FORMAL METHODS THE FOR ACCURATE DEFINITION OF SOME FUNDAMENTAL TERMS IN PHYSICAL GEOGRAPHY

Andrew U. Frank
Dept of Civil Engineering
University of Maine at Orono
103 Boardman Hall
Orono, ME 04469
(207)581-2174
FRANK@MECAN1.bitnet

Bruce Palmer
Digital Equipment Corporation, Marlboro
DVINCI::Palmer.decenet

Vincent B. Robinson
Hunter College
City University of New York
XXGHC@CUNYVM.bitnet

ABSTRACT

A previous paper presented a method for the extraction of geomorphological features using symbolic processing. The programs used therein can also be interpreted as formal definitions for these features.

Geographers have tried in the past to find suitable definitions for features such as ridges, valleys, hills, etc., but no complete agreement has been reached. A problem encountered with definitions written using a natural language (e.g. English) is the inherent ambiguity of the language used. Different individuals may understand such definitions differently. Moreover, the definitions proposed are often incomplete or contradictory, so that counter examples could easily be found.

In order to avoid these problems, we suggest that definitions be formulated in a formal system, such as first-order predicate calculus. The surface of the earth can be modelled with irregular triangles, and such a model can be expressed as a theory in first-order logic. Definitions for features of interest can then be studied in a formal setting.

Given formal definitions for features, it should be possible to mathematically prove that the features have specific properties, e.g. that streams never cross ridge lines. Our analysis (re-)discovered the scale-dependency of present definitions of stream and watershed lines, which seem to defy objective definitions. We thus concentrate on à number of similar geomorphologic features, which can be rigorously defined.

We conclude with the recommendation that research to formally analyse the problem of scale-dependency should be carried out. Hierarchical (and similar) nesting of features is important for most spatial analysis operations, but the understanding of these processes is presently mostly informal.

1 INTRODUCTION

A number of interesting studies have been undertaken in physical geography and geomorphology in the past. Various researchers attempted to capture some of the distinctions between different types of landscapes, so evident to human observers but at the same time so difficult to express in formal terms.

The goal of such studies has been to identify the landscape types (e.g. the major processes that formed the present situation) not by a detailed study of the area, but from easily observable macroscopic features. This method would be of great interest in cartography, as it would aid in the automatic generalization of topography [PFALTZ 1976],[WOLF 1986], planning etc.

It has further been proposed by [PEUKER 1973] and others that the network of watersheds and stream lines would be a very good, compact description of a landscape useful for other processing tasks.

An attempt was made to separate the seemingly random element in landscapes from the systematic components attributable to geomorphologic processes ([SHREVE 1966], [SHREVE 1969] cited in [MARK 1977]). This requires a method of counting features in the landscape which might be used to characterize it, followed by the use of statistical comparisons to identify types of landscapes [MARK 1977] [MARK 1982]. However, the identification of the features was based on subjective classification, and results were usually not comparable. Two problems were discovered in these attempts.

First, the features which were to be measured needed to be clearly defined. Only features with clear and objective definition can be reliably measured or counted, and such measurements compared to the findings of other researchers. If features are not well defined, it cannot be ascertained whether differences in results are caused by actual differences in the features observed (i.e. in the landscape) or are the consequence of varying methods used to define the objects measured.

Second, it was observed that the methods did yield different results depending on the type of raw data used to define the terrain. For example, the representaion of stream networks on traditional topographic maps are the result of subjective cartographic processes, and are not always proper descriptions based on geomorphological definitions. Results computed from maps of different scales showed differences which are clearly unrelated to the landscape, but rather measure the cartographic generalization process. Similar variations were found in digital terrain models, depending on the density of the grid and the process used in data collection [O'CALLAHAN 1984].

This paper advocates the use of symbolic processing for geographic problems. Numerical analysis, as an important tool, should be supplemented with methods based on symbolic processing. Both the theoretical basis and computer support neede in such an approach are available today. In [PALMER 1984] an example of practical use was given. In this paper we will discuss the fundamental properties of this method and explore approaches to their solution. We will demonstrate methods for the objective definition of terms which can be used to extract features from a formal model. Unfortunately, the scale-dependency of definitions for some geomorphologic features, which currently precludes an objective definition, and we will point out a number of similar feature definitions which avoid some of the problems.

2 MODELS OF REALITY

Generally, the geographer's work relies on models of the landscapes to be studied. Such models may be often informal, i.e. expressed as narrative in a natural language or as a cartographic drawing, and thus subject to interpretation by human beings, or they may be formal, expressed in some formal language and therefore usable for further formal processing, such as statistical operations. Today's geography has long since abandoned the realm of collections of landscape descriptions as found in the diaries of the first explorers, but modern quantitative methods in geography are only possible if formal models are available.

Formal models of reality may be constructed in many different ways, representing many different properties of reality. A formal model consists of two parts: a theory consisting of a collection of expressions in a formal language (which may be extremely simple) and an agreed-upon interpretation of formal expressions which link the formal symbols to reality.

For example, a list of point coordinates (x, y and height) together with rules for their processing, can be interpreted as a formal model for the surface of the earth. A model is formal if it is expressed in symbols which do not depend on human interpretation. Examples range from digital surface elevation models to drainage network descriptions in terms of graphs, not to mention the arbitrary complex knowledge representations used for expert systems. The information contained in a formal model is accessible to objective treatment according to formal rules, which yields a deduced formal model. Electronic data processing in general deals with formal models (the input data) and their transformation using formal rules (the programs), but formal models are also more generally useful as they permit information transformation in terms of objective data processing, and avoid the pitfalls of subjective interpretation.

The best known language for the expression of a formal model is first order logic. First order language is both a simple and a powerful tool to express all sorts of different facts and rules, and its deduction methods are well understood [GALLAIRE 1984]. A formal model is considered a theory, and the deduction of information is performed as an attempt to prove a proposition in this theory.

Geography is generally dependent on other disciplines for the original collection of data describing reality. Formal models (data) available from organizations that collect facts of geomorphologic interest are often dependent in their content and classifications on the primary objectives of the collecting organization and the methods it uses (for some details regarding digital elevation models see [O'CALLAGHAN 1984]). In order to reduce these influences, it is preferable to select models describing facts of reality which are less subject to interpretation than others. The surface of the earth is determinable with less personal judgement than is the classification of a watercourse. Unfortunately, models of terrain elevation are quite accurate in the determination of the surface, but record this elevation only for selected points, leaving areas in between undetermined.

Such methods necessitate the use of interpolation to fill the gaps, but this interpolation does not really add information. The interpolated surfaces may, depending on interpolation method and actual surface form, be quite different from reality. This problem requires our particular attention.

3 DEFINITIONS

Verbal definitions of terms can be written in a natural language using other terms which describe physical reality. This is customary in geography as in other sciences, but has the substantial drawback that such definitions either use terms which are not exactly defined but are assumed to be generally understood, or end up in circular definitions. As noted beofre, definitions written in natural language (e.g. English) depend on human beings for interpretation, and are thus often not capable of capturing fine nuances in meaning. In physical geography most terms are defined in this way and seem to be very clear unless they must be applied to a real (rather than a textbook) example.

Operational definitions, i.e. definitions which describe a procedure for the identification of a feature, are sometimes more formal, as they rely on the use of a formal model of reality. For example, in [KRUMMBEIN 1970] the definition of drainage basin the use of sufficiently detailed contour lines is emphasized.

Using an arbitrary map, it is sometimes even difficult for a specialist to find and delineate watersheds and like features using the operational definitions. Problems arise, however, if a statistic of all saddles for a given area or all valleys is to be collected: when is a small bump in a ridge a saddle?

Additionally, discrepancies can be found in definitions for basic terms such as 'ridge line' and 'drainage divide'. For example, in [WERNER 1978] ridges are a subset of the drainage divides, whereas they are not in [BAND 1986].

3.1 Formal Definitions

Formal definitions avoid the problem of dependency on generally accepted terms or circular definitions. They are based on formal theories, (e.g. Euclidian geometry) and consist of a small number of initial terms on which all other terms build. New terms are defined as formal expressions in the theory. They are only meaningful with respect to the interpretation of the theory which links it to features of reality.

Such definitions read differently from traditional natural language definitions. They are not dependent on interpretation and subjective judgement for application - instead they define features only in terms of the formal system previously established. The interpretation of the formal system as a model for reality is, of course, subjective.

Formal definitions are not a new invention. All previous researchers into the use of computerized methods for the extraction of geomorphologic features from a model (e.g. digital terrain models) have used formal definitions. However, most often, they did not consider their formulations as definitions, nor did they discuss this aspect of their work in particular. In [O'CALLAGHAN 1984] but also in [BAND 1986], [TORIWAKI 1978] and [HARALICK 1985] geomorphologic features are extracted from a regular grid digital surface elevation model. The authors give more or less detailed descriptions of the programs used to detect such features as ridges. These programs are expressed in a formal language (most often FORTRAN) and operate on a formal system (namely the input data). Therefore the programs are formal definitions (in terms of the model represented by the input data) for the features they extract. The procedural languages used were not designed for definitions and are therefore not specifically suited to the task: they describe how to find a feature, not what it is. Further, most publications provide only natural language renderings of the procedures used, and do not publish the original program; this is understandable in the light of the major thrust of these papers. However, it is difficult to judge the implications of the definitions used in [TORIWAKI 1978], for example, in terms of geomorphology, and often the methods rely on the appropriate, subjective selection of some values.

In [PALMER 1984] a similar method for the extraction of geomorphological features from a terrain model was given, based on an irregularly triangulated surface model. It was defined and programmed using the PROLOG [CLOCKSIN 1981] language. PROLOG is firmly based on mathematical logic and (pure) PROLOG can be considered an executable implementation of a subset of first order logic (i.e. Horn clauses) [KOWALSKI 1979]. A program written in (pure) PROLOG can always be considered a description of its result (as well as a procedure to arrive at it). The PROLOG rules given in [PALMER 1984] can therefore be considered formal definitions of the features they extract.

It was somewhat surprising how small the programs to find ridges or streams were, and it became evident that the definition reading of the program text is an objective formulation for the definition of these features.

4 FORMAL MODELS FOR SURFACE ELEVATION

In order to use formal definitions for geomorphologic terms, we have to select a formal model for the surface of the earth, i.e. a formal theory with an agreed-upon interpretation.

4.1 Level Of Detail Included In Model

The surface of the earth may be considered a continuous surface. The major problem we must deal with is the amount of detail to include. This is often linked to the scale of a map, as generalization is customary and necessary in cartography to produce maps of smaller scales.

The level of detail considered (or, as it is more commonly known, the scale of the model) is sufficient to differentiate between sciences. Several examples may be cited in which the same object is studied by different scientific disciplines, the sole difference in approach being the level of detail considered. For example, historians and political scientists use different levels of time scaling. Different levels of detail require different levels of aggregation and different methods for research [MORRISON 1982].

Geography has some implicit understanding of the level of detail at which it studies the surface of the earth. Large rivers such as the Mississippi are clearly included, whereas the detailed topography of my wife's cabbage patch is as clearly excluded.

It seems eminently advisable, that geographers study ways to define the degree of detail necessary for geomorphologic studies. Some studies of the distribution of the size of geographic features raise hopes that there may be gaps in this distribution which may indicate appropriate levels for the exclusion of smaller details [BUTTENFIELD 1984].

Unfortunately the level of detail in a digital terrain model can only be grossly described by the density of the points in the model. Additionally, the method used for collection seems to affect the model [O'CALLAGHAN 1984].

4.2 Terrain And Surface Elevation Models

As mentioned above, the surface of the earth can be modelled as a continuous surface. In the following discussion we exclude folds (overhanging cliffs), as these occur only rarely.

It is not possible to deal with an infinite representation of the surface: thus a representation using a finite number of discrete points must be selected.

Formal surface elevation models consist of parameters which describe the surface (usually heights for specific points) and rules for the use of these parameters (usually interpolation methods to determine the values for points other than the given ones). The selection of the points used to determine the surface is either accomplished using a regular grid [HOLROYD 1985] or is based on special points which are thought to characterize the surface. The selection of special points can be done either using formally defined methods on a previously established formal model (preferably using a regular grid) or by subjective decision on the part of the person doing the data collection. Some more advanced models even include lines.

Studies have shown that the irregular grid - using representative points - requires far less information (fewer measurements) for the same quality of representation than a regular one does [PEUKER 1976],[HEIL 1978].

The form of the surface between reference points must be assumed. For irregularly distributed points a triangulation can be formed (preferably a Delaunay triangulation [Lee 1980]). A triangle of the surface is thought of as a continuous surface of undefined form at this level of resolution. It would be possible to observe finer details of the surface form, but this detail is not available at this level. (We also generally assume that the triangle has uniform attribute values for other observable properties).

For the remainder of the paper an (irregular) triangulation will be assumed. This does not limit the generality of the results, as a regular grid is only a special case of an irregular collection of points, and can therefore also be triangulated.

Most operations applied to surface elevation models are influenced by the density of the points (so-called scale-dependency, but the term resolution-dependency would probably be more appropriate). A relation to fractal dimension seems apparent [MANDELBROT 1975][BUTTENFIELD 1984]. Research towards the identification and formalization of the influence of scale-dependency on results seem to be needed in order to advance quantitative methods of physical geography.

4.3 Terminology Used

The basic entities used in our formulation will be

- nodes
- edges connecting two nodes with a straight line (edges have an orientation)
- triangles, defined by three nodes and the three edges connecting them

It may be noted that these are exactly the simplices of dimension 0, 1, and 2, and that some of the properties subsequently discussed may be expressed in a dimension-independent manner [FRANK 1986].

For the sake of brevity we will assume a most general position of the nodes and exclude the treatment of special cases. For example, we will assume that we have no horizontal edges, i.e. no two connected nodes have the same height. We will also not include treatment of the surfaces of lakes and the sea [MARK 1982].

The triangulation network is defined by the following predicates:

- node, with attribute values identifier, position (x, y coordinates) and height, e.g. node (1A, 17, 11, 62);
- edge, with an identifier for itself, identifiers for start- and end-nodes, and identifiers for the left and right area, e.g. edge (a, 1A, 2A, A, B). We assume that this correctly determines a triangulation of the area of interest;
- triangle, with identifier and values for other attributes, e.g. triangle (A).

In addition, we include a predicate which determines the position of an area with respect to an edge:

```
edge_area (edge1, Left, area1) IF
            edge (edge1, p1, p2, area1, area2).
edge_area (edge1, Right, area1) IF
            edge (edge1, p1, p2, area2, area1).
```

and a predicate to find all edges connected to a node:

```
edge_start (node1,edge1) IF
          edge (edge1, node1, node2, area1, area2).
edge_end (node1, edge1) IF
          edge (edge1, node2, node1, area1, area2).
edge_node (node, edge) IF
          edge_start (node, edge) OR
          edge_end (node, edge).
```

As a formalism, we will use a PROLOG-like, extended language based on predicate calculus. Generally, constants will be capitalized and variable names will start with lower case characters.

The following discussion uses only metric properties of the space. The use of coordinate values for the description of the node position is not necessary, and is included for illustration only.

5 DEFINITION OF FEATURES

Many traditional features cannot be formally defined as their identification depends upon human judgement. Valleys, basins, etc. are examples of such features. The approach taken here is to find features which can be formally defined in the context of the surface elevation model formed by an irregular triangulation.

The use of other than traditional feature designations may seem surprising to some geographers, but this approach may yield non-subjective statistics of landscape features.

5.1 Flow Over Edges

Polyhedrons formed of triangles would be the most general geometric concept. Their edges are either convex or concave. Given that many geomorphologic processes are influenced by gravitation, we introduce the direction of gravity (the local vertical), which gives a special orientation to the surface.

Given the downslope direction for a triangle, it can be determined, for each edge in the triangle, whether the downslope crosses the edge in an inside-out or in outside-in direction. (This replaces the flow numbers used in [PALMER 1984] by a boolean value, where positive flow is 'in', negative flow is 'out'). This grossly models the potential direction of water flowing under the influence of gravity over the edge.

For each triangle, the direction of downslope can be computed (it is the projection in the direction of gravity of the normal on the triangle onto the x-y plane of the triangle) and we can detect whether it crosses an edge from left to right or right to left. (Having asumed the most general positions, we have excluded the case of the two being parallel). For a triangle A,B,C in a Cartesian coordinate system with origin at C, the normal onto the triangle is a x b (where a = CA, b=CB). The projection of the normal in direction of the gravity g onto the side of the triangle a is a x g (where g is the local gravity vector). The cosine between the normal onto the projection of a and the projection of the normal onto the triangle is therefore

cos a = (a x b) . (a x g),

which is

cos a = ay by bz - az by ** 2 - ay bx ** 2 + ax bx bz

(with a = (ax, ay, az) and b = (bx, by, bz)). We have flow across the edge from left to right if this quantity is positive, and flow from right to left if it is negative.

This computation is included in the predicate

```
flow_direction (edge1, area1, Left_Right) IF cos a > 0.
flow_direction (edge1, area1, Right_Left) IF cos a < 0.
```

and we can deduce a predicate for flow

```
flow (edge, area, In) IF
        (edge_area (edge, Left, area) AND
         flow_direction (edge, area, Right_Left))
     OR (edge_area (edge, Right, area) AND
         flow_direction (edge, area, Left_Right)).
```

```
flow (edge, area, Out) IF
        (edge_area (edge, Left, area) AND
         flow_direction (edge, area, Left_Right))
     OR  (edge_area (edge, Right, area) AND
         flow_direction (edge, area, Right_Left)).
```

A predicate edge_flow gives the direction for the flow on both sides of the edge:

```
edge_flow (edge1, direction1, direction2) IF
         edge (edge1, node1, node2, area1, area2) AND
         flow (edge1, area1, direction1) AND
         flow (edge1, area2, direction2).
```

We now differentiate between three types of edges:

- cofluent edges, where the flow from both adjoining triangles is outward (this is sometimes called a valley)

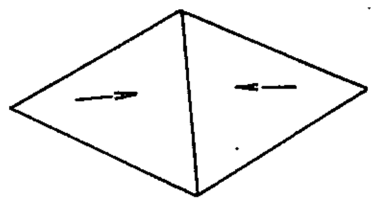


Figure 1: a cofluent edge

- difluent edges, where the flow from both adjoining triangles is inward (this is often called a ridge or watershed)

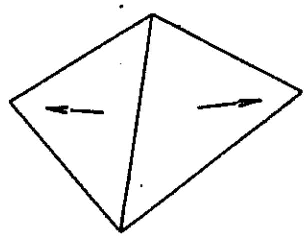


Figure 2: a difluent edge

- transfluent edges, where the flow is out of one and in to the other triangle.

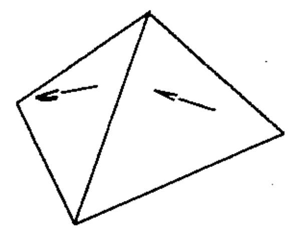


Figure 3: a transfluent edge

We use a predicate edge_type to differentiate between these cases:

```
edge_type (edge1, Cofluent) IF
          edge_flow (edge1, Out, Out).
edge_type (edge1, Difluent) IF
          edge_flow (edge1, In, In).
edge_type (edge1, Transfluent) IF
          edge_flow (edge1, In, Out) OR
          edge_flow (edge1, Out, In).
```

Transfluent edges are only needed for the modelling of the surfaces. They do not contribute to the identification of major morphological features, so we can disregard them in what follows. For more detailed studies, it might be revealing to classify the transfluent edges by the size of the break in steepness and retain those with large values. However, the separation of edges with important breaks from those with smaller changes would require the establishment of an arbitrary cut-off point.

We will use the co/difluence graph, consisting only of cofluent and difluent edges, for the remainder of this paper. (This is similar to approaches used for raster representations, where first pixels with convex or concave neighborhoods are detected (e.g. using the algorithm from [PEUKER 1975]) and then only these pixels are subject to further processing [BAND 1986]).

5.2 Flow At Nodes

Each edge also has a direction induced by the vertical, which is defined as positive from top to bottom.

```
edge_delta (edge, h) IF
        edge (edge, point1, point2, area1, area2) AND
        node (point1, x1, y1, z1) AND
        node (point2, x2, y2, z2) AND
        subtract(z1, z2, h).
```

(This formula uses coordinates, and is deduced from the projection of the edge onto the gravity vector).

```
edge_flow_direction (edge, Down) IF
        edge_delta (edge, h) and h > 0.
edge_flow_direction (edge, Up) IF
        edge_delta (edge, h) and h < 0.
```

Two predicates may help to see if the slope on an edge is towards or off the node.

```
node_slope_to (node, edge) IF
        (edge_start (node, edge) AND
         edge_flow_direction (edge, Up))
     OR (edge_end (node, edge) AND
         edge_flow_direction (edge, Down)).

node_slope_off (node, edge) IF
        (edge_start (node, edge) AND
         edge_flow_direction (edge, Down))
     OR (edge_end (node, edge) AND
         edge_flow_direction (edge, Up)).
```

We can now define subcategories of nodes in a similar fashion to the operations on edges. This yields:

1. peak nodes, where all edges lead away from the node.

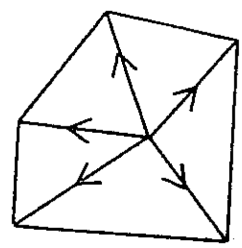


Figure 4: a peak node

2. pit nodes, where all edges lead towards the node.

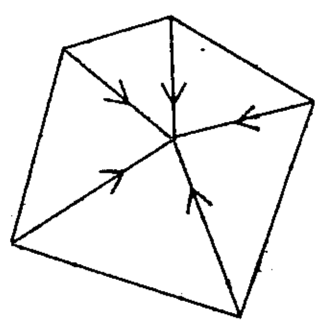


Figure 5: a pit node

3. intermediate nodes, where some edges lead toward, some away from the node.

```
peak (node) IF
        FOR ALL edge_node (node, edge):
        node_slope_off (node, edge).
pit (node) IF
        FOR ALL edge_node (node, edge):
        node_slope-to (node, edge).
```

```
intermediate (node) IF
        NOT pit (node) AND NOT peak (node).
```

To show that all peak and pit nodes are part of the co/difluence graph, we consider the slope of all edges around a peak (pit). One of them has the smallest slope value, thus both its neighbor edges have a steeper slope, and this edge is a difluence (cofluence) edge.

Considering either the cofluent or the difluent edges with the nodes between them, we first look at the situation where two edges join in a node. The slope of both edges may be either towards the node ('pit'-like situation) or off ('peak'-like). Moreover, both edges may slope in the same direction, either towards or off.

First we define a predicate to find two edges of a given type joining in a specified node.

```
node_edges (node, type, edge1, edge2) IF
        edge_node (node, edge1) AND
        edge_type (edge1, type) AND
        edge_node (node, edge2) AND
        edge_type (edge2, type) AND
        NOT equal (edge1, edge2).
```

Using this rule we can define the predicates for the different types of connections between two edges and a node.

```
node_pit (node, type) IF
        node_edge (node, type, edge1, edge2) AND
        node_slope_to (node, edge1) AND
        node_slope_to (node, edge2).
node_peak (node, type) IF
        node_edge (node, type, edge1, edge2) AND
        node_slope_off (node, edg1) AND
        node_slope_off (node, edge2).
node_trans (node, type) IF
        node_edge (node, type, edge1, edge2) AND
        (node_slope_to (node, edge1) AND
        node_slope_off (node, edge2))
     OR (node_slope_to (node, edge2) AND
        node_slope_off (node, edge1)).
```

If for a node in a difluent graph (ridges) all edges form a peak, we have a real peak (this definition is equivalent to the previously given one).

```
peak (node) IF
        node_peak (node, Difluent) AND
        NOT node_pit (node, Difluent) AND
        NOT node_trans (node, Difluent).
```

For a node in a difluent graph (ridges) where some edges form a peak, some a trans join, we have a ridge junction:

```
ridge_junction (node) IF
        node_peak (node, Difluent) AND
        node_trans (node, Difluent).
```

On the other hand, looking at the confluent graph (streams), we find real pits and stream junctions.

```
pit (node) IF
        node_pit  (node, Cofluent) AND
        NOT node_pit (node, Cofluent) AND
        NOT node_trans (node, Cofluent).

stream_junction (node) IF
        node_peak (node, Cofluent) AND
        node_trans (node, Cofluent).
```

Saddles occur in difluent graphs (ridges) as pits in difluent nodes.

```
saddle (node) IF
        node_pit (node, Difluent).
```

It is worth noticing that the dual in the confluence graph, a peak in a cofluent node, is not necessarily a saddle (two streams may touch across a peak).

6 STRICTER ASSUMPTIONS

The above characterizations seem relatively weak compared to such traditional geomorphologic terms as valley, catchment basin etc. Indeed, considering a 'normal' landscape with clearly visible relief, it seems easy to define additional terms.

A number of restrictions apply, however, for a general system that is usable for any landscape. It seems that the attempt to apply more meaningful terms restricts a given landscape to conform to a artificial model, and thus obscures the individual and typical characteristics of the landscape under the general concepts applied.

A number of assumptions often held by geographers cannot be observed on discrete models of terrain. It is often assumed that there is always a stream between two watersheds, and that there is always a watershed between two streams. This is correct for a continuous surface, but cannot be carried over to a discrete model:

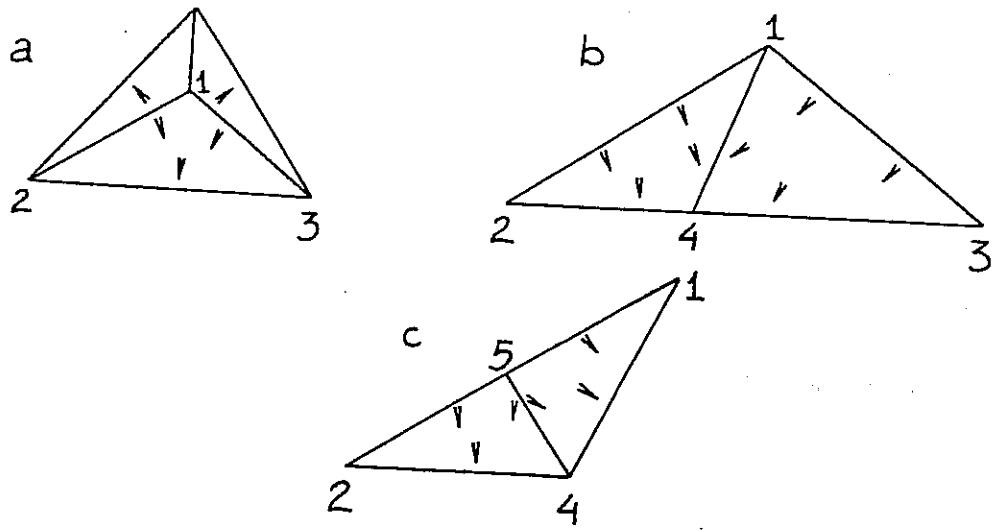


Figure 6: (a) a peak,
(b) with a stream inserted between two watersheds and
(c) a further watershed between two streams.

Assume a peak with three ridges formed by three triangles (as in Figure 6a). According to the above consideration, there must be three streams starting at this peak (otherwise we would not have a stream between any two ridges). This, however creates a stream junction at the bottom of the peak (Figure 6b) where no ridges are included between the streams. Including three ridges (watersheds), we get figure 6c. Unfortunately we have again reached the original situation, by creating a ridge junction where no streams lie between the ridges.

We can conclude from this that some simple assumptions about watersheds and streams cannot be maintained in a discrete model, and may lead to infinite recursion (this was already known to [WARNTZ 1975]), especailly we cannot find strict rules for joins of cofluent and difluent edges in stream junctions and peaks.

Generally, triangulation is a tool suitable for technical applications [JETT 1979], [GRAYMAN 1982]. In an arbitrary triangulation, peaks may have cofluent edges between the difluent ones, but need not. The same holds true for stream junctions, where we may or may not find difluent edges. It follows that the difluent edges do not form closed cycles which clearly delimit catchment basins.

In [PFALTZ 1976], ridges and course lines are defined such that course lines run from pits to passes. The theory lacks provisions for junctions of ridges and course lines; unfortunately this type of analysis is not very useful in fluvially eroded landscape (for a detailed critique see [PEUKER 1973] and [MARK 1977]). In an arbitrary triangulation we cannot guarantee that difluent lines will intersect cofluent lines at saddle points.

Usually the watershed lines are connected and complete, but the streams are only drawn without the top reaches. "Hydrophysically, a drainage channel represents points at which runoff is sufficiently concentrated that fluvial processes dominate over slope processes. If the spatial concentration of surface runoff is simulated, then those points at which this runoff exceeds some threshold can be considered to be the drainage network" [O'CALLAGHAN 1984]. This definition includes other elements than the surface elevation model, and makes the definition of a drainage channel dependent on rain intensity etc. However, it may help to avoid the problem of infinite recursion, as it sets a threshold for the minimal size of objects considered and thus determines the level of detail included in the model.

To avoid the problem of infinite recursion, we are interested in the so called fixed point of the (recursive) operation. The fixed point of this operation is exactly the continuous surface we try to discretize. This does not solve the problem, but shows that it was introduced by the discretization. We can avoid infinite recursion by relaxing therequirements, such that we demand only a difluent edge between two cofluent ones (but not the reverse). This clearly destroys the duality between the two graphs, and produces closed cycles of difluent edges (but not of cofluent ones). The assymmetry between di- and cofluent edges can probably be justified with hydrophysical arguments.

7 CONCLUSION

We come to the following conclusions:

- Formal definitions are only possible based on formal models of reality;
- The definition of formal models for terrain surface depend on the level of resolution (scale) and there seems to be a scarcity of good, formal criteria for the level of detail to include;
- A number of low level concepts can be formally defined. Such features can be automatically extracted and statistics made following this extraction, without resorting to subjective judgement. Some of these concepts are quite near to traditional terms, but not identical;
- On the other hand, a number of traditional concepts (catchment basin, mountain range etc.), with their associated assumptions, cannot be formalized on a discrete model of the surface without further assumptions.

It seems as if traditional concepts are dominated by some 'typical' landscape types. Some of the assumptions held in the literature about their properties are consequently difficult to maintain if applied to other landscape types. Alternatively, these other landscape types, if described by traditional concepts, are pressed into a mold which does not fit them.

We close with recommendations for future work:

The effect of the level of detail on geomorphological modelling should be studied. An attempt should be made to identify levels of generalization at which all finer details can be described by aggregate properties of the lowest level of detail included. Unfortunately, geometric forms of hierarchy are presently not well understood.

This resolution-dependency is not only, as found here, present in the morphologic features of a landscape, but also applies to other features, e.g. political units (town, county, state etc.), railway and street networks etc. The understanding of such structures is generally important, as we conjecture that many of the problems human beings vastly outperform computer algorithms may be solved using a hierarchical approach. For example, the problem of finding the fastest path between '30 Grove Street, Orono, Maine' and '3 Iron Way, Marlboro, Massachussets' is nearly unsolvable for a computer (the shortest path algorithm is approximately of the order n ** 2 [SEDGEWICK 1984] (where n is the number of edges in the network). Any human, however, can easily solve this problem by finding the connections to the interstate network and then solving for the shortest connection between Exit 50 on I-95 and Exit 49 on I-495.

It should be noted that this paper contains several generalization steps, going from a complex surface to its approximation by an irregular triangulation, from the triangulation to the co- and di-fluent edges, and so on. However, a more systematic understanding of these operations would be very desirable. Another method of generalization can be seen in the determination of drainage basins (for edges of channels or ridges alike): the selection of an edge as the root of a larger tree generalizes all the drainage basins contained by its subtrees.

The proposed concepts for the description of morphologic features should be tested on actual landsçape descriptions. We would expect that the relative frequency of these would be different for different landscape types, or that in landscapes influenced by more than one process (e.g. glacial and fluvial) we would see statistical signatures for each process (e.g. two different peaks for the frequency of certain feature classified by size). This would somehow conceptually organize geomorphologic processes into classes of different 'scales' and link this concept to the above generalization problem (a hint appears in [MARK 1977] and [BUTTENFIELD 1984]).

8 ACKNOWLEDGMENT

Part of this work was supported by a research grant from Digital Equipment Corp. to the the University of Maine. Cooperation between the authors was possible using the facilities of BITNET for rapid exchange of ideas; this was a great help. Special thanks go to Linda Langley and Max Egenhofer for help with the preparation of the manuscript.

9 REFERENCES

Band, L.E., Topographic Partition of Watersheds with Digital Elevation Models, Water Resource Research, Vol. 22, No. 1, 15-24, January 1986

Buttenfield B., Line Structures in Graphic and Geographic Space, Ph.D. theses, University of Washington, 1984

Clocksin, W.F., Mellish, C.S., Programming in PROLOG, Springer Verlag, New York, 1981

Frank, A., Kuhn, W., A provable correct method for the storage of geometry, Proceedings of the 2. International Symposium On Spatial Data Handling, Seattle, 1986

Gallaire, H., et al., Logic and databases: a deductive approach, ACM Computing Surveys, Vol. 16, No. 2, June 1984

Grayman, W. M., et al., Use of integrated spatial data and modeling capabilities for urban runoff analyses, International Symposium on Urban Hydrology, Hydraulics, and Sediment Control, Lexington, Kentucky, 1982

Heil, R.J., Brych, S.M, An approach for consistent topographic representation of varying Terrain, Proceedings of the Digital Terrain Models Symposium, ASP/ACSM, St.Louis, Missoury, 1978

Haralick, R., et al., Extraction of drainage networks by using the consistent labeling technique, Remote Sensing of Environment, 18, p.163-175, 1985

Holroyd, F.C., Addressing systems for digital pictures, Tesseral Newsletter No. 2, Reading, UK, 1985

Jett, S.C., et al., Geographic information systems in hydrologic modeling, Proceedings, Hydrologic Transport Modeling Symposium, New Orleans, Louisiana, 1979

Kowalski, R., Logic for Problem Solving, North Holland, 1979

Krummbein, W.C., Shreve, R.L., Some statistical properties of dendritic channel networks, ONR Contract Nonr-1228(36), Task No. 389-150, Tech. Rept. No. 13, Dept. Geol., U. California, Los Angeles, 1970

Mandelbrot,B.B., Stochastical models for the earth's relief, the shape and fractal dimensions of the coastlines, and the number-area rule for island, Proc. Nat. Acad. Sci., 72, p.3825-3828, 1975

Mark, D.M., Topographical Randomness Of Geomorphic Surfaces, 1977

Mark, D.M., Goodchild, M.F., Topologic model for drainage networks with lakes, Water Resources Research, Vol. 18, No. 2, p.275-280, April 1982

Lee, D.T., Schachter, B.J., Two algorithms for constructing a delaunay triangulation, International Journal of Computer and Information Sciences, Vol. 9, No. 3, 1980

Morrison, P., Morrison, Ph., Power of Ten, Scientific American Library, New York, 1982

O'Callaghan, J.F., Mark, D.M., The extraction of drainage networks from digital elevation data, Computer Vision Graphics and Image Processing, 28, p.323-344, 1984

Palmer, B. Symbolic feature analysis and expert systems, Proceedings of the International Symposium on Spatial Data Handling, Zurich (Switzerland), 1984

Peuker, T.K., Geographic Data Structures, Report after year one, ONR Contract N00014-73-C-0109, Dept. Geography, Simon Fraser University, Burnaby, B.C., 1973

Peuker, T.K., Douglas, D.K., Detection of surface-specific points by local parallel processing of discrete terrain elevation data, Computer Graphics and Image Processing 4, 1975, p.375-387

Peuker, T., et al., Digital Representation of Three-Dimensional Surfaces by Triangulated Irregular Networks (TIN), Tech. Report No. 10, Office of the Naval Research, Geography Program, 1976 (revised 1981)

Pfaltz, J.L., Surface Networks, Geogr. Analysis, 8, p.77-93, 1976

Sedgewick, R., Algorithms, Addison-Wesley, 1984

Shreve, R.L., Statistical law of stream numbers, J. Geol., 74, p.17-37, 1966

Shreve, R.L., Infinite topologically random channel networks, J. Geol., 75, p.178-86, 1967

Shreve, R.L., Stream lengths and basin areas in topologically random channel networks, J. Geol., 77, p.397-414, 1969

Toriwaki, J., Fukumura, T., Extraction of structural information from grey pictures, Computer Graphics and Image Processing 7, 1978, p.30-51, 1978

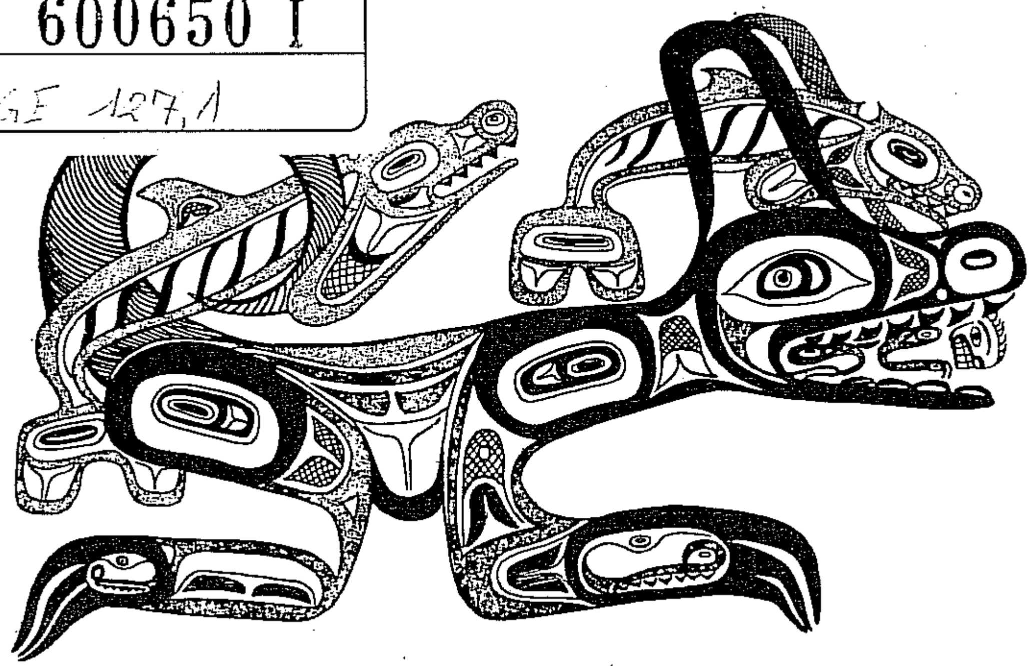
Warntz, W., Stream networks and contour mapping, Journal of Hydrology, 1975

Werner, C. Towards a General Theory of Maturely Eroded Landscapes, University of California, 1978

Wolf, G.W., A Mathematical Model of Cartographic Generalization, unpublished, 1986

PROCEEDINGS

SECOND
INTERNATIONAL SYMPOSIUM ON
SPATIAL DATA HANDLING



July 5–10, 1986
Seattle, Washington
U.S.A.