before the present;
- as an adverb, it means that an event is comparatively moving behind in time (*earlier* than) or space (*backwards*);

- as a preposition, indicates that an event is comparatively moving beyond in time (*later* than) or space (forwards)

The term *present* can only either be used as noun, to indicate the portion of time between the past and the future, or as an adjective, to indicate the current position of an event along the temporal line. Same happens with the term *future*, which used as noun indicates the time yet to come, while used as adjective serves to indicate the coming position of an event.

Explicit movement in generic directions is also conveyed by means of terms such as *toward, backward, forward,* and *onward*. The first term requires the specification of a referenced event, that is, an event whose position indicates the sense of direction; the other terms implicitly convey this sense.

Qualitative systems of orientation

The accuracy of meanings of many orientation terms in natural language depends largely on the reference frame being used, which can be based on one of three possible orientation (perspective) systems (Levelt, 1996; Levinson, 1996):

- Relative. When the system of orientation is relative to an observer (to a single external event).
- Intrinsic (sometimes called deictic). When the system of orientation is referred to the intrinsic axes of orientation of the events existing in a place (to all the internal events).
- Absolute. When the system of orientation is neither relative to an external event nor to the intrinsic axes of the internal events, but rather to an abstract coordinate system.

These orientation systems, when referred to geographic space could be called homocentric, topocentric and astrocentric, respectively, but we will use the terms provided above to avoid confusion when talking about orientation in other spaces. Figure 4.1 provides an illustration of these systems.

Levelt (1996) mentions that not all languages use the three systems, with western languages favoring the use of the three, although only one can be used at a time. Carlson-Radvansky (1999) says that, in humans, all three systems are initially activated when faced with a spatial situation requiring the designation of a spatial relation of orientation. Multiple activation is followed by the selection of a single reference frame, with selection accompanied by inhibition of the non-selected frames. In my experience, this may happens even when there is a predefined system of reference we are working with. For instance, when working with maps the default reference system would be absolute, but the interpreter of a map may sometimes switch to a deictic or intrinsic system to ease the interpretation of orientation relations, which then can be translated more formally into terms of the absolute system. In other cases the system of choice is constrained by the environment. For example, in a free-gravity environment, the system of choice is intrinsic, because being outside the Earth, and with observers being part of the environment, the sides of the spaceship are used as frame of reference for locating spatial events.

Each orientation system has advantages and disadvantages for spatial reasoning, making necessary to known their characteristics to choose appropriately. Regarding logic properties, asymmetry applies for orientation relations using the relative and absolute systems, and for specific cases of the intrinsic system; transitivity and reflexivity apply only for cases of the absolute system and for cases in the vertical axes in the relative and intrinsic systems; additivity works only in the absolute system and for one specific case where the events can be considered as points forming a rectangular triangle.

Orientation relations can be expressed in any qualitative system using three axis or sets of directions. Each system has two axes for horizontal orientations and one for vertical orientations in the following way:

- Absolute: North - South, East – West, and above - below (mean sea level).
- Intrinsic: front - rear, left - right, and top – bottom.
- Relative: front - back, left - right, and up – down.

Notice that in the case of the absolute system, the terms for the vertical axis use mean sea level as the origin, whereas the cardinal directions for the horizontal axes use the Equator and the Greenwich Meridian as their origin. Therefore, in the absolute system, the meaning of the orientation terms is dependent on the position of these three characteristics of our planet. In the case of the intrinsic system the terms refer not to an origin, but to the shape, function, or characteristic use of spatial events, that is, to the “sides” of spatial events. In the relative system, the meaning each term conveys depends on the orientation of the external event (the observer) with respect to the position of the internal spatial events.

The way directions are assigned in qualitative systems, respect to specific axes, arises some exceptional situations, where it is not possible to describe orientations in three axes, but only in two of them. This can be exemplified with the absolute system: for two points located at two distinct places on the Equator, or on another parallel, it can be said that they may be located west or east respect to each other, or above or below the sea level, but cannot be orientated in the north – south axis. The same situation arises for two distinct points located on the Greenwich or another meridian: it is only possible to establish orientation relations on the north – south or the above – below (sea level) axes. A similar situation occurs with two distinct points located at the same height or depth above or below sea level, for which only east – west and north – south orientations can be given.

We must be aware that these types of situations are limitations of qualitative systems of orientations only, and not of quantitative systems, where the specification of a direction with values of 0 or 180 is possible for those situations where places are coincident on one or several direction axis.

Quantitative measurements

Likewise proximity, measurements of directions as expressions of orientation relations have a binary character, that is, they are obtained between two events. That only applies if the events are points themselves or parts of the events represented as points. When events are considered as linear, areal or volumetric entities, the value of an orientation relation may take the form of a range of directions, the average direction, or the most frequent direction (the mode, in statistical terms).

Directions can be expressed quantitatively through the concept of angle. This concept allows measuring the direction between events by establishing one event as the reference event and setting a specific direction as the origin of the reference framework.

An angle is created whenever two segments intersect at one point at least. The point where the segments intersect is called the vertex of the angle and the two segments are called the sides of the angle.

The angle between two places (points) is formed by the intersection of a segment representing the line of direction between the two points and another segment (usually vertical or horizontal) that defines the axis of reference from which the angle is measured. The usual notation to represent an angle between points A and B is $\angle AB$.

Angles, either measured in degrees or radians, may be used in many quantitative orientation frameworks:

- Bearings Angles, expressed in 90 degrees horizontally, as in wind and geologic lineaments diagrams (rose diagrams)
- Azimuthal Angles, expressed as azimuths are measured in the range of 0 to 360 degrees of horizontal direction, as when Aspect is measured.
- Zenithal Angles, expressed in 90 degrees in the vertical, either measured from above (from a spaceship) or from below (terrain slope).

4.3 EXPOSURE

Definition

An exposure relation is defined as:

The degree of potential for an event to be reached from other events.

The relation arises from the property of space which affords the measurement of concentration of events in the form of matter, energy, concepts or other category of substance (substance is here defined as the essential part of an event, not necessarily material).

Events, physical or conceptual, have a specific amount of substance which tends to concentrate with more or less density. This density determines if an event allows another two events to reach each other passing through the space occupied by the former, or to act as an obstacle or barrier.

The concept of Exposure

When exploring or measuring exposure the emphasis is given to the potential, in terms of possibility or probability, of reaching an event from another event, without necessarily establish a contact or connection, although these last two concepts may imply the existence of exposure. In other words, adjacency (contact) or connectivity (connection) relations may imply the existence of an exposure relation, but only connectivity actually indicates the existence of exposure (see figure 3.2).

In this context, it can be said that exposure implies a passive contact, as when, for example, visual contact between two or more events is established, without this necessarily implying the existence of other types of interaction.

Once established the possibility of existence of an exposure relation, the degree of exposure depends primarily on the concentration of substance of the different events existing in a particular space, and secondly, on the space curvature and position of the different events in such curved space. Concentration of substance constitute a first obstacle to exposure, space curvature can be a second obstacle but also, under certain circumstances, an agent for exposure.

Geographically, there are three types of exposure of great interest:

- Visual exposure is the most common form of expressing the relation. It does not necessarily refer to human vision, for it can be applied to any possibility of contact along a line of sight between any two events (instruments, for example).
- Acoustic exposure, both as noise or acoustic pollution, but also as a perceptual characteristic of environments, is also another form of expressing the relation among events in geographic space. Again, it does not necessarily refer to human hearing, but to any situation were sound waves produced by an event can reach another event.
- Touch exposure, concerning the possibility of events to become adjacent or connected. The important thing is the potential for touching each other, not if actually they touch or

connect each other. In this category of exposure we may found many physical and conceptual forms of the relation: exposure to a disease (the potential for an organism to become infected), the exposure to ideas (for instance by reading a book), the exposure to pollutants (the potential of aquifers to be contaminated), etc.

Barriers

A fundamental concept for the determination of this type of relations is that of barrier. A barrier may be physical, chemical, economical, ideological, etc., but its effect is that of allowing or not allowing an interaction between two or more events. Barriers can be total or partial, and in this last case a degree of exposure can be established.

It is assumed that it is the concentration in events of energy, matter or concepts which dictates whether an event is barrier or not. But actually it is the interaction between the concentrations of several events that determines the barrier quality of an event. Some assumed barriers may not be so in certain situations: glass may act as barrier for certain classes of material events but not for events formed by specific kinds of electromagnetic energy. In the same context, particular religious beliefs may act as barriers for some persons but not for all persons.

Thus, the magnitude / intensity of an exposure relation are controlled by the interaction between the densities / frequencies of barrier-events and other events.

Logic Properties

The relation has the following properties:

- It is not always symmetric; the value of the relation is not the same when measured from two different events: if A is exposed to B, it does not necessarily follow that B is also exposed to A.
- It is not reflexive; the value of the relation is not necessarily the opposite when measured from one or the other event, instead, for many types of events and situations the value is usually the same.
- It is not transitive; the value of the relation cannot pass on in between three or more related events: if event A is exposed to event B and event C is exposed to event B, it does not necessarily follow that A and C are also exposed.
- It is sometimes additive; values of exposure among events can be added to give an aggregated value of exposure for a set of events.

Topologic Properties

As the other relations in this group, it is not topologic, because changes in scale or in the number of dimensions of the space, where events are located, also change the value of the relation.

For example, two persons can be visually or acoustically exposed to each other in a space of a few meters of extension, but they might not when the space has extensions of more than a few kilometers. Of course they may use some electronic device to transmit their images and voices over the air, but then the relation of connectivity would allow only and indirect exposure, not a direct one. Another example, concerning the shift in dimensionality, is that two events located in a flat terrain and separated by a long distance, may appear to be exposed in a map, but in 3D real space, the curvature of the Earth may impede such an exposure.

Qualitative Expressions

Expressions corresponding to nominal values of the relationship are given by adjectives in three ways:

- adjectives denoting total exposure relations are *exposed*, and *visible*, while *hidden* is the term for the opposite value of both.
- another set of adjectives give the idea of a negative value of the relation, in the sense of not being exposed, as in *impassable*, *insurmountable*, *impenetrable*, *inaccessible*, etc.
- a last group of adjectives implicitly indicate that some events or parts of them may or may not be exposed depending on their relative position, such as, *underground*, *flat*, *superficial*, *abrupt*, *prominent*, *sunken*, *uneven*, etc.

Verbs indicate an action that may lead to a change in the value of the relation for some events, as in *to block*, *to flat out*, *to arise*, etc.

In general the expression of an exposure relation in qualitative form starts by indicating the existence of exposure, and once established, it can be further qualified as partial or total, depending on the arrangement of events, and then, if partial, a last qualification can be done to assess how much of an event is exposed, using then quantitative measures.

Quantitative Measurements

Exposure can be measured numerically in two ways:

- For partial exposure, it is feasible to measure the amount (length, extension, volume) of an event that is exposed to other event(s).
- For total exposure, the number of exposures of an event with respect to other events.

A typical application of the first type of exposure is what is obtained when using the GIS function known as the calculation of viewsheds. The result of such function is a representation (and an amount) of the area in a region that is visually exposed to an observer situated somewhere within the region. In the second type of measurement, it is possible to use the viewshed function to calculate a visibility gradient by using several locations for the observers instead of one. For each portion of the region under study, the gradient indicates the number of observer locations from which a particular portion of the region is visually exposed.

It is possible to combine both types of measure in an exposure index.

CHAPTER 5

NEUTRAL RELATIONS

When properties of space and properties of events interact without dominance of either set of properties, a new set of relations arise. In each of these three new relations a property of space neutralizes the dominance of the corresponding property of events.

This neutralization creates relations that participate of the characteristics of both sets of properties. We know that properties of events are absolutely topologic, but properties of space are not, hence neutral relations are only relatively topologic, that is they are invariant to certain transformations of space but not to all.

This double nature is highly relevant to spatial analysis, because neutral relations allow us to investigate equally the structure of space but also its organization. That is, they serve alike to express the potential for interaction given by the structure of space and the actual interaction deriving from the organization of space.

5.1 ADJACENCY

Definition

An adjacency relation is defined as:

The condition of events of being in contact with each other

The relation arises from the interaction between the property of space called distance and the capacity of events for connecting with each other. Given its structural nature, adjacency could also be defined as the minimum degree of proximity between events, but the relation implicitly excludes the existence of distance-time (separation) in between the adjacent events. On the other hand, its organizational origin could lead to think that a connection exists where contact between events occur, however it only specifies the existence of contact but not of a connection, thus implying but not endorsing the existence of a connectivity relation.

Adjacency and the concept of boundary

If events were not bounded, there would not be adjacency, for differentiation between events would not exist. However, it is necessary to recall that such differentiation is only artificial and derives from the idea that events have limits or boundaries. Evidence of the artificiality of the concept of boundary and the relativeness of the relation of adjacency is given by modern Physics. We now know that physical events, whatever its nature, can be never in contact, for there always exist space, atomic or subatomic, between the particles that compose events. Physical events never touch each other; they only appear to do it, depending on the scale of observation. As strange as it may sound, there is no differentiation between events, only a continuum with sudden transitions that our limited senses interpret as boundaries.

But if we accept the idea of boundary then we can sometimes regard them as part of a single event or some other times as an event in itself shared by other events. From a geometric point of view, boundaries can be represented in a dual manner, as lines or regions, which have implications in the ways adjacency relations can be measured and represented. This duality in the conceptualization of boundaries determines that adjacency relations can only be measured relatively in respect to the kind of boundary that events appear to have. Also, measures of adjacency are relative to the artificial conceptualization of boundaries as lines, regions, or volumes.

Same happens with virtual or conceptual events such as social or economic events. The American writer Ambrose Bierce once defined a boundary as 'an imaginary line between two nations, separating the imaginary rights of one from the imaginary rights of the other'. Such definition implies what in reality occurs: there is a continuum between events, because space where they take place is a continuum. It is only because of our sensing limitations, or because of our needs to differentiate events, that we create boundaries and therefore adjacency relations apparently exist.

The concept of adjacency

We say that events are adjacent when they are in contact. However, the meaning of contact can only be defined relatively in regards to the scale of observation. Thus, whenever we want to establish the existence of an adjacency relation and its value, we must first define what we should understand by contact.

Since, from an absolute point of view, physical events will never be in contact, for adjacency to exist it is necessary to assume that events are in contact when they appear to touch at a specific scale. That scale is of course the scale at which analysis or observations are carried out. For instance, at a small scale, types of soils can be represented as adjacent polygons, giving the idea that the boundary between them is a single line or a discontinuity implying that soil types abruptly change from one to another. However at a larger scale, it will often be found that they are not truly adjacent, but that they are adjacent to a transitional stripe of land of varying width (or volume), which is neither entirely of the types that were supposed to be adjacent at the smaller scale of observation.

Contact among social or economic events is even more difficult to define, and therefore adjacency might be uncertain, except when events are represented as polygons. But this kind of adjacency relation is only artificially imposed, having only geometric meaning, an only as a potential for interaction. Other kinds of adjacencies, those that correspond to actual interactions of some sort, might be concealed, or not well represented by a simple line in a map. In these cases contact must be defined by functional means. A common representation for these kinds of adjacencies is an adjacency matrix, in which binary values define contacts in spaces other than the physical.

Thus, adjacency may be physical or functional, but in either case the concept of adjacency entirely depends on our concept of contact. Once it is correctly defined, then adjacency can be properly established.

Logic properties

The relation has the following logic properties:

- It is not symmetric for functionally defined adjacency, the value of the relation is not the same when measured from different events; if event A has *x* degree of adjacency with event B, it does not necessarily follows that B has also *x* degree of adjacency with A (*x* degree means a certain degree of contact).
- It is not reflexive, the value of the relation is not the inverse when measured from either event; if event A is adjacent to event B, then B is adjacent to A as well.
- It may be transitive in some cases. Usually the value of the relation does not pass on in between three or more related events; if event A is adjacent to event B, and B is also adjacent to event C, then it does not necessarily follow that A is adjacent to C. However, when a relaxed concept of contact is admitted, we can say that if event A is adjacent to event B and event B is adjacent to event C, then A and C are indirectly adjacent, with B acting as border event, and therefore there is transitivity.

Topologic properties

The relation has some topologic properties, but it is not an absolute topologic relation. It is invariant to translation, rotation, and warping of space, but it is not invariant to changes of scale and number of dimensions in which the events occur.

This is totally true of physical events, and applies to some conceptual events. For example, a zone of high risk can be adjacent to a zone of moderate risk at certain scale, but if the scale is changed to a more detailed one it is possible that the adjacency between high and moderate risk zones is intermingled with zones of adjacency to low risk zones. Such situation may occur, for example, when, at a small scale, the part of the valley of a river which corresponds to its main course is classified as a high risk zone of flooding whereas the adjacent plain is considered of moderate risk. If we model flood risk at a larger scale it may be observed that both areas are separated by a levee, which is of course more elevated than the river’s main course and the plain, hence corresponding to a lower risk zone adjacent to both the high and to the moderate risk zones, and therefore modifying the adjacency relations of the former zones.

Adjacency is then considered as a relatively topologic relation. This affects real geographic events only, not their cartographic representations, since the representations have a fixed scale, and it is not possible to change it.

Invariance to warping of space is a property exploited in a type of cartographic representations known as cartograms, where the shape of cartographic elements in a map, usually polygons, is distorted in accordance with the value of a variable measuring a characteristic of events, but preserving the actual adjacencies among them.

Qualitative expressions

As with other relations, there is a temporal as well as a spatial way of expressing the value of the relation. The temporal term to denote adjacency is *immediate*, that is, contiguous in time, the spatial term is *contact*.

Nouns denote the parts where events are adjacent or not adjacent. *Border* and *coast* are instances of parts of events where adjacency with other events occurs. Terms such as v*oid* or *crack*, designate parts of events where adjacency with other events may not occur or it is lost.

Adjectives indicate the form of the relation. *Tangential*, *bordering*, *contiguous*, *loose*, *scattered*, are terms that tell us the way adjacency is taking or not taking place. The last two terms, for example are used to imply a lack of contact.

Verbs and adverbs are used to create, destroy, or modify the value of the relation. For instance, *to land* is used to establish that contact between two events (for example, tires and ground) took place from a situation where they were not adjacent. Similarly, *separately* implies that events participating in a process may not interact because there is no adjacency.

Some terms give a very precise value of the relation, and others carry a certain degree of uncertainty. *Next* or *conterminous*, denote a clear relation of contact. *Along* is an ambiguous term because it does not indicate where the adjacency takes place, as in *the road goes along the river,* that includes the possibility of the road being adjacent to either river margin or only in some parts along their length (Christopher Habel, cited in Jackendoff, 1996), or the existence of a relaxed adjacency where the road does not necessarily must touch the river for them being adjacent.

Quantitative measurements

As with other relations, adjacency can be expressed as a binary value indicating the existence or the absence of the relation. This simple measure is sufficient in many cases, for others it might be necessary to use values based on the number or length of adjacencies.

Binary values giving the value of the relation are conveniently represented in tables, either as a two-column table, showing for each event in a map the identifiers of adjacent events, or as a multi-column table where for each event an indication of the condition of adjacency or not adjacency with all of the other events is given (an adjacency matrix). These binary values give a simple idea of the potential or the actual capacity to interact with other events.

A more advanced expression of adjacency can be created from the binary values, by counting the number of times an event is adjacent or not adjacent to others. The number of adjacencies, then, can be interpreted as an indicator of the degree of potential or actual capacity to interact. Simple or accumulated binary expressions of adjacency are often employed in cellular automata algorithms.

The number of adjacencies can be replaced or complemented by a measure of the length of space where adjacencies occur. The length of adjacencies can be used as an absolute value or as a proportion of the total possible length of adjacencies, that is, including the length of non-adjacent portions.

Of course, an index combining the number of adjacencies with the length of adjacencies can be built. It is even possible to give weights to some portions of the adjacencies, for instance when adjacencies occur with specific types of events, modifying in this way the value of the relation. This kind of expressions of adjacency is frequently used in the construction of (vegetation) fragmentation indexes.

In some instances, the definition of contact is relaxed to allow two events to be adjacent if both are adjacent to a third event, where this last event acts as a 'border' event for the former. Then, indirect adjacency can be expressed as an amount indicating the number of 'border' events existing between two events. This is analogous to the *nth*-neighbor where *nth* indicates the order of the relation, thus allowing the existence of relations of *nth*-orders, with the first order corresponding to a direct contact. This type of conceptualization of adjacency relations is very useful in establishing autocorrelation among events, and used in indexes such as the Joint-Count statistic or Moran's I statistic, with spatial lags based on indirect adjacency.

In any case, it must be remembered that binary or multivalued expressions of the relation may serve either to indicate potential or actual interactions.

5.2 CONTAINMENT

Definition

A containment relation is defined as:

> *The condition of an event of being surrounded in all possible directions by another*

The expression '*all possible directions'* in the above definition is dependent on the number of dimensions of the space where the relation is measured. For instance, in 2D space, all directions may refer to all azimuthal or cardinal directions, while in 3D space it includes above and below directions besides left, right, in front and back directions. In a structured space such as the Euclidian space, the number of possible directions can be infinite; by contrast, in an organized space, as that of the subway lines, there are usually only two possible (and opposite) directions.

The number of possible directions is the basis to define the existence of containment relations.

Containment and explicit order

The concept of order appears briefly in adjacency relations, and then only in cases where indirect adjacency is allowed, but is not a requirement for that relation to exist. However, in containment relations order is an inseparable aspect, so that whenever there is containment there is always a main event and a subordinated event.

The relation of containment requires the differentiation of events as containers and contained. These types of events are complementary rather than opposite. Order in them manifests in the way a third, not contained, event can interact with a contained, event: it must interact firstly with the container event and secondly with the subordinated event. Also, it dictates the sequence a subordinated event must follow to interact with a third non-contained event: it firstly must interact with the main event.

When there is multiple and nested containment, the order relations may become more complex, but in all cases interactions can only take place in the order specified by a particular arrangement of containers and contained events.

The concept of containment

We say that an event contains another if it includes all parts of it. That usually entails that both events are clearly defined so that we can exactly determine if one is surrounded by another in all possible directions. However it is not strictly necessary to explicitly define the spatial extensions of both events to establish a relation of containment.

In both, temporal and spatial contexts, containment can be assumed to exist even for ambiguously defined events. For example when in a temporal context we say that an event has occurred or can occur within two days, we are implying that the event is contained by another event, of temporal nature, that spans two days of extension (duration), and that the location of the contained event is somewhere / sometime in that extension, without the need to explicitly define the characteristics of both events. Same happens in conventional spatial contexts, for instance, the expression ‘within two kilometers’ implicitly states that there is some kind of event that is contained by another event which spans an extension of two kilometers, without necessarily defining how the two kilometers are distributed over a territory, or the exact extension and limits of the contained event.

The concept of containment is universal: it can be used with events of the same or different kind, in temporally or spatially dominant contexts and in structured or organized spaces. This includes both physical and conceptual events, as shown in relations indicating nested containment of portions of the earth surface, such as nation – state – municipality – locality, in which the physical components of conceptual events maintain containment relations although often there are no physical boundaries, and therefore order defining containment is only relative to their conceptual components.

Containment can be sometimes conceptualized as a case of coincidence; however, there are two important differences:

- Coincidence between two events can sometimes be partial, whereas containment requires that the whole extension of an event totally coincides with another.
- Coincidence only indicates if two events share the space in which they occur. Containment, additionally, indicates the order in which an event is included in other events.

It also can be taken as a case of exposure relations, in the sense that subordinated or contained events are not directly exposed to non-contained events, with container events acting as obstacles or barriers. However, again, order is not important for exposure relations, and is not explicitly expressed, whereas in containment order is always resultant from the differentiation in containers and contained events.

Lastly, containment is related to aggregation relationships, in which an emergent, high-order, event can be seen as the container of many, low-order contained events. A very good example of this is our physical body which at the same time that is a combination of properties of several organs, it also surrounds them in all directions and it contains or includes them. But again, the relevant fact is that there is order in the interaction of organs with exterior events (substances, phenomena, etc.).

Logic properties

The relation has the following logic properties:

- It is asymmetric; the value of the relation is not the same when measured from the contained or the container event. This is clearly expressed in the order imposed on the potential and actual interactions.

- It is reflexive; the value of the relation is exactly the inverse when expressed from the container or the contained events.
- It is transitive in both directions of order; if event A contains event B, and event B contains event C, then A contains C, albeit indirectly. Conversely, if event C is contained in event B, and B is contained in A, then C is contained in A.
- It is additive in the ascending order; if A contains B and B contains C, then B is once contained and C is twice contained.

Topologic properties

As a relation participating of the nature of structural and organizational properties, containment has some topologic properties. Is invariant to translation, rotation, scaling, and warping of space, but it is affected by the change in the number of dimensions of that space. It is therefore only a relatively topological relation.

As an example of how the number of dimensions may have an effect on the value of the relation, consider the case where a forest encloses a lake and the lake contains an island. With containment relations as established in this manner, apparently the only way to have access to the island from the forest would be to cross the lake (that is, assuming we cannot fly). This would be totally true if we are analysing the situation in two dimensions. But in a more realistic model of the same situation, adding a third dimension, completely changes the value of the relations, leaving the lake as the only enclosed event, because in 3D it is obvious that the ground where the forest grows, is also the same ground of the island, therefore allowing the construction of a tunnel under the lake to access the island from the forest.

Again, topologic properties are very closely related to the representations of events. To avoid falling into a topological fallacy, mistaking representations as the real events, care must be exercised when building models of the reality, specifically regarding the number of dimensions in which events are conceptualized to occur.

Qualitative expressions

Natural languages offer many terms for expressing containment relations. Most of them are verbs, a few nouns, and even less adjectives and prepositions. This is a different situation to that of adjacency, and more similar to the one of coincidence. The explanation for the abundance of verbs expressing containment relations is that order in containment necessarily results from a process, i.e. from an interaction where events became something they were not, in this case a duality of container-contained.

In general, verbs, such as *to surround*, *to encircle*, *to unwrap*, *to come out*, denote the transition of passing from one state of containment to another, that is to become containers, or to lose the state of being contained.

Nouns express whether the events are containers or are contained, as in the terms *box* (a container) or *cavity* (a contained event). Adjectives, such as *included* and *external* designate the current state of containment of an event in the relation, much in the way as adverbs, such as *inside* or *outside,* and prepositions as *within* or *without*.

Quantitative measurements

A first measure of containment can be provided by the number of occurrences of one event that is contained in another. An extension of this is to express those events contained as a proportion of the total number of events that could be contained, or as a proportion respect to all events contained by other events. A further variation consists in counting the number of containers for a specific type of event, that is, how many events contain at least one event, in a non-nested fashion. Complementary, it could be interesting to know how many nested containers have a single event.

Measurements can be oriented to measure order rather than the condition of containment. In complex containment situations, valuable information may be given by the order in which events contain each other. This usually results in the identification of levels of containment.

5.3 COINCIDENCE

Definition

A coincidence relation is defined as:

The situation where an event apparently coexists in space with other events

Strictly speaking no physical event shares its space (and time) with other events. If it appears to do it is only because the representation lacks the dimensionality or the detail enough to appreciate that each event occupies its own space. However, a relaxation of this fact is allowed, so that if the representations of two physical events appear to occupy the same space in a specific number of dimensions, it is said that they are co-incident. For example, soils and vegetation actually do not share any space, but their corresponding representations as 2D maps, when overlaid, may give the appearance that, in some portions or in the entire space, certain types of soils always or often coexist with certain types of vegetation. This apparent coexistence is what is called a coincidence relation.

Since the space of conceptual events is defined by the functions that these types of events perform or are assigned, it is quite possible that two or more events, related to another event that defines the space of observation, coexist. For example, population and economic activities coexist in spaces defined by administrative units, such as cities, municipalities, etc.

Coexistence in time and space

All events that are more or less coetaneous are somehow in coincidence, since they occur more or less at the same time. It is not strictly necessary that they are simultaneous to say that they coexist, but it is a requirement that they do coexist in space for a relation of coincidence to take place. So, coincidence is coexistence of events in space without them necessarily occur at the same time, although usually they do happen at approximately the same time.

This apparent paradox can be explained if we recall that we can artificially bring any two different classes of events to coexist, simply by overlaying their representations. Thus coincidence does not require that the real events do coincide in the same space-time. They may of course, but the *sine qua non* condition is only that they apparently coexist in space, even though it is only a coexistence of representations.

The concept of coincidence

In everyday language we usually use the term coincidence to indicate a situation that was not expected, in which two or more events are found in the same space and time. In this sense, the concept of coincidence has a meaning of fortuitous, casual, accidental, or ruled by chance. But remember that chance is a term we apply to situations where causes are unknown. In fact, every coincidental situation, even if seemingly fortuitous, has an explanation, causes and mechanisms of co-occurrence. The term coincidence in science can be applied to any situation where two or more events co-occur or appear to co-occur, irrespectively of the knowledge we may have of its causes of co-occurrence.

When we overlay maps, we are creating or reproducing coincidence relations that may correspond to true relations. The purpose of such operations is usually to assert if two events are frequently found in the space common to both, or if such coincidence exists no more. In the first case, for events of different kinds, we are interested in knowing whether there is an association between them (although a real association relation is more than a coincidence relation), usually in the form of cause and effect. In the second case we explore coincidence relations among events of the same class, but different time, to see if the coincidence still exists or has changed.

In any case, coincidence relations are used to denote either a potential or an actual interaction.

Logic properties

The relation has the following logic properties:

- It is asymmetric; the value of the relation is not necessarily the same when measured from two coincident events. For event A the coincidence with event B may be total, that is, all instances of A coincidence in spaces occupied by B, but not necessarily all instances of B coincide always with instances of event A.
- It is not reflexive; the value of the relation is never the inverse when expressed from any event in coincidence with another
- It is transitive only if events are collinear. If event A coincides with event B, and event B coincides with event C, then A coincides with C only if A, B, and C take place in the same direction.
- It is additive only if events are collinear; if A coincides with B and B coincides with C, then the number of coincidences at A, B and C is three, and complementarity, A, B, and C coincide twice.

Topologic properties

As with the other two relations of this group, topologic properties are only partial. Invariance occurs in the cases of translation, scaling and warping, but not in the case of change of dimensions and rotation. It is a relatively topological relation.

A coincidence of two events, as observed in a map, may change to no coincidence when both events are observed in 3D and the space is rotated so that the angle of vision is orthogonal to that of the map. This is the case, for example of the apparent coincidence of a road crossing a river on a map. It may be that in reality the road crosses the river by a bridge 90 meters above the river level, so that there is no sharing of space between the river and the road, no coincidence in 3D.

It is important to be aware of the types of events for which we want to establish a relation: real events or representations of events.

Qualitative expressions

Since coexistence of events has an underlying cause, there is always a process creating coincidence relations, and therefore most of the terms used to express a relation of coincidence are verbs reflecting these processes. There are fewer nouns and even less adjectives adverbs and prepositions.

Verbs such as *to cross*, *to overlap*, *to match*, *to place*, indicate that an event enters in coincidence with another trough some kind of action. Nouns are used to name the places where coincidence occurs as in *intersection*, *region*, *portion*, *somewhere*, in the same way as adverbs of the type of *wherever*, *on top of*, *anywhere*. Adjectives such as *occupied*, *covered*, *crossed*, denote the existence of a coincidence relation, or the absence of it as in their opposites *uncovered*, *vacant*, *uncrossed*, etc.

Quantitative measurements

The common way of obtaining the value of a coincidence relation is through overlying the cartographic representations of real events. There are many ways of performing overlay operations. Some depend on the data structure of digital maps to overlay.

For vector maps, the most common overlay operations are intersection and union. The results in both cases are new maps where the coincidences between overlaid events are described in the associated tables. For intersection, it is possible to calculate the amount of space, i.e. the area, in which two events are coincident. For union, besides the amount of area in coincidence, it is also possible to calculate the amount of area of no coincidence, and therefore the possibility of expressing coincidence as a proportion of the total area of possible coincidence.

In the case of raster structured maps, there are more possibilities, although the measurements of coincident areas might be less precise than with the vector operations. But, in addition to the amount of area in coincidence, the identification of coincidences can be used to establish whether a particular set of events or a particular set of values in coincidence are appropriate to perform certain uses of a space, and to what degree are they appropriate . This is usually called a suitability analysis. This type of analysis can be carried out with vector representations as well, but in case that the values of events are measured in a continuous scale, the operations with vector maps would be not appropriate, and in sometimes even impossible to perform.

Coincidence relations between events might also be used to build models of the co-variance of the events, such that it could be possible to forecast the values of an event given the values of another.

CHAPTER 6

ORGANIZATIONAL RELATIONS

Organizational phenomena have a remarkable impact on the structure of space, to the point that they can create a paradox. Consider the following. It is incontrovertible that whenever a distance has to be traversed there is a certain amount of time involved in the movement. Thus, if I move 10 cm from my current position, anybody would take it for granted that I moved, and that I also spent some time in that movement. But is it true from an absolutely point of view?

The answer to the above question is that it depends. Although all parts of my body did move, I, as a person, still occupy some part of the space I was before, and for a location in that previous space I have not moved at all, I am still there. In fact for a certain range of distances (that depend on a person´s physical complexion), it can be said that a person does not move even when parts of the body have moved. From the point of view of a cell in my body I have not moved. For that cell, I am omnipresent. Wherever it moves I am there.

The reason to this apparent paradox is that an organization such as me (an organism) is more than the sum of its parts, or in this case, more than the sum of the individual actions of its parts. Depending on the way it is organized, any organization can be in many places at one time. This has profound implications, the most important being the capacity to interact in many places without the need to move. This also implies that organizations can act (almost) instantaneously within the range of distances spanned by the extent of their 'organisms'. As we shall see, the range of places an organization can cover without movement is limited mainly by the type of organizational relations that dominate the organization.

The ultimate implication is that organizational relations create a space in which the effects of the properties of space are reduced or nullified, and therefore, the properties of events dominate the way spatial interaction occurs.

6.1. CONNECTIVITY

Definition

A connectivity relation is defined as:

The condition of events of creating and maintaining links between them

This is the most basic of the organizational relations. The relation is a product of the capacity of events for connecting. Whenever a connection is created, space structure is modified into an organized space where interactions take place preferably through the space occupied by the connection rather than through the rest of the space.

This has an effect on the space properties of distance, direction and concentration, thereby modifying the relations of proximity, orientation and exposure among events.

Links, flows and connections

Organization is not an accidental process. It has a purpose. It arises from the impulse of translating idea into form. Forms are concentrations of matter or concepts. The forms of these concentrations are regulated by the idea of how to assemble matter or concepts into complex events. In order to assemble matter or concepts it is necessary to create elements that facilitate interaction of matter or concepts. Those elements are called links and the resulting assemblage of matter or concepts using links is an event.

An event is an organized form, that inherits the capacity to create links from the organization process, and by propagation they can produce other similar or more complex events. Inheritance creates links and flows in a top-down or bottom-up direction; propagation creates links in a left-right or front-back direction. Inheritance-and-propagation-produced links are organized so that the event can perform its functions in the best possible way. Functioning requires interaction among events or parts of events, and this interaction takes the form of flows.

Whenever there is a link there is also a flow. Flows are interchanges of matter, energy, concepts, ideas, goods, money, etc. among events. Links and flows are the basic elements that events use to organize space. Links are also events in themselves, and their main function is to conduct flows. When a link carries a flow we call it a connection.

The concept of connectivity

Connections are used to allow and facilitate interaction. When an event is connected to another, interaction among them is taking place mostly along the connection. This happens because the connections are so designed as to optimize the effort of interacting, by optimizing distance, direction and concentration. When these properties of space are thus modified, they no longer dominate the ways interactions occur. Instead, events impose their own properties to define how interactions should take place.

Connectivity is then the fundamental tenet of organized interaction. We could delve into why (apparently) non organized events create links to get organized, but we are still far from understanding such fundamental questions. Meanwhile we have advanced in the understanding of how interaction occurs through connections. A key fact we have learned in that respect is that sometimes is the number of connections that matters, and some other times it is the quality and quantity of the flow they carry what is relevant to the organization.

Logic properties

The relation has the following logic properties:

- It is asymmetric. The value of the relation is not necessarily the same when measured from one of two connected events. If event A has a connection with event B, it does not necessarily follow that the flow in the connection is the same in both directions.
- It is not reflexive. The value of the relation is never the inverse when expressed from any event connected with another.
- It is sometimes transitive. If event A is connected to event B, and event B is connected with event C, then A may be connected with C through B.

- It is not additive. If A is connected with B and B is connected with C, then A has one connection, B has two connections, but C has only one instead of three.

Topologic properties

As with the other two relations of this group, invariance to all transformations of space is the constant. That makes the relation an absolute topological relation.

Why is that transformations of space such as translation, rotation, scaling, etc., do not affect these relations? The reason is that topological transformations affect only the structure of space, where structural and neutral relations regulate interactions, but not organized space, where the properties of events dominate and dictate how interactions take place.

Qualitative expressions

Most connectivity terms are nouns. The explanation for this relative abundance is that there are many types of events that represent connections. Examples of these are *limit*, *bridge*, *track*, and *path*. Notice that these terms do not necessarily refer to physical events but they may be used to denote conceptual or virtual events as well.

Verbs are second in abundance. These indicate mostly the qualities of flows, but are also used to express actions that create or modify connections, as in *stream*, *emanate*, *lead*, *cut*, *open*, *join*, etc. Adjectives are the fewer, they specify different states of flows and connections, as in *flowing*, *continuous*, *joined*, *unconnected*, etc.

Quantitative measurements

Based on the nominal quality of being or not being connected, the simplest measurement that can be made is the number of connections an event holds. This can be extended into a connectivity matrix, analogous to an adjacency matrix, where besides the number of connections for all events, we can calculate this number as proportions of other aspects of the relations, such as the proportion of connections of an event in relation to the total number of connections existing in all events, or the proportion of events that have at least *n* connections in respect to the total of events, and some other measurements.

In a more geometrical context, we can use the length of connections as indicators of the characteristics of interactions. More complex measurements can arise from the representation of interconnections as graphs, such as routes, circuits, diameters of networks, paths, depth and width of trees, and many others related to the optimal configurations of graphs.

Flows are also subject to measurements, the most common are quantifications of magnitude, that is, the amount of flow per length or area unit, and of the intensity of flows, usually as velocity or the amount of flow per time unit. Other measurements may focus on flow directions and diversity/purity of flows.

6.2 AGGREGATION

Definition

A relation of aggregation is defined as:

The condition where events are combined to create another event.

The relation is a product of the capacity of events for combining. Connectivity is a precondition for aggregation relations. But in aggregations the quality and quantity of the flow they carry is more relevant than the number of connections.

This is the most common relation in events, because most events in the real world are aggregations of some sort. Examples of this kind of events are: an ecosystem, a termite mound, some parts and types of cities, a mountain, a forest, some types of industrial clusters, a watershed, the traffic, some types of agricultural systems, to name a few.

Emergence without (apparent) purpose

When events of different kind interact combining their characteristics and functions under principles of self-organization, a new event is created. This event is called an emergent event. Emergents are events of higher order than those that combined to form it. The emergence is not only of the self, but also of properties or functions. These events have two categories of properties, those that derive from the individual properties of the events that form them, and those that sprout from the combination of individual properties.

The process of self-organization that leads to the combination of events of dissimilar nature and to the emergence of an event is not yet fully understood. A plausible explanation is that random fluctuations in the interactions among events trigger organization, and this is reinforced by positive feedback. The term *random* is used here in the sense of an ‘unknown law’ not as ‘chance’ or ‘luck’.

Emergent events formed in this way are called aggregations and the events forming them are called aggregates. Spatially, these aggregates are relatively close to each other, and therefore the aggregations span a relatively limited extension, for example, forests, animal bodies, hurricanes, etc. However, there are aggregations that occupy vast extensions (from our perspective), such as planets, solar systems, and ultimately the universe.

In aggregations, organization is not guided by any of the events forming the emergent event. Moreover, the low-order events are totally unaware of the high-order, organized, event they form. It is a decentralized organization, in which all low-order events play important roles, in proportion to their levels of interaction. This gives the emergent event a strength that manifest especially when is subject to stressful conditions or damage resulting from its interaction with other events. Aggregations are highly resilient.

Paradoxically, the strength derived of organization being shared by all events, is also the weakness of emergent events, because in order to attain such coherence they have to maintain strong dependence bonds among them. If one these strong bonds cannot be repaired or regenerated when damaged, then the entire event may disappears. Aggregations are highly sensitive.

An aggregation where the aggregates are mostly inanimate or biological (rock, soil, air, water, heat, plants and animals) usually has not severe problems to preserve its functioning as aggregation, even though some changes may occur due to sudden alterations of one of these elements: overpopulation of some species, an extraordinary storm, a gradual increase in air temperature along many years, etc. Self-organization mechanisms such as feedback and auto-regulation are able to cope with the adaption actions required, because the aggregates do not interfere or modify these two mechanisms. However, when some external elements purposely or unintentionally (but drastically, in intensity and / or magnitude), modify or interfere with these mechanisms, the aggregation enters a state of disequilibrium producing a reaction which sometimes seems chaotic or catastrophic, but that is nonetheless directed to recover a new state of order or a new dynamic equilibrium. Aggregations with a high number of strong interdependences are the most sensitive to changes in the events that form them (including connections between events), risking disappearance if the number and/or nature of changes go beyond critical thresholds.

Combination of characteristics and functions of low-order events follow laws. These laws are not instructions but only conditional rules that dictate the limits and directions of interactions among events, which ordered into a process, can create a pattern or an aggregation. In this process there is no apparent purpose, but the fact that we cannot perceive it does not mean that it does not exist.

The concept of aggregation

Aggregation can be conceptualized as the process of bringing together several events of different kind, to create a new event. Thus, aggregation can be regarded as a process, but also as an event in itself and as relation as well. As events, aggregations are made of other events, the aggregates. As processes, aggregations are sequences of interactions among aggregates. As relations, aggregations are ordered combinations of characteristics and functions of aggregates.

Another form of conceptualizing aggregations is as systems. Any organization is also a system. Systems have structure, which is given by the elements that compose it; behavior which may be deterministic (as most natural aggregations) or stochastic (as most social aggregations); and interconnectivity, which is the way processes are performed by the elements in order to produce a specific behavior. However, this view is too mechanistic, it conceptualizes organization in terms of inputs, outputs, control, feedback, environment and boundaries, failing to capture complexity and offering no explanation for emergent behavior.

The state of being a 'Part Of' is the characteristic condition of aggregates, while the characteristic condition of aggregations is the state of being more than the sum of its parts, that is, a ‘Whole’.

This part-whole duality is the building block of aggregations, persistently manifested throughout many levels of aggregation.

Organization in levels is an important characteristic of all aggregations, being more perceptible in complex ones. At each level aggregates maintain aggregation relations resulting in new events, which in their turn also maintain aggregation relations and create new events, only perceivable in higher levels. Actually, our whole perceived universe is one continuous process of aggregation, whose beginning and end we have not yet found. Because of intellectual limitations we choose to separate this continuous process into seemingly distinct events, at different levels of aggregation, to facilitate its comprehension. Unfortunately, in doing this artificial breakdown we obtain a fragmented view of the whole, which we use to guide our interactions with aggregations, frequently with undesirable consequences.

Aggregation results in the organization of space. Organized spaces facilitate interaction by optimizing distances, directions, and concentrations of events, so that the functioning of aggregations tends to be optimal, a condition which is known as dynamic equilibrium.

Aggregations are dynamic events. Consider a city as an example of a spatial relation of aggregation. The city where we were born and grew surely experimented changes as we grew up, but those changes took place in such a way that the city did not lose its identity. The city expanded, became more populated, perhaps more noisier and polluted, some buildings disappeared and some others were build, some streets were widened or lengthened, some green areas appeared some other disappeared, land use changed from residential to commercial, public services increased and improved, public administration expenditures raised, tax collection built up, a political party governed for a certain period only to be displaced by another one, and, in general, quality of life may have improved or worsened, but the city remained the same high-order event, yet many of its aggregates may now be different and the number of emergent functions is perhaps greater. Identity remains because non-essential spatial events changed to the point of disappearance while key spatial events were only allowed to change to a certain controlled extent.

The above example introduces another important feature that characterizes aggregations, that of key events. In all aggregations there exist a set of events that, if modified beyond certain thresholds, may disrupt the aggregation relation to the point of making aggregations become instable in an irreversible way.

Properties

Since interactions among events in an aggregation relation may adopt the form of any of the other spatial relations, it makes no sense to talk of logic properties anymore. Instead, we can describe how properties held by individuals are shared in the aggregation:

- All properties held by the aggregates as parts are *inherited* to the aggregation
- Some properties deriving from the combination of the aggregates are back - *propagated* to the aggregates

Propagation and inheritance are the mechanisms through which properties are shared in the aggregation. All properties of the aggregates are passed on to the aggregation, but only some properties of the aggregation are passed on to the aggregates.

In a forest, for example, some properties of the individual plants that form it, such as the presence of a photosynthetic process, are inherited to the forest, in such a manner that the overall photosynthetic capacity of the forest equals the sum of all the individual photosynthetic processes. The inherited photosynthetic capacity of the forest is an emergent property of the aggregation. Other examples of these types of properties are the evapotranspiration capacity of the forest, the average number of roots per area unit, the total amount of water in soil and biomass at a particular time, etc. All of them are measures of the accrued values of single attributes of the individuals, reported either as averages or totals, or in some other form. Although they belong to the aggregation as a whole, they are not truly new properties, and that is why they are considered as examples of what is called weak emergence.

Some other emergent properties concern only to the forest, such as the mean height of trees, the levels of clustering or dispersion of trees, or the percentage of different plant species composing it. These are not shared with the aggregates. However, there are other properties which are imposed on the vegetation events, for instance, the level of vulnerability to plant diseases. The forest vulnerability to a particular disease is not only the simple sum of the individual health states of each tree or plant, but depends also on the aggregated spatial interactions (proximities, orientations, exposure with respect to dominant winds, orientation respect to the sun, coincidence with certain types of soils, etc.) derived from the specific spatial distribution of all plants suffering a disease and all plants in healthy condition within the forest. This integrated (aggregated) condition is propagated back to the individuals, such that part of their individual level of vulnerability to a disease is contributed by the overall level of vulnerability of the forest. Emergent properties that belong only to the whole, whether or not are back propagated to the aggregates, constitute examples of what is called strong emergence.

The existence of aggregation relationships among two events may lead to dependencies causing effects in two events located far away from each other. For instance, a flood in a factory leading to a stop in its production process, may cause that dependent events, such as stores or consumers, located far away from the factory, experience negative effects in their own performance. This is an example of how vulnerability propagates from factories to stores and consumers (the aggregation is formed of economic units, what is commonly known as an economic *cluster*). Even though, the only event directly vulnerable to floods is the factory, because of the aggregation relations consumers and stores maintain with the factory, vulnerability propagates to these events.

Qualitative expressions

There are relatively few terms to express specific aggregation relations. One reason is that most events are aggregations of some kind, hence sharing features common to any aggregation, what lessens the need to have many specific terms. Another reason is that interactions in aggregations may take the form of any other relation, therefore expressing these interactions requires terms that are already used for other relations.

Nouns such as *division*, *piece*, *congregation*, are used either to name generic parts or wholes. Verbs serve to indicate actions that result in aggregation or in disaggregation, such as *to concentrate*, *to gather*, *to take apart*. Adjectives designate states in the aggregation, as in *complete*, *joined*.

Quantitative measurements

Given the vast amount of events of different kinds that may hold aggregation relations, it would not be convenient to describe specific forms of measuring aggregation. Instead, we can say that in general there are four categories of measurements:

- Those that measure the number of different types of aggregates.
- Those that measure the magnitude and direction of interactions between the aggregates .
- Those that measure the number and characteristics of levels in the aggregation.
- Those that measure the number of different functions of the aggregation.

The first two focus on the characteristics of the aggregates and their interactions. For example, if we count the number of types of events that intervene in the aggregation, we can say something about the diversity of the aggregation. In general, the more diversified the aggregation, the more sensible to changes, but also the more resilient it is (compare a tropical forest with a temperate forest in terms of diversity, sensibility and resilience). We can also choose to measure spatial interactions. For example, by measuring connectivity relations, in the form of the amount and directions of flows of some kind, we may establish the dominance of some events over others, and then we can start understanding dependencies. Measuring interactions also signifies measuring proximity relations, inclusion relations, coincidence relations, exposure relations, etc.

The last two categories of measurement concentrate on the aggregation as an emergent event. Measuring the number of different levels of aggregation within an aggregation, gives us an indicator of its complexity. The more levels it has, the more complex the emergent event is, and the greater the number of specialized functions it performs. Also, the differences in the number of events at each level may indicate where within the aggregation are located the key events, those that perform a critical function or a great number of functions and therefore are more important for the performance of the aggregation as a whole. In general, levels with lots of events of many types, indicate that the emergent event at the next higher level is a key event.

6.3 ASSOCIATION

Definition

An association relation is defined as:

The condition where events are grouped to create another event.

Another type of organization results when high-order events decide to group themselves or impose grouping to other events. The mere fact that the events decide to group or to impose grouping in order to create a higher-order event implies the existence of some level of intelligence. Intelligence comes from Latin *intelligere* that means 'to perceive', 'to realize', 'to understand'. A pack of lions, a school of fish, or a flock of ducks, are examples of the lowest, biological, types of associations. The corresponding level of intelligence is called instinct, and relates more to the *intelligere* concept of perception. Human society, a federation of states, a mall, and a cluster of industries, are examples of the higher, social and economic, associations. *Intelligere* here takes the form of intellect, corresponding to the notion of understanding.

As in aggregations, connectivity is a precondition for associations. But unlike aggregations, in associations it is the number of connections what is more relevant to the association, while the quality and quantity of the flow they carry is less relevant.

Awareness and purpose

The Wikipedia definition for awareness is: '… a state or ability to perceive, to feel, or to be conscious of events…'. According to this definition, only animals, including us, have the capacity of being aware. If plants and rocks are aware, we do not know yet.

If the organization is formed by events that decided to be organized, there is usually full awareness of the relation, although the events may not be fully aware of all the functions the organization performs. If the organization is formed by events that were brought into association by the decision of an external event, then, this is the only event aware of the organization.

If awareness exists, then the desire of purposely interact with the perceived events arises. If purposeful interactions are effected in an organization, the organization starts to change from one where awareness does not exists to one in which events are perceived and interactions do not fully follow laws. Awareness in organizations is bidirectional: events may be aware of the organization they form, but also the organization may be aware of the events that form it. There is no awareness in aggregations, if awareness appears in some events in an aggregation these events cease to be part of the aggregation. Although they may still interact with other events in the aggregation, they are external events interacting with the aggregation, which may compromise the stability of the aggregation because the external events may act on purpose and interfere with feedback and auto-regulation mechanisms.

This transformation of an aggregation through awareness of the aggregates might not be a totally undesirable characteristic for some types of aggregations such as car traffic. As a spatial aggregation of streets, cars, and drivers, traffic initiates without awareness (or unified purpose) on the part of the drivers (and none on the part of cars and streets, of course). As soon as a traffic jam occurs, drivers start to show an increasing level of awareness about the traffic, including their own roles in it, and since this situation is not beneficial for them as individuals, they start to work in association with other drivers to act purposely to eliminate the jam (for instance, by willingly let other drivers pass first), until it fades away. Once the problem is solved they reenter a condition of aggregation as part of the traffic. Awareness in organized events is the beginning of associations; awareness is the precursor of purpose.

The concept of association

The events that decide to group together in order to form an organization are called associates. The resultant organization is the association. The characteristic condition of associates is to be a 'Member of', while the characteristic condition of associations is a 'Group of'. In a concentration of spatial events, if there is no purposeful interaction, they do not form an association but only a concentration of spatial events (a bunch in the worst case; a set in the best case).

This relation is also a process, and an event in itself. As events, associations are made of other events, the associates. As processes, associations are sequences of interactions among associates. As relations, associations are ordered groupings of characteristics and functions of associates.

Likewise aggregations, associations are emergent events too. However, they are not derived from a self-organization process, but rather from an ordered process directed by some agent or agents. Unlike aggregations, events in a single level of an association are all of the same kind. This restricts the number and types of interactions to a range of those possessed by the associated events. Compared with aggregations, associations are very simple organizations. In fact, most associations are organizations composed of two levels. Their behavior is non-deterministic, either possibilistic (as most natural associations) or stochastic (as most human associations). The associates in the association act as individuals (members of) not as fragments (parts of) as in an aggregation. Although the low-order events can control the interaction and therefore the properties of the high-order event, unlike in aggregations, in associations they do it purposely for the benefit of the association, and not only for their own, at least in theory.

The reasons to create associations correspond to the EEE principles of planning:

- Efficacy, or the capacity to produce a desired effect or achieving a goal.
- Efficiency, or the capacity to realize the effect / goal with minimal waste, expense or effort.
- Equity, or the capacity to fairly distribute among the associates the benefits or the costs of achieving the effect /goal.

A spatial association is not necessarily created to meet the three requirements. Even if it does, it may not be able to fully meet them, because the achievement of the three requires a great deal of intelligence. The required levels of intelligence to create an association based on those principles are related to the order in which the EEE Principles are mentioned, with the first requiring the least degree of intelligence, and the last demanding the highest level of intelligence. Most associations are initially created to meet the first principle only. When the levels of awareness increase, they may include the second and the third.

Purposefulness marks a fundamental difference between aggregations and associations. While in an aggregation, an effect may be produced as result of the aggregation process, and this effect may even be produced with efficiency, and perhaps, with equity, there is no intention of doing it so, whereas in associations these principles are precisely the *raison d'etre*.

The EEE principles are not goals of aggregations, though in most of them one can observe that the function performed by the emergent event might be done in a remarkably efficient way, this is entirely a result of auto-regulation and feedback mechanisms and not of a conscious effort of the aggregates to produce a specific effect. And of course there is no the minimum intention of being equitable. In aggregations, efficacy, efficiency and even equity, are the byproduct of physical, chemical, social or economic laws which guide and constrain the interaction of the aggregates.

In Chapter 3 we manifested our amazement at the fact that constellations (totally random events) had such a great importance in defining spatial relations for navigation purposes. This is a wondrous example of how intelligence can decide to create an association of events (stars) that have no immediate or actual relations by themselves, in order to derive a benefit from such association, in this case to comply with the principle of efficacy, by achieving the goal of helping to travel to remote places with some certainty. Notice that, in this example, association is artificially imposed on a set of events that by themselves would not associate. Also notice that interactions among them are also artificial, and reduced to interactions defined by proximity and orientation relations on the apparent plane of the sky.

Unlike an aggregation, an association does not have key constituents but only events that might have more or less influence on the functions of the association, but do not determine its existence. The existence of an association is a function of all of its constituents; only if all constituents disappear or decide to dissolve the association, it ceases to be.

Properties

Since interactions among events in an association relation may adopt the form of any of the other spatial relations, it makes no sense to talk of logic properties anymore. Instead, we can describe how properties held by individuals are shared in the association (notice that they are the opposite of aggregation properties):

- Some properties held by the associates as individuals are inherited to the association

- All properties deriving from the grouping of the associates are propagated to the associates

Propagation and inheritance are the mechanisms through which properties are shared among the association and the associates, but unlike aggregations, only selected properties of the associates are passed on to the association, while all properties of the association are passed on to the associates. The opposite of what happens in aggregations.

For example, in a mall, which is a spatial association of businesses, only the individual properties related to the capacity for attracting customers are inherited to the mall as a whole, in such a manner that the total capacity of the mall for attracting customers equals the added capacities of the individual businesses. This inherited overall capacity for attraction is an emergent property of the association. Other properties of the associates in a mall such as the amount of annual revenue, or the physical size of the department stores are not as relevant for the association and therefore not inherited to the mall, because the purpose in creating the association is to attract as many customers as possible. Because this emergent property is not truly a new property, it is considered as an example of what is called weak emergence.

Notice that in the example of the mall, the only desired principle is efficacy: to meet the goal of attracting the greatest number of customers. Also notice that this is a simple association with only two levels of events, the businesses and the mall. Lastly, notice that businesses do not interact too strongly among themselves, dependencies are loose, and there are no key constituents.

Some other emergent properties are truly new properties, such as the amount of space available for businesses, the connectivity of the site where the mall is located, the area of influence of the mall (the market area), the level of competition with other malls, etc. All of these properties are shared with the associates, affecting them positively or negatively. The sharing of benefits and costs implied by these mall properties is accomplished through propagation. Emergent properties that belong only to the association and that are back propagated to the associates constitute examples of what is called strong emergence.

Unlike aggregations, the advantage of not having key constituents is manifested when some businesses leave the mall because of convenience or lack of enough revenue: in spite of losing constituents, only the function of the mall might be affected somewhat by losing capacity of attraction, but the association continues to be without risk of disappearance. Associations have low sensibility to changes, but also low resilience.

Qualitative expressions

For the same reasons that in aggregations, there are very few terms to express the relation, in associations the number of terms is also reduced to those that describe common features of these organizations.

Nouns such as *division*, *section*, *meeting*, are used either to name generic parts or wholes. Verbs serve to indicate actions that result in association or in disassociation, such as *to distribute*, *to assemble*, *to split*. Adjectives designate states or positions in the association, as in *equal*, *single, last, bigger*.

Quantitative measurements

Given the diversity of association relations, it would not be convenient to describe specific forms of measuring association. Instead, we can say that in general there are four categories of measurements:

- Those that measure the number and characteristics of associates
- Those that measure the degree of compliance with the EEE principles
- Those that measure the number and characteristics of levels in the association
- Those that measure the number of different functions of the association

The first category focuses on the amount and characteristics of the associates. Since in each level of associations all events are of the same kind, the number of events that intervene in the associations becomes the first indicator of the potential to meet the goals of the association. As complementary measures, we can weight this potential by the individual characteristics and capacities of the associates. For example, in a mall, the number of businesses may give an initial idea of the potential to attract customers. But to have a better idea, we can weight this potential by the diversity of commercial services and/or by the individual capacity of businesses to attract customers: big department stores usually have more resources for publicity campaigns and therefore better capacity to attract customers.

The second category targets the forms in which associations carry out their main functions. For instance, in the public administration of a territory, the government -a spatially distributed and multilevel association of country, state, municipal and community authorities and public servants - may perform the functions of governance and provision of public services with varying levels of efficacy, efficiency and equity. Measurements of these levels can indicate how well the association is meeting its goals, but also where in the association there might be the situations responsible for the measured levels of performance. For example, whether there is deficiency or excess of associates or resources to perform the intended functions, whether there is the need to reorganize spaces in order to optimize the provision of public services, etc.

The last two categories of measurement concentrate on the emergent properties of the association. Measuring the number of different levels in an association, gives us an indicator of its complexity. The more levels it has, the more complex it is the emergent event, and the greater the number of specialized functions it performs. In contrast to aggregations, associations perform a very limited number of functions. For that reason, the number of functions as such is not a very useful measurement, however when related to the number and characteristics of associates it may become and indicator of the load upon each associate.

6.4 FINAL NOTES ON SPATIAL ORGANIZATION

As they are being created, connections, aggregations and associations, modify the original structure of space to conform to specific structures, which are optimized for interaction. The basic forms of these optimized structures are networks and hierarchies. As organizations evolve it is common to find combinations of these two structures.

Networks are structures used to organize events in a single level. Hierarchies are structures used to organize events into multiple levels. Hierarchies are higher order structures than networks because unlike them, order relationships are explicit in the hierarchy's levels. In addition, spatial events situated at every level of the hierarchy, are emergent events with emergent properties. In contrast, the spatial events located at the ends of connections in a network are only connected events, they are not emergent events.

When the connected events in a network create an emergent event, a new, higher, level arises, thus producing a two-level structure in the form of a hierarchy, the simplest form of an organization. The first level is the single high-order event created through the interaction of low-order events, and the second level is a network of low-order events. This is the building block of all organizations.

The structuring by levels represents an advanced characteristic of an organization because at each level functions are specialized, and specialization can only derive through evolution, therefore we could say that hierarchies are more evolved than networks. This does not mean that hierarchies are better structures than networks, just that they are more complex; we cannot apply the qualifier "better" to structures because in order to do that we would need to know the purpose or function to be performed by a particular form of organization. This means that an organization needs both structures in order to accommodate different types of interaction.

Networks accommodate same-order interactions. Hierarchies support different-order interactions. Same-order interactions are those that take place by events of the same level of organization. Different-order interactions are those that occur among events at different levels of organization.

As the organization becomes complex, the events located at successive and equal levels in a hierarchy become events connected in a network structure, in a cyclic manner.

CHAPTER 7

A GEOGRAPHIC APPROACH TO VULNERABILITY ASSESSMENT

SAVE: A GEOGRAPHIC APPROACH TO VULNERABILITY ASSESSMENT

(Spatial Analysis of the Vulnerability Environment)

Luis M. Morales-Manilla

Centro de Investigaciones en Geografía Ambiental, Universidad Nacional Autónoma de México,
Morelia, MICH. C.P. 58190

Abstract

The assessment of vulnerability has been approached from many disciplinary points of view, including geographic ones. However, no approach has obtained universal recognition as a comprehensive solution, nor have those proposed by geographers agreed on their perspective and methods. The lack of success in the first case can be partly explained because vulnerability is truly a complex event requiring interdisciplinary work. In the second case, there is no doubt that geographers have contributed substantially, but the absence of a common geographic perspective can be hardly justified by vulnerability's complexity, leaving unclear what the contribution of geography might be. Despite recent advances in the comprehensive conceptualization of vulnerability, comprehensive methodologies for vulnerability assessment, if any, suffer of a number of drawbacks. This article presents the design of a geographic model for the assessment of vulnerability, based on the concepts of place, spatial relationships, and pattern. I show here that these concepts can be successfully used to build a comprehensive geographic model that integrates biophysical and socioeconomic elements of vulnerability. In a vulnerability assessment using the proposed model, the idea of place serves to define the study units; the notion of spatial relationships is the guiding principle for the vulnerability analysis and the selection of vulnerability indicators, while the concept of pattern provides the objectives of the assessment. Moreover, the use of spatial relationships as, or to build, vulnerability indicators, has the advantages of providing across-scale, multi-hazard, and place-independent indicators.

Key Words: vulnerability, pattern, spatial relationships, place.

7.1 INTRODUCTION

Whatever its origin, natural or human induced, Global Environmental Change (GEC) is a reality. The understanding of GEC is important because it may pose hazards to the human-environment system (HES) that can threaten our established or desired lines of development. But whether it is feasible or desirable to manage or to adapt to GEC, we must have knowledge on those aspects of the GEC and the HES that can potentially impact an established or desired line of development, namely:

- The type and characteristics of hazards posed by the GEC, including natural and anthropogenic, and their environmental interactions.
- The type and characteristics of the vulnerability of the HES, and its environmental - development interactions.

- The levels (in terms of magnitude and intensity) and evolution (in terms of variability and trends) of risk conditions caused by the combination of hazards and vulnerability.

Sometimes, knowledge on some critical characteristics of hazards such as their possible impact, or their time of occurrence, is not easy to obtain (e.g. earthquakes, tsunamis). In other cases, even with full knowledge of the hazards we face, we can hardly act because the amount of energy/time needed to control them is beyond our current capabilities (e.g. hurricanes, vulcanism). Some other times, the hazards surpass our capacity (or will) to organize ourselves and face them as a global and whole society (e.g. famine, poverty, terrorism). Dealing with hazards, although possible, may prove a difficult task. On the other hand, when dealing with vulnerability we can accomplish more towards the end goal of reducing risk.

The fast-growing literature on the subject (Musser 2002; Janssen 2006) shows that there is a concern for the vulnerability assessment of some, or all, of the following types of vulnerable (spatial) events:

- People, considered both as individuals and as groups (age groups, gender groups, income groups, ethnic groups, etc.).
- Economic activities, taken as the activities people engage in to produce capital or means for subsistence (agriculture, commerce, manufacture, services, etc.).
- Infrastructure, that is, the physical - functional assets used to support development (roads, buildings, institutions, organizations, etc.).
- Biophysical events, comprising all events of non-human origin present in nature (forests, rivers, mountains, wildlife, soils, etc.).

Specific conditions of development in the HES for any of these types of spatial events, coupled with specific conditions of GEC dynamics, mainly hazards, create the vulnerability environment. The vulnerability environment of a given spatial event can be assessed in terms of its exposure, sensitivity, and resilience to a hazard or group of hazards (Turner *et al.,* 2003). To do this, it is necessary to develop comparable metrics across types of hazard, scales, peoples, and places. This goal has proven difficult to achieve.

Vulnerability has been defined as the likelihood of the HES, or of any of its components, to suffer harm derived from exposure and sensitivity to a hazard, and the incapacity to recover and adapt once the hazard has caused an impact (IPCC, 2001b; Turner *et al.,* 2003). According to this definition, any vulnerability assessment would ideally seek to establish the degree of impact that the HES or its components can experience, given defined magnitudes of exposure and sensitivity, and to assess their ability to respond to that impact, with the ultimate goal of devising ways to reduce vulnerability. Currently, there are several approaches to the assessment of vulnerability (Cutter, 1996; Alwang et al 2001; Füssel and Klein 2002; Adger 2006) due in part to different disciplinary views and vulnerability subjects, and in part to the need to study and mitigate the effects of specific hazards which seem to be more acute or chronic for our society. Even within a single discipline, different conceptualizations exist (Hinkel and Klein, 2006).

Since vulnerability is a geographic event, geographers have been particularly active in the field. We could cite the works of Burton et al (1978), Butzer (1980), Timmerman (1981), Liverman (1990), Dow (1992), Blaikie et al (1994), Kasperson et al (1995), Cutter (1996), Watson et al (1996), Kates et al (2001), Turner et al (2003), and O'Brien et al (2004), to mention some of the most relevant contributions made by geographers. While it is clear that geographers have contributed substantially to the field, this contribution has been directed more towards the conceptualization of the event rather than to the creation of a common methodological approach to the problem of its assessment. Nevertheless, some outstanding attempts have been made to resolve this issue such as the "Common Methodology for Assessing the Vulnerability of Coastal Areas to Sea-Level Raise" (IPCC CZMS, 1992), the "Hazards of Place" model (Cutter 1996, 2000), or the more recent DIVA Model (Ionescu et al, 2005; Hinkel, 2005).

Without denying the exceptional value of conventional geographic research in advancing vulnerability science, it is also manifest from its many and distinct contributions that there is no a clear methodological perspective regarding vulnerability assessment. In my view, perhaps the most important reason for this failure has been the lack of a truly geographic theory supporting the approaches devised so far. Although some theories such as social constructivism or human ecology (political ecology) stand behind some of the most consistent attempts, these are not fully geographic theories. Endowments, entitlements, actors, livelihoods are concepts with a strong geographic character but shared with, or imported from, other social disciplines.

Thus, in spite of the many contributions geographers have made to vulnerability science, this lack of commonality in approaching the problem of vulnerability assessment lefts unclear what exactly is the contribution of geography as a discipline. Should we be satisfied with the traditional disciplinary approach to vulnerability assessment, or should we be concerned instead with the development of a common geographic approach? If what geographers do in the field of vulnerability increase our knowledge of the event, why should we bother with a common approach?

Besides the fact that by proceeding without theoretical guidance we are not building disciplinary identity in the field, more important and less obvious is the fact that in doing science without worrying about where it fits in science' greater design, we are not working to simplify complexity, we are only adding to this complexity. Vulnerability is a complex event, and although any piece of knowledge may increase our understanding, we might be forgetting that to understand is not to solve, and that only structured knowledge can help to solve complex problems. The mere gathering of pieces of knowledge does not lead to a solution; we must provide a structure, a theory, to accommodate all available parts of the solution. This has the additional advantage of making clear what we are missing, that is which parts of the structure are not yet there and require our attention. In the case of vulnerability assessment, an appropriate structure can be constructed with fundamental geographic concepts as we show here.

The other part of the answer to the above questions rests in recalling that the ultimate goal of knowledge is to use it to solve real-world problems, that is, to make concepts operational.

Operationalization of concepts requires systematization and standardization of knowledge. So, while we may continue increasing knowledge, we must also work towards a universal way of using such knowledge in solving the complex problem of vulnerability assessment with a standard methodology. Geography should be able to provide a systematic approach to that problem, by making use of geographic concepts standard to any geographic inquiry.

7.2 THE NEED FOR A COMPREHENSIVE APPROACH TO THE ASSESSMENT OF VULNERABILITY

Whether discipline, subject, or hazard oriented, some of the approaches developed to date for conducting vulnerability assessments are sound at the conceptual level, fewer are sound at the operational level as well, but none of the last can be qualified as comprehensive, in the sense of having wide applicability. Usually, when an approach becomes operational limitations for wide applicability arise because of any, or all, of the following:

- Data quality / availability for required variables are not adequate in all cases.
- The concepts in the conceptual model are too general, and functional relationships between them are undefined (Hinkel and Klein, 2006).
- Indicators used are not appropriate to measure the vulnerability of any type of event.
- The approach is scale, or hazard, or place specific.
- Local public officials find difficult to understand and apply the approach by themselves.

Hinkel and Klein (2006) argue that the design of a generic vulnerability assessment capable of providing comparable results across systems but that at the same time is specific enough to take into account the unique configurations of these systems should comply with two elements: (i) a common domain-independent conceptual framework of vulnerability and (ii) a well-defined process that specifies how the framework’s general concepts can be specialized to account for an specific case of assessment.

Most existing approaches are based on one of two conceptual models of vulnerability: the risk-hazard (R-H) and pressure-and-release (PAR) models. Turner et al (2003) have succinctly enumerated the advantages and deficiencies of both models, concluding they are not sufficiently comprehensive. As an improvement to those models they proposed an expanded model which addresses the coupling of the human and environmental systems and the nested scales of their interactions. Also, it explicitly decomposes vulnerability in exposure, sensitivity, and resilience, which in other models are considered as distinct or equal to the vulnerability concept, but not as part of. So far, this is one of the most comprehensive conceptual models of vulnerability, also complying with the domain-independency requirement of Hinkel and Klein’s first element. Other conceptual models have been recently presented as “comprehensive”, but are domain-specific (e.g. climatic change, Füssel 2005).

But equally important than to have a comprehensive conceptual model for vulnerability assessment, is to use a methodological approach just as comprehensive, in the sense indicated by

Hinkel and Klein's second element. Besides the indication of how to proceed from the general concepts to a specific situation, the approach should also comply with the following requirements (Cardona 2003; Cutter 2003; USGS 2005a; IIASA 2006; Adger 2006):

- The approach should capture the complexity of the HES's interactions, integrating biophysical and socioeconomic elements and factors of vulnerability, including issues of risk perception and local values.
- The approach should be transferable to any global or local vulnerability condition, and not be place, or scale, or hazard specific; this task requires the development of generic metrics, which must also incorporate the relativity of perceptual values.
- The approach should help public officials to reduce risks in their jurisdictions using indicators that are relatively easy to comprehend and apply, and that, in addition, could be monitored with limited or existing resources.

I will call these requirements the Integration Requirement, the Independency Requirement, and the Simplicity Requirement, respectively. The degree to which a given approach complies with them can be used to measure its "comprehensiveness".

Since the need for a comprehensive methodological approach can be considered as fully justified by the above demands, a fundamental question is whether geography can or cannot meet the challenge. This question can be extended to ask if indeed a single discipline can do it, which in turn leads us to ask whether it is feasible to devise such an approach, given the complexity of the field and the state of the art in vulnerability science. We will seek to answer these three questions by answering the first.

Looking at the three requirements, we can say that geography has the potential to meet the Integration Requirement, given the traditional human-natural focus in the study of geographic space. In fact, it is the only discipline of which this can be said (here we disregard Environmental Science or Sustainability Science as candidates, because they are an amalgamation of several disciplines rather than an integration), although the ever increasing geographic studies where only human or only natural elements are considered could make us doubt of the assertion. However, we can assume that geographic research focused solely on either the human or the natural part of geographic space is a researcher's choice issue (or a researcher's limitation issue in any case), that in no way lessens geography's potential for studying the human-nature interactions in an integrated fashion. Notwithstanding this potential, geography has yet to develop a sound methodology to achieve such integration.

The Independency Requirement seems also a 'natural' for geography, if only because scale, hazard and place are geographic concepts. However, what the requirement specifically calls for is scale-independent, hazard-independent, and place-independent vulnerability assessments. These are complex issues, demanding respectively the knowledge of:

- Scaling mechanisms. Ecology and geography, mostly, have contributed substantially to this end, although there is still a long way to fully understand across-scale linkages that determine systemic vulnerability.
- Generic modeling of interactions. The challenge is to go from the understanding of hazard-vulnerability specific interactions, to generic interactions as determined by the components of vulnerability. Particularly sought is the development of 'universal' metrics to express exposure, sensitivity, and resilience interactions of hazards and vulnerable events.
- Place as a functional unit. By emphasizing place functional commonalities instead of place physical differences, place could be used as the standard unit of study in vulnerability assessments, where place should not be entirely defined *a priori*, but as a result of the investigation of the functioning of a particular portion of geographic space.

Geography's strong tradition in network and regional analysis can serve to address the first and third issues, but more work is needed to understand and model generic hazard-vulnerability interactions at several scales.

The Simplicity Requirement becomes especially important when a vulnerability assessment must provide real world answers to allow for the specification of policies to reduce vulnerability. It is all well to use sophisticated science, or even ill-structured knowledge, if the assessment is an academic exercise carried out in order to advance the understanding of vulnerability. But whenever the end-goal is not fully academic, and end-users are not only scientists but politicians and the common citizens as well, science must be simplified or knowledge better structured to allow policy design and implementation. Without providing a full answer to this demand, as it is mostly the domain of decision theory and public policy theory, geography can offer some help in the simplicity of the indicators it can deliver, since, as we use many spatial concepts in everyday language, these have the possibility of being intuitively grasped. If complexity cannot be avoided in the study of vulnerability, at least we should strive at the specification of indicators that are easy to understand and monitor.

The potential of geography in meeting these three challenges must be interpreted as the possibility of the discipline to provide a better approach, rather than considering geography as *the* discipline that can provide the ultimate solution to the problem of vulnerability assessment. No single discipline can, although some are expected to contribute more than others.

I present in this article the guidelines to develop a generic geographic (spatial) model to the assessment of vulnerability that takes into account, without fully complying, the mentioned requirements. The model uses three fundamental concepts as the basis for the study of the interaction between nature and society: place, spatial relationships, and pattern. I have named the model as SAVE, Spatial Analysis of the Vulnerability Environment, to emphasize its spatial and analytical character and the main thematic focus.

7.3 THEORETICAL FRAMEWORK

It is necessary to remark that the concern here is with the theoretical concepts behind a generic spatial approach to vulnerability assessment, not with a conceptual framework of vulnerability. In fact, the conceptual framework on which the SAVE approach is based is that of Turner et al (2003), because of its comprehensiveness and the threefold structure of vulnerability it promotes.

Since the model pursues the goals of being standard and common to geographic research on vulnerability assessment, it must be based on geographic concepts possessing those qualities. In particular, it should use concepts lying at the very core of the discipline, reflecting the spatiality of geography's approach to problem-solving.

Three of such fundamental concepts are place, spatial relationships and patterns. These notions are integrated under the concept of space adopted here, where geographic space is deemed as a physical-functional space-time continuum where geographic events, natural and human-created, interact at specific places, with these interactions expressed as spatial (spatio-temporal) relationships, and resulting in specific distributions of events, i.e. patterns.

Place

As used in the SAVE approach, *place* is a generic concept, initially unbounded, used to make reference to a location, which in terms of extension, can range from the global, to the regional, to the local. Thus, a place can be the planet or a continent, a state or a region, a watershed or an ecosystem, a city or a village, even a home or the spot where a particular person, asset, economic activity, or biophysical event exists or takes place at specific times.

In making place one of the two the fundamental concepts of the SAVE model, we are following Golledge (2001), who maintains that place-based reasoning should be at the core of any problem-solving approach within the discipline. Also, we are building on the "Hazards of Place" model (Cutter 1996), in the sense that the notion of place allows for the integrated view of social and biophysical vulnerabilities.

The notion of place is fundamental to the approach because it provides the spatial units where the methodology is applied. A place, in the context of the approach, can be thought as formed by a set of interlinked places, usually nested, and the space occupied by the place can be a physical space, or a functional space (social space, perceptual space, economic space, etc.), but it is often a combination of both types. In particular, the SAVE approach defines a place by focusing on the interactions of specific hazards, human groups, human activities, assets (infrastructure), and biophysical events, which exist in the same location or in different but related locations. Thus, this notion of place is not that of an empty space, but rather, one of a space created by the interactions of these events, much in the way of Lefebvre's social constructivism (Lefebvre 1991), but expanded to include an "environmental constructivism" as well.

This embodies the idea of place as a functional unit where the emphasis is on the functional aspect of geographical space instead of on its physical differentiation. This concept of place could be used to delineate a more realistic unit of study in any vulnerability assessment, where the place of study should not be defined *a priori*, but as a result of the investigation of the functioning of a particular portion of geographic space. Since, in following this guideline, we could end with an impractical situation in which to fully understand the vulnerability of a very concrete place we should investigate the vulnerability of the entire planet (although for some places and vulnerable situations this is the only possibility), any (or the combination) of two mechanisms could be used to constrain the place of interest to a reasonable extension:

- Thresholding: implies establishing the degree of relevance of functional linkages among related places using some social, economic or physical proxy for vulnerability and setting relevance thresholds.
- Modeling: involves the creation of models of the systems that contain the place of interest, with the most external system being the most general model; the key point in modeling is to include relevant variables in each model.

Whichever is the mechanism employed to constrain the place of interest, to take advantage of this notion of place we must overcome the current practice of defining *a priori* what the specific area of study of our assessment is. While we can specify an initial area of interest, we should allow for some flexibility in the determination of the boundaries of the final area. Currently, the procedure is either to explicitly choose a natural unit, or worse, a political / administrative unit, as our only area of interest, or to define such area implicitly, by choosing a biophysical (e.g. a forest) or socioeconomic event (e.g. the poor) as the subject of our vulnerability assessment, in such a way that the area of interest corresponds with the physical locations and extent of the subject. Although, in some cases, any of both practices can be justified, we might be sacrificing the reliability of our assessment since, in a global world, the vulnerability of a place depends on the vulnerability of other places.

Spatial Relations

Spatial relationships are concepts of wide use in disciplines such as geography, geology, geophysics, astronomy, architecture, archeology, ecology, physics and mathematics (and their derived sub-disciplines, e.g. geomorphology, climatology, meteorology, soil science, etc.), and even in some that we do not consider ordinarily as space-aware disciplines, such as medicine, sociology, anthropology, political science, chemistry, engineering and computer science. Their importance to science lies in that they are useful to indicate interactions between spatial events at any scale, allowing the study of apparently dissimilar phenomena, such as floods or famine, using the same spatial approach.

The Theory of Space-Event Interaction, TSEI (Chapter 2) states that there are nine basic types of spatial relationships derived from the interplay of the structural properties of the space (distance, direction, concentration) and the organizational properties of events (capacity for connecting, capacity for combining, capacity for deciding). All nine types of relationships are common to any physical or functional space and to any place, with only the form of expression or calculation, or their values, being specific to a particular space or place. The TSEI establishes the occurrence of the following types of generic spatial interactions:

1. Proximity. When distances determine the magnitude-intensity of an interaction.
2. Orientation. If directions are influential in the existence of interactions.
3. Exposure. If concentration of matter, energy, or concepts, acting as barriers in geographic space, determine if events interact with each other.
4. Adjacency. When contacts between events define possibilities for interaction.
5. Containment. When containment of events by other events define some sort of order (or disorder) in the interactions between geographic events.
6. Coincidence. When coexistence of events in the same portions of an *n*-dimensional space establish possibilities for interaction.
7. Connectivity. When interactions take place through connections and flows.
8. Aggregation. If there are strong, unconscious, interdependent interactions among a set of events resulting in a high order event with emergent properties.
9. Association. If a set of events have the capacity of deciding how to interact to create a higher order event with emergent properties.

The first three types of relationships are termed structural because they result from the dominance of the properties of space (the TSEI establishes that the space without events has structure but no organization). The next three relationships are named neutral because there is no dominance of space or event properties. The last three relationships are called organizational because they result from the dominance of the properties of the events (the TSEI establishes that the events without space may have organization but no structure).

It is important to keep in mind that, all these types of relationships may take place in different kinds of spaces, but when speaking of geographic space structural relationships tend to dominate the physical part (although there are some exceptions), while organizational relationships might be more relevant in the functional part (with exceptions as well), with neutral relationships participating in both.

All types of relationships occur between two or more geographic events or parts of events. When considered in a dynamic setting, events and their relationships can tell us how spatial processes take place and therefore the arrangement of patterns. This is especially important for the SAVE model, where vulnerability is treated as a spatial event, albeit complex, presenting a spatial pattern resulting from a set of spatial processes, which in their turn are composed of sequences of spatial relationships occurring between a variety of spatial events. Hence, vulnerability patterns can be studied in terms of the spatial relationships taking place at three different levels of analysis:

- Between GEC and the HES (generic level)
- Between the environmental dynamics and development (thematic level)
- Between vulnerable events and hazardous events (specific level)

The concept of spatial relationships makes possible the design of a multi-hazard, across-scale, and integrated (social-economic-biophysical) model to the assessment of vulnerability, owing to the fact that they can be used as generic indicators, applicable to any particular vulnerability condition, of any vulnerable event located in any place, and at any scale. The notion of spatial relationships is also central to the SAVE approach because it provides the framework for the analysis and synthesis of vulnerability patterns.

Because these relationships exist in any place, the concepts can be used to build place-independent indicators. Also, all spatial relationships can occur between any types of spatial events, hence they can serve to indicate multi-hazard events and multi-vulnerable events interactions. In addition, the organizational relationships (connectivity, aggregation, and association) can help us to undertake the problem of specifying across-scales linkages.

Organizational relationships, because of their systemic nature, are a good mechanism to address the interactions between the global processes (e.g. climate change, oil prices, country-wide environmental policies, etc.) and the local processes (e.g. proximity to a river, exposure to hazardous substances, individual capacity to recover after a disaster) leading to vulnerability. These types of relationships help us, as well, to explain and account for the links and aggregated effects of vulnerability across systems at the same scale.

The use of organizational relationships may also help us to avoid pitfalls of data / properties aggregation such as the Modifiable Areal Unit Problem or the Ecological Fallacy. If a number of spatial events (the aggregates) hold an aggregation or an association relationship, a new higher-order spatial event (the aggregation) is produced, usually at a larger ontological scale whose spatial structure is not arbitrarily configured and modifiable, but defined by the interactions of the aggregates. This emergent event, in addition to some properties derived (propagated) from the aggregates, has its own higher-order properties, which may or not be passed on (inherited) to the aggregates. When these types of relationships are taken into account, the scale problem of a vulnerability assessment ceases to be because the vulnerability conditions propagate or are inherited through different levels of emergent events, from the lowest to the highest, based on the direction and magnitude of their interdependencies.

The importance of organizational relationships in a vulnerability assessment can be exemplified within a city and its hinterland, where the vulnerability of retail commerce may depend to some extent on the physical vulnerability of the people involved in the activity or on the physical vulnerability of the factories where the goods are produced. If, for example, the factories located outside the city (but within its hinterland) experience the effects of a severe flood, production may stop and certainly distribution of goods to the retail stores may not be possible. Even if retail stores in the city did not experience the flood themselves, their vulnerability to this hazard may be high because of the spatial relationships they hold with the factories. In other words, even if the exposure levels of individual units of retail commerce are low or null (as determined by proximity or coincidence relationships with the flood), sensitivity levels, measured in this case as the degree of dependence on other components of the system or on other systems (through connectivity and spatial aggregation relationships), may be high, with resilience levels varying according to the capacity of diversifying commercial activities and the possibilities of receiving help from commercial organizations or government institutions (as indicated by spatial association relationships).

Turner et al (2002) stated that "(vulnerability) prescription, based largely on the perturbation-stress or spatio-temporal characteristics of exposure will surely miss the mark in regard to impacts across systems". This assertion holds true when the spatio-temporal conception is reductionist, as when a GIS is used to map vulnerability without a theoretical framework of space behind. But when this concept of the spatio-temporal includes organizational relationships, there is no possibility of missing impacts across systems, as showed in the above example.

The design of the SAVE model was conceived with the main objective of developing and testing a comprehensive geographic methodology to identify, describe, explain, predict, and design (reduction) vulnerability patterns, useful across a range of hazards, vulnerable events, and places. Based on the two theoretical concepts described above, the design is guided by the following hypothesis:

- Spatial relationships describing the interactions between the patterns of environmental dynamics (such as climate change, land cover change, or the occurrence of hazardous phenomena), and patterns of development (population growth, land use change, urbanization, economic activities, etc.) at specific places, can serve as indicators of vulnerability levels to multiple hazards.
- These indicators can then be used to design policies aiming at the reduction of vulnerability, by modifying the values of key spatial relationships between vulnerable and hazardous events.

Thus, the model focuses on the interactions of environmental change elements and development elements in a place. Both sets of elements produce together a diversity of vulnerability scenarios where particular vulnerability levels can be measured using spatial relationships as generic indicators.

7.4 PROBLEM-SOLVING FRAMEWORK

When facing a vulnerability assessment, scientists and, especially, public officials are often confronted with the problem of translating a conceptual framework into operational procedures for the assessment. Disregarding the appropriateness and completeness of the conceptual framework to use, two questions require answer for a successful application of an approach to the assessment of vulnerability: How to proceed? What will be the outcome? The first question is a methodological one, involving both, the steps to follow, and the means to analyze and synthesize knowledge in a systematic way. The second question is one of concern for the utility of the assessment outcomes, specifically about the sufficiency and appropriateness of the results to solve a particular vulnerability problem (usually, the reduction of vulnerability).

The SAVE approach could remain a theoretical approach to vulnerability, unable to answer those questions, if not supported by a problem-solving framework. The SAVE approach is a theoretical model but also a practical methodology. In compliance with Hinkel and Klein’s (2006) second element, this framework describes the tools to structure and apply the theoretical concepts. It does it focusing on five concrete goals, where the key notion is that of pattern, defined as the spatio-temporal distribution of events.

Those goals are related to four basic steps of the scientific method, with the addition of a fifth goal related to planning / engineering (in their broadest sense):

- Search of vulnerability patterns.
- Description of vulnerability patterns.
- Explanation of vulnerability patterns.
- Prediction of vulnerability patterns.
- Design of vulnerability (reduction) patterns.

The diagram in Figure 7.1 helps to explain the situation of these goals within the problem-solving framework. The central portion of the diagram corresponds to those pieces of knowledge fundamental to any geographic inquiry, from the most elemental (values) to the most complex (patterns). Vertical arrows in this portion of the diagram indicate that an element is part of the next level of elements. Thus, for instance, interactions between spatial events define spatial relationships; sequences of spatial relationships create spatial processes, and the outcomes of processes are patterns. At both sides of the diagram we see the two conceptual devices that allow us to handle spatial knowledge. Horizontal arrows indicate the possibility to apply both devices to the same piece of knowledge. Together, spatial analysis techniques and GIScience tools form a powerful instrument to analyze and synthesize spatial knowledge. At the bottom of the diagram we have the five fundamental types of spatial problems (Chapter 1) corresponding to the goals of the SAVE approach. Below these we can place the name of any geographic phenomena, vulnerability in this case, to have a complete problem-solving framework.

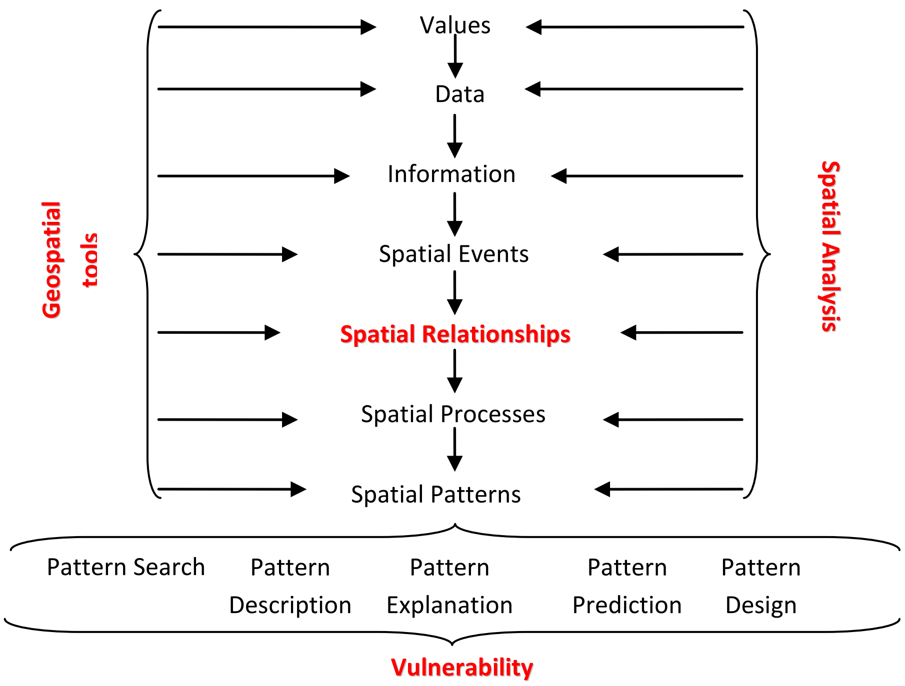


Figure 7.1 Problem-solving framework of the SAVE approach.

In this framework, the notion of place is implicit in the concept of spatial events, since all events occupy a place. The events refer to the five types of spatial events considered when assessing vulnerability: hazards, people, economic activities, infrastructure, and biophysical events. Since the events might initially have indeterminate boundaries, the places they occupy might as well have fuzzy boundaries. Such boundaries become less uncertain as relationships among events are being found and specified during the application of the approach.

According to its goals, the SAVE approach is structured in five phases:

- Phase 1. Search of vulnerability patterns. When a vulnerability pattern exists but it is not directly observable and needs to be revealed.
- Phase 2. Description of vulnerability patterns. When the pattern is observable, either directly or as result of the previous phase, but needs to be described.
- Phase 3. Explanation of vulnerability patterns. When we need to know the causes and mechanisms producing a specific vulnerability pattern.
- Phase 4. Prediction of vulnerability patterns. When it is necessary to have an idea of the future state of a vulnerability pattern.
- Phase 5. Design of vulnerability (reduction) patterns. When we wish to modify or create a vulnerability pattern, to make it acceptable or compatible with development and environmental change, either by decreasing exposure and sensitivity or increasing resilience.

The five phases can be executed in sequence, but it is very likely that as a result of work in the first phase, some knowledge pertaining to the other phases can be simultaneously attained. Each phase entails the application of the spatial knowledge elements of the problem-solving framework. Using those elements, the general procedure is to organize phase work in three stages:

1. Problem Structuring. Organizes knowledge, identifies and structures relevant elements to include in analysis.
2. Problem Analysis. Finds spatial relationships, applies analytic techniques to transform and generate knowledge.
3. Problem Synthesis. Assembles knowledge, about spatial relationships and vulnerability, into patterns.

The general workflow to follow in every phase is shown in simplified form in Figure 7.2. The schema can be applied to any of the five phases; however, each one introduces adjustments to carry out specific tasks as needed. The vulnerability patterns (event, component, and overall) in the analysis / synthesis stages, must be regarded as the vulnerability patterns to be found, described, explained, predicted, or designed, depending on the phase of the approach.

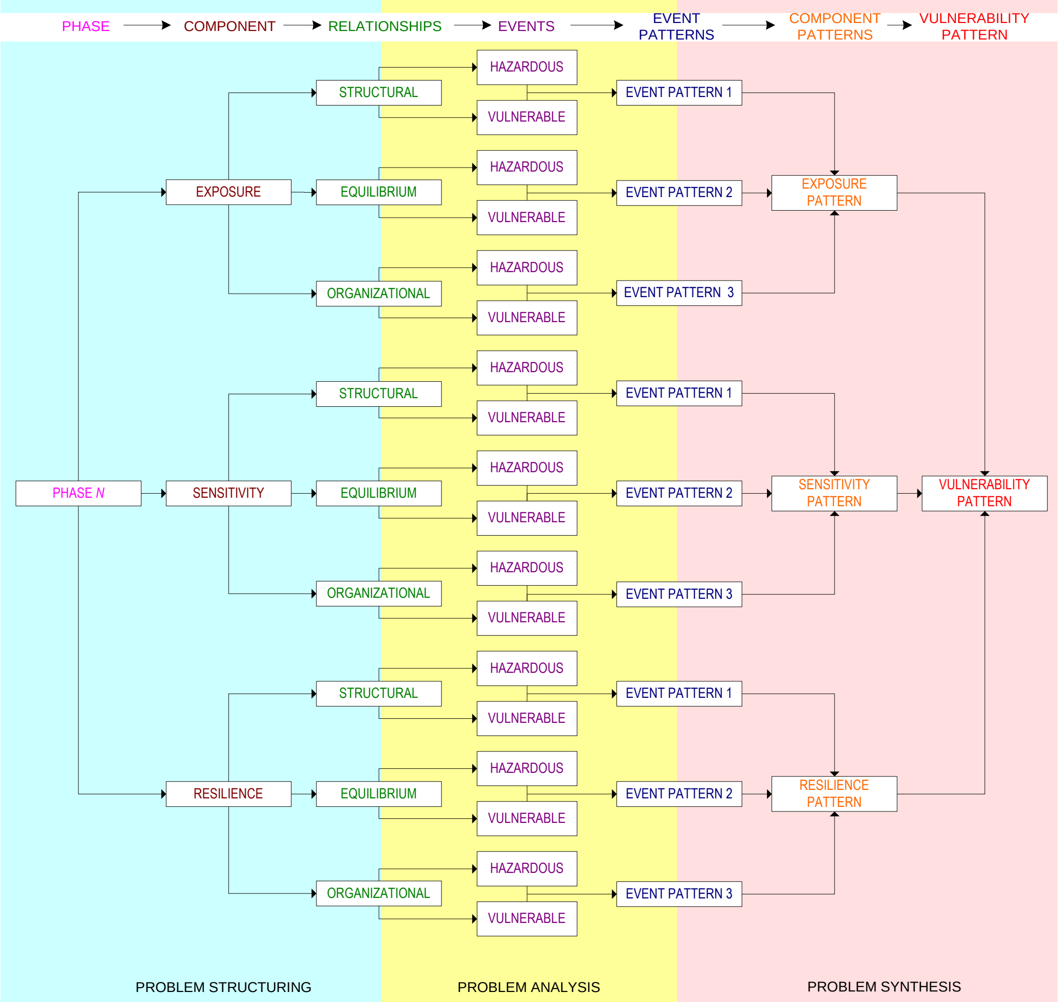


Figure 7.2 Simplified schema for phase work. Spatial relationships are grouped, interactions are represented for one hazard (or factor /agent of hazard) and a single vulnerable event, and the event patterns represent the overall pattern of an event's vulnerability under a group of relationships. Event patterns derived from a single relationship between one hazard and one vulnerable event (not shown here) are the simplest vulnerability patterns; component patterns are the result of the aggregated vulnerability patterns of single events for each component; the final vulnerability pattern results from the aggregation of the overall vulnerability component patterns.

It is necessary to remark that although all spatial relationships are investigated, only those found or considered relevant to the objective of each stage are finally used for the purposes of a specific phase. Given the nature of each vulnerability component, it is foreseen that, in general, the structural and neutral relationships can be more relevant in the analysis of exposure, while the neutral and the organizational relationships become more important for defining sensitivity or resilience, as these last components focus on events, relationships, and processes that are more or less related to the organization of space (human or human-related events), whereas the first component focus more on events reflecting the structure of space (biophysical events and natural hazards, except when the hazards are of anthropogenic origin).

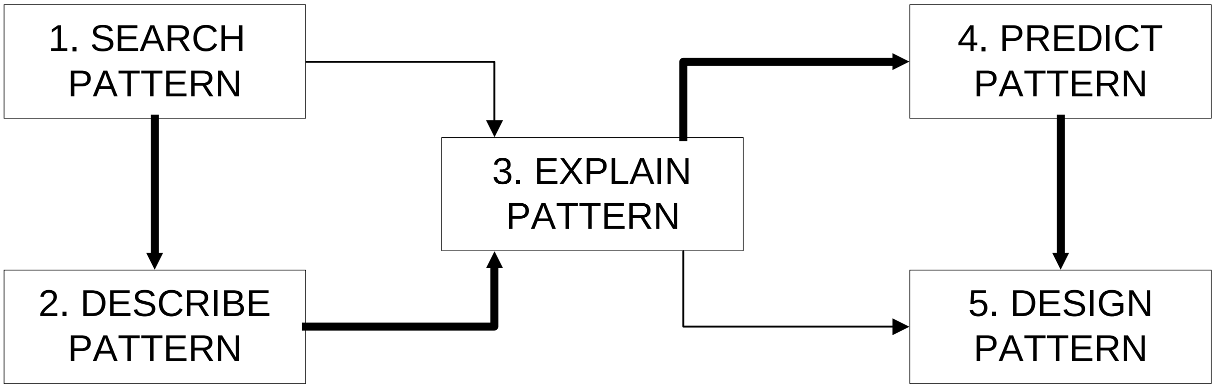


Figure 7.3 The SAVE approach workflow. The thick lines and phase numbers indicate the recommended path.

The execution of the five phases is necessary when a full assessment of vulnerability is wanted, but the SAVE approach allows to skip some phases if research interests dictate it so. In some instances the interest could be focused on knowing only the existing vulnerability patterns, for which the execution of the first two phases is sufficient. The explanation and the prediction of a vulnerability pattern are needed when we want to understand how the interplay of environmental and development elements lead to specific vulnerability patterns, or what could be the future state of a vulnerability pattern given a trend or a particular scenario of development-environmental change. The last phase is only required when we wish to make decisions to create or modify a vulnerability pattern to make it acceptable or compatible with current / future environmental and development conditions. Figure 7.3 shows the workflow in the execution of the five phases. A brief description of each phase follows.

Phase 1. Search of Vulnerability Patterns

This is the most intensive phase; hence, it will be described in more detail than the others. The objective is to find current vulnerability patterns, using spatial relationships as indicators of vulnerability levels. Specifically, the goal is to know the spatio-temporal distribution and the different types and levels of vulnerability using spatial relationships to model the interactions between and within hazardous events and vulnerable events.

In general, the procedure to carry out this phase requires the identification of all spatial events related to hazards (factors and agents of hazard, and hazards themselves) and all the spatial events for which we want to know their vulnerability levels (in contrast to hazards, factors and agents of vulnerability are modeled as attributes of vulnerable events, not as standalone spatial events).

Next, it is necessary to establish, for each group and type of spatial relationships, a set of initial interactions, either hypothetical or proved:

A. Among all types of hazards, including the hazards themselves as spatial events, but also among the spatial events serving as factors and agents for the hazards.
B. Among all types of vulnerable events, including people, assets, activities and biophysical events, taken as thematic events, not single instances. Examples of thematic events are people, agriculture, roads, riparian vegetation, etc. Thematic events may be as specific as needed, for example, roads might be considered as two types of thematic events, paved and unpaved roads.
C. Among all hazards (including factors and agents of hazards) and vulnerable events.

These interactions can be initially specified as qualitative or semi-quantitative statements that give an idea of the magnitude and direction of the interaction according to a specific type of spatial relationship. For instance, in the exposure component of vulnerability, a specific proximity relationship between people and a river (as an agent of a flood hazard) could be established as follows: *the nearer the people to a river, the higher the exposure level to a flood*. This is, of course, a generalization, because given the characteristics of the flood, the distribution of people, and terrain morphology, at a certain distance from the river the exposure level of the people becomes null, but it provides an example of how spatial relationships can be used to initially describe an interaction between a vulnerable event (the people) and a hazard-related event (the river, an agent of hazard in this case).

The next step is to convert those statements into formal measures of the interactions. Mostly, this can be accomplished using a GIS, but depending on the complexity of interactions, it can become a difficult task, especially when measuring interactions in time or those determined by organizational relationships, which may require a dynamic representation. Measures does not necessarily have to be numeric and crisp, in some instances qualitative measures (class values) or fuzzy measures can be sufficient, or even desirable. The measures are then to be converted to vulnerability levels using a set of rules and a standardized scale. The result of measurement and its conversion to levels is a set of vulnerability patterns for each thematic event, up to a theoretical maximum of *n* thematic events times *m* spatial relationships identified, times 3 vulnerability components, where each event pattern specifies the corresponding vulnerability levels. Vulnerability levels can also be expressed as numeric / nominal or crisp / fuzzy values.

The basic vulnerability patterns of single event – single spatial relationship can then be aggregated into event patterns by group of relationships up to a maximum of nine groups of event vulnerability (3 vulnerability components times 3 groups of relationships). This step may be very informative and has to be executed to understand the degree of influence of the structure of space or the organization of space in producing vulnerability. This might be important for policy making because:

- If vulnerability levels are mainly determined by the structure of space, policies for vulnerability reduction should be aimed to improve place planning by, for example, enforcing land use regulations, executing public works to protect society from hazards, or conducting studies to safely locate new infrastructure or economic activities.
- If the organization of space controls vulnerability levels, policies should be directed to improve the functional capacity of the place by, for instance, eliminating / reducing harmful dependencies, reinforcing diversity and strength of beneficial connections, or building organizations where they are absent.

The next step of vulnerability investigates the overall levels of its three components. Thus, for example, the overall vulnerability pattern for the exposure component is obtained from the aggregation of either the single-event vulnerability patterns or the group vulnerability patterns for that component. Likewise, the vulnerability patterns for the other components are obtained. These component patterns inform the assessment about whether the location or the intrinsic characteristics of spatial events are the responsible for the vulnerability levels, specifically:

- If vulnerability levels are higher for the exposure component than for the other two components, then the location of events mainly determines the vulnerability levels of a place, therefore relocation activities or structural measures are indicated to reduce those levels (this is equivalent to modify the structure of space).
- If vulnerability levels are higher for the sensitivity or the resilience components, then the characteristics of events are mainly responsible for the vulnerability levels, hence changes in those characteristics would be appropriate to reduce vulnerability (this is equivalent to modify the organization of space).

The last step of the phase is the construction of an overall vulnerability pattern of a single thematic event, as derived from the aggregated vulnerability of the components. Rather than simply adding together the vulnerability levels of the three components, interactions between the three vulnerability patterns are investigated and the aggregated vulnerability of the place of study is derived through a weighted model of such interactions as indicated by the type, magnitude and direction of the vulnerability levels.

The entire procedure has to be repeated for as many thematic events as considered in the assessment. Also, the previous steps can be applied considering one hazard at a time or considering the overall effect of all existing hazards in a place. This second variant is recommended because it takes into account the interactions among hazards, which results in a less reductionist model of vulnerability, although at the same time it may complicate work. Indeed, a distinctive characteristic of a vulnerability assessment is its focus on the recipient (place, according to the SAVE model), which allows for the assessment of the effects of multiple perturbations and stresses and their interactions (IIASA 2006). Phase work can be carried out for one or multiple hazards and for one or multiple types of vulnerable events existing in one or multiple interconnected places.

The outcomes of the vulnerability patterns found in this phase usually take the form of vulnerability maps linked to databases or the form of models of vulnerable events (dynamic computer models using hierarchical and network structures). In addition, modeling aggregated vulnerability may require some weighting mechanism capable of dealing with qualitative data and subjectivity. Also, some way to handle uncertainty in geographic information is needed. Recommended candidates for these two tasks are the Analytic Hierarchy Process (Saaty 1977) and fuzzy logic (Zadeh 1965).

Phase 2. Description of Vulnerability Patterns

The objective is to use spatial relationships to characterize found or observed vulnerability patterns. An accurate and exhaustive description of the overall vulnerability and its components complete the results obtained in the previous phase. This phase can be initially skipped, and work can directly proceed from Phase 1 to Phase 3, and further, but in the end it is necessary to perform it, especially if the results need to be communicated outside the research group conducting the assessment.

In general, the procedure to follow is to document every step followed in the first phase and describe in full what is depicted in the vulnerability maps and models.

Phase 3. Explanation of Vulnerability Patterns

Since, following Luers et al (2003) any vulnerability assessment should not only identify the systems at risk but also help to know why, this phase focuses on the understanding of the processes leading to vulnerability. Again, using spatial relationships as the organizing concept, an explanation of the causes or factors of vulnerability, and of the mechanisms by which those factors operate upon the vulnerable events, must be given to facilitate the understanding of the phenomena. This is a very important phase because such understanding should help us to detect the key spatial relationships that can become indicators of vulnerability.

As this phase must investigate the processes leading to a particular vulnerability pattern, it is essential to identify which points / parts in a process are critical, so that if modifications are made at these points / parts the resulting pattern may be different, either leading to an increase or to a decrease in vulnerability levels. Such critical points / parts may consist of single or multiple events (vulnerable and hazardous events, or factors / agents of hazard and vulnerability), or specific mechanisms, or both. In particular, two types of changes in critical points/parts must be investigated:

- Instantaneous. Changes in points / parts that may trigger relatively instantaneous / simultaneous changes in a vulnerability pattern when elements of a process are created, modified, or destroyed; where the effect of modifying an interaction or the characteristics of an event participating in an interaction is immediate on other interactions or events, therefore modifying the vulnerability pattern at once (or seemingly at once).

- Gradual. Changes in points / parts that may cause changes in a vulnerability pattern occurring after some time of the modification of a process; where the effect of changing an interaction, or the characteristics of an event participating in an interaction, requires some time to manifest, either because the modification must reach a specific threshold to produce an effect, or because the affected interactions or events evolve according to their own specific times, therefore modifying the vulnerability patterns after some time, with the possibility of keeping this modification active until the effects die off or another process interrupts it.

Changes of the first kind are related to changes in specific and generic vulnerable events (people, economic activities, infrastructure and biophysical events), or in specific hazards, or factors / agents of hazard that determine exposure (e.g. a particular river in a flood hazard). An example of this kind of change would be the enforced and permanent abandonment of buildings at some type of risk. This action, when actually carried out, may imply a modification of up to three types of vulnerable events: people, infrastructure, and economic activities, with the effects on their existing vulnerability patterns being almost immediate: less people would be exposed to the hazard when people moves to safer buildings; the at-risk buildings could be demolished, implying that less infrastructure would be vulnerable; and possibly, if some kind of economic activity took place in the buildings (commerce or services), it would also experience a reduction in vulnerability by moving the activity to a safer place.

Changes of the second kind are represented by changes in social, economic, and environmental processes, whose times of occurrence range from short to very long, but where a change may trigger changes in the process itself or in other processes or elements of a process, affecting vulnerability patterns after some time of the initial change, with this time determined by the specific shape and rate of evolution of the process. For example, a global reduction in greenhouse gases emissions would bring about, in due time, some changes in the specific occurrence of particular hazards that have been predicted to intensify as greenhouse gases emissions increase, thereby helping to reduce the exposure to those hazards.

Investigation on changes should include suggestions about their magnitude, direction, and interactions. Once critical points / parts and their most appropriate changes are identified and characterized, these become candidates for policy making.

Phase 4. Prediction of Vulnerability Patterns

The goal is to investigate vulnerability dynamics. Work starts with the definition of the current state of vulnerability as a benchmark to be used when forecasting trends, creating scenarios and predicting outcomes of specific vulnerability changes. Also, this phase should suggest the appropriate timing for monitoring vulnerability indicators.

Focus on forecasting, prediction or scenario-building is somehow determined by the needs of the assessment. Forecasting vulnerability trends gives us information on the possible outcome of the impact of a future hazard if nothing is done to modify existing vulnerability patterns, and/or if there is no modification in the behavior of hazards. Prediction applies to specific changes in a process to investigate their possible effect in a vulnerability pattern. On the other side, the creation of scenarios implies accounting for the overall effects of policy implementation, usually coupled with GEC or planned change in development conditions of the place.

Success in modeling the dynamics of vulnerability requires knowledge of the underlying processes, as derived in Phase 3, and knowledge on the rates of environmental and development changes, either derived from historical data or modeled by some empirical function.

Phase 5. Design of Vulnerability Patterns

This phase is policy-making oriented and aims at creating or modifying a vulnerability pattern that helps to reduce vulnerability levels, and in the end, to design a dynamic pattern of vulnerability compatible with environmental change and development.

The phase focuses on the procedures to evaluate, design and monitor (in this order) those indicators of vulnerability (spatial relationships) whose modification may bring about the greatest reduction in vulnerability levels, either by decreasing exposure and sensitivity or increasing resilience, or all of them. Once the key relationships were identified (in Phase 3), the phase uses knowledge on the mechanics of vulnerability (obtained also in Phase 3) and knowledge on the dynamics of vulnerability (obtained in Phase 4), to evaluate the feasibility / desirability of changes in these interactions such that the best possible combination is considered for policy making, within the context of the assessment.

In order to guide the policy-making process, the evaluation of the feasibility / desirability of modifying key spatial relationships must consider a structure appropriate for decision making. A hierarchical structure, with the first level and node corresponding to the goal of the evaluation, and the second level represented by the three vulnerability components (exposure, sensitivity and resilience) as nodes of the hierarchy, is deemed as initially suitable for the problem, although other structures could also be explored. The idea of this suggested structure is to evaluate whether modifying the exposure, or the sensitivity, or the resilience, or which combination of modifications, is the best overall decision to reduce vulnerability. Decision support techniques (multi-criteria decision making methodologies, including cost-benefit analysis) that allow public opinion and promote the involvement of all stakeholders interested in the reduction of vulnerability, are called for at this point, and need to be incorporated as part of the SAVE approach. Since the approach is place-oriented, not a community-based approach, it is necessary to bear in mind that if the place of assessment corresponds to a community, there is a potential for public participation (Turner et al 2003) that might be more difficult to achieve in places with larger extensions.

Once the evaluation processes has been completed, and decisions have been taken, the design of feasible / desirable vulnerability patterns must start by converting the selected changes into modifications in the characteristics of vulnerable events or into modifications of the structure of space where those vulnerable events take place. This is first done in the models of the affected patterns (tables, maps, computer models), running again some of the analysis performed in Phase 1, with the modified data, to observe the outcome. If the results agree with the idea of acceptable or compatible (with environmental change and development) vulnerability levels that decision makers have, the next step of the approach is to help decision makers to convert proposed changes into vulnerability reduction policies. The selected changes to key spatial relationships can be converted to policies by specifying concrete actions to change specific magnitudes and directions of relationships.

The final step of this phase of the SAVE approach is to assist local officials in the monitoring of vulnerability levels by means of indicators. Knowledge on the mechanics and dynamics of vulnerability, obtained in Phase 3 and Phase 4 respectively, coupled with knowledge on the dynamics of development and environmental change, can be used for this purpose. In particular, two types of indicators should be monitored for every place investigated:

- Those derived from the identification of critical points / parts of processes, which were not subject to policy making, but that, nevertheless, may trigger drastic increases in vulnerability levels if are allowed to reach specific thresholds.
- Those derived from the introduction of changes in the vulnerability patterns through the implementation of policies, where monitoring is needed to assess if expected reduced levels of vulnerability are met at the expected times.

A collateral issue in this phase deals with how to isolate the terminal users of the approach, public officials wishing to reduce the risk levels in their jurisdictions, from technical concepts. In other words, how to tell an official that a specific place or group of people, or a particular physical asset or biophysical event, has low or moderate vulnerability level and under which conditions this level can become high, without explaining him the theory behind fuzzy sets or the mathematical procedure for synthesizing weights from purely expert opinions. That could be accomplished through a computerized decision support system. Such a system should have a strong geographic information management component. Indeed, geographic information systems (GIS) technology is especially important to the SAVE approach because it allows to recognize and analyze spatial relationships between spatial events (Gustafson, 2005; USGS 2005b).

7.5 CONCLUDING REMARKS

The SAVE approach can be seen as a methodological improvement to vulnerability assessment. Despite recent advances in the comprehensive conceptualization of vulnerability, comprehensive methodologies for vulnerability assessment, if any, suffer of a number of drawbacks, including their specificity to place, hazard, or scale, their lack of sufficient integration between natural and human interactions, and their limited applicability to any vulnerable event because of the chosen indicators. The SAVE model attempts to overcome these restrictions by promoting the use of generic concepts such as place, spatial relationships, and pattern.

The SAVE approach can also be regarded as a true geographic approach to the problem of vulnerability assessment. By using fundamental geographic concepts such as place, spatial relationships and pattern, the approach focuses on the spatiality of vulnerability, and since the spatiality is a generic characteristic of any vulnerability situation, it is possible in this way to provide generic indicators, and consequently, facilitate the design of policies of wide applicability to reduce vulnerability.

If we look at the research agendas of some leading institutions and world initiatives in the field of vulnerability (IHDP, 1999; ESFS, 2004; USGS, 2005a; IIASA, 2006), the design of comprehensive methodologies that integrate natural and human systems interactions, that have wide-applicability by avoiding being place, or hazard, or scale specific, and that are easy to understand and apply by public officials wishing to reduce vulnerability, appears always as one of the central topics.

In that context, the SAVE approach contributes to "improve the scientific basis for vulnerability…." (USGS, 2005a; Goal 4) and to "…develop concepts and methodologies for the purpose of addressing the complexity of social-economic-ecological systems…" (Linerooth-Bayer, 2006; Conceptual and Methodological goal), because it represents an innovative method for the assessment of vulnerability, that incorporates sound scientific concepts in a systematic framework.

By using spatial relationships as indicators, the SAVE approach provides a base to the "…development of standards and metrics for assessing vulnerability and resilience to hazards…" (USGS, 2005a, Sidebar 4.1), and also to "… characterize (e.g., through indices) risk, vulnerability and resilience in ways that are useful for policy negotiations, processes and decisions" (Linerooth-Bayer, 2006; Assessment goal).

The SAVE approach can help to "Develop and implement a monitoring program that provides perspectives at multiple scales of vulnerability…" (USGS, 2005a; Strategic Action 4.5), and facilitate the identification of at-risk areas by helping to choose the appropriate timing for monitoring vulnerability indicators, forecasting trends and predicting possible vulnerability outcomes, within the context of Phase 4 of the approach. "The challenge is to adopt a frame broad enough to encompass the systems underlying global change and sustainable development, yet narrow enough to provide insight to the relevant stakeholders and policy process" (Linerooth-Bayer, 2006). The SAVE approach meets this challenge through its systemic view of place where the same set of spatial relationships can be used to devise global and local indicators of vulnerability.

The SAVE approach covers several other topics, such as the assistance to managers in determining the effectiveness and feasibility of mitigation and risk management under a variety of scenarios, provision of methods for incorporating uncertainty, and the role of geospatial information in mitigation analyses (USGS, 2005a; Strategic Action 4.6).

CHAPTER 8

LAND ATTRACTABILITY AS A SUITABILITY MEASURE FOR AGRICULTURAL PURPOSES: THE AGRILOCAL MODEL

LAND ATTRACTABILITY AS A SUITABILITY MEASURE FOR AGRICULTURAL PURPOSES: THE AGRILOCAL MODEL

Luis M. Morales-Manilla

Centro de Investigaciones en Geografía Ambiental, Universidad Nacional Autónoma de México,
Morelia, MICH. C.P. 58190

Abstract

A complementary method to estimate land suitability for agriculture, focusing on the attractiveness of land to farmers, is here presented. Conventional methods in land use planning for agricultural purposes (*sensu* agro-ecological zoning or land capability approaches) overemphasize the importance of technical knowledge. Farmers' knowledge and preferences, especially how attractive the land is to them, are rarely taken into account in these planning exercises. Because of this oversight, many land use plans fail to reach the implementation stage, in spite of their technical merits. One reason this knowledge is not used more often is that it is not formal and is difficult to obtain from farmers. The method presented here estimates the probability that a portion of land will be attractive to farmers. Departing from the conventional method of interviewing farmers to obtain the information on attractability, this method proposes to infer it from the spatial distribution of farmland in a region, as depicted in land use maps. Discerning the patterns of several spatial relations between the historic and current distributions of farmland and local development / environmental factors of agriculture (roads, tenure, settlements, slopes, soils, rainfall, etc.) provides the information needed to build a suitability model. Because it includes estimates of farmers' preferences, the results of this model could have a better chance of being accepted and implemented by local farmers than those of a conventional approach.

Keywords: agriculture, land attractiveness, land suitability, farmer knowledge, spatial relations

8.1. INTRODUCTION

The availability of land suitable for agriculture is fast decreasing as population increases and existing farmland degrades or is used for other purposes. Knowledge about the availability of land is of high strategic value for sustainable development. Indeed, in its list of unknowns regarding land-use planning for agriculture, the FAO puts the following question first: Exactly how much potentially arable land is available, and where is it? (FAO, 2007).

In planning contexts, the conventional approach to allocating new land for agriculture is to conduct a suitability study. This may take the form of a land capability study (LC) (USDA, 1973; Emery, 1986; Agronomic Interpretations Working Group, 1995) or an agroecological zoning study (AEZ) (FAO, 1978; Sivakumar and Valentin, 1997), where the land suitability is estimated taking into account various biophysical parameters, with the implicit goals of finding land where crop yields or profitability may be higher or where agriculture can be sustainable. Land suitability studies based on the AEZ approach have found a high degree of correlation between the obtained suitability values and the observed distributions of farmland (Fischer, 1999). The LC methodology approaches land evaluation by means of a land classification useful for guiding conservation practices (Norton, 1939). However, despite their technical soundness, often the results of these or similar methodologies do not reach the implementation stage, the main reason being that they do not take into account farmers' knowledge in the form of preferences and responses to environmental / development conditions in a local way.

There are many examples in other fields (mining, road construction, and biological conservation, to name a few) where overlooking local actors' opinions, traditions or concerns has resulted in failed implementation of projects. To my knowledge, however, few examples of this problem have been reported in agriculture, perhaps because political circumstances encourage discretion. The World Bank (1994) reports on a study of soil conservation in El Salvador where one reason for the failure of extension programs was that proposed soil conservation schemes were influenced by pressure from the extension services, rather than from what farmers considered desirable. Barbier (1997) asserts that failure to take into account the factors determining farmers' land management decisions are a source of failure in extension projects. Bennett (2000) mentions that in some cases, heavy investments have been made in LC and AEZ studies, including national efforts, only to find that the results are rarely used in practice because factors other than the biophysical were ignored.

Three studies may not seem enough to show the importance of this concern in agriculture; but the problem may begin to take on greater urgency as food security situations become more pressing. Therefore, to resolve this deficiency in current methodologies, I suggest that another type of land suitability study be conducted first or complementarily: one where the suitability reflects the land potential according to the farmers' local preferences and practices, i.e. the attractability of land.

However, farmers' knowledge is hard to use in planning contexts because it is not formal and might be difficult to obtain. A method to overcome at least some aspects of these inconveniences is to infer such knowledge from the way agriculture has developed over time. I claim that given a set of development and environmental conditions, agriculture tends to take place locally in the best possible form, with those conditions being reflected in the spatial distribution of farmland. I present here a set of methodological steps to extract local knowledge from this distribution and use it to build a land attractability model.

This article first emphasizes some differences between formal (agronomic) and non-formal (farmer) knowledge, highlighting the importance of spatial knowledge in the form of spatial relations. It next describes the method to extract non-formal knowledge using spatial relations to build a suitability model that estimates land attractability, using proof-of-concept data from a region in Southern Mexico. It then discusses how results can be interpreted in a land suitability context, giving indications for using the model as a diagnostic and predicting tool and pointing out some limitations arising from its spatial nature, as well as its value in agriculture planning. Finally, it concludes with remarks on possible future research using the model.

8.2 THEORY

Ideal practices versus actual practices in agriculture: technical knowledge versus local knowledge

In agriculture, as in other economic activities, there is a difference between what theoretically should be done to efficiently carry out the activity and what is actually done. Factors contributing to these differences include the experience, knowledge, and the amount and quality of information that farmers have when making decisions about the use of land; but they also arise from local development / environmental conditions.

Whichever of these may be the source of deviation from an ideal practice, I argue that agriculture tends to take place locally under the best possible practices, distributions of farmland being an excellent indicator of such practices. In the same way that ideal agricultural practices indicate the optimal theoretical conditions for agriculture, actual practices dictate the location and types of land considered locally as optimal, despite what technical knowledge might suggest. Girard et al (2001) say to this effect that “Farmers’ practices … govern the organization of the land areas of each farm through optimal use of soil types, land exposition, slopes of the various spatial entities and through creating interaction and complementarity among them”.

“Ideal” practices, as suggested by experiments or theory, might not always be adequate for predicting where agriculture can be successfully accomplished. Some reasons are:

- Knowledge of all environmental conditions required by a specific crop may be incomplete.
- Experimental information can be valid for a limited range of environments. Extrapolation might not be advised or possible.
- Ideal conditions used as parameters in a land suitability model may not be adjusted to reflect real conditions under which local farmers act.
- Models may be complete, accurate, and reality-adjusted, but information on the model’s parameters might not be available, may lack quality, or may be out of date.
- Ideal conditions may suggest practices, which may be effective, efficient, and equitable, but still are not implemented because farmers are reluctant to change their practices or because the alternative practices represent a burden to the farmer’s economy or involve radical departures from their way of life.

On the other hand, local practices implicitly carry on the expert knowledge of factors that farmers regard, or have traditionally regarded, as important to the development of successful agricultural activities (Girard et al., 2001).

The container of farmers’ expert knowledge

Interviews are the conventional method to extract farmers’ expert knowledge, but this approach needs to involve the majority of farmers in order to extract significant trends in their knowledge.

Depending on the farmers' population and dispersion, such universal engagement may be untenable due to financial or time-related restrictions. Moreover, the issue of mistrust and lack of cooperation has to be dealt with when interviewing people. Also, interpretations of questions and answers may vary depending on how the questions are designed and interviews conducted. In addition, ambiguity and qualitatively oriented answers can further complicate the extraction of knowledge.

I suggest instead looking at the spatial conditions under which agriculture has locally developed, i.e. farmland spatial distribution as observed in maps and remote sensing images. This distribution contains farmers' expert knowledge, specifically in the underlying spatial relations producing such distribution. Analysis of these relations needs to be done locally or regionally, because even while similarities in the spatial distribution of agriculture under like socioeconomic or environmental constraints do obtain (Duvernoy, 2000), the opposite is also true: farmers' practices in response to similar socioeconomic and environmental conditions are diverse (Milleville and Dubois, 1978).

Spatial relations and agriculture

Spatial relations are used to indicate existing or desirable spatial interactions among the geographic events composing an agricultural system. The values of these relations are established by the farmers' response to development and environmental conditions, outlining a sort of local optimum for agriculture.

Spatial relations play a significant role in constraining or promoting economic activities. A well-known example is Ricardo's theory of land rent, where proximity to markets is an essential relation. It is also acknowledged that proximity to roads promotes the accessibility to new land (Barbier, 1997). Farmers pay great attention to proximity relations, prioritizing the distance to important landmarks such as housing, roads, water points, etc., much in the way that herders manage the use of grazing lands according to the distance the animals must cover to reach the grazing areas (Girard et al 2001).Of course, proximity is just one the relations that seems to be of importance to farmers. There are nine types of spatial relations (Chapter 3) each including many instances, that farmers' may also use, as in the following examples:

- *Proximity*: Lands proximal to roads are more preferred for agriculture than those located far from them.
- *Orientation*: In hilly terrain and middle latitudes (in the Northern hemisphere), south-facing slopes are preferred for vineyards because they may receive more solar radiation than slopes facing other directions.
- *Exposure*: Hedges along the borders of agricultural fields may prevent exposure to wind erosion or freezing winds.
- *Adjacency*: Land adjacency to a river determines possibilities for irrigation.
- *Containment*: To affect the rights of farmland enclosed within administrative units, or to implement taxation policies, it is first necessary to interact with the container unit (state, municipality, community, etc.)

- *Coincidence*: The coincidence of farmland with areas of well conserved soils assures higher yields than coincidence with degraded soils.
- *Connectivity*: A well-developed network of rural roads connecting to a highway (physical connectivity) may offer a way to export products. A network of production places and markets (functional connectivity) may lead to an economy of scale in agriculture.
- *Aggregation*: In a regional agricultural system (a form of spatial aggregation), sustainability is determined by the aggregated (interdependent) interactions of its components: the conservation of soil in farmlands, the amount and distribution of rainfall in the region, the economy of farmers, the stability of internal and external markets, the absence of natural disasters, etc.
- *Association*: The participation of farmers in an agricultural cooperative or a cooperative farm (two forms of spatial association at different scales) may enhance efficacy, efficiency and equity in agricultural production.

Relation values for individual farmers may be very disparate, but when these values are extracted from the spatial distribution of farmland, a pattern emerges, showing the relevancy of a relation, its preferred values, and its local importance. We can use this information to build a suitability model that estimates how attractive land is to farmers.

8.3. THE AGRILOCAL MODEL

Key characteristics of the model

AGRILOCAL is a suitability model to estimate the attractability of land, where the model's parameters are spatial relations representing local agricultural preferences. The attractability estimates in the AGRILOCAL model are a combination of relevance – preference – influence measures of spatial relations held between the land under consideration and other geographic events participating in an agricultural system. Attractability estimates are interpreted as the degree of possible success in using new land for agricultural purposes according to the local development / environmental context. Thus, there cannot be a single AGRILOCAL model applicable to all places.

The procedure to build an AGRILOCAL model can be summarized in three steps:

- Find relevant spatial relations between farmland and other development / environmental factors participating in the activity.
- Specify probability functions to indicate the degree of preference represented by the values of each relation.
- Derive weights to establish the degree of influence of each spatial relation on the decision of using the land.

To build a specific AGRILOCAL model the following general form can be used:

$LA = WR_i (f\,PR_i) + WR_{i+1} (f\,PR_{i+1}) + \ldots + WR_{n-1} (f\,PR_{n-1}) + WR_n (f\,PR_n),\ i = 1 \ldots n$

Where: LA = land attractability according to the AGRILOCAL model; WR_i = weight of relation i; $f\,PR_i$ = probability function for relation i; n = number of relations in the AGRILOCAL model.

In order to show the construction of an AGRILOCAL model we will explore only four spatial relations: proximity of current farmland to roads (PFR), current farmland coincidence with terrain aspect (FA), current farmland coincidence with terrain slope (FS), and proximity of current to previous farmland (years 2000 and 1976, PPF).

Step one: finding relevant spatial relations

Relevancy implies an appraisal of the importance of certain spatial relations in the decision to use land for agricultural purposes. Although interviewing farmers would be the most direct means of acquiring this knowledge, in the model it is obtained by observing how much the actual distribution differs from a random distribution for the same relation.

The number and type of parameters in the model depend entirely on the relevancy of spatial relations. The procedure can be summarized as follows:

A. Prepare a list of proved or hypothetically relevant spatial relations.
B. Prepare a set of maps showing the actual distribution of every relation.
C. Prepare a set of maps showing a random distribution for every relation.
D. Remove size effects in both sets of distributions.
E. For every relation, compare the frequency distribution of values in the actual and random distributions to determine if there are statistical differences indicating whether a relation is relevant or not.

The procedure to establish the relevancy condition starts, in Task A, by preparing a list of hypothetically relevant relations, such as those mentioned above.

Next, in Task B, prepare maps showing the actual distribution of each relation. For example, if the relation under investigation is PFR, we will need a map of roads and a map of farmland (or will have to extract them from the land use / land cover map) at the appropriate scales and dates. Ideally, the road map should slightly antedate (by a year or two for most regions) the farmland or the land cover / land use map, since the hypothesis is that roads are a factor in the expansion of farmland. Then a map of proximity of land to roads is calculated. Also ideally, the proximity to roads should be calculated measuring distances on the road network and, additionally, account for any frictional or barrier effects caused by slope, land cover, water bodies, or certain types of infrastructure where roads do not exist. Finally, the map of proximity to roads has to be overlaid on the farmland map to obtain a map containing the distances of farmland to the nearest road. This last map has the values we need to determine the relevancy of the PFR relation (Fig. 8.1). We proceed likewise for the other relations, overlaying the map of current distribution of farmland on the maps of proximity of land to previous farmland (1976), land slope, and land aspect.

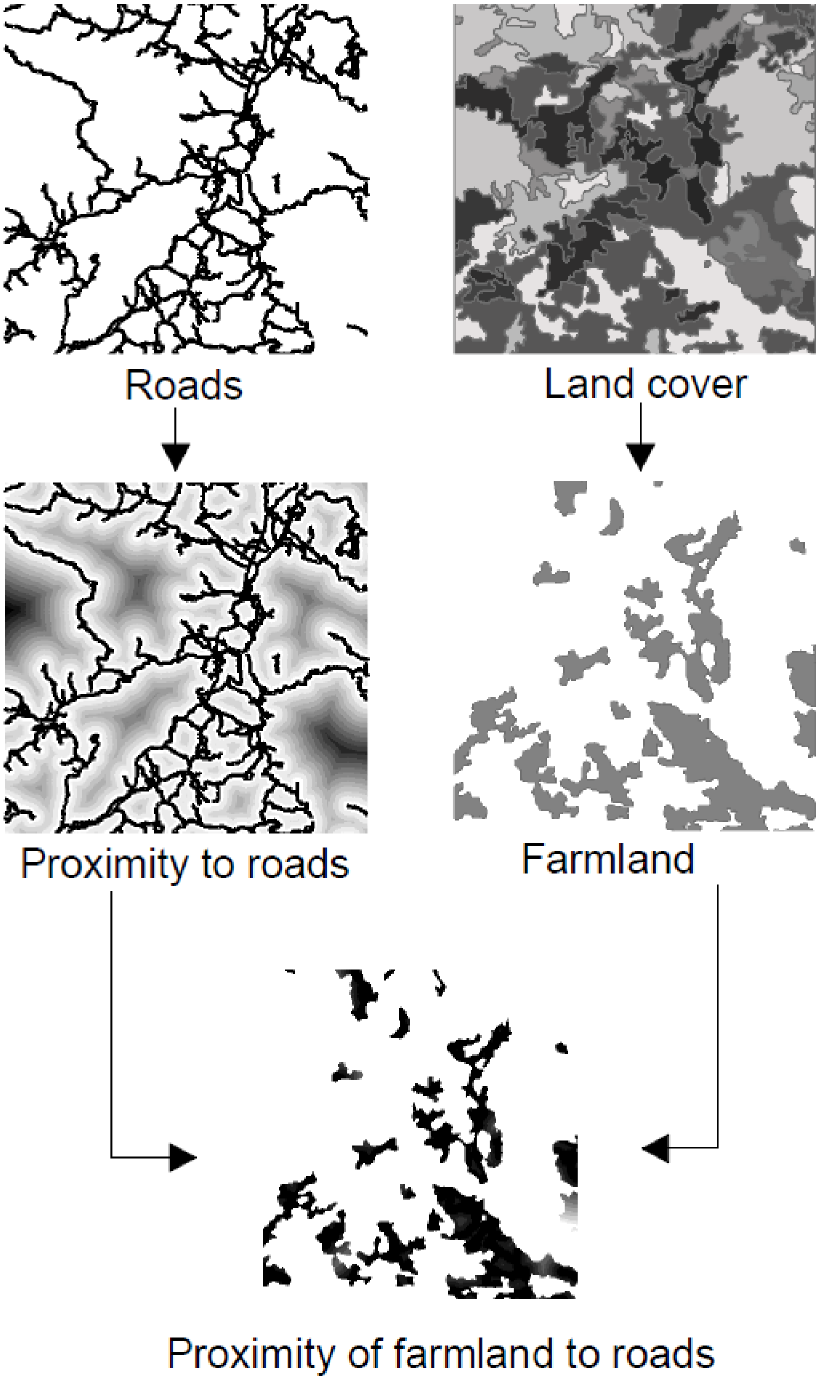


Figure 8. 1 Maps used and process followed to obtain the pattern of the spatial relation 'Proximity of Farmland to Roads' (PFR).

Very likely, in the PFR example, we will obtain a distribution of distances like that illustrated in Fig. 8.2a, where shorter distances show a higher concentration of values, with the concentration decreasing sharply as distances increase.

In Task C, we need to create maps of random distributions for the same relations. First, we create a map of random distribution of farmland by randomly selecting (with uniform distribution) a certain number of spatial units (raster cells) in the study area map, such that the sum of their area is equal to the area of farmland in the farmland map. In the PFR example, this map of random distribution of farmland is then overlaid on the proximity to roads map in order to assign a distance value to each spatial unit in the random map, thus creating the random map version of the relation. We proceed likewise for the other relations.

In Task D, before delivering any final conclusion on the relevancy of a relation, we should find out if the observed frequencies are due to a size effect – that is, if some of the higher frequencies might simply result from there being more land in the study area that falls within that class of values than within the other classes. To remove the size effect, an area-normalization procedure must be applied to each frequency class, in both the actual and the random distributions of the relation. For any relation, the formula is:

$$SNF_i = \left(\frac{NCIE_i}{NCIOE_i} \right) \tag{1}$$

Where: SNFi = size-normalized frequency of class i; NCIEi = number of cells of class i within farmland; NCIOEi = total number of cells of distance class i in the study area.

Finally, in Task E, the size-normalized frequency values are used to perform comparison tests. When applying these tests to the actual and random distributions, if both distributions are normal we use parametric statistics such as t (Student) and F (Fisher) to establish the significance of the similarity of both distributions in terms of concentration and dispersion; if one or both distributions cannot be considered normal, however, non-parametric statistics such as W (Wilcoxon) and D (Kolmogorov-Smirnoff) should be used. If, according to the statistics employed, the actual and the random distributions are significantly different in concentration and dispersion (at α = 95% or $p<0.05$), then we can conclude that the relation is relevant; that is, the distribution of the relation is strongly dictated by farmers' preferences and does not follow a random-like behavior. If the difference is significant only for either the concentration or the dispersion, then we can say that there is some indication of the relevancy of the relation, and we must decide whether or not to include it in the model.

In Table 8.1 we show the outcome of the comparison tests for the actual and random distributions of the PFR, FA, FS, and PPF relations. Since in all cases the distributions are not normal, we calculated D and W and their corresponding p values.

Table 8.1 Results of the tests of dissimilarity for the actual and random distributions of four spatial relations.

Relations	D statistic	P value for D statistic	W statistic	P value for W statistic	Relevancy
Proximity of farmland to roads	0.666667	**0.009655	117	*0.010193	Relevant
Farmland aspect	0.444444	0.338843	52	0.331386	Not relevant
Farmland slope	0.388889	***0.000567	1893	**0.007489	Relevant
Proximity to previous farmland	0.588235	**0.005576	212	*0.021014	Relevant

* moderately significant; ** highly significant; *** extremely significant (α = 95%)

For the PFR, FS, and PPF relations, both D and W have a $p<0.05$, meaning that the relations are relevant with a degree of significance higher than 95%. For the FA relation both D and W have a $p>0.05$ result, meaning that the relation is not locally relevant for the practice of agriculture, because its actual distribution is not significantly different from a random distribution. For visual comparison purposes, we provide graphs (Fig. 8.2) showing the size-normalized forms of the actual and random distributions of the four relations.

The PFR relation appears to be very relevant, showing a sharp decrease in frequency between the first and second distance classes (Fig. 8.2a); and the differences between both distributions are highlighted by the logarithmic scale of the Y axis in the second graph (Fig. 8.2b). In the FA relation, the slight concentration of aspect values under the SE-S-SW cardinal directions is only a size effect, as it is also present in the random distribution (Fig. 8.2c); the similitude of both distributions is evident when using the logarithm of frequency (Fig. 8.2d), showing that FA is not a relevant relation and therefore must not be included in the model. The differences in the frequency distributions for the FS relation indicate that it is a relevant relation (Fig. 8.2e); the logarithmic scaling of frequency (Fig. 8.2f) indicates a significant departure in behavior after the thirteen degrees of terrain inclination, meaning that farmers mostly prefer to use slopes with values lower than this. The PPF relation is also relevant, as observed in the differences in the frequency chart (Fig. 8.2g) and its logarithmic version (Fig. 8.2h).

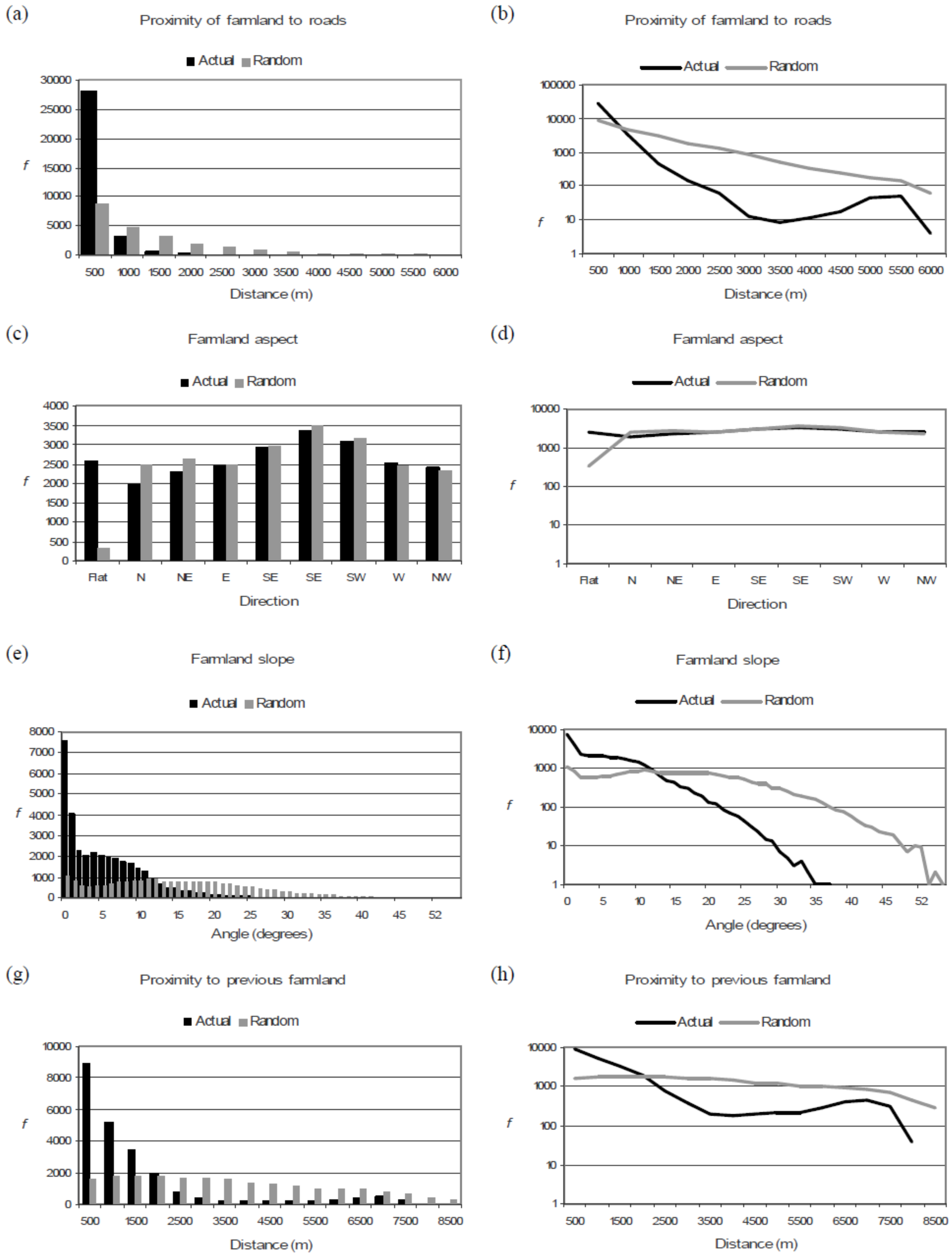


Figure 8. 2. Comparison of the actual and random distributions of four spatial relations. The graphs on the left show the differences in frequency; on the graphs at right, frequencies have been drawn on Y axis with a logarithmic scale to stress differences.

So far, considering the three relevant relations, the AGRILOCAL model has the following form:

[illegible] (2)

Where: LA = land attractability; PFR = proximity of farmlands to roads; PPF = proximity to previous farmlands; FS = farmland slope.

Step two: specifying preference functions

The values of each relation are converted into a degree of preference by transforming them into empirical probabilities. The assumption regarding preference is that the higher the frequency of values, the more preferred is a value or range of values (a class) for agriculture.

The procedure to model preference is as follows:

A. Convert size-normalized frequency values into probabilities for each class.
B. Find a function describing the relation between the probability values and the corresponding lower values of the class intervals.

In Task A, for each class, the frequency values are converted into probabilities by assigning the highest probability value (1.0) to the highest frequency value and linearly scaling down the rest.

In Task B, we fit a function to the probabilities assigned to each class. If there are several possibilities, we recommend choosing the one with the simplest form and the highest R^2 value.

For the PFR relation, the chosen function has the form $y= a+b*e^{-x/c}$; and specifically, the equation ($R^2 = 0.9999$) is:

[illegible] (3)

Where: PPFR = probability of using a portion of land for agriculture according to the PFR relation; DR = distance to the nearest road.

The function for the PPF relation has the same form as the former, and specifically, the equation ($R^2 = 0.9923$) is:

[illegible] (4)

Where: PPPF = probability of using a portion of land for agriculture according to the PPF relation; DF = distance to the nearest farmland.

For the FS relation, the function has the form $y=1/(a+bx^2)$, and specifically, the equation ($R^2 = 0.9980$) is:

[illegible] (5)

Where: PFSA = probability of using a portion of land for agriculture according to the FS relation; SA = slope angle.

We combine equations (3), (4) and (5) to obtain an unweighted (or equal-weighted) form of the model (constants are shown with reduced precision):

[illegible] (6)

Where: LA = land attractability; DR = distance to the nearest road; DF = distance to the nearest farmland; SA = slope angle.

If this form is used for exploratory purposes, the result must be divided by the number of relations in the model (three in this case). Distances in all equations are expressed in meters, and angles in degrees.

Step three: establishing the degree of influence

Relations may have different degree of influence upon decisions to use the land. Influence measures could be obtained from farmers, converting their answers into weights, but in the method they are extracted from the distributions of the relations.

The assumption regarding influence is that the more different the values of the actual relation are from the values of its random equivalent, the more influential the relation is. To measure this difference we will use the Kolmogorov – Smirnoff D statistic since it can measure the maximum difference between the two cumulative distributions, thus providing a maximum weight for each relation. The procedure to derive weights is:

A. Apply a Kolmogorov-Smirnoff test to compare the actual and random distributions of the relations and find the values of the D statistic.
B. Transform the D values into normalized weights.

In Task A, in our example, the D values for the PFR, PPF, and FS relations were already calculated in Step 2 (Table 8.1); they are 0.6666, 0.5882, and 0.3888, respectively.

Task B requires that in order to set the weights of the relations we transform their D values in the following manner:

$$WR_i = D_i / \sum_{j=1}^{n} D_j \quad , \text{ i = 1 … n} \qquad (7)$$

Where: WRi = weight of relation i; Di = Kolmogorov-Smirnoff D value of relation i; n = number of relations in the model.

The equation (7) has the effect of normalizing the weights so that their sum is equal to 1. Using this equation, the resulting weights for the PFR, PPF, and FS relations are 0.4056, 0.3579, and 0.2365, respectively.

Thus, the final, weighted form, of the AGRILOCAL model, considering three-relations, is:

[illegible] (8)

Where: LA = land attractability; DR = distance to the nearest road; DF = distance to the nearest farmland; SA = slope angle.

This model can be used in two modes: predictive and parametric. Both complement each other in the investigation of land attractability. The former provides information on the attractiveness of land and the possibilities for expanding agriculture. It represents the regional trend. The latter is used to explore anomalies in the attractability trend that may indicate areas of opportunity, i.e. land that could be used for agriculture because it has high attractability with respect to a single relation, but requires improvement of the attractiveness values of other relations.

Synthetic mode of the AGRILOCAL model

In this mode the model is used to estimate the probability that a portion of land will be used for agriculture, based on its attractability as calculated in equation 8. Up to three AGRILOCAL maps can be prepared (Fig. 8.3).

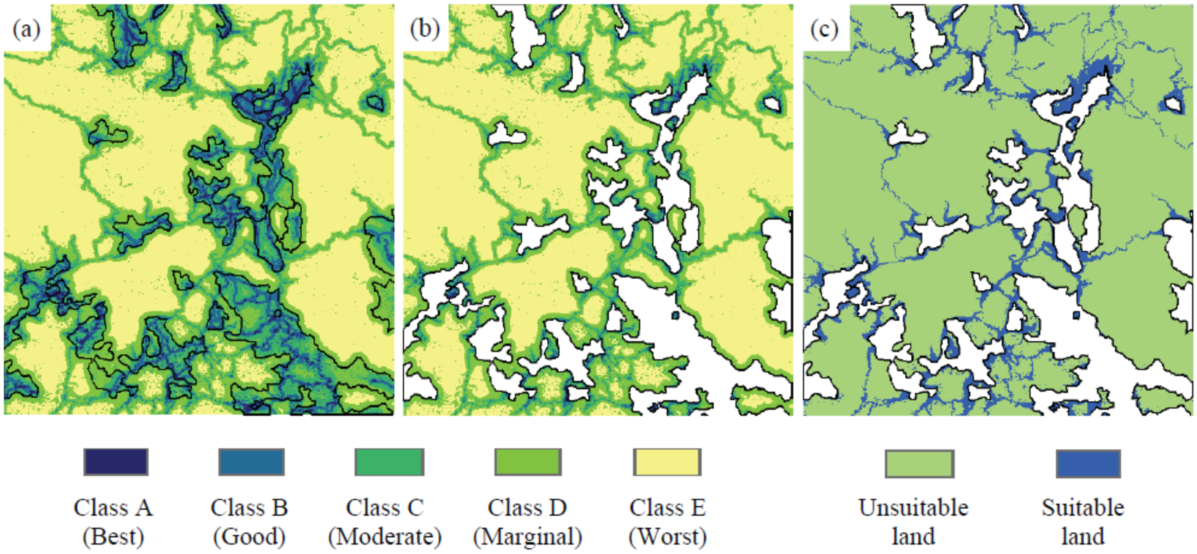


Figure 8.3 Maps showing the theoretical attractability of land derived from the AGRILOCAL model described by equation (8). Dark green and blue colors show the most preferable land; light green to yellow colors show the less preferable land.

A first AGRILOCAL map shows the attractability of all land, including current farmland (Fig. 8.3a). It is useful to have a surrogate idea of the quality of land as it appears to local farmers. In the map, the areas in yellow represent unattractive land for local farmers because of some constraining factor (in the example the main constraints are the slope and road accessibility). If we overlay a map representing farmland as transparent polygons, we would see that in general all areas with the highest attractability are already being cultivated, possibly corresponding to the first years of the activity; but it is apparent also that there are sizeable areas of low attractability that have been incorporated as farmland, perhaps at a later time.

A second map shows in color the location of land that has not yet been used in agriculture (Fig. 8,3b). It is useful for observing how much land meeting the farmers' preferences remains. As seen on the map, there are few areas that can successfully be incorporated to agriculture. These areas would be still fewer if we set aside the narrow strips of land adjacent to roads. Two of the three relations in the example model are proximity relations where the shorter is the distance the higher is the preference. That is why the map shows land with high attractability near roads and previous farmland.

Optionally, a third map could group portions of land with values above a specific threshold to display the most attractive land (Fig. 8.3c). We classified attractive land into two classes, unsuitable and suitable, to illustrate this. According to the map, the areas in blue should be used first for new farmland, for they represent areas where local development / environmental conditions render them attractive.

Parametric mode of the AGRILOCAL model

Individual parameter maps can be generated from the un-weighted functions of each relation to see how land is rated on every single parameter, i.e. the parametric attractability. (Fig. 8.4).

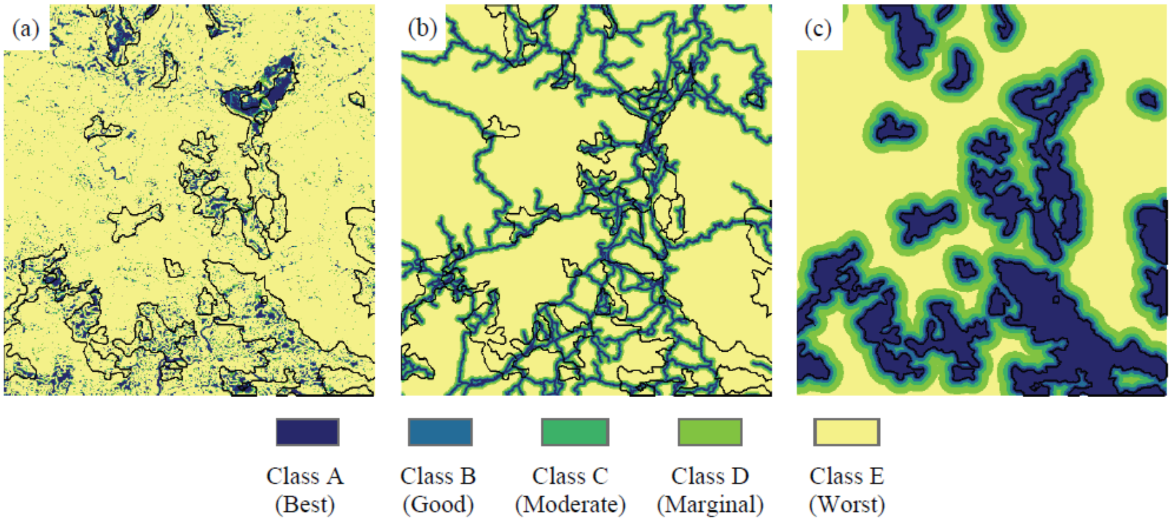


Figure 8. 4 Land attractability according to each individual parameter: a) slope; b) proximity to roads; c) proximity to previous farmland (1976).

This mode requires the calculation and analysis of residual maps for each parameter. The maps can be calculated by subtracting the values of an individual parameter map from the full model map, as generated in the predictive mode. The analysis focuses on the detection of anomalies that could indicate potential targets for new farmland or problems with current farmland. Two characteristics of the residuals are important to detect the anomalies: magnitude and direction (positive / negative). There are five possible cases of interest:

1. If there is some land still attractive for agriculture under a specific parameter, expect a fair number of relatively small-extent and large-value negative anomalies (-1.0 to -0.4) outside the areas of current farmland. These zones are land that farmers would consider as preferable or acceptable for the practice of agriculture if the only factor to consider were the relation embodied by the parameter under investigation. These zones can be regarded as potential areas for new farmland, provided their size and the conditions in other relations are rendered attractive.
2. Zones with large-value negative anomalies (-1.0 to -0.4) inside current farmland areas correspond to portions of land where the practice of agriculture is primarily dictated by a specific parameter; that is, farmers mostly use that land because its parametric attractability is very high.
3. Zones with relatively small negative or positive values (-0.4 to 0.4), inside or outside current farmland, indicate an acceptable agreement in quality values between the parametric and the full model attractability; they should not be considered anomalies but “normal” deviations around the predicted attractability values.
4. Zones with large positive (0.4 to 1.0) values inside current farmland areas indicate problems related to the presence of less-than-acceptable local values for a particular relation. The significance of a problem rests on the anomaly’s value and the parameter’s weight: the higher the value of both the more severe is the problem. Information on these anomalies could serve to promote non-agricultural uses for this land or to apply particular conservation measures, depending on the type of parameter.
5. Zones with large positive (0.4 to 1.0) values outside current farmland areas suggest the presence of land where the values of the other parameters are relatively high but not those of the explored parameter. In general they are attractive areas, but are severely limited by the parametric attractiveness under consideration.

If the study area is large, a further use of anomalies would be to separate it into smaller regions where anomalies tend to be homogenous, in magnitude and direction, so that more specific AGRILOCAL models are built for these areas.

As an example of parametric analysis, we show the map of residuals for the FS relation (Fig.8.5) and provide the results of the corresponding analysis of anomalies.

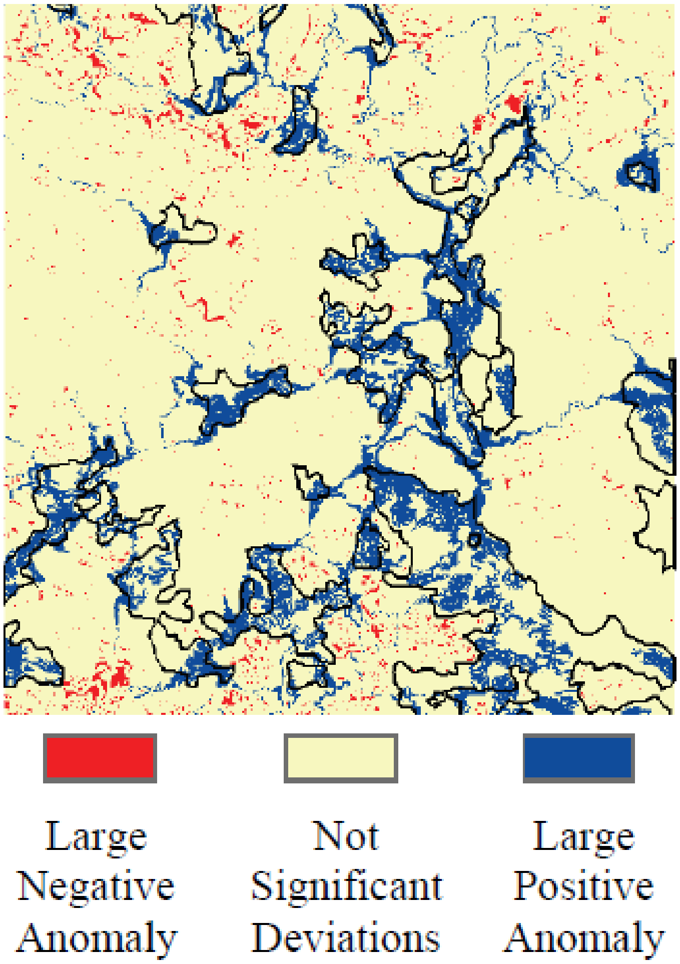


Figure 8. 5 Map of anomalies derived from the map of residuals for the slope parameter.

The anomalies of relation FS belong to cases a, c, d, and e; there are no cases of type b. Large negative anomalies (in red in Fig. 8.5) are exclusively outside farmland areas. These are land with slopes considered preferred or acceptable by local farmers, but the absence of roads and its position, relatively far from the border of existing farmland, decrease its attractability. The zones in light yellow comprise most of the land and indicate that slope is within the predicted attractability: zones outside farmland have unsuitable slopes according to farmers; zones inside farmland have slope values that conform to the farmers' preferences. The zones in dark blue inside farmland specify land where the slope does not have what farmers consider preferred values. However, it has high attractability values for the other two parameters, explaining its farmland status. These areas could be more prone to erosion than the rest of farmland, and could be targeted for soil conservation programs. Finally, the zones in dark blue outside farmland areas specify land that, despite its high attractability according to the other parameters, is severely limited by unsuitable slopes according to farmers.

8.4 DISCUSSION

Attractability as a measure of land suitability is a missing element in land evaluation studies. The model here developed contributes to fill this gap by calculating attractability values estimating farmers' actual preferences. The results of the model show that it is possible to accurately estimate values for many spatial relations, using a farmland distribution map as the basic input. However, the model has some limitations (all of them controllable at varying extents) that it is important to be aware of: a) the validity of the modeled relations is highly dependent on the dynamics of land use in the region; b) the model is sensitive to the number and nature of parameters included; c) the results can be strongly affected by problems inherent to many spatial models, mainly edge and size effects. Limitations notwithstanding, the model's utility in land evaluation for agricultural purposes goes beyond estimating land attractability; it also can be used to investigate aspects of the sustainability of existing agricultural systems, complement the results of LC and AEZ studies, and build and evaluate trend and intervention scenarios for agricultural expansion, among other foreseeable applications.

Validation of the AGRILOCAL results

To test the results' validity we can compare the land attractability (predicted with the AGRILOCAL model) against the actual distribution of farmland. Multiplying the map of farmland by the full model map will give us the attractability values within current farmland areas. In general, we can say that results are valid if most land (≥50%) with high attractability values (0.6 – 1.0) is also actual farmland. Portions of low-attractability values should be absent or minimal (Table 8.2), although a deviation from this norm, i.e. presence of large portions of farmland with low-attractability values, would not necessarily indicate a failure of the model; instead, it may suggest that agricultural practices have diverged farther from the preferred or acceptable local practices as the agricultural frontier has expanded.

The results (Table 8.2) confirm that the model is predicting reasonably well the expansion and current situation (year 2000) of agriculture in the study area. For comparison, we also show similar results for a previous situation (year 1976). From 1976 to 2000 farmland expanded almost 5 times, but this increase has been at the expense of using lower quality land. As expected in the model, in 1976, most farmland was of best and good quality (61%), with nearly a third (31%) being of moderate quality, and only few hectares (7%) of marginal quality. No land of worst quality was used because there still remained some land of Class A and fair amounts of Class B and C. Also, most of the best (74%) and almost half of the good (48%) quality lands in the entire study area were already being used for agriculture. The situation had changed drastically by 2000: while most of the best (85%) and good (73%) quality land were already being used, half (50%) of the moderate quality and even a noticeable portion (14%) of the marginal quality land of the entire study area had become farmland.

Table 8.2 AGRILOCAL results for two dates of farmland data*.

Land Class (Quality)	AGRILOCAL values (Attractability)	Land in study area (ha.)	Farmland in 1976 (ha.)	Farmland in 2000 (ha.)	Farmland in 1976 as percentage of land	Farmland in 2000 as percentage of land	Percentage of total farmland in 1976	Percentage of total farmland in 2000
A (Best)	0.8 - 1.0	3,376.00	1,395.50	2,875.75	74.62	85.18	20.5	9.89
B (Good)	0.6 - 0.8	12,315.25	2,784.75	9,036.25	48.33	73.37	41.01	31.08
C (Moderate)	0.4 - 0.6	23,844.75	2,124.50	11,961.75	15.28	50.17	31.29	41.15
D (Marginal)	0.2 - 0.4	39,957.00	485.50	5,197.75	1.63	14.06	7.15	17.88
E (Worst)	0.0 - 0.2	74,538.50	0.00	0.00	0.00	0.00	0.00	0.00
	Totals	151,031.50	6,790.25	29,071.75	4.49	19.25	100.00	100.00

* Farmland data sources:1:250,000 Land Use / Land Cover Map (INEGI, 1976), and National Forest Inventory 2000 (SEMARNAT, 2001).

The portions of marginal and moderate quality land in farmland together accounted for more than half (58%) of total farmland. That is, most farmland no longer exhibited what farmers considered the preferred values. While the data confirms that most of the land with high predicted attractability values (0.6 to 1.0) was farmland (roughly 11,800 of 15,600 hectares, 75%), as expected in the model, it also shows that the agriculture frontier had moved into lower-quality lands with respect to the preferred practices. The remaining percentage of high quality land (Classes A and B , 25%) not used in agriculture can be explained by considering other existing land covers and uses for this area, such as urban areas and remnant tropical forest.

Limitations of the model

The model might be sensitive to the number of relevant relations included. If more or fewer relations are included, the spatial distribution of the predicted attractiveness of land may be significantly different. This cannot be avoided, as it is related on one hand to the availability and quality of spatial information, and on the other to the number of spatial relations thought to be relevant for agriculture in a region, although this last situation can be somewhat alleviated by interviewing local farmers.

When applied to the modeling of the attractability of land for agriculture as a general activity, the model provides no information about the attractability of land for specific crops, though this could be overcome by using crop-specific information, i.e. with a farmland map only showing land where the crops of interest are being cultivated.

A change in the scale, the extent, or the place of the analysis affects the form of the parameters and the outcome of the AGRILOCAL model. Also, the form is affected by the temporal interval of the information used to build the model. Since these issues can be controlled in the modeling process, we discuss here their main effects.

Temporal limitations

As in any other economic activity, the practice of agriculture evolves in response to many environmental and development factors, and so do farmers' preferences. Therefore, the outcome of an AGRILOCAL model may have a time-limited utility. Since not all spatial relations evolve at the same pace, it could be hard to estimate this utility time-frame, but in places with a very dynamic transformation, a model may have a reduced life. This limitation can be partially overcome by using at least three (or more) evenly spaced, dates of land use maps, to ascertain whether a significant shift in preferences has taken place recently, and consequently, a model built with the last two dates will be more representative of current preferences.

Despite this limitation, a major difference with respect to conventional approaches is that in them, a portion of land with an estimated suitability will never change its status unless a drastic change in the environmental conditions occurs, whereas in an AGRILOCAL study the attractability of land varies with the evolution of development and environmental conditions, with some formerly unsuitable land becoming suitable and vice versa. The AGRILOCAL model is then capable of recording the dynamics of the activity, provided that appropriate information is available.

The effects of scale and resolution

In many spatial models, scale and resolution are considered to be equivalent concepts. Kok and Veldkamp (2001), evaluating the impact of spatial scale on land use patterns, found that although the concept of resolution is linked to the concept of scale, conducting an analysis in terms of one or the other of these can lead to different conclusions. Even when the scale of analysis is not modified, changing the resolution of the data may bring changes in the results. In spatial analysis, the general effect of using a coarser-than-appropriate resolution is a reduction of heterogeneity (King, 1991), which may be significant depending on the relation investigated. Using a finer resolution than necessary usually does not result in significant changes other than increasing the physical size of the computer files used to store data. In terms of the AGRILOCAL model, a change in resolution does not affect the relevancy (number of parameters) and influence (weights) of spatial relations because although changes in the form of the distribution of the actual relation are noticeable, similar changes occur in the corresponding random distribution, thus preserving the p values that determine relevancy and the D values used to derive weights. On the other hand, alterations of the model's preference values may be significant at all resolutions, with the general effect of increasing concentration and minimizing dispersion when resolution is coarser, so that the classes with higher frequencies will have higher probabilities, and hence higher levels of preference, as resolution decreases.

A good rule is to use a resolution where the size of a spatial unit (a square cell in a raster grid) has a length equal to ¼ of the length of a side of the minimum mapping unit (MMU), according to the scale of analysis or of the original dataset. Cartographical standards (Anson and Ormeling, 1994) establish that, for any map at any scale, the physical length of the side of a MMU should be 2 millimeters; therefore, the cell size of a map in the analysis should be equivalent to 0.5 mm.

The effect of extent

Since one of the uses of the model aims at extracting trends in the spatial distribution of agriculture, varying the extent of what is considered "local" has a more substantial effect than changing scale-resolution. King (1991) says that larger extents tend to include more processes and that previously considered important processes may lose significance. In the AGRILOCAL model, if the extent is small, very likely the relevancy of relations, preference of values, and weight of relations will differ radically from those corresponding to a much wider region where the smaller extent is inserted. These effects are tied to a more general problem in spatial analysis known as the edge effect.

Relevancy is the first aspect affected, with the number and type of relevant spatial relations in the model depending on the size of the extent. Some relations might not be affected, however; for instance, coincidence of agriculture with certain values or classes of slope (flat or gentle slope) is as relevant at the plot scale as it is at the continental. Some others, such as the proximity to roads, might be severely affected. If, for example, in the case of small extent, a road is relatively far from agricultural areas (because other roads outside the extent are nearer or adjacent to the same agricultural areas), the relation will appear to be of little relevancy; conversely, if a road crosses agricultural areas within such a small extent, the relation will appear as highly relevant.

Preference and weight can also be affected. In the above example, proximity values indicating adjacency or nearness may not exist in the first case, resulting in high probabilities associated with relatively long distances from a road (Fig. 8.6).

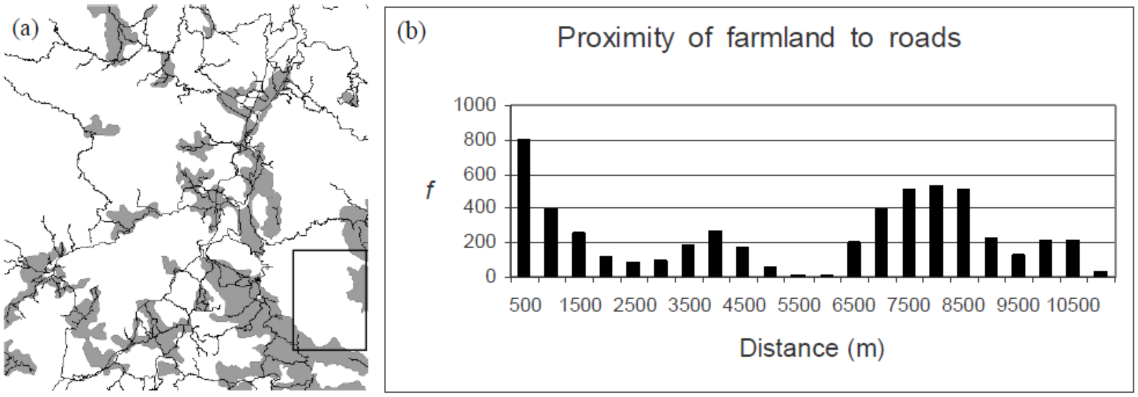


Figure 8. 6 Pattern of a spatial relationship (PFR) corresponding to a smaller extent (rectangle inside map) within a larger region. a) spatial distribution of farmland (year 2000) and roads; b) frequency distribution of distances of farmland to the nearest road corresponding to the area in the rectangle in a.

Another effect is that the smaller the extent the higher the possibility of obtaining an atypical result (from the regional point of view). Moreover, it is very likely that, for some relations, no regular trends can be observed because these only occur when a larger set of measurements is taken. In addition, there might appear to be more than one trend when in reality this appearance might be a product of an edge effect (compare Fig. 8.6b with Fig. 8.2a).

As a general rule, to avoid edge / extent effects as much as possible, the selected extent should include a buffer area of reasonable size. If care is taken to avoid edge effects, results for a small extent are valid for that extent, although they cannot be extrapolated to a larger area. However, the opposite is not entirely true: results obtained with a larger extent may be valid for smaller extents within the larger region, with some restrictions.

The effect of place

It is common that differences in place entail differences in the spatial patterns of agriculture, and consequently, different relevancy, preference, and influence of spatial relations. Results obtained in one place may not be extrapolated to other places even when the scale and extent are the same. This may be held to be true even if the environmental and socioeconomic conditions of two or more places are similar (Milleville and Dubois, 1978). But whereas results cannot be extrapolated to other places, they are comparable across them, provided that the same analytical framework for determining relevancy, preference, and weights is used in the construction of the specific models.

Utility of the AGRILOCAL model

The basic hypotheses of the model assume that farmers' expert knowledge is contained in the spatial distribution of agriculture, and that this knowledge is used to carry out the activity in the best possible way, given a set of local constraints and favorable factors. This might also be interpreted in the sense that farmers' practices somehow contain a measure of the sustainability of agriculture. This 'implicit' sustainability could also be reflecting the effect of environmental as well as development factors and constraints. For instance, proximity relations of farmlands to roads or to human settlements reflect economic concerns, while other spatial relations are linked to environmental concerns.

It could be even argued that this assumed sustainable behavior of farmers also includes the effect of regional, national and even world conditions. But although external conditions may be embedded in the patterns of agriculture, modeling them would require the inclusion of those spatial relations related to the organization of space, namely connectivity, aggregation, and association. Because of their complex nature, these relations are difficult to express in simple measures, and a way to explicitly include them in the AGRILOCAL model has yet to be explored. Nonetheless, this alleged sustainability is what makes the model valid for estimating the location and amount of land attractive for agricultural purposes.

Our emphasis on land attractability may give the false idea that the model's usefulness is restricted to the investigation of this characteristic. Instead, the model can also be successfully used in a variety of other problem-solving and geographic contexts:

- In countries where there is still land to expand the agriculture frontier, the model could be used to discover where and how much land is attractive and to assess the corresponding levels of attractability in terms of local practices. (This remains the primary use of the model.)
- In countries where the option of bringing new land to production has reached its limits, the model may be used to investigate the spatial conditions of farmers' practices and determine which portions of farmland are beyond those levels considered to be commonly acceptable or optimal (the model's parametric analysis of anomalies would serve this purpose well). This information may be used to investigate the sustainability and risk levels of agriculture under a variety of scenarios.
- In countries where LC or AEZ studies have already been conducted, the model may be useful to optimize their results by investigating which local practices might be taken into consideration to make land found to be suitable also attractive. A complementary idea would be to make land attractability a conscious goal of regional infrastructure projects. For instance, when building a road or a bridge, it would seem wise to carry out analysis not only to improve the connectivity of the transportation system but for the benefit of agriculture as well.
- The model can also be used to create trend and intervention scenarios of expansion of the agricultural frontier, in order to assess if perhaps it would affect the environment or alter development relations (e.g. through deforestation, loss of soil fertility, or aquifer overexploitation). Results could be used to avoid negative impacts by, for example, requiring

that road construction facilitate access to new lands, or forbidding settlements in areas where locally-relevant relations would render the land attractive but leave it exposed to unfriendly environmental practices.

Concluding remarks and future research

Consideration of local farmers' practices is of great importance in any planning exercise, not only so that local knowledge can be incorporated into the processes of land evaluation, enhancing the accuracy of results, but also so that agricultural plans can successfully reach the implementation stage. Although the various models for land evaluation should be regarded as complementary rather than competing, the AGRILOCAL model with its emphasis on attractability might be better suited than other models to answer quick demands for new agricultural land, as well as to improve the long term strategic planning of agriculture.

In a technical sense, the success of the AGRILOCAL model may be ascribed to its capacity to infer preferable, acceptable or desirable agricultural practices from the spatial distribution of farmland, while other models, though also using a spatial framework, focus on the characteristics of biophysical requirements of crops rather than on the spatial relations of agriculture.

The AGRILOCAL model has enormous potential to assist in research and development tasks involving the practice and planning of agriculture. Important future research topics include investigating:

- the specificity of AGRILOCAL models for different types of crops or different varieties of a single crop. This would reveal farmers' knowledge about how best to use land for different types of cultivars.
- the specificity of AGRILOCAL models for different geographic settings, involving a variation or similarity of development and environmental factors and constraints. This could provide information on the specific responses of farmers facing different development and environmental scenarios.
- the evolution of the spatial patterns of agriculture in a single region to assess whether there are temporal trends in addition to spatial ones, and use this information to improve modeling the possible future states of the activity in a region.
- how well the relevancy and weights of the spatial relations, obtained using different quantitative methods, match the relevancy and weights that farmers explicitly indicate.
- to what extent the values of one relation can be extended in order to assess them as preferable or acceptable for agricultural practices when the values of other relations are also present or absent.

We conclude that in order to determine the availability of new land for agriculture, in the sense that the FAO gives to this concept, the best possible approach would be to combine the LC, AEZ and AGRILOCAL methodologies. Our specific recommendation is to use first the AGRILOCAL approach, in a regional planning context of the activity, and then conduct detailed LC and/or AEZ studies on those areas estimated as attractive for agricultural activities from the local perspective. Proceeding the other way around might produce results of limited practical use.

CONCLUSIONS

As work on this dissertation evolved, I was able to confirm the actuality of the three main issues that, in my view, have delayed the progress on true geographic thinking and reasoning, and have hindered the (re)unification of the discipline: scarcity of theoretical developments *about* geographic space, lack of structure and definition in the use of fundamental spatial concepts (such as those presented in this work), and insufficient dissemination of theoretical findings among old and new practitioners of the discipline. These three issues are closely related, with the first and the last being of major consequences to the advance of the discipline. Within the context of these issues, I can offer these concluding remarks:

1. The systematization of knowledge is a task necessary to facilitate its use in scientific contexts. The generation and utilization of knowledge without regards for its position in the cognitive structure of science may create confusion instead of making knowledge more simply to understand and use.

2. Theory in geography suffers from being mostly perspective-oriented (regionalism, exceptionalism, Marxism, environmentalism, postmodernism, information science, behaviorism, structuralism, historicism, constructivism, etc.) relegating fundamental theory about geographic space as unimportant. Consequently, few efforts have been recently done in this last direction.

3. In view of the scarcity of fundamental theoretical developments, practitioners of geography usually adopt and adapt theories from related disciplines, but although some could claim that the discipline is becoming richer by this integration of external theories, the truth is that it is losing identity as a discipline in its own.

4. Since theory about geographic space is inadequate and incomplete, the language of geography has been built with insufficient attention to structure and semantics. As a result, it is poorly defined, ill structured, and prone to misuse.

5. Concepts such as spatial relations, although widely used in geography and in other spatially-aware disciplines, are used empirically and not formally, and when formalisms have been employed these have had a limited disciplinary scope, often tied to narrow fields of knowledge in mathematics or in information science, instead of being derived from a comprehensive conception of space-time.

6. Since language is inseparably united to cognition, it is anticipated that language contains indications about the cognition structure that gave rise to the concept of spatial relations, both, as for the nature and the number of spatial relations.

7. Regarding the nature of ‘spatial relations among events’, the expression itself suggests that there are two primary categories of cognition, that of *space* (space-time) and that of *event* (object, entity, feature, thing, etc.), that if combined somehow should result in a

third notion relating space with events. Space and events are two different views of the same thing rather than two different entities.

8. One way of relating both primary notions is through their respective affordances or fundamental ‘properties’. On one side, space affords possibilities: the possibility to measure distances, the possibility to measure directions and the possibility to measure concentration, being these the fundamental properties of space. On the other side, events afford capacities: the capacity for connecting, the capacity to integrate into higher-order events, and the capacity to decide how to group together, being these fundamental properties of events.

9. Concerning the number of spatial relations, if we consider that each spatial property has a counterpart in an event property, the interaction among opposed and complementary properties can only result in the dominance of the property of the space, the dominance of the property of events, or in a neutral situation. Since for each of these possibilities of interaction there are three properties, the number of spatial relations must be nine.

10. Those relations where the properties of space dominate are called structural because they describe the way the fabric of space determines the location and characteristics of events. Those relations where the properties of events dominate are called organizational because they describe the way the structure of space is modified by the organization of the events. Where neither set of properties dominate, the relations are called neutral, because they can be used equally to describe the structure or the organization of space.

11. When using these concepts in spatial problem solving contexts, it is not necessary to investigate the magnitudes of all relations but only of those that are relevant to a particular problem solving situation. In some cases the interest could be in the investigation of the structure of space, in others is the organization of space the focus of the inquiry, and thus the appropriate relations have to be considered.

12. The use of these concepts must be done within the context of a geographical methodological framework, identifying their specific roles in problem-solving, as mid-tiers between spatial events and spatial processes.

13. Spatial relations are concepts useful to describe, measure, estimate, and calculate the magnitudes of interactions among spatial events, and when variations of these magnitudes are taken into consideration, they are also useful to understand processes.

14. Since space is a unique concept, spatial relations function as indicators of interaction among events in *any* space, being highly suited to explore and solve geographic situations, where events in social and economic spaces interact with events in natural spaces.

15. Dissemination of these findings is a necessary following step. This must be done concurrently among those that are already practitioners of geography and those that are being introduced in the field. Of course, preferred ways of disseminating knowledge need

to be different: publications, seminars and professional meetings in the first case; lectures, dissertations and practical work in research projects in the second.

REFERENCES

- Abler, R. F. 1987. The National Science Foundation Center for Geographic Information and Analysis. International Journal of Geographical Information Systems. 1(4):303-326
- Adger, W. N. 2006. Vulnerability. Global Environmental Change 16 (3):268-281
- Agronomic Interpretations Working Group, 1995. Land Suitability Rating System for Agricultural Crops: 1. Spring-seeded small grains. Edited by W.W. Pettapiece. Tech. Bull. 1995-6E. Centre for Land and Biological Resources Research, Agriculture and Agri-Food Canada, Ottawa. 90 pages, 2 maps.
- Alwang, J., Siegel, P.B. and Jorgensen, S.L. 2001. Vulnerability: a view from different disciplines. Social Protection Discussion Paper Series No. 0115, Human Development Network, The World Bank.
- Anson, R., Ormeling, F., 1994. Basic Cartography. Vol 1. Butterworth – Heinemann.
- Barbier, E.B., 1997. The Economic Determinants of Land Degradation in Developing Countries. Philosophical Transactions: Biological Sciences, Vol. 352, No. 1356, Land Resources: On the Edge of the Malthusian Precipice?. pp. 891-899.
- Bennett, A. J., 2000. Environmental consequences of increasing production: some current perspectives. Agriculture, Ecosystems and Environment, 82:89–95.
- Béra, R. and Claramunt, C. (2003) Relative adajacencies in spatial pseudo-partitions. COSIT'03 Doctoral Colloquium. Available online at http://www.spatial.maine.edu/~cosit03/dc_home.html
- Bierwisch, M. (1996). How much space gets into language? In: Language and space. Boom, P., Peterson, M.A., Nadel, L., and Merrill F. Garrett (eds.). The MIT Press, 597 pp
- Blaikie, P., Cannon, T., Davis, I., Wisner, B., 1994. At Risk: Natural Hazards, People's Vulnerability, and Disasters. Routledge, New York, NY.
- Burton, W. 1979. Logical and physical data types in geographic information systems. Geo-Processing. 1:167-181
- Burton, I., Kates, R.W., White, G.F., 1978. The Environment as Hazard. Oxford University Press, New York, NY.
- Butzer, K.W., 1980. Adaptation to global environmental change. Professional Geographer 32 (3), 269–278.
- Cardona, O.D. 2003. The need for rethinking the concepts of vulnerability and risk from a holistic perspective: A necessary review and criticism for effective risk management. In: Mapping Vulnerability: Disasters, Development and People. G. Bankoff, G. Frerks, D. Hilhorst (Ed.), Earthscan Publishers, London.
- Carlson-Radvansky, L. A. (1999). *Spatial Cognition and Computation, 1, 365-379*
- Cutter, S.L. 1996. Vulnerability to environmental hazards. Progress in Human Geography 20,4: 529-539
- Cutter, S.L., Mitchell, J.T., and M. S. Scott. 2000. Revealing the vulnerability of people and places: a case study of Georgetown County, South Carolina. Annals of the Association of American Geographers. Vol. 90, 4: 713-737
- Cutter, S.L. 2003. The vulnerability of science and the science of vulnerability. Presidential Address. Annals of the Association of American Geographers. Vol. 93, 1: 1
- De Blij, H.J. (1987). Confusion of Innovations. President's Session, Association of American Geographers. 1987 Meeting in Portland, Oregon. Available online at http://deblij.net/opinion04231987.html (last verified: 12/sep/2006).

- Dow, K., 1992. Exploring differences in our common future(s) — The meaning of vulnerability to global environmental change. Geoforum 23 (3), 417–436.
- Duvernoy, I., 2000. Use of a land cover model to identify farm types in the Misiones agrarian frontier (Argentina). Agricultural Systems 64:137-149
- Eastman, J, R (2001) IDRISI, Release 2. Guide to GIS and Image Processing, Volume 2, May 2001.
- Egenhofer, M.J. and Herring, J.R. 1991. High-level spatial data structures for GIS. In: Maguire, D.J., Goodchild, M.F., Rhind, D.W. (eds) Geographical Information Systems. Principles and Applications. Longman. 1:227-237
- Eliot, J. 2000. The nature and measurement of spatial intelligence. College Park, MD: Institute for Child Study, College of Education, University of Maryland
- Emery, K.A., 1986. Rural Land Capability Mapping. Soil Conservation Service of NSW.
- ESFS. 2004. Hazards – minimizing risk, maximizing awareness. Available online at http://www.yearofplanetearth.org/downloads.htm (last verified: 9/oct/06)
- FAO, 1978. Report on the agroecological zones project. World Soil Resource Report No. 48. Rome: FAO.
- FAO, 2007. Quantifying land resources in developing countries - FAO's agro-ecological zones studies. Land and Water Development Division, Food and Agriculture Organization of the United Nations. Available online at: http://www.fao.org/docrep/T8300E/t8300e00.HTM (last verified: 8/JAN/07)
- Feuchtwanger, M.1989. Geographical logical database model requirements. Proceedings Autocarto 9. 599-609
- Fischer, G., 1999. The LUC land productivity assessment. Options Magazine, Summer 1999. IIASA. Available online at: http://www.iiasa.ac.at/Admin/INF/OPT/Summer99/the_luc_land_productivity_assessment.htm (last verified: 8/JAN/07)
- Freeman, J. 1975. The modeling of spatial relations. Computer Graphics and Image Processing 4:156-171
- Füssel, H. M. & Klein, R.J.T., 2002. Assessing vulnerability and adaptation to climate change: an evolution of conceptual thinking. UNDP Expert Group Meeting on "Integrating disaster reduction and adaptation to climate change". Havana, Cuba. June, 2002.
- Füssel, H.-M., 2005: Vulnerability in Climate Change Research: A Comprehensive Conceptual Framework. Breslauer Symposium Paper 6 and FAVAIA Working Paper 2, University of California International and Area Studies, Berkeley, CA, USA, ii+34 pp. Available online at http://portal.pik-potsdam.de/research/current/topik6/favaia/ workingpapers (last verified: 4/oct/06)
- Girard, N., Bellon, S., Hubert, B., Lardon, S., Moulin, C.H., Osty, P.L., 2001. Categorising combinations of farmers' land use practices: an approach based on examples of sheep farms in the south of France. Agronomie, 21:435–459
- Golledge, R. (2001). What is what Geographers do? Directions Magazine. July 02, 2001 issue. Available online at http://www.directionsmag.com/columns.php?column_id=41 (last verified: 6/sep/2006)
- Golledge, R,. G., (2002) The Nature of Geographic Knowledge. Presidential Address. Annals of the Association of American Geographers, 92(1):1-14
- Goodchild, M.F., and K.K. Kemp, eds. 1990. *NCGIA Core Curriculum in GIS.* National Center for Geographic Information and Analysis, University of California, Santa Barbara CA. Units 12 and 15.

- Goodchild, M. F. 1992. Geographical Information Science. International Journal od Geographical Information Systems 6:31-45
- Grunbaum, A. 1963. Philosophical problems of space and time. A. A. Knopf, New York 1963
- Gustafson, E.J., January, 2005; available online at www.ncrs.fs.fed.us/people/Gustafson
- Hinkel, J., 2005: DIVA: an iterative method for building modular integrated models. Advances in Geosciences, 4, 45-50.
- Hinkel, J. and R. Klein (2006). Draft chapter forthcoming in "Managing Coastal Vulnerability: An Integrated Approach", edited by L. McFadden, R.J. Nicholls and E.C. Penning-Rowsell, Elsevier Science, Amsterdam, The Netherlands.
- IHDP. 1999. Global Environmental Change and Human Security Project Science Plan. IHDP Report No. 11, 1999. Available online at http://www.ihdp.uni-bonn.de/html/publications/reports/report11/gehssp.htm (last verified October 2006)
- IIASA. 2006. Risk and Vulnerability Program Research Plan 2006-2010. IIASA. Available online at http://www.iiasa.ac.at/Research/RAV/RAVPlan.pdf
- INEGI, 1976. Cartografía del uso y cubierta del suelo a escala 1:50,000. México
- Ionescu, C., R.J.T. Klein, K.S. Kavi Kumar, J. Hinkel and R. Klein, 2005: Towards a Formal Framework of Vulnerability to Climate Change. NeWater Working Paper 2 and FAVAIA Working Paper 1, Potsdam Institute for Climate Impact Research, Potsdam, Germany, ii+20 pp.
- IPCC. 2001a. Summary for policy makers, Third Assessment Report of Working Group I of the Intergovernmental Panel on Climate Change (http://www.ipcc.ch/pub/spm22-01.pdf).
- IPCC. 2001b. Climate Change 2001: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change (http://www.grida.no/climate/ipcc_tar/wg2/index.htm).
- IPCC CZMS, 1992: A common methodology for assessing vulnerability to sea-level rise. Second revision. In: Global Climate Change and the Rising Challenge of the Sea, Report of the Coastal Zone Management Subgroup, Response Strategies Working Group of the Intergovernmental Panel on Climate Change, Ministry of Transport, Public Works and Water Management, The Hague, The Netherlands, Appendix C, 27 pp.
- Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In: Language and space. Boom, P., Peterson, M.A., Nadel, L., and Merrill F. Garrett (eds.). The MIT Press, 597 pp.
- Jammer, M. 1993. Concepts of space: The history of theories of space in Physics. Dover Publications Inc., New York 1993.
- Janssen, M.A., Schoon, M.L., Ke, W., and Börner, K. 2006. Scholarly networks on resilience, vulnerability and adaptation within the human dimensions of global environmental change. Global Environmental Change 16: 240-252.
- Jenness, J. 2004. The Nearest Features Extension. http://www.jennessent.com/arcview/nearest_features.htm
- Kasperson, J.X., Kasperson, R.E., Turner II, B.L., 1995. Regions at Risk. United Nations Press, New York, NY.
- Kates, R.W., Clark, W.C., Corell, R., Hall, J.M., Jaeger, C.C., Lowe, I., McCarthy, J.J., Schellnhuber, H.J., Bolin, B., Dickson, N.M., Faucheux, S., Gallopin, G.C., Grubler, A., Huntley, B., Jager, J., Jodha, N.S., Kasperson, R.E., Mabogunje, A., Matson, P., Mooney, H., 2001. Environment and development—sustainability science. Science 292 (5517), 641–642.
- King, A.W., 1991. Translating models across scales in the landscape. In: Turner, M.G., Gardner, R.H. (Eds.), Quantitative Methods in Landscape Ecology. Ecological Studies, Vol. 82. Springer, Berlin, pp. 239–288.

- Kok, K., Veldkamp, A., 2001. Evaluating the impact of spatial scales on land use pattern analysis in Central America. Agriculture, Ecosystems and Environment, 85:205–221.
- Lefebvre, H. (1991). The production of Space. London: Basil Blackwell.
- Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial description. In: Language and space. Boom, P., Peterson, M.A., Nadel, L., and Merrill F. Garrett (eds.). The MIT Press, 597 pp.
- Levinson, S.C. (1996). Frames of reference and Molyneaux's question: crosslinguistic evidence. In: Language and space. Boom, P., Peterson, M.A., Nadel, L., and Merrill F. Garrett (eds.). The MIT Press, 597 pp.
- Liverman, D.M., 1990. Vulnerability to global environmental change. In: Kasperson, R.E., Dow, D.G., Kasperson, J.X. (Eds.), Understanding Global Environmental Change: The Contributions of Risk Analysis and Management. Clark University, Worchester, MA.
- Luers, A.L., Lobell, D.B., Sklar, L.S., Addams, C.L., Matson, P.A., 2003. A method for quantifying vulnerability, applied to the agricultural system of the Yaqui Valley, Mexico. Global Environmental Change 13, 255–267.
- Milleville, P., Dubois, J.P., 1978. Réponses paysannes à une opération de mise en valeur de terres neuves au Sénégal, in: CNRST/ORSTOM (Eds.), Maîtrise de l'espace agraire et développement en Afrique Tropicale, Logique paysanne et rationalité technique, Actes du colloque de Ouagadougou (4–8 déc. 1978), 1978, pp. 513–527.
- Minkowski, H. (1908/9). "Raum und Zeit". Jahresberichte der Deutschen Mathematiker-Vereinigung: 75–88
- Minsky, M. 1988. The society of mind. Simon and Schuster, New York. March 15, 1988
- Molenaar, M. 1989. Towards a geographic information theory. ITC Journal. 1:5-11
- Musser, L. 2002. Vulnerability Bibliography. Center for Science and Technology Policy Research, University of Colorado-Boulder.
- Norton, E.A., 1939. Soil Conservation Survey Handbook. Physical Surveys Division, USDA.
- Nystuen, J. D. 1963. Identification of some fundamental practical concepts. In: Berry J. L. B. & Marble, D. F. (eds.) Spatial Analysis. A reader in statistical Geography. Prentice Hall. Englewood Cliff. 35- 41
- O'Brien, K., R. Leichenko, U. Kelkar, H. Venema, G. Aandahl, H. Tompkins, A. Javed, S. Bhadwal, S. Barg, L. Nygaard, and J. West: 2004. Mapping vulnerability to multiple stressors: climate change and globalization in India. *Global Environmental Change* 14:303–313.
- O'Keefe, J. and Nadel, L. (1978). The hippocampus as a cognitive map. Oxford, Clarendon Press.
- O'Keefe, J. (1996). The spatial prepositions in English, vector grammar, and the Cognitive Map Theory. In: Language and space. Boom, P., Peterson, M.A., Nadel, L., and Merrill F. Garrett (eds.). The MIT Press, 597 pp.
- OCG (1999) OpenGIS simple features specification for SQL, Revision 1.1. Available at www.opengis.org
- Peuquet, D.J. 1984a. Knowledge-based engineering for spatial database management and use. In: Proceedings of the First International Symposium on Spatial Data Handling. 2: 372-391
- Peuquet, D.J. 1986. The use of spatial relationships to aid spatial database retrieval. In: Proceedings of the Second International Symposium on Spatial Data Handling. 459-471
- Peuquet, D.J., Zhan, C. 1987. An algorithm to determine the directional relationship between arbitrarily shaped polygons in the plane. Pattern Recognition. 20 (1) 65-74
- Pullar, D. 1987. Query language for spatial relations. Technical Papers ASPRS – ACM Annual Convention. 5:180-192

- ___. 1988. Data definition and operators on a spatial data model. Technical Papers ASPRS – ACM Annual Convention. 2:196-202
- Robinove, C. J. 1977. Principles of logic and the use of digital geographic information systems. U. S. Geological Survey Professional Paper No. 977. U.S. Gov. Printing Office, Washington, D.C. p. 19.
- Saaty, T.L. (1977), "A scaling method for priorities in hierarchical structures", *Journal of Mathematical Psychology*, 15 (234-281).
- SEMARNAT, 2001. Inventario Nacional Forestal 2000. Instituto de Geografía, Universidad Nacional Autónoma de México y Secretaría de Medio Ambiente y Recursos Naturales. México.
- Sivakumar, M.V.K., Valentin, C., 1997. Agroecological Zones and the Assessment of Crop Production Potential. Philosophical Transactions: Biological Sciences, Vol. 352, No. 1356, Land Resources: On the Edge of the Malthusian Precipice?, pp. 907-916.
- Smith, T.R. and D.J. Peuquet, 1985. Control of spatial search for complex queries in a knowledge based geographic information system. In: Proceedings of the International Conference of the Remote Sensing Society and the Centre for Earth Resources Management. 439-452
- Swinburne , R. 1968. Space and time. Macmillan, London, 1968
- Taylor, E. J., 2001. Migration: new dimensions and characteristics, causes, consequences and implications for rural poverty. In: Stamoulis, K. G. (ed.). Food, Agriculture and Rural Development. Food and Agriculture Organization of the United Nations. Available online at: http://www.fao.org/DOCREP/003/X9808E/X9808E00.HTM (last verified: 8/JAN/07)
- The MathWorks (2005). http://www.mathworks.com/access/helpdesk/help/toolbox/fuzzy/a1033563774.html
- Timmerman, P., 1981. Vulnerability, resilience and the collapse of society. Environmental Monograph 1, Institute for Environmental Studies, Toronto University.
- Turner, B. L. II, P. A. Matson, R. E. Kasperson, & J. E. X. Kasperson. 2002. Vulnerability in Human-Environment Relationships. AAAS Symposium: Science and Technology for a Transition Toward Sustainability, Feb. 17, 2002, Boston, MA
- Turner, B. L. II, R. E. Kasperson, P. A. Matson, J. J. McCarthy, R. W. Corell, L. Christensen, N. Eckley, J. X. Kasperson, A. Luers, M. L. Martello, C. Polsky, A. Pulsipher, and A. Schiller. 2003. A framework for vulnerability analysis in sustainability science. PNAS, vol. 100, 14 (8074-8079). Available online at www.pnas.org/cgi/content/abstract/100/14/8074.
- UNESCO, 1999. UNESCO IIP GIS Project. Training Module on the Applications of Geographic Information Systems (GIS) for On-line Governance and Accessing Public Domain Information. Module G. Available on-line at http://gea.zvne.fer.hr/index.html
- USDA, 1973. Land-Capability Classification (LCC). USDA, SCS. Agriculture Handbook No. 210.
- USGS. 2005a. Geography for a changing world. A Science Strategy for the Geographic Research of the U.S. Geological Survey, 2005-2015.
- USGS, 2005b; available online at *http:// erg.usgs.gov /isb/pubs/gis_poster/index.html)*
- Vickers, G. (1981).Some implications of systems thinking. In: Systems Behavior. The Open University Press, 1981, 332 pp.
- Watson, R.T., Zinyowera, M.C., Moss, R.H. (Eds.), 1996. Climate Change 1995: Impacts, Adaptations, and Mitigation of Climate Change. Cambridge University Press, Cambridge, UK.
- World Bank, 1994. El Salvador-natural resources management study. Agriculture and Natural Resources Operations Division, Latin America and the Caribbean Region. Washington DC: The World Bank.

- Youngman, C.1978. A linguistic approach to map description. In: Dutton, G. (ed.) Harvard Papers on Geographic Information Systems. Vol. 7
- Zadeh, L., Fuzzy sets. Information Control 8, 338-353, 1965