Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989

STATUS AND PROBLEMS OF GEOGRAPHICAL INFORMATION SYSTEMS, A CONCEPTUAL APPROACH

M. Molenaar, Wageningen, the Netherlands

1. INTRODUCTION

When I received the invitation to give the first lecture of this Photogrammetric week, I considered it as an honour which should ofcourse be accepted. The subject gave enough freedom to fill three quatres of an hour. There is so much going on in the field of GIS that one could easely spend a full year travelling from one meeting on this topic to another. Lots of activities, but .... when looking through the proceedings of these meetings and through the journals one gets this foggy uneasy feeling. Questions arise like: "What is going on? Where are we now and what is the progress?" Of course there is progress, if we compare the possibilities of GIS now with ten years or even five years ago, the growth is tremendous. A large number of GIS applications have been realised in many disciplines. The expression "G.I.S." is "salonfähig", which means that managers and moneysupliers are using the word in the official and (more important) unofficial communication channels like parties and receptions. They are even able to pronounce it after their second or third drink, but one gets the impression that they often use it without a proper idea of its meaning. Most important is however that they are willing to invest money in GIS, for research and for the development of new applications. It is not necessary to be clearvoyant to predict a further growth in the nineties. But how long will that last, won't there be a feeling of deception after a while, like there has been in the field of remote sensing. That too was a promising field in the seventies, the expected potentials were great, it was almost treated as a kind of magic with its own priesthood. The deception came though when it seemed that the potentials could not be realised as easily and quickly as expected. Now remote sensing people have to fight hard to get their research funds in competition with others. They are no wonder children anymore. This does not mean that remote sensing is not a powerfull tool, but rather that it takes much more time and hard work to realize its potentials, to extract from r.s. data the information required in the different application fields.

A similar situation may occur in the field of GIS. At this moment people are impressed by the progress which has been made with the present tools for data handling and display. In a few seconds maps can be displayed on a graphics screen, which took many hours for drafting even five years ago. Queries can be applied to the stored spatial data and map overlays can be made. Again there are great potentials, but how to realise them and how to use them. Somewhere we seem to be at a loss when we try to define the status of GIS now and its future development. So once I had accepted the invitation for this presentation, second thoughts came up. How to review the status and problems of such an unstructured field in three quarters of an hour. How to describe the status , without spending all the time on the description of existing systems and applications. The instruction which

Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989

I got after accepting this task was not to paint such a beautifull scenery giving the illusion that we are close to entering paradise. A more realistic sketch was required. This does not mean that we should turn the positive into the negative, we just should stop walking with our heads up in the clouds and face the problems which we will have to deal with. These considerations confronted me with the first problem in this field: "how to describe the status?" To do that a structural approach is required, for such a structural approach a theoretical framework is required and for such a theory some basic concepts have to be formulated. The rest of this lecture will be dedicated to these problems and related topics.

2. GEOGRAPHIC INFORMATION

Many expressions are found in literature as equivalents of "geographic information" or "geoinformation". Sometimes they are treated as full equivalents, sometimes as having a different meaning. In this lecture "geographical-" or "geoinformation" will be used in a general sense. In this section attention will be paid to some structural and semantic aspects of geoinformation. I will not try to give a short concise definition and I will try to avoid to refer to specific applications otherwise then in examples.

Geoinformation is used to describe objects, phenomena or processes at the earth's surface. To date this is mainly done in the form of a state description at a certain moment or for processes as a series of state descriptions. These may refer to physical aspects of the terrain, or administrative aspects or landuse etcetera. These aspects are given as thematic attributes in relation to geometric data. In the most primitive form the thematic attributes are directly linked up with positional data as in raster structured GI systems. The positional data serve as a vehicle to link different types of thematic data or to link data obtained at different moments. Figure 1 gives a scematic presentation of this datastructure

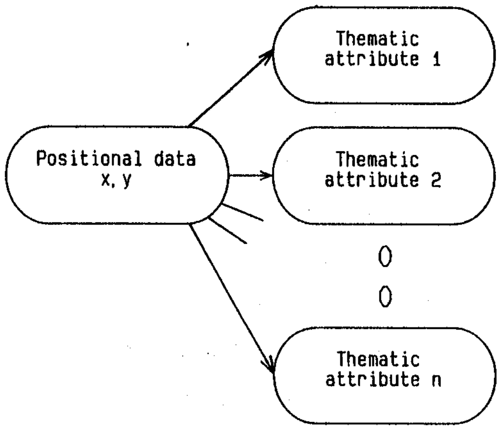


figure 1: datastructure for position bound thematic data

The analysis of a terrain situation can be done through the processing of the stored thematic data. This can be done in two ways: the processing can be based on a functional model which describes the relationship among the attributes per location:

$$f(x,y) = f(attr.1_{xy}, attr.2_{xy}, \ldots\ldots, atrr.4_{xy})$$

or it is based on a functional model which relates attributes at different locations, in which case the coordinaters determine a weighting function for the attributes.

$$F(u,v) = f(\ldots., gi(u-x,v-y)\ attr.i_{xy}, \ldots.)$$

The datastructure of fig.1 gives a low information level in the sense that spatial patterns of the attributevalues can only be found through data processing.

A datastructure with a higher in formation level can be obtained by the introduction of terreinobjects or terrainfeature data. In that case the thematic attributes are not linked directly to the positional data, but to the terrainfeature data. This implies that the terrainfeatures are thematically characterised by a set of thematic attribute values, their geometric description is given by the geometric data. This datastructure is scematically represented in fig.2, where the terrainfeatures are represented by feature identifiers.

The fact that terrainfeatures have a geometric- and a thematic description implies that there are two basic ways to define sets of features. The sets defined through the geometric characteristics will be called "feature types". The sets defined through the thematic characteristics will be called "feature classes".

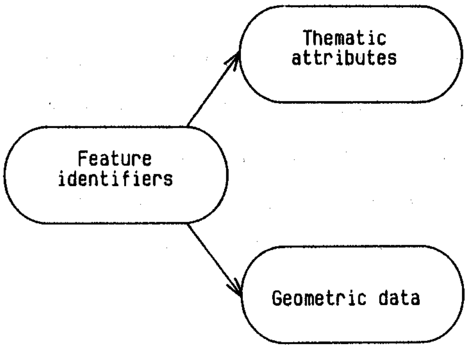


figure 2: a feature oriented datastructure

2. Feature types

Three feature types can be distinguished in case of a two dimensional terrain description. These are the areafeatures, the linefeatures and the pointfeatures. Areafeatures are extended in two dimensions. Linefeatures are extended in one dimension, their geometric description refers to their linear structure. Pointfeatures are represented by just one point with a coordinate pair.

The decission to which type a certain terrainfeature belongs determines how it is stored in the data base. The decission should depend on the context in which the terrain description or mapping is made.

A town may be treated as an areafeature in one context and as an pointfeature in another. Similarly a road may be linefeature in one context, but an areafeature in another.

2.2 Feature classes and superclasses

Feature classes are defined through their thematic aspects. Each class has a class name or a class label and a list of attributes which give the thematic characteristics of the terrainfeatures belonging to the class. As an example we can define different classes of linefeatures, such as: roads, railroads and rivers. The road may have the attributes: road type, pavement, traffic density, last maintenance operation, responsible authority. The railroads may have the attributes: number of tracks, maximal velocity, traffic density etc. The rivers may have the attributes: width, depth, current, max shipsize etc.

For each feature belonging to one of these classes the relevant attributes have to be evaluated, hence each feature is represented by an identifier and a list of attribute values. The relationship between classes and features and attributes is represented in fig.3.

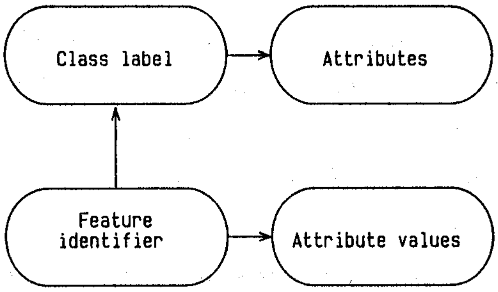


figure 3: feature-class-attribute relationships

The figure can be extended by the definition of superclasses which are sets of classes. In a landuse classification system the classes "pasture land" and "arable land" can be combined into the superclass of "farmland". Similarly "moorland", "forests" and "marshland" may form a superclass of "wastelands". It may be usefull to characterise each superclass by means of a list of superclass attributes (s.c. attributes) which have to be evaluated for each class. This datastructure is given in fig.4.

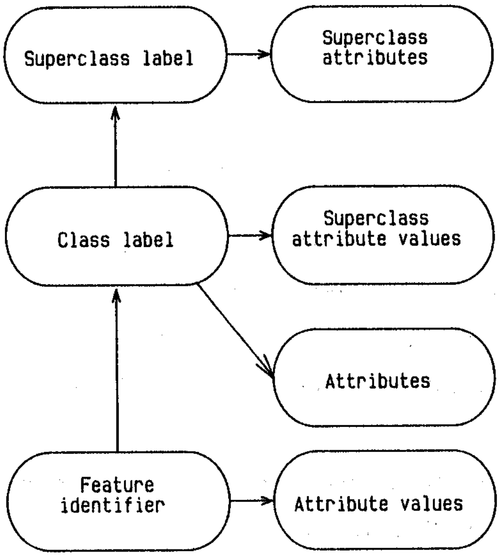


figure 4: datastructure for classes and superclasses

This procedure can be repeated in a upwards direction adding higher levels of superclasses. In this way a classification hierarchy is created where at each level the classes are specified by a valuelist for the attributes of the classes at the next higher level and an attributelist which should be evaluated at the next lower level. At the lowest level are the objects or terrainfeatures which are specified by their feature identifiers and their attribute valuelist, no new attributes are introduced here.

2.3 The connection of thematic data and geometric data

It is at this lowest level, through the terrainfeatures, that the link between thematic and geometric data is made. This new link is represented in fig.5.
As stated earlier, this link is one of the characteristic aspects of geoinformation. The geometric data can be structured in two different ways.

2.3.1 Rasterstructured data

Firstly a structure similar to figure 1 can be realised. There the thematic attributes were directly linked to the positional data. Now there will be an intermediate link through the feature identifiers (see fig.6).

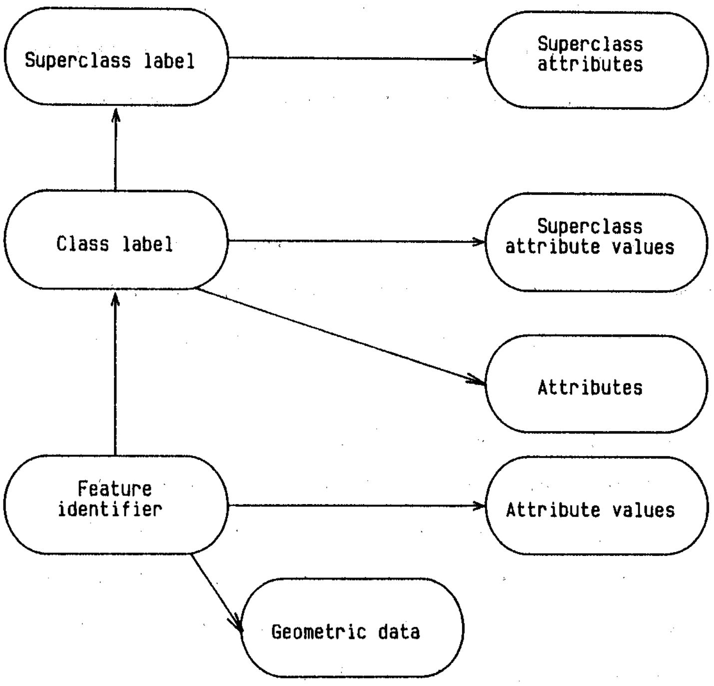


figure 5: linking thematic data and geometric data

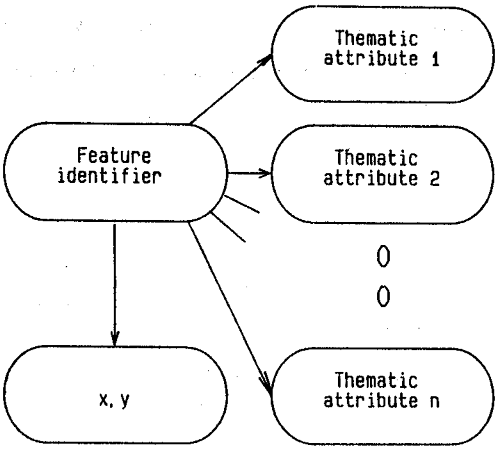


figure 6: linking thematic data to positional data

This is in fact the case when terrainfeatures are mapped in a rasterstructure. The link between rasterelements and features can be made in two ways. In a rasterelement oriented approach each rasterelement is labeled with the identifier of the terrainfeature it represents. The geometry of a specific feature can only be found by inspection of the labels of all the rasterelements and selection of those which have the required label value.

In a terrainfeature oriented approach the feature-position link is a pointer from the feature to all the relevant rasterelements, so that the geometry of the feature can be found directly. Quadtrees are an example of such a datastructure [5,13].

The spatial structure of a rastermap must be found through the rastertopology, which is build up through the connectivity of neighbouring rasterelements (see fig.7).

A	A	A	A	A	B	B	B
A	A	A	A	A	B	B	B
A	A	A	A	A	B	B	B
B	B	B	B	B	C	C	C
B	B	B	B	B	B	B	C
B	B	B	B	B	B	B	B

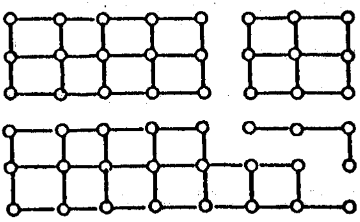


figure 7: rastermap rastertopology

Neighbouring rasterelements are connected if they have the same labelvalue. Terrainfeatures are represented by connected rasterelements. Spatial relationships among terrainfeatures should be found by inspection of neighbouring rasterelements which are not connected. Such a spatial anlysis is rather cumbersome.

On the other hand it is rather simple to make overlays in a rasterstructure of maps with e.g. different thematic contents if the maps are based on the same grid. In that case one only has to inspect the labels of the rasterelements in the different maps. This gives the datastructure of fig.8.

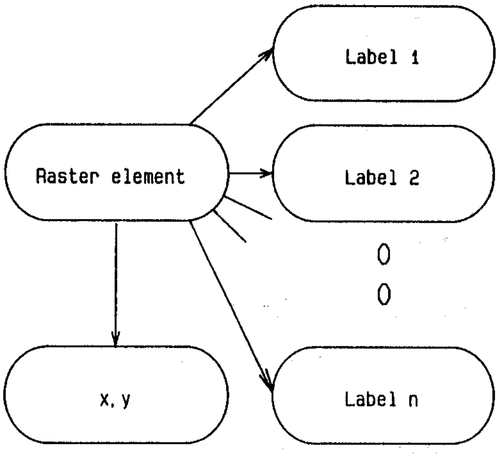


figure 8: datastructure for rasteroverlay

2.3.2 Vectorstructured data

The alternative is the vectorstructured terraindescription, which we will call a vectormap. This represents the linear characteristics of the terrainfeatures, that is in the linear struture of linefeatures, the boundaries of areafeatures and the position of pointfeatures. In this case too, there are several posibilities for linking the geometric data to the feature identifiers. Fig.9 b, c and d give three posibilities for the geometric description of the situation in fig.9.

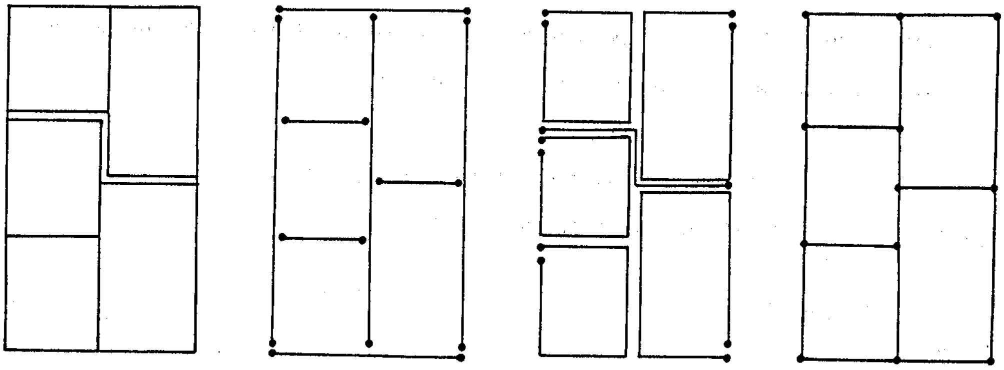


figure 9: a. original situation b. unconnected line-elements c. polygons per object d. graphstructure

Figure 9b gives a correct representation of the geometry of figure 9a, but it is not possible to define a unique link between the geometric elements and the identified terrainfeatures. In figure 9c such a direct link is possible, but it will be cumbersome to analyse the spatial relationships among the terrainfeatures. The geometric elements in figure 9d have been chosen so that well defined links between feature identifiers and geometry can be made, which also allow the analysis of the spatial relationships among the features. The consequences of this choice have been explored carefully in earlier publications [1,7,8,11,12]. From [7,8] it becomes apparent that a powerfull datastructure can be develloped with a limited number of elementary datatypes, with only nodes and edges or arcs as geometric elements. The resulting datastructure is given in fig. 10.

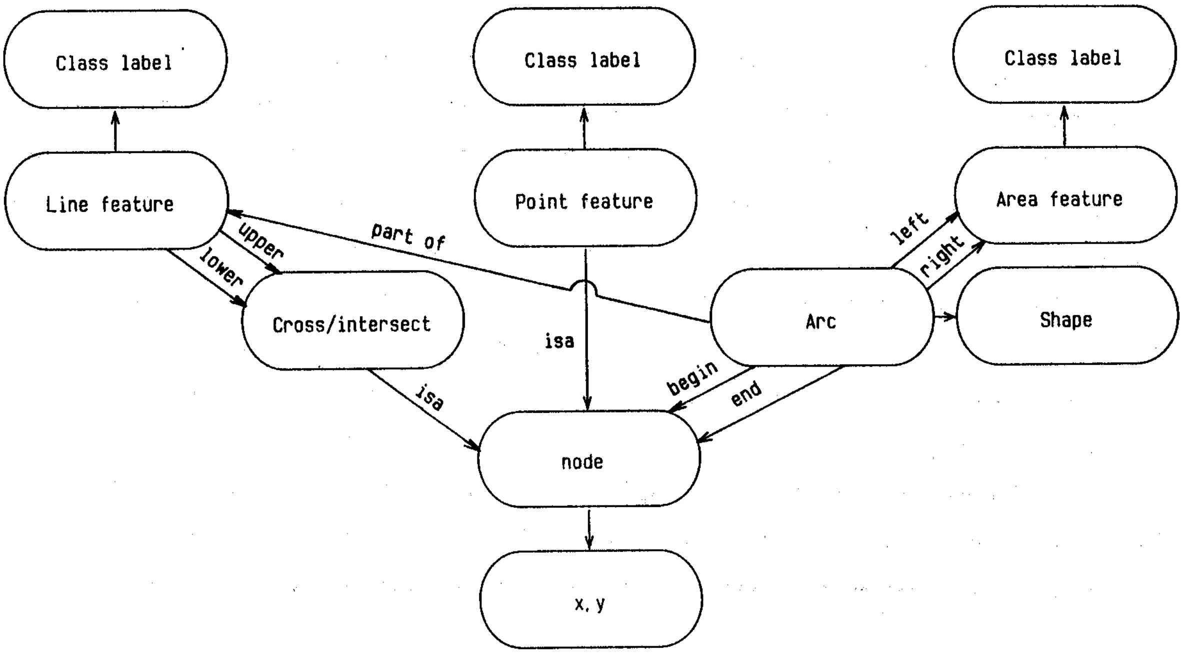


figure 10: formal datastructure for vectormaps

The diagram of fig. 10 is called a formal datastructure (f.d.s.) because no reference has been made to an actual databasestructure. The f.d.s. can serve as a mathematical model on for the design of a database according to the relational, hierarchinal or network model.

Analysis of this diagram shows that many topological relationships among the terrainfeatures can be derived from this f.d.s. These relationships are given in fig.11.

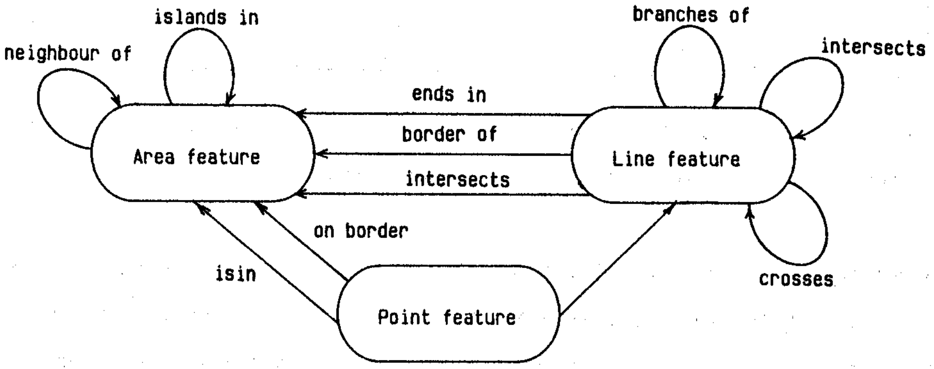


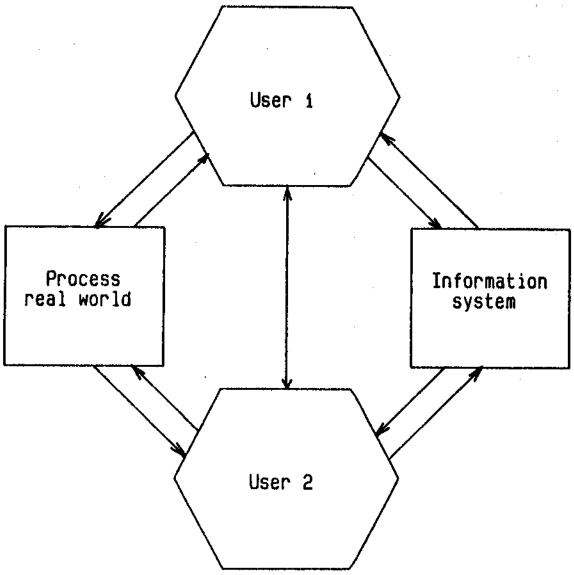
figure 11: topological feature relationships

Hence a vectormap facilitates the analysis of spatial relationships among terrainfeatures [7,8]. The analysis of overlays of vectormaps is rather cumbersome though. Therefore vectormaps and rastermaps appear to be complementary.

3. INFORMATION SYSTEMS

We spend quite some attention to the structural aspects of geoinformation, and not so much to the semantic aspects. That is because the structural aspects are most important for the definition and design of information systems. To date a large number of geoinformation systems and software packages exists. Many of them have been designed in a specific environment for particular applications. Others have been develloped for more general purposes. These latter systems often went through an evolution process in which each stage gave new extensions to meet the users requirements based on the experiences of the earlier stages. Due to this situation it is very difficult to compare the different G.I.Systems on general criteria. Another complicating factor is the fact that there is no unanimity in the use of the expression "information system". That is why a conceptual definition will be given here of "information systems" in general, followed by a more restricted definition of "geoinformation systems". For this definition I will follow Wintraecken [15], which till now is only available in the dutch language. (The publication in english is expected by the end of 1989 [16].)

People are active in the real world, where they monitor and controll processes of different kinds. To do so they should have knowledge of the status of such a process and its expected develloppment. Part of that knowledge is transferable to other people, who are also involved in the management of the same process (or may be of a closely related interacting process). This transferable knowledge will be called "information". The information transfer or exchange will be called "communication". The communication among persons may be directly or intermediate through an information system (fig.12).



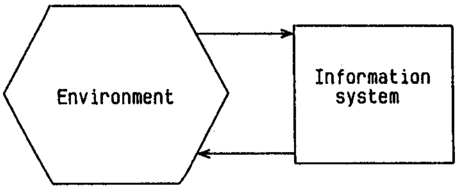


figure 12: communication network with informationsystem

figure 13: simplified communication

In the latter case information has to be transfered to the information system, which should be able to store it, and to give answers to requests for information from the users. Fig.12 shows the interaction between users and the information system. The term "users" should be interpretated in a wide sense, i.e. not only persons communicate with the system, but also machines or other information systems. The interacting outer world will be called "environment", which makes a simplification of fig.12 possible ([15] ch. 2.1)

The information in the system refers to the state of the process and its state transitions. The fact that the process state changes in time implies that the information content of the system must be updated continuously. Consequently the system should be able to store information in an information base. Secondly the system should be able to communicate with the environment to receive information, to receive requests for information and to give back information and to give answers to the requests for information. Hence it should be able to receive messages from the environment and to send messages back to it.

Therefore the second component of the system should be an information processor which serves as an interface between the environment and the information base. Additionally the processor should be able to derive new information from the original information.

An information system functions according to formal rules, which means that for a given state of the information base the system always gives the same output in reaction to the same input. In other words the behaviour of the system is predictable and repeatable ([15] ch. 1.2). To control the behaviour and the functioning of the system it has a third component, the grammar, which describes the messages it may exchange with the environment. These messages should only refer to the content of the information base, which can only be changed through these messages. In this way the grammar gives rules for the allowable states of information base and its state transitions ([15] ch. 2.3). With these three components we can fill in the picture of the informationsystem of fig.13.

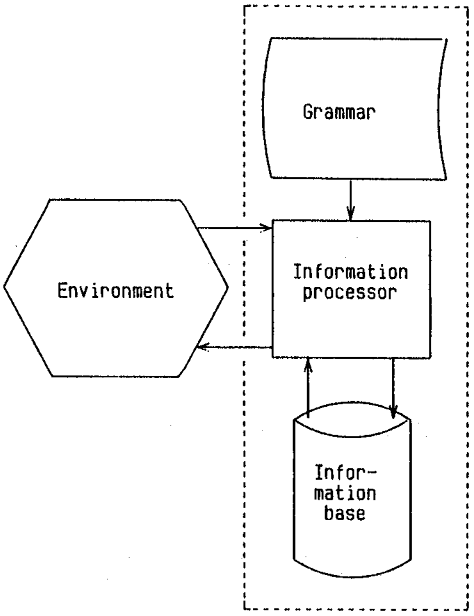


figure 14: internal and external communication links of an information system

The environment communicates with the system through the information processor, which checks wether this communication is according to the rules given by the grammar. If so, then the processor communicates with the information base according to the instructions of the environment ([15] ch. 2.4).

With fig.14 we have a general description of an information system. This is a conceptual description because no reference has been made to any technical realisation. The terminology used here follows [15], where the following expressions are used: "Knowledge" is all somebody knows about a certain topic or area. "Information" is the transferable knowledge. "Communication" is the exchange or transfer of information. "Data" is the representation of the information in a communication medium (information system). "Grammar" is a description of the conventions concerning the data which may be used to transfer information and concerning their meaning ([15] ch. 1.1).
In this sense a database management system is a component of an information system; it deals with the storage of data, data retrieval and the updating of the database.

After this general discussion we can define Geographical Information System as Information Systems handling geographical information.

4. TOWARDS A GEOINFORMATION THEORY

The conceptual description of geoinformation and information systems in the previous chapters should facilitate an abstract discussion of geoinformation systems without direct reference to specific applications or a technical realisation. These concepts should serve as a stepping stone for the formulation of a geoinformation theory. In this chapter we will first spend some attention to the role such a theory could play, after that we will sketch some of its characteristics.

4.1 The role of a geoinformation theory

The theory could serve several aims. It could help us to structure the whole field of geoinformation systems in the sense that the common aspects and the differences of existing GISystems can be made more clear. In a similar way the common aspects and differences of the applications of GIS can be made more transparant. The theory will probably also give criteria for the user to structure the information processing for his particular applications and to choose a suitable geoinformation system or to build one.

The theory will also have an effect on the training of students. It will help to set up courses with well structured curicula. Nowadays the curicula often give the impression that they have been put together through a series of adhoc decissions. Consequently they show little structure. They often look like variants of traditional surveying courses. We should realise that most people working with GIS nowadays had their training in other disciplines. So, however enthousiastic we are developing the potentials of GIS, we are still a generation of amateurs. We should take care that new generations enter this field with a more professional attitude. Therefore they need a training on a sound theoretical base.

A third reason to develop a geoinformation theory is that it will help us to extend and generalise the functionality of the present systems. The better we understand the basic principles of geoinformation and of the systems handling this information, the better we will be able to develop more powerfull systems to assist us with the analysis of spatial phenomena and processes and with activities in the field of spatial monitoring, planning and design.

4.2 Queryspaces

If the theory is based on the concepts of the chapters 2 and 3, it will give a formalistic framework for the definition of grammars for geoinformation systems. According to chapter 3 a grammar primarily defines the communication between an informationsystem and its environment. This communication consists of two sets of messages from the environment to the system:

a. questions about the information stored in the system
b. instructions to change the content of the information base

and the response of the system to these messages.

The question under a can ask about the occurence of a specific data item or a set of data items such as: > is there a university in town

or

> give all the forest area's in Germany

But also relationships among data items can be investigated: e.g. in a raster overlay one could ask:

> give all the rasterelements where the vegetation is forest, the altitude is > 500m and the population is < 5 persons/km^2

Or in a vector structured database one could ask:

> through which countries runs river Rhine and which are the capitals of those countries

Whether a system will be able to answer such questions depends on the data items and on the relationships among them stored in the database. Additionaly it depends on the ability of the system to derive new information from the stored information. This problem is directly related to the formal datastructure which is supported by the system. In fact the grammar gives implicitely the formal datastructure and with that it defines the structure of the query space. This is the set of all queries (questions under a) which can be handled (see [7], [8] and [9]). Hence we see that the queryspace of a system depends on the actual data in the database and the formal datastructure. These two facts determine together which statements can be formulated by the system in answer to questions from its environment.

4.3 Context and contexttransformations

A complicating factor for the geoinformationsystems is that, although the general structure has been given in chapter 2, the actual data definition is not always sharp. The definition of certain classes of terrainfeatures may only have sense within a certain context. A context can be given by the mapping discipline, the aim of the mapping, the aggregation level (related to mapscale), the time of the mapping etc. E.g. the definition of a soil unit will depend on whether the mapping is for landuse planning, erosionstudies or another aim, it will also depend on the aggregation level of the mapping. But even within a well defined context the definition of terrainfeatures and also their attributes will often have a limited accuracy. E.g. how sharply can the border of an urban area be defined, how precisely can a riverbed be located, or how sharply can different soilclasses or vegetationclasses be distinguished. These observations lead to two conclusions:

- Some data and consequently the statements derived from them are only valid within a certain context.
- Statements may have a limited accuracy even within a specified context.

A geoinformation theory should help to define contexts and it should give rules to handle and evaluate data inaccuracy.

Hence the theory should give general rules how a grammar should specify the context of the information handled by a system. Several factors have been mentioned which define the context of the information externaly. Internaly, that is in the system, this will effect the actual attribute values which are used. In case of a feature oriented approach it will also effect the feature definition and the feature classification system and the geometric description of the features.

Further more the theory should give rules and methods to transform data from one context into data of another context. This will be partly equivalent to the transformation of a system with a particular grammar into a system with another grammar.
The theory should specify under which conditions such transformations are possible and how they should be performed. One family of context transformations is the change of aggregation level, e.g. the transformation of a raster structured landuse mapping with a cell size of 1 km^2 to a cell size of 25 km^2, or the transformation (conceptual generalisation) of a 1:25.000 topographic mapping to a 1:100.000 scale. An other family of context transformations is the change from one feature classification system to another through the feature attribute values.

4.4 Single valued maps

In chapter 2 rastermaps and vectormaps have been presented as two alternative and complementary terrain descriptions. The expression "map" was used in a wider sense, it is not restricted to mapsheets or other graphical representations, but it refers to a terraindescription with a geometric component, independent whether it is an anologue graphical map or a database in some form.

In earlier publications the concept of single valued vectormaps, s.v.v.m., was introduced [7] [8]. The concept has been developed in anology to single valued rastermaps (s.v.r.m.). The latter concept is easy to understand. A single valued rastermap has only one attribute or label, which is evaluated for each rasterelement of the map. Hence the map gives only one thematic attribute such as height, or soil type. Simple queries can be formulated for such a map, whereas multivalued rastermaps are generated by overlaying several s.v.r. maps. If a s.v.r.m. gives a feature oriented terrain description then a well defined feature classification system is required, which should be unique and complete. It should be unique in the sense that the classes are multually exclusive, hence each rasterelement belongs to one feature with one class label. It should be complete in the sense that all rasterelements belong to some feature in this classification system.

For single valued vectormaps a similar definition can be formulated. The most complete definition has been given in [8]. It refers to the formal datastructure of fig.10 and it states a.o. that for each geometric element of a map (nodes and arcs) there is at most one occurence of each of its link types to the feature identifiers, whereas the left and the right links occur exactly once for each arc. S.v.v. maps too require an unique and complete feature classification system, unique in the sense that each node represents at most one point feature and each arc can be part of at most one line feature, whereas it has only one area feature to its left and one area feature to its right hand side. All features belong to just one feature class. It is complete in the sense that all area segments of the map belong to some area feature. Note that not necessarily all nodes represent point features and not all arcs should belong to a line feature.

Again a multivalued vectormap should be constructed through the overlay of several s.v.v. maps. Further research showed that a s.v.v.m. has a well defined query space [7,8] and it allows the formulation of consistency checks for maps updating operations [4]. It seems that the concept of single valued maps is a very powerful tool for map definition and the design of databases for geoinformation systems.

4.5 Some other aspects of the geoinformation theory

Some important aspects have been mentioned of a geoinformation theory which still have to be developed. We have certainly not been complete. Data accuracy has been has been mentioned but it has not been elaborated. In section 4.2 two types of messages have been mentioned, we only elaborate on the queries under a, but not on the messages under b, for updating the information base. Other topics to be dealt with by the theory are the integration of raster and vectordata and the generalisation to three dimensional terraindescriptions (not 2.5 dimensions).

The fact that the theory should serve as a framework for the definition of grammars for geoinformation systems, means that it indirectly defines the exchange of messages between these systems and their environments. This implies that it is basically a linguistic theory, in the sense that its object is the definition of data items, datastructures, data classification systems and contexts. In fact the theory tells the users which statements a geoinformation system can formulate about the terrain and how it formulates them. It also tells the users how to communicate with such a system. Therefor the theory should also absorb knowledge from the application fields of GIS. But for its care it should most likely lean heavily on modern theoretical developments in computer sciences, such as database theory [3] [14], the theory of computation [17] and discrete mathematics [6] and artificial intelligence [2] [10].

5 FINAL REMARKS

You may have noticed that, although I sketched many problems which we are facing in this field, I have failed until now to describe the status of geographic information systems. The reason for that failure was given in the introduction: the lack of a geoinformation theory. I have the feeling that my position is comparable with that of a photogrammatrist fourty or even thirty years ago, who was asked to give a status report on photogrammetry. In that time there were many applications of that rather new technique. Maps were produced, many different types of equipment were available, several components of the photogrammetric process had just been developed or were being developed.At that time the field was too fragmentaric to define its staus. It was hardly possible to compare the different photogrammetric methods because a central theory was lacking. It was only at the end of the sixties and in the seventies that aerotriangulation and other components of the photogrammetric process, were properly understood and theoretically described in the context of a well formulated mathematical model. It was in the eighties that this model was completed with a theoretical model for quality control. At that time, in the seventies and the beginning of the eighties, with the help of these theoretical tools. it was possible to give status reports on the performance of photogrammetry.

To my opinion the status of GIS is very much comparable with photogrammetry in the fifties and the early sixties. Ofcourse we use more modern tools and techniques and our problems are different. But strange enough it is due to these modern tools and techniques that we are in a similar situation. At that time photogrammetry used equipment which was newly developed and which worked well. But because a proper mathematical description was lacking the operation of that equipment and therefor the photogrammetric process and its potentials were not properly understood. Only after the appropriate mathematical tools were available, a complete theory could be developed. Consequently many of the instrumental and manual operations could be replaced by computional procedures, making photogrammetry a more powerful tool.

With GIS we have a similar situation in the sense that we are also using newly developed equipment with which we can process data according to rules which are not yet properly understood. Much of the data analysis is done in visual interaction with the system and in that interaction we make decissions and perform operations which are not properly modeled yet. In the previous chapters I have tried to make clear that we should find the proper tools to formulate an appropiate theory for this field. At the end of chapter 4 I gave an indication where some of those tools might be found. For those members of the audience who, like me, had their training in photogrammetry, landsurveying and geodesy, this means hard work to become acquainted with these new fields. I personaly consider it as an fascinating privilege that in the first part of my career I could be a witness of the completion of the theoretical model for photogrammetry and that I can now participate in the development of a new theory for geographic information systems.

We will have to learn another way of thinking, quite different from what we learned at university. Consequently we have to reconsider seriously how we want to train the new generations. Do we still want to give them the traditional professional education with its fundamental training in physics, mechanics, calculus, linear algebra and statistics. Or should we rather prepare them to find their way in these newly developing fields which require a fundamental training in the modern branches of computersciences and discrete mathematics. I can not give you a definite answer to this question, but I ask you to consider it carefully and I am tempted to choose for the new option. To my opinion this choice is one of the most urgent problems in the field of geoinformation systems.

REFERENCES

1. Broome, F.R. : Mapping from a topologically encoded database 1
Auto Carlo London, 1986
2. Charniack, E. and D. McDermott: Introduction to Artificial Intelligence
Addison-Wesley, Reading, (Mass.), 1985
3. Date, C.J. : An introduction to Database Systems, vol I
Addison-Wesley, Reading (Mass.), 1986
4. Kufoniyi, G. : Editing of a topologically structured data
M.Sc.Thesis ITC, Enschede, 1989
5. Yang, H. and D. Fritsch : A Raster Oriented Approach for Spatial Data in Geographic Information Systems
Geo Informationssysteme, Anwendungen, Neue Entwicklungen
Wichmann, Karlsruhe, 1989
6. Lio, C.L. : Elements of discrete mathematics
Mc Graw-Hill, New York, 1983
7. Molenaar, M. : Knowledge representation in vectorstructured databases
Geo-Informationssysteme, Anwendungen, Neue Entwicklungen
Wichmann, Karlsruhe, 1989
8. Molenaar, M. : Single valued vectormaps-a concept in GIS
Geo-Informationssysteme, vol.2 no.1. 1989
9. Molenaar, M. : Towards a geoinformation theory
ITC Journal 1989-I
10. Rich, E. : Artificial Intelligence
McGraw Hill, New York, 1983
11. Roessel, J.W. van : A relational approach to vector datastructure conversion
Int. Symp. Spatial Data Handling, Zurich, 1989
12. Roessel, J.W. van : Design of a spatial datastructure using the relational normal form
Second Int. Symp. Spatial Data Handling, Seattle, 1986
13. Samet, H. : Region Representation: Quadtrees from Boundery Cades
Communication of the ACM vol.23 no.3, 1988
14. Ullmann, J.D. : Principles of Database Systems
Computer Science Press Inc., Rockvill, Maryland, 1982
15. Wintraecken, J.J.V.R. : Informatie-analyse volgens MIAM Controldata Academic Service, Den Haag, 1985
16. Wintraecken, J.J.V.R. : NIAM Information Analysis Method
Kluwer Acad. Publ. Group, Dordrecht (1989/90 ?)
17. Woody, D. : Theory of computation
John Wiley and Sons, New York, 1987

ABSTRACT

The fact that the field of GIS is very heterogeneous and fragmentric makes it difficult to give a status report on its development. Therefor it is desirable to structure this field by means of geoinformation theory. The formulation of such a theory should start with the formulation of some basic concepts for geoinformation and for informationsystems. The major part of this article has been dedicated to these problems. It treats some important structural aspects of geoinformation, such as the relationships between terrainfeatures, their thematic attributes and their geometry. A general structure for feature classification systems has been given and different solutions for linking thematic data to geometric data. Furthermore a discussion of the general characteristics of informationsystems leads to the identification of three components: the information base, the information processor and the grammar.

With these basic concepts an outline is sketched of an information theory for geodata. In this sketch problems are identified which have to be solved in the near future, such as the definition of the context of geoinformation, the definition of context transformations, the evaluation and handling of data inaccuracy and the definition of curicula for courses to train new generations of GIS professionals.

Prof.dr.ir. M. Molenaar

Wageningen Agricultural University

Department of Surveying Photogrammetry and Remote Sensing

P.O.Box 339

6700 AH Wageningen

The Netherlands