Monograph:
Gotts, N.M. (1992) Human Wayfinding in Path-Networks: A Survey of Possible Strategies. Working Paper. Institute of Transport Studies, University of Leeds , Leeds, UK.

Working Paper 364

Institute of Transport Studies
University of Leeds

Published paper
Gotts, N.M. (1992) *Human Wayfinding in Path-Networks: A Survey of Possible Strategies.* Institute of Transport Studies, University of Leeds. Working Paper 364

Working Paper 364

May 1992

Human Wayfinding in Path-Networks: A Survey of Possible Strategies

N M Gotts

CONTENTS

Human Wayfinding in Path-Networks: A Survey of Possible Strategies

1. Introduction.

This paper is derived from an unpublished D.Phil thesis:
"Unplanned Wayfinding in Path-Networks: A Theoretical Study of Human Problem-Solving" (Gotts 1988).
It is a modified version of those parts of the thesis most relevant to the design of wayfinding aids such as variable message signs (VMS) and in-vehicle route-guidance information (IVRGI) systems. This paper, however, does not itself discuss such wayfinding aids; rather, it is a step toward understanding the problem-solving strategies and the internal representations of the environment human wayfinders use in the absence of aids of these types. The paper is theoretical in nature, but suggestions are made for empirical tests of some of the hypotheses put forward.

Without an understanding of common wayfinding strategies and the internal representations underlying them, efforts to design genuinely useful wayfinding aids are bound to be groping in the dark: at best, they may be shown to be useful in a specific network, to a specific group of people, under given circumstances - with no guarantee that this usefulness will transfer to other networks, travellers, or circumstances.

Understanding the route-choice behaviour of road-users requires consideration of the information about the travel environment available to them. An assumption of perfect information about the layout of the travel environment and the traffic conditions within it is a convenient simplification - but is also wholly unrealistic. We cannot be confident that this lack of realism is unimportant in transport modelling.

2. Path-networks.

In order to understand the range of possible human wayfinding strategies, we must consider the problems which the travel environment presents, and the informational resources it makes available. The investigation of human knowledge of travel environments and of the constraints that different environments impose on the choice of wayfinding strategy must go together. Both are necessary if we are to intervene effectively to help road-users. This section begins these investigations by developing the notion of a "path-network": an idealised version of real networks of roads, corridors, or similar features of the environment. The section examines the range of structural characteristics such networks can display. The relevance of these levels of structure to wayfinding strategy will emerge in the course of the paper.

Path-networks, nodes and links.
The corridors of a building, the streets of a city, the set of intersecting tracks cleared of vegetation by the regular passage of large animals through a jungle, and the navigable parts of a river-system are all clear examples of "path-networks", as the

term is used here. A path-network is a set of relatively long, narrow path-segments or "links", to which travel of some sort is confined by physical obstacles or social convention. For some purposes, links can be thought of as one-dimensional. At each end of each link is a "node" - which will be shared with one or more other links unless it is a dead end, past which travel cannot continue. We can distinguish a special class of path-segments, "loops", which are generally not considered as links: a loop sets out from a node and returns to the same node, while the ends of a link are distinct nodes. The path-segment marked "5" in Fig.1 is a loop, while the rest are links. Fig. 1 also shows that several links (6, 7 and 8) can form an unbranched sequence. In a real road-network, such a sequence might be regarded either as several successive links, or as a single link with bends in it. Fig.2 shows two similar examples: in 2a, the shaded part could be regarded as a single link or as a chain of four links, while in 2b the shaded area could be thought of as one loop or a cycle of four links. There is therefore not necessarily a <u>unique</u> representation of a real road-network as a collection of nodes, links and loops.

The fact that real path-segments are not truly one-dimensional, but have a non-negligible width, can also give rise to doubts about the exact node/link structure that should be assigned to a path-network. Fig.3 shows an example, in which it is unclear whether the part of the network surrounded by the dotted line is best regarded as one node, or two joined by a short link.

Human path-networks can be considered to exist at a wide range of scales, and what is treated at one scale as a complex path-network - such as a city - can appear on another to be a single node - as when the cities of a country are considered along with the network of main roads or rail lines that links them. In a rather similar way, what is at one level of detail treated as a single link may at a more detailed level turn out to be a "bunch" or "braid" of smaller links, running roughly parallel, like the navigable channels of some rivers, or the lanes of a multi-lane highway.

A link in a path-network may be bidirectional - meaning that it can be traversed in either direction - or unidirectional, like a one-way street. Street-maps which fail to indicate one-way streets are a bane to the driver or cyclist, often leading to the choice of a route which turns out to be impossible to follow.

Connectivity.
It is useful here to distinguish between "connected" and "unconnected" networks. In a connected network, there is at least one continuous chain of links from any node to any other. Among connected path-networks, "trees" form an important subset. A tree has no loops, and no "cycles" - a cycle being a sequence of links forming a circular route, like links 2, 12 and 3 in Fig.1. There is only one possible route between any two nodes in a tree.

Circular ordering of exits.
At any road-junction, the exits can be arranged in a "circular order", clockwise or anticlockwise around the junction. The exit

to be taken from a junction on a particular journey is often specified in terms of its circular-order relationship to another exit, or to some other landmark situated at the junction. The exit to be chosen may be specified relative to the path-segment by which the traveller is expected to enter the junction ("egocentrically"), or in a way independent of the traveller's entry-point ("allocentrically").

> "Take the third exit" is egocentric, while
>
> "Take the next exit after the one going to the Clock Tower" and
>
> "Take the exit to the left of the fire station" are allocentric.

The direction in which exits are to be counted (clockwise or anticlockwise) must be specified or understood.

Plane path-networks and mosaic regions.
Some path-networks lie in a single continuous surface, in which all their nodes are embedded. The corridors of a single storey building and the streets of a town without bridges or tunnels are good examples. As the second example indicates, the surface need not be entirely flat, so long as it does not curve back on itself like the surface of a sphere. Such a network will be called a "plane path-network".

Unless it is a tree, a plane path-network divides the surface it lies in into a mosaic of regions, each bounded by a circuit of links. Minimal and non-minimal mosaic regions can be distinguished: minimal regions are not divided into two by any loop, link, or chain of links. The minimal regions of an example network are shown by numbering in Fig.4. Knowing the mosaic-topology of a path-network is potentially useful in wayfinding: mosaic regions, minimal or non-minimal, can be used to specify location - as inside or outside that region. This issue is discussed in section 8.

Two plane path-networks can have the same node-link topology, while differing in the way they are embedded in the plane. Fig.5 shows three examples (the letters attached to nodes in 5c are intended to indicate how the nodes and links should be matched). In all these pairs, despite identical node-link topology, the two networks differ in the circular order of exits around one or more junctions.

Paths And Path-Continuity.
We do not always represent a street intersection simply as an ordered circular list of the links leaving it: frequently, some pairs of links will be perceived, and remembered, as continuous with each other - that is, as belonging to the same "path". A "path" in the sense used here is a series of one or more links, which is conceived of and treated by a wayfinder as a single linear element. A path may form a closed curve, like the Circle Line of the London Underground.

A path-structure can be imposed on any street junction, but there would be a much greater degree of consensus among travellers at some than at others. Most commonly, two links are perceived as part of the same path for one of two reasons. The first is that no other links leave the junction (i.e. the links form part of an unbranched sequence). Second, two links may be seen as part of the same path because entering the junction by one of the two and leaving by the other involves little or no change in direction. These criteria were used in constructing Fig.6. This shows two networks that are identical in node-link topology, link-order at nodes, and mosaic-structure. However, most people would assign them different _path-structures_ - that is, would group their links into a different collection of paths. Factors such as width, colour, and the presence of indications such as white lines across roads, may also cause two links to be regarded by a traveller as parts of the same path, even if there is a sharp angle between them and there are other path-segments leaving the junction. Two path-networks with different node/link topologies cannot have exactly the same path-structure: a full description of the path-structure will specify the number and order of links in each path, and the nodes at which paths meet.

Three important consequences for wayfinding follow from the conceptual merging of links into paths. First, the imposition of an extra level of structure on the network should make its node/link structure and the exit-order at nodes easier to remember, because of the human propensity for "chunking" information. Second, paths are useful in specifying location - a point further developed in section 8. The location of something may be specified as to one side or the other of a path. In a plane path-network, paths also allow particular non-minimal mosaic regions to be picked out as significant and easy to remember: those bordered by a small number of _paths_, even if by a large number of _links_.

The third consequence of joining links into paths is to produce a system of default choices at junctions. Fig.7 illustrates the point. The route marked can be specified in a set of instructions such as:

> "Follow the path through A until you reach B, then turn right and follow the path until you reach C".

These instructions should be easy to obey. Yet they mention explicitly only three of the nodes on the journey, all the rest being covered by the phrase "follow the path". This instruction has to be accompanied by information about when to stop path-following, but can be obeyed at road-junctions in completely unfamiliar cities, so long as the local properties of the junction allow a clear choice of exit on this basis. Such instructions can also make use of the ordering of junctions, or of some other type of element such as traffic lights, as in:

> "Go left at the second junction after the first set of traffic lights."

"Following the path" is not always possible. If you approach a T-junction or fork along the stem, for example, it may not be possible to pick out one exit as that which continues the path you are on. Furthermore, continuity between path-segments may not be symmetrical, as Fig.8 shows. If you were obeying an instruction to "follow the path", and approached V1 along E2 or E3, you would almost certainly continue along E1 without hesitation. Approaching along E1, however, neither E2 nor E3 is an obvious continuation of the same path. It seems best to say here that two paths merge at V1.

Major roads.
In the road-networks of most towns of any size, some streets are of greater importance than others. Importance may be assessed by width, by the flow of traffic, or by the precedence of one street over another at a junction - as shown by the way traffic through that junction is regulated. These criteria tend to go together, creating a fairly well-defined set of <u>major roads</u>. Such a set generally forms a single, multiply-connected network covering the city, and this can be used as the basis of longer-distance travel.

The subset of major roads in a city street-network may, like the whole network, form a mosaic of regions. Regions bounded by major paths, and crossed only by non-major paths or none, are the minimal mosaic regions of the major-path network.

3. A taxonomy of information useful in wayfinding.

This section switches the emphasis to the types of information about the travel environment a road-user can have. The taxonomy given is certainly not complete; it is based on the content of pieces of information - what they are about - rather than on their source or their form. Any of the types of information discussed here could come from any of three possible sources:

(1) From that part of the travel environment which is visible, or otherwise accessible to the wayfinder's senses at any particular time (the "local environment", as I shall call it). This local environment may include signposts, street names and other fixed-position wayfinding aids.

(2) From portable wayfinding aids such as maps and sets of directions, which can give information about both the local and non-local parts of the travel environment.

(3) From memory. As an area becomes more familiar, the contribution available from this source increases, and may therefore decrease dependence on the other two. Clearly, it is less easy to assess this source than the others.

Any of the classes of information discussed could also be expressed in various forms - verbally or diagrammatically, for example. All the specific examples discussed here are in verbal form, but I do not mean to imply that a wayfinder must be able to verbalize a piece of information in order to use it.

Two principles of classification that cut across each other form the basis of the taxonomy. The first of these separates pieces of information that explicitly specify, suggest, or limit the wayfinder's choice of route from those that do not. This distinction parallels that between roadside and in-vehicle information systems that give route guidance or recommendations as such, and those that merely provide information about the state of parts of the network relevant to route-choices, leaving the road-user to draw appropriate conclusions. A piece of information that does make explicit recommendations on the choice of route I shall call a "route-choice rule". The second principle classifies pieces of information by their degree of "place-specificity". This does not give rise to a straightforward dichotomy between place-specific and non-place-specific pieces of information. Nor is it always possible to say whether one item of information is more place-specific than another, because the ranges of places in which two pieces of information can be used may overlap without either containing the other. Rather, pieces of information can be regarded as forming a partial order with respect to degree of place-specificity.

A route-choice "rule" could also be called a "conditional instruction" - conditional, that is, on the wayfinder's destination, current location, and possibly other conditions such as the time of day, and traffic density. Examples could be drawn from almost any set of directions for a journey that one person gives another - e.g:

> "To get to Rob's place from here, go down Lewes Road toward Brighton, past the traffic lights, and take the next left" (rule 1).

The most place-specific route-choice rules are of very narrow applicability: they can be used on only a narrow range of journeys. These most place-specific rules include those that specify particular roads and turnings to be followed in a street-network (like the example given above). Another variety of highly place-specific route-choice rule specifies an intermediate destination to be adopted, e.g:

> "If you are in Brighton, and want to travel to London by car, first get to the start of the M23" (rule 2).

This rule shows that the specific places referred to by a highly place-specific rule may be very large - like London and Brighton.

Less place-specific are rules that can be used at or within any place of a particular type, for example:

> "If you are seeking the sea in a seaside town, take downhill exits at junctions where possible", and

> "If you are at a road-junction, and your destination is visible in a direction between those of two of the roads leaving it, take one or the other".

Finally, it is possible to find route-choice rules that are completely non-place-specific, in the sense that they could reasonably be used in any travel environment whatever, such as:

> "If you find yourself back at a point you reached before, do not make the same sequence of moves again".

The class of route-choice rules can also be subdivided in ways that do not concern place-specificity. In particular, we can distinguish two subclasses that will be referred to as "where-next" rules and "intermediate destination" rules, of which rule 1 and rule 2 above are respective examples. The instruction half of a "where-next" rule, as the name suggests, tells the wayfinder which way to go from her current location; following or obeying such a rule during travel will result in a change in position. Of the rules given above, only rule 1 belongs to this subclass: rule 2 is an intermediate destination rule, while the remainder impose constraints on where the wayfinder should go next, but will fail, in at least some circumstances, to specify just which way the wayfinder should leave the current location - that is, where she should go <u>next</u>. It is possible for both where-next rules and intermediate destination rules to be less place-specific than rules 1 and 2. For example:

> "If there is only one exit from the current junction that you recognize, take it", and

> "If your destination lies within a region you are currently outside, and there is only one entrance to that region, adopt that entrance as an intermediate destination".

The range of types of information other than route-choice rules that could be relevant to route-choice is practically unlimited. The following list of types is intended to include all those discussed in this paper. Some are specifically related to wayfinding, while others are primarily related to other types of problem, or to no type of problem in particular. Items of information within each of the first five classes listed vary in their place-specificity, and the examples given of each of these classes are arranged in order of decreasing place-specificity. All items in the last two classes are non-place-specific.

(1) Wayfinding advice. Examples of this category include:

> "If you are lost in Central London, keep a lookout for the Telecom tower", and

> "Keep track of the direction in which you are travelling relative to your destination".

(2) Information about the costs and difficulties of travel on particular routes or in particular areas. This might be regarded as a sub-category of wayfinding advice, but is so important that it deserves to be considered independently. Examples, again in order of decreasing place-specificity, include:

"It takes fifteen minutes to walk from my home to Brighton station",

"It is easy to get lost in the Hampton Court maze", and

"Wayfinding is particularly easy in cities where the streets form a rectangular grid".

(3) Episodic knowledge: records of current or previous journeys and the wayfinding problems encountered in the course of them, recording information such as:

"At the last three junctions encountered on the current journey, the leftmost exit has been taken",

"The last time I travelled from Wells to Brighton, I took a wrong turning a few miles from the start of the journey", and

"Every time I have made a journey through a foreign city without a map, I have got lost".

(4) Factual information about the spatial properties and relationships of places, that may be useful in wayfinding. Examples here include:

"North Street meets Queens Road at Brighton Clock Tower"',

"No corridor in the Arts E building at Sussex university is more than 100 metres long",

"Many North American cities are built on a rectangular grid plan", and

"Major roads are generally wider than minor ones".

(5) Rules for tasks that are themselves closely related to wayfinding, such as map-reading and place recognition. These, like route-choice rules, may be more or less place-specific. For example, a place recognition rule might enable the wayfinder to identify a particular junction, as in:

"If you are in Brighton, and there is a fire station visible, identify the junction it is at as Preston Circus",

or it might be used to identify a general type of place, as in:

"If you see a tall building with a cross on top, identify it as a church".

(6) Information about the properties of space itself, and about general or abstract spatial entities such as straight lines, triangles, and flat surfaces. This can be indirectly useful in route-choice. Such information includes, for example, the knowledge that:

"Four right-angle turns on a flat surface make up a complete rotation", and that

"If X is larger than Y and Y larger than Z, then X must be larger than Z".

Our representations of specific places may actually be incompatible with what we know about the properties of space - properties to which we are aware any specific place must conform. Sometimes such inconsistency leads to the discovery of errors in stored place-specific information, but pieces of place-specific information that would be seen to describe spatially impossible arrangements if considered together may coexist as long as they are not compared.

(7) Finally, general principles of problem-solving are relevant to wayfinding problems, as to any others. For example, we know that if we have solved a particular problem before, it is often worth trying to remember how it was done. Furthermore, if we have not solved the problem we are presented with before, it is often worth considering whether it can be broken down into sub-problems of which some or all have been solved before.

Route-choice rules, and the rules used in wayfinding-related tasks discussed under point 5 above, can be classified as "procedural knowledge", or "knowledge <u>how</u>", while the other types of information discussed above can be classified as "factual knowledge", or "knowledge <u>that</u>". This distinction is important in the next section, and in later parts of the paper.

<u>4. Cognitive maps</u>.

In this section, I am concerned with factual knowledge of the travel environment only. This is a broad category - including knowledge about distances and directions, structural relationships between places, their appearance, and travel conditions along particular links, among other possibilities - but not route-choice rules. The standard term for a person's factual knowledge of the medium-to-large-scale environment is their "cognitive map"; this term has been used with a wide variety of meanings. Some of these uses give a misleading idea of how people's internal representations of places are likely to be organized.

Some users of the term "cognitive map" intend it to indicate only that people build up complex symbolic representations of their surroundings (e.g. Kuipers (1977)). For others, however (such as Tolman (1948), O'Keefe and Nadel (1978)), the adoption of the term expresses the belief that something is built up inside the brain or mind that resembles (in some way) both a real map such as those of the Ordnance Survey, and space itself (since a map is a spatial entity). Generally, this type of "cognitive map" is supposed to contain information about the metrical relationships between each pair of places represented (particularly straight-line distance and direction), and to allow its possessor to scan it and read off information as if it were indeed printed on paper. It is this sense of the term that is misleading, as I shall argue.

There are strong objections to the kinds of "cognitive map" theories of wayfinding, depending on accurate knowledge of metrical relationships between places, developed by Tolman, O'Keefe and Nadel, and others. Four objections will now be developed in turn. The first three are theoretical, the fourth considers empirical evidence in the light of these theoretical considerations.

(1) Limitations of maplike representations.

The fact that we have either explicit or tacit knowledge of many properties of space and abstract spatial entities may make it plausible at first sight that we have access to an internal "psychological space" or "mental space", with spacelike properties. In fact, such knowledge does not imply the possession of any "quasi-spatial" mental objects with the property in question, any more than a physicist's knowledge of thermodynamics implies the existence of a "mental gas" composed of multitudes of "mental molecules" that respond by moving faster when the physicist imagines heating it.

It cannot be the case that all our knowledge of the spatial properties and relationships of places is held in any very "maplike" form. This can be shown by considering the limitations that maps on paper have. It is characteristic of maps that each part of the area mapped is represented by a part of the map, and conversely, the contents of each part of the map represent what is to be found in a specific part of that area. This means that maps are very poor at recording partial states of knowledge. For example, I may know that junction B lies on London Road, between junction A and junction C, without knowing whether A or C will be encountered first when travelling into Brighton. There is no way of recording this partial information on a conventional map, because each element mapped must be assigned a particular area of the map to represent it - but placing A, B and C on London Road in either order risks a misrepresentation of the real situation.

A map is a specialized format in which some forms of information can be stored and used with great convenience, while others cannot be expressed at all.

(2) Now would maplike internal representations work?

If it is claimed that some of our place representations are maplike (or spacelike), it has to be specified:

(a) In what ways these representations are maplike.

(b) How these maplike properties are implemented or supported.

Taking question (a) first, it must be specified whether it is possible for the cognitive map to be scanned and rotated like a real map, and whether it has a scale. Can distances and directions can be measured across it? Can the mapper mentally draw a pair of straight lines or a circle on it and count (for example) the Chinese restaurants between the lines or within the circle? Can a piece of "mental string" be used to measure travel distances along winding paths on the map in the way real string can be on real maps? If the answer to any of these queries is affirmative, question (b) above must be raised. A "cognitive map" needs a "cognitive map-reader", as maps do not read themselves. This is not in itself an insuperable objection to the existence of such maps, but the need for such a map-reader makes it clear that attributing problem-solving skills to the possession of an internal map is <u>insufficient</u> as an explanation of human abilities (the interpretation of real maps is far from being fully understood).

Turning to question (b), O'Keefe and Nadel (1978, p.223) tentatively propose that we have sheets or arrays of neurones upon which maps are "printed". However, they do not address the immense problems of explaining how these maps are produced, stored and accessed. Such arrays do not in themselves make possible operations like scanning, rotating, or drawing circles, all of which would require that the arrays, or some process working on them, perform <u>computations</u>.

(3) Limited usefulness of metrical information in path-network wayfinding.

Let us now consider the problem from the other end: how <u>useful</u> in path-network wayfinding is the information about straight-line distances and bearings between places that can be gained from a scale map or model (and that the "cognitive map" approach to wayfinding stresses)? This question can be illuminated by considering different possible forms of external, publicly-available maps as aids in solving a wayfinding problem.

Fig.9 shows three representations of the same imaginary city-street network. 9a shows both metrical and structural information, while 9b shows only the relative positions of the nodes, and 9c shows only the node/link structure of the network: it is a list of all pairs of adjacent junctions in the network formed by the paths, which can be checked by reference to 9a. (Few wayfinders are likely to be given "maps" like 9b or 9c: they are used here simply for the purpose of comparing the helpfulness of different types of information.) 9a actually shows more than 9b and 9c combined, as it portrays both additional structural properties of the path-network beyond its node/link structure (the mosaic of two-dimensional regions it forms), and the metrical properties of the links and paths. Neither of the other representations includes these types of information. It would be all too easy to assume that the wayfinder has an internal representation "like" Fig. 9a, without specifying what information is supposed to be immediately available to the wayfinder, what further information could be inferred, and what is absent.

Imagine yourself approaching point A from the left edge of the diagrammed area, with a compass and one of the three representations in Fig.9. No junction is visible from any other. The problem is to get to point L by the shortest route, while sticking to the path-network. Assume that you have no prior acquaintance with the area, that every junction in the area mapped is clearly labelled, and that each exit from each junction bears the name of the next junction in that direction - as in 9a. Information about node/link structure and metrical properties of a junction can therefore be gathered from the local environment when that junction is reached during travel.

(1) 9a would make the problem simple, for anyone able to read it. The shortest route is easy to find. (Of course, in real travel, we may want to find - for example - the quickest, least costly, least congested or most scenic route.)

(2) 9b is almost useless. Using the map and compass, you can work out which exit leads most directly towards the destination: the exit labelled (B). In the absence of non-local node/link-topological information, it is reasonable to take this path since it will take you closer to the destination as fast as possible, and <u>in general</u> straight-line distance and travel distance correlate positively in path-networks. In this case, however, following it wastes time, and you will have to return to your starting point in order to reach the destination. You should find the destination eventually, as its bearing will be calculable from any other junction, using the map and compass. Notice that this "map" cannot be used effectively for planning: the wayfinder has insufficient information to plan.

(3) 9c is more helpful than 9b (and can be used for planning), but less helpful than 9a. The compass is of no use, but the list of links reveals a number of possible solutions (two, if no routes going through the same node more than once are counted). The list does not show which is the shorter of these two, nor which requires least time or effort. One possible strategy is to use the route with fewest nodes along it. This will work reasonably well in networks where most links are of about the same length. Another way of describing this strategy makes clearer the fact that it is guaranteed to yield a solution: always choose an exit that reduces (by one) the minimum number of links between your own position and the destination. In the layout shown, this strategy does not produce the <u>shortest</u> <u>route</u>, but you can be sure when you set out that you will not need to return to point A in order to reach L.

(4) Empirical studies.

Empirical studies suggest that we do not in general possess metrically accurate information about large places; that information about path-network structure is more readily acquired during travel than information about direct distances and directions; and that the latter is not necessary in path-network wayfinding.

Gärling and his colleagues (Gärling et al 1981) tested Gärling's hypothesis that the location of landmarks (meaning their direct distances and directions from each other) would be acquired before their order along a path. This hypothesis, as Gärling notes, is directly opposed to that of Siegel et al (1975) and Cousins et al (1983): that the appearances of landmarks, and their order along a path, would be learned earlier by those using the path than the direct distances and directions between the landmarks. The experimental results supported Siegel's hypothesis (as the arguments above would lead us to expect).

Estimates of the straight-line (or "direct") distances and directions between places can be obtained from human subjects in numerical form, or, if originally non-numerical, can readily be translated into numerical terms (as when subjects are asked to represent the ratio of two distances by marking off line-segments of proportionate lengths) . The numerical data can then be treated as the input to statistical processing. There is a good review of work in this area (Bartram and Smith 1983). I make only a few points here.

Studies of distance estimation in particular have been plagued by inconsistent results. For example, Lee (1970) found that estimates of distance toward the centre of town were overestimated relative to those away from the centre, Golledge et al (1969) and Briggs (1973) found the opposite.

More recent work has led to a number of conclusions that are of interest here. Byrne (1979) found a tendency to distort angles between streets, recalling them as nearer to being right angles than they are. According to MacEachren (1980), the estimated distance along a path is correlated more highly with travel time than with actual distance - although actual and estimated distance also correlate strongly. Sadalla and Staplin (1980) discovered that the estimated distance along a route is positively correlated with the number of nodes along the route - but not with the amount of information encoded about each node. Distance estimates may depend on whether a subject is asked for the distance from A to B, or that from B to A (Cadwallader 1979) - in other words, they may be non-commutative. Distances are always commutative in a metric space, so Cadwallader's subjects were not arriving at their answers by measuring internal representations with metric properties.

Heft (1979) found that subjects learned a filmed trip easily, despite minimal opportunity to maintain overall orientation. Ability to make the right choices at junctions correlated negligibly with ability to keep track of location in a global framework. This should not be surprising: in a path-network, what you <u>need</u> to know is which exit to take at each junction. Maintaining global orientation sometimes helps, but the demands of the task do not make it essential.

Pailhous (1970, 1980) studied the wayfinding skills of novice and expert Paris taxi-drivers. He came to the conclusion that his subjects (both novice and experienced) distinguished a "primary network" of main roads from a "secondary network" of minor ones, treating travel on these two networks differently. These two networks, he believes, are common to all who travel in towns, although sometimes less clearly distinguished for other people than for taxi-drivers. The primary network, he believes:

> "supports ... behavior governed by an algorithm" (Pailhous 1980, p.315),

according to which the driver chooses at each junction the exit that makes the smallest angle with the bearing of the destination from that junction - which requires accurate knowledge of the metrical relationships between places. In the secondary,network, the drivers used intermediate destinations and local "tactics". Both experts and novices tended to favour travel on the primary network, and to minimize travel on the secondary network (novices more so). According to Pailhous, all drivers knew all the bearings between points in the primary network to a good approximation, but "distortions" appeared in their knowledge of the secondary network.

Pailhous does not give a diagram of the primary network his taxi-drivers used, nor discuss its layout. He notes (Pailhous (1980), p.316) that the use of the "least angle" algorithm requires that the primary network should include no gaps or dead ends. In fact, even with this restriction, the algorithm would not work well. Fig.10 shows the results of applying it to travel problems in particular path-networks. In this figure, "S" stands for "starting point" and "D" for "destination". Fig.10a shows that the algorithm would have to be modified to forbid the traveller making U-turns. 10b shows the need to find some way of preventing it leading the wayfinder around in circles. Fig.10c shows that in a rectangular grid, it would lead to the adoption of zig-zag routes, which might be as short as any other in length, but would be slow for a car because of the excessive number of turns involved. Also, the algorithm takes no account of traffic conditions, while the taxi-drivers with whom I travel frequently do so explicitly in choosing the best route.

In short, if Pailhous's drivers use the methods he believes they do, it is difficult to see how and why they do so.

Chase, who studied taxi-drivers in Pittsburgh (Chase 1983), and whose conclusions contrast strongly with those of Pailhous, comments:

> "If taxi-drivers have access to a bird's eye metric view of the city, they certainly cannot draw it." (p.396).

He found that expert and novice Pittsburgh taxi-drivers did not differ significantly in their ability to draw maps of the important parts of the city or of selected intersections, name as many neighbourhoods as they could, place 20 well-known neighbourhoods on an outline map of Pittsburgh, or make estimates of the distances between places within it. On the other hand, they did show a greater ability both to describe good routes, and to find them in practice. This difference was particularly marked when the best routes involved using relatively obscure streets. These results indicate considerable limitations on the accurate metrical knowledge available from memory even to wayfinding experts.

To conclude this section: the view that wayfinding (including path-network wayfinding) depends upon the possession of "maplike" internal representations that encode accurate information about metrical relations between places is not supported either by empirical evidence or by sound theoretical arguments: there is no reason to believe that accurate metrical information about place relationships is of central importance in path-network wayfinding, nor do people generally appear to possess it.

It may be said that although the evidence and arguments put forward in this section make it unlikely that human wayfinders possess a map with a metric defined over it, their internal representations may still be "maplike" in some weaker sense. Yet the argument about states of partial knowledge in point 1 above shows that these internal representations must also differ from maps in being able to record a greater number of possible states of knowledge. To give a second example, a map, even if not intended to be metrically accurate (like the London Underground map), forces determinate answers to questions such as:

"Where does the Bakerloo line meet the Circle line?",

whereas a user of the system, while knowing that the two do indeed intersect, may be unable to say where.

This section's primary aim is critical, but it is reasonable to ask what alternatives there are to maps. One possibility is that information about the travel environment is stored in a way which makes no use of the particular characteristics of spatial properties and relationships at all: as an indexed database of assertions about particular places in which all classes of information are treated in essentially the same way, for example. However, there is a possible halfway house between maplike representations and an unstructured database, which can be called a "structural representation" of the travel environment. This would include all the types of factual information that need to be stored, but would be organized around information about the <u>structural relationships</u> between places. In the category of "structural relationships" I include whole/part relationships, those of overlap, adjacency and ordering (as in the order of nodes along a path), and the relationship between an area and its boundary (e.g, a mosaic region and the circuit of links surrounding it). Insofar as maps make the representation of such relationships particularly easy, any system of representation which lays stress on them could be considered "maplike", but unlike a map, the type of structural representation proposed here would allow for the storage of states of partial knowledge about structural relationships.

By saying that the structural representation would be "organized around" information about structural relationships, I mean that information involving structural relationships would not simply be stored, but would be used to direct the processes of inference set off when a new item of information is added, and if necessary the resolution of contradictions between old and new information. The implications of any new piece of information about one place might, for example, be checked by running through the structural links it has to other places (its parts, larger places it is part of, its neighbours, etc.), and in each case making use of a "standard package" of inferences to be drawn and consistency checks to be made, given the type of structural relationship and the type of new information being added. Structural information might generally be given precedence in the case of a contradiction being uncovered: for example, if region A is believed to be part of region B, but A is recorded as being at least 5 sq.km., while new information comes in that the area of B is less than 3 sq.km, something is wrong somewhere - and I hypothesize that the metrical information would be questioned and rejected more readily than the whole/part relationship.

5. To plan or not to plan?

There has been a widespread assumption among researchers into human wayfinding that it is usually, or even always, planned: that the wayfinder works out the route in advance or runs through it mentally (though just what counts as planning is in general not defined). This section both documents the existence of this assumption, and introduces some arguments against it.

To begin, let us ask what purposes the planning of a journey can serve. First, planning may enable the wayfinder to choose which way to go next, by revealing the consequences of each possible choice. Second, it can enable her to be reasonably confident that the intended destination can in fact be reached, within any given constraints of time or resources. Third, it can allow the wayfinder to make the journey with less effort: if you plan your route, you may avoid errors you would otherwise have made.

Gärling and his colleagues at Umeå university believe that detailed journey planning is normal. For example, in Gärling, Böök, Lindberg and Säisä (1984), the following statement is made:

> "Rather than observing people actually travelling in the environment, the present research program takes as its point of departure the assumption that travel is planned."

This assumption, although frequently made, is quite unjustified, for a reason actually noted by Säisä et al (1984):

> "Much time and effort is not likely to be invested in planning of everyday activities."

Pailhous (1970, 1980), asked his Paris taxi-driver subjects to give verbal descriptions of routes to places he named, comparing the performance of experienced and novice drivers. He did not study his subjects' performance during actual journeys, but claims that the plans made in advance were seldom "questioned" during travel (p.315), although it is unclear how he knew this without studying his drivers during real travel. He states that the non-visibility of the destination from the starting point on most journeys makes planning necessary. The claim that planning is always necessary when the destination is not visible from the starting point, however, is quite implausible: do you need to plan a route from your kitchen to your front door when the doorbell rings?

Chase's findings, also from a study of taxi-drivers (Chase 1983), contrast markedly with Pailhous's. Unlike Pailhous, he studied actual travel as well as advance route-planning, and found that about 25% of routes were improved when the actual journey was made, suggesting that even wayfinders who are very familiar with an area cannot run through routes mentally as well as they can follow them on the ground; locally-available information remains important, and therefore the reliability with which advance planning will produce the best possible route is limited. Even Chase, however, assumes (Chase 1983, p.404) that a "global plan" is formed, and the details are then <u>sometimes</u> filled in during travel without further planning.

Passini (1980a, b), reports studies of wayfinding in Montreal, in which subjects were given moderately complex wayfinding problems to solve involving routes through buildings, the streets, and the Metro. A problem-solving protocol was taken during each trip, and a post-test interview conducted. The subjects had been in Montreal only a couple of weeks - and so were presumably not very familiar with the settings, although Passini says only that with one exception all subjects were equally familiar with them.

Under these conditions, the subjects did indeed plan - but they did not plan the whole route in advance, then follow the plan step by step. Rather, the plan appeared to be developed in a continuous and flexible fashion. As Passini puts it:

> "The data show that an original task does not lead to an assessment of all the subtasks involved because not all environmental factors are known or predictable. This applies also for subjects who know the setting relatively well."

Kuipers (1979, 1982, 1983) suggests that knowledge of routes is stored in the form of "views" and "actions" - "V"s and "A"s. These are used both in real travel, to recognize visible places and direct physical movement, and in recalling the route. In Kuipers (1983) the views and actions are described as:

> "egocentric descriptions of sensorimotor experiences" (p.217).

They enable their possessor both to recognize views and perform actions on the one hand, and to <u>recall</u> these same views and actions on the other.

Vs and As can be linked in two ways:

> "The link V->A has the meaning that when the current view is V, the current action should be A to follow the route. The link (VA)->V1 has the meaning that if the action A is taken in the context of view V, the result will be view V1." (p. 217).

Kuipers proposes that when there is a complete set of V->A links covering a route between two places, the route can be followed; and that when there is also a complete set of (VA)->V1 links, it can also be "reproduced in the absence of the environment" - that is, "rehearsed" or run through mentally. (Running through a route mentally is one way of planning a journey, although not the only one. The wayfinder could plan "top down" instead of in linear fashion, selecting one or more high-level subgoals, then replanning at successively more detailed levels. More mixed, opportunistic strategies are also possible.) Kuipers' idea is that if only the V->A links exist, the real world will produce the next view automatically, but the possessor of the links will be unable to generate the sequence of Vs and As by herself. However, Kuipers is confounding two things here:

(a) Whether a representation of an action has associated with it a representation of the <u>effect</u> of the action (as it does in the case of (VA)->V1 links but not in that of V->A links).

(b) The differences between the recall and use of stored information about a place that is possible for a person when <u>within</u> that place, and when located elsewhere.

To see that this distinction needs to be made, and that Kuipers has not done so, consider a single "view" (on some route from A to B that has been previously followed) at which a decision must be made on the action to take. When encountering this view during a journey from A to B, the wayfinder may be unable to decide what action to take; may be able to do this, but unable to predict what new "view" will result; or may know what to do, and be able to describe the new view at some level of detail in advance of taking the action. However, even if this last can be done during actual travel, it is quite possible that the wayfinder will be <u>unable</u> to recall what will be seen after the action, or even what action should be taken, when <u>imagining</u> the journey while located elsewhere. This is so because when the traveller is actually located at the place corresponding to her stored "view", she is likely to be presented with many features of the place that she can always <u>recognize</u>, but could not reliably <u>recall</u> in their absence. The recognition of one or more of these features may act to trigger the memory of the action required, and/or of the view that will be seen once that action is taken.

Hayes-Roth and Hayes-Roth (1979) studied the protocol provided by a human subject given a map of a town and a set of "errands" to plan to perform within a given time. They then attempted to explain this protocol within an artificial intelligence model. The subject's planning turned out not to be the type of neat-and-tidy start-to-finish or top-to-bottom arrangements that theoretical treatments of planning often suggest. Instead, the planner "jumped about", sometimes planning a later part of his journey before an earlier, or letting a low-level detail change a previous higher-level decision. This appears to contradict suggestions by Gärling's group (Gärling et al 1983, 1984) that a more-or-less uniform procedure is used in planning multi-destination journeys.

There is much fascinating material in the Hayes-Roths' article. However, there is an implicit assumption that all travel is planned - the plan produced is very detailed. The Hayes-Roths note that the subject greatly underestimated the time that various trips and activities would take - but this suggests he would not normally have planned in the sort of detail they required and if he had, it would have done him little good.

A strong argument against the ubiquity of planning in advance of wayfinding is that human wayfinders are sometimes aware of planning a journey in great detail, while at other times - particularly in familiar areas - we decide to go to a particular destination and simply set out, without being aware of any intervening process of planning. A few years ago, I was at point A on Fig.11, and wanted to walk to point B. Without being aware of any intervening problem-solving, I set off for the road-junction between Upper Lewes Road and Wakefield Road (C), which is on my normal route from what was then my home (H) to B. This intermediate destination was adopted as soon as it came to mind, without conscious consideration of its merits. In fact, going via points D and E would certainly have been quicker, and less effort, but did not occur to me even though I have gone to B from D on a previous occasion. This is one of the possible costs of unplanned wayfinding - habit may lead to an inferior route being chosen.

It could be claimed that even when we are not aware of planning, covert planning has nevertheless occurred. This is an unnecessary and implausible hypothesis. It is unnecessary because there are unplanned wayfinding methods that should not be too difficult for people to acquire and use, and that are superior to methods involving planning in many circumstances (as argued in sections 6-8). It is implausible because we appear to have considerable ability to report our own planning processes in relation to wayfinding. Since we are often able to report planning for wayfinding, it is clear that this is not a process of which we cannot become aware, unlike many of the stages of vision and language understanding, for example.

There are considerations which suggest that planning may sometimes be disadvantageous. Planning itself takes up time and effort, the expenditure of which may be unnecessary. If the journey you have to make is a thoroughly familiar one, you know which way to set out from where you are, and there is no reason to expect any difficulty in making the journey, why waste time and effort on planning? Also, planning may not produce a useful result: it may make the subsequent journey no easier, or make so little difference that the effort involved in planning is not worthwhile. It is often <u>easier</u> to make a journey without planning than to plan it in advance in any detail, because during actual travel the local environment provides cues not available during advance planning. These disadvantages will often outweigh the advantages of planning.

Because human travel is so widely assumed to be planned in advance, little attention has been paid to the possibilities of unplanned wayfinding. This section has argued that actual travel along a route may be possible although running through it mentally is not - both because the real world provides automatic sequencing of route-information during actual travel, and because recognition is easier than recall. It is shown in sections 6-8 that methods not involving planning <u>could</u> in principle be used to solve a wide range of wayfinding problems in path-networks, in both familiar and unfamiliar areas.

Some of the studies discussed here - particularly those of Chase, Passini, and the Hayes-Roths - suggest a need to distinguish what might be called strategic planning (advance planning of an entire journey or set of journeys) from tactical adjustment during travel, often requiring a degree of replanning in response to travel conditions in the parts of the network being traversed, or to locally available information reminding the wayfinder of possible alternative routes. Decisions about when and in how much detail to plan may themselves involve planning: for example, a driver may think:

> "When I see what the traffic is like toward the centre of town, I can decide whether to go round the outskirts - in which case I'll need to plan my route at that point."

This kind of "metaplanning" has been little studied, and is likely to be difficult to get to grips with, being two steps from the primary activity (travel) concerned. Its existence does, however, show clearly that we cannot assume that if an activity can be planned, then it always <u>is</u> planned - an assumption made at least by Gärling's group in the case of travel. Since planning can itself be planned, the assumption that activities that can be planned always are leads to an infinite regress of planning, planning to plan, planning to plan to plan, and so on.

6. Compiled knowledge.

This section outlines how a driver's wayfinding system could develop toward a set of highly-specific rules for selecting exits at particular junctions, without the need for pre-travel planning. It makes use of the concept of "compiled knowledge" (Neves and Anderson 1981, Chandrasekaran and Mittal 1982, Hart 1982).

"Compiled knowledge" is knowledge that is stored as a collection of autonomous rules or procedures, where each rule or procedure responds to a particular situation that may be encountered during problem-solving, and produces a specification of the action to be taken. Many members of the class of "route-choice" rules described in section 3 could function as compiled knowledge. Exactly which ones would do so depends on how strictly the definition above is interpreted.

Neves and Anderson (1981) are interested in human skill acquisition. They start from the observation that people get better at a task with practice. They focus on a geometry task: that of providing justifications for each line of a given geometric proof. Giving their subjects a set of geometric postulates to be used in constructing the justifications, they observed the improving performance of subjects faced with a sequence of geometric proofs to justify. Neves and Anderson propose a theory of skill acquisition to account for the gradual improvement people show on such tasks with practice.

Neves and Anderson's analysis of their subjects' improvement in performance identifies three stages in skill acquisition: encoding, in which a set of facts required by the skill (in this case, the postulates) are committed to memory; proceduralization, which produces a representation in the form of rules that directly specify steps in the task to be performed (compiled knowledge, as the term has been used here); and composition, which involves combining pairs of rules that are executed in sequence into single rules, thus reducing the number of rule-cycles a task demands. Neves and Anderson refer to proceduralization and composition together as "knowledge compilation". They usually call rules "productions".

These authors propose that all incoming knowledge is encoded declaratively, as a collection of facts. These facts are used by general interpretive procedures to guide behaviour. In the wayfinding domain, this occurs in a system such as Sugie's CARGuide (Sugie et al 1984), when Dijkstra's algorithm (Dijkstra 1959) is applied to place-specific facts about the network's topology and travel costs, to yield travel instructions. It is not clear why <u>all</u> incoming information should first be encoded declaratively - the first information relevant to wayfinding acquired in a particular city might often be, for example, a sequence of highly task-specific rules for making a single journey.

Neves suggests that the greatest benefit of declarative information is the flexibility with which it can be used; the major drawback is that its application is slow. He suggests that the time required to execute a procedure depends on the number of productions applied in performing it - but not on the total number of productions in memory. Proceduralization and composition both supposedly speed procedures up by reducing the number of production-cycles required.

According to Neves and Anderson, proceduralization cuts out the use of productions to retrieve factual information from long-term storage (envisaged as a "semantic net"); the productions that proceduralization gives rise to are independent of any other stored knowledge. Proceduralization supposedly occurs as a byproduct of the use of factual or declarative knowledge in problem-solving. As a result of the retrieval of factual knowledge by a general production, a more specialized version of the production can be created, including the retrieved information.

In the wayfinding domain, the equivalent process would be one of retrieving and using factual knowledge about a path-network, in order to work out which exit to take at the junction you have reached on a journey from one point to another; then storing the resulting direct representation of what to do at that junction for use on subsequent occasions - in other words, the construction and storage of where-next rules.

So far as composition is concerned, there is an important difference between the domain of geometric proofs and that of wayfinding. Neves and Anderson assume that composition, by reducing the number of production applications a task requires, will accelerate performance. In wayfinding, however, it may be that the disadvantage of the inflexibility composition brings about is not balanced by the advantage of an increase in speed: the speed of travel along well-known routes is not usually limited by the speed of route-choosing procedures. Combining rules or productions used to make choices at successive junctions is therefore not likely to speed the wayfinder.

Chase (Chase and Chi 1981, Chase 1983) distinguishes the use of "automatic procedures" from that of "inference rules" in route-choice. From his description of "automatic procedures", they appear to correspond to what I have called where-next rules, and to constitute compiled knowledge for wayfinding:

> "At choice points along a well-known route, perceptual features from the environment automatically retrieve the appropriate choice of route from the long-term memory knowledge base." (Chase 1983, p.394).

The sequence in which different types of information about a path-network are acquired has been found to depend on mode of travel (specifically, whether a person drives themselves around it or goes by bus (Appleyard 1969); on whether information is acquired from maps or travel (Thorndyke and Hayes-Roth 1982); and on whether real or simulated (filmed) travel is the source of information (Thorndyke and Goldin 1983).

As a person becomes familiar with an everyday travel environment, she is likely to spend less and less time studying maps of it, or making deliberate journeys of exploration. The only types of travel-related problem likely to be encountered frequently in such an area are the wayfinding and place-recognition problems tackled in the course of everyday travel. The main use of acquired information about the most familiar environments will therefore be in wayfinding and place recognition in the context of everyday travel, and information that is not useful for these purposes is likely to be accessed infrequently. Everyday travel and its requirements will also have a predominant effect on any continuing knowledge acquisition, to the extent that learning occurs in response to difficulties encountered in the course of problem-solving. The types and specific items of information most useful in everyday place recognition and wayfinding will therefore tend to be acquired, retained and used more than those of less or no such use. Much of the information about a familiar area that is known to the wayfinder may not be used in wayfinding within it - including some that was used in wayfinding when the area was less familiar. Appleyard (1969) reports the interesting finding that the sketch-maps of residents of Ciudad Guayana in Venezuela who had been in the city less than a year were more complex than those of 5-10 year residents - suggesting a winnowing of readily accessible information over time.

In a highly familiar area, where many journeys are made repeatedly, the wayfinder has both the incentive and the opportunity to produce compiled route-choice knowledge: that is, to acquire route-choice rules which can be used to guide travel directly, rather than via any process of planning, or of interpreting an internal or external representation of the travel network expressed in factual or diagrammatic terms. Having ready access to stored rules that tell you which way to go at each point in the journey will save the time and effort required to generate such a sequence of choices anew each time the journey is made.

7. Wayfinding in highly familiar path-networks.

7.1. Section introduction.

This section assumes that a human wayfinder in a highly familiar area will indeed have developed a set of route-choice rules for use in travel, and will rely almost entirely on these rules during travel, seldom needing to plan a route in advance or pause to work out which way to go. For convenience, this model of human wayfinding in familiar areas is referred to as the "Rule-Follower".

As a human wayfinder begins to learn her way around a path-network, her ability to plan routes will generally increase as information is acquired. However, this ability may well plateau before the wayfinder is able to plan any required route in detail - that is, to recall all the junctions where decisions must be made, and the exit selected at each of these junctions. The set of route-choice rules stored for a given problem may be sufficient to support actual travel (so the wayfinder will feel no need to alter it) without being able to support <u>imagined</u> travel. Once a path-network is highly familiar, little planning will be needed, and the ability to plan routes in the area may decline through disuse.

Certainly, a model of a human wayfinder as a non-planning Rule-Follower cannot explain all human wayfinding skills: after all, people can make use of maps to find routes, and can sometimes imagine journeys in considerable detail. However, a Rule-Follower model is taken here as the simplest plausible type of system that could be the <u>basis</u> of this skill in human beings.

The rest of the section describes a type of Rule-Follower wayfinding system that could model the sort of system human wayfinders are likely to approximate to as they approach complete familiarity with a path-network. Subsections 7.2 and 7.3 discuss the kinds of rules such a wayfinding system might make most use of, while 7.4 considers how some of the hypotheses presented about the wayfinding methods used in familiar areas could be tested.

7.2. Junction-specific exit-selection rules.

These rules, which will sometimes call simply "junction-specific rules" or "exit-selection rules" when the context allows, are the core of the proposed Rule-Follower model. A junction-specific exit-selection rule instructs the wayfinder to take a particular exit from a particular junction, given certain conditions. To give an example of such a rule, expressed in English:

> "If your destination is Preston Circus, you are at The Clock Tower, and the Dyke Road exit has been identified, take that exit."

Three types of condition occur in this rule, and are discussed here: conditions on the destination the wayfinder is seeking; conditions demanding that some place or object should currently be visible, and have been recognized; and conditions on the wayfinder's current location. A real wayfinder might have exit-selection rules with other types of condition - specifying the local traffic conditions, or the time or resources that must be available to the wayfinder, for example.

So far as conditions on the wayfinder's destination are concerned, such a condition might specify one particular destination; a list of individually specified destinations; or any destination <u>other than</u> one or more specified exceptions. Fig.12 illustrates how each type of destination condition might be used. The wayfinder is pictured at a junction (J) with three exits. The wayfinder should take E1 if the destination is A, E2 if it is one of B, C and D; and E3 if it is any other node.

The other two types of condition are not independent: recognizing a place always conveys some information about location relative to the place recognized. However, it is useful to distinguish the two types of condition. First, knowledge of location has sources other than what is currently visible and recognized - if I walk around an unfamiliar building and then go inside, I know I am inside the building I walked around, even if I recognize nothing. Second, place recognition conditions generally have a second function in exit-selection rules: ensuring that identification of the rule's target exit will be possible.

What a rule needs in the way of conditions of these two types is bound up with the way that its instruction's target is specified. A distinction was made in section 2 between two ways in which the exit to be taken can be specified: egocentrically and allocentrically. An egocentric exit specification uses the current position and heading of the wayfinder as a starting point in specifying the instruction's target. Two examples should make this clearer. The instructions: "Take the exit to your left", and "Take the narrower of the two exits to the right" are both egocentric: the referents of the (underlined) exit specifications depend on the location of the wayfinder. Allocentric exit specifications are simply those that are not egocentric. They either specify the exit directly (in verbal directions, this might be done by naming or describing it); or use as a starting point some fixed feature of the environment other than the target, as in: "Take the exit directly opposite the 'Blue Boar Inn'".

Egocentric target specification tends to demand greater specificity in the associated conditions on location than allocentric alternatives. For example, if you remember that to get from A to C you must turn left at B, this piece of information may not be remembered, or may be used wrongly, when you approach B from another direction, but still wanting to get to C. If the original information had been encoded in terms of the exit by which to leave B, without reference to the turn involved, no such problem would arise. If an exit-selection rule's instruction specifies its exit by reference to the turn the wayfinder habitually makes in order to take that exit, it is thereby useful in a narrower range of circumstances than one that specifies the exit allocentrically.

Egocentric target specification is unlikely to occur much in exit-selection rules used in familiar buildings or neighbourhoods, because in such areas place recognition is easy (most junctions and exits encountered are normally recognized without deliberate effort), while it is common for pairs of routes that start in different places to come together at some junction along the way to a common destination - the circumstance in which allocentric exit specification has an advantage.

Egocentric exit specifications are most likely to remain in use in familiar areas when the use of a particular exit from a junction always follows an approach to that junction from the same direction, so that the egocentric specification can always be used when that exit is required.

If target exits in a completely familiar path-network are all specified allocentrically (and assuming that the number of possible destinations that can be listed in the conditions of a rule is unlimited), then a wayfinding system could encode all the knowledge needed to make the right choice at a given junction in no more than one exit-selection rule for each exit it uses. Consider a single junction in such a network. For each exit, either the wayfinder never needs to use it, or there is some set of destinations for which it is the best exit to take. (The best exit for a destination may be contingent on factors such as traffic conditions. This would not increase the required number of exit-selection rules, only the complexity of their conditions.) The rule for each exit would simply list the destinations for which it is to be taken under each possible combination of other factors; if an exit is never used, of course, no rule is required. Functionally, the rules used at a junction would correspond to a set of signposts, one pointing to each of the exits that is ever to be taken.

A Rule-Follower wayfinder wholly dependent on junction-specific exit-selection rules would be very vulnerable to the loss or corruption of single rules, or of the information necessary for place recognition, or to changes in the travel environment. In the absence of a rule covering its current location and destination, it would have no means of deciding which exit to take. Ways of providing for the "graceful degradation" or robustness of a wayfinding system's performance in the face of limitations on the data or resources available are discussed further below, but two initial and very simple possibilities are mentioned here.

The first of these is the possession of more than one junction-specific exit-selection rule suitable for a particular destination, current location, and set of traffic conditions. Even in completely familiar path-networks, travel does not necessarily become stereotyped: alternative routes from one place to another may remain in use. Each may sometimes be easier or more pleasant to take than the alternative(s), or there may be nothing to choose between them in functional terms.

The second possible way of giving the Rule-Follower described greater robustness is to provide it with an additional <u>non-place-specific</u> exit-selection rule, not to be used if any junction-specific rule is available:

> "Take an exit chosen at random".

Given this "arbitrary exit-selection rule", the wayfinder would perform a random walk in parts of the network where either its recognition knowledge or specific exit-selection rules were inadequate, until it reached a point where one of its specific rules could be applied. This would allow the wayfinder to continue using the recognition knowledge and exit-selection rules it does have.

Of course, the measures proposed above would not completely save the Rule-Follower from the consequences of its rules becoming lost or corrupted. As it lost more and more rules, an increasing proportion of its time would be spent wandering around the path-network until it happened upon the destination, or the beginning of an intact sequence of rules leading to it. Worse still, certain forms of rule-corruption could lead to the wayfinder getting trapped in a loop. Suppose that the correct way from A to Z is via the junctions B, C, and D. If the rule used at C got altered so that it directed the Rule-Follower back to A, it would cycle around A, B and C indefinitely - because it does not maintain any form of record of the current journey. The uses of such a record ("episodic knowledge") to avoid getting trapped in this way, and to undertake systematic searches, is beyond the scope of this paper.

7.3. The default path-following rule.

This section discusses a way to make use of the path-continuity phenomenon described in section 2. The suggestion made here is that a Rule-Follower wayfinder depending mainly on junction-specific exit-selection rules could also have a non-junction-specific exit-selection rule saying:

"If it is possible to do so, follow the path".

This rule will be called the "default path-following rule". It should be used in preference to the arbitrary exit-selection rule, but should not be used if the conditions of any junction-specific rule are met. As noted in section 2, it is not always possible to follow the path through a junction you are approaching - for example, when approaching a T-junction or fork along the "stem" - and in such circumstances this rule will not be used. On the other hand, it is frequently possible to "follow the path" even if the wayfinder recognizes neither the junction nor any of its exits. This is the key to its effort-saving properties.

Consider again the path-network shown in Fig.7. If a Rule-Follower wayfinder is restricted to the use of rules like those described in 7.2, the journey indicated by arrows will involve the use of 8 junction-specific exit-selection rules: one at the starting-point, and one at each junction along the route shown (including the two-exit "dog-leg" junction where the first change of direction occurs). At all these intermediate junctions except B, however, the wayfinder follows the path when taking this route.

Suppose that the wayfinder always wants to follow the path through at least some of these junctions when they are encountered. For example, if the wayfinder never needs to visit anywhere along the path sloping up to the right from A, then the path should always be followed through A (either as shown or in the opposite direction). If the Rule-Follower wayfinder had a version of the default path-following rule, then it would not need a junction-specific exit-selection rule for junction A, nor even need to recognize the junction, or the exit taken, when passing through.

The sort of occurrence that would be expected if human wayfinders used a default path-following rule is very common in travel by car: driving along most roads in or near towns means passing through numerous junctions with other roads, most of which even regular users of the road may be unable to enumerate or recognize if they never use them. A less than complete familiarity with the network may therefore make no difference to the wayfinder's ability to reach the destinations she uses.

If a junction-specific rule fails to fire at the appropriate time (e.g, because either the junction or the exit goes unrecognized), then the default path-following rule can lead the wayfinder astray. Like all default systems, it can lead to inappropriate action if it is not overruled when it should be. I have often missed a turning from a major road into a minor one when not paying attention to where I am going, sometimes in very familiar areas. This sort of error would certainly be expected if human wayfinders use a default path-following rule.

If we assume that human wayfinders do tend toward the sort of Rule-Follower wayfinding system described here, what would be the effect of failure to retrieve a junction-specific exit-selection rule and consequent use of the default path-following rule? When such a junction-specific rule failure leads to the wrong exit being taken, the human wayfinder is likely to take care to remove its cause by re-learning the appropriate junction-specific rules, and refamiliarizing herself with the junction and exits concerned. Where the default rule selects the right exit, this re-learning will not happen. Junction-specific exit-selection rules that are acquired early in the course of becoming familiar with an area, but later fail to respond when they should without adverse consequences, may fall into disuse and become difficult to access. The default path-following rule may thus contribute to a winnowing of the information the wayfinder has acquired.

7.4. Putting the Rule-Follower model to the test.

The Rule-Follower model of human wayfinding in familiar path-networks could certainly be criticized with justice if put forward as the whole truth rather than an initial approximation to the truth. Nevertheless, I maintain that it makes a good starting point for empirical investigation, and may be close to the truth in many instances (i.e. for many wayfinder/path-network pairs). Its most serious weakness is probably the lack of an episodic memory. I shall now show that it can suggest specific questions for empirical investigation, and in some cases generate predictions of the results.

The Rule-Follower model leads us to expect some very general features of human wayfinding performance in familiar path-networks. The first of these has been a consistent theme: the amount of planning human wayfinders do will decline as a path-network becomes more familiar, ending up at a negligible level in the most familiar. One way of testing this hypothesis is to train a subject to different levels of wayfinding performance in similar path-networks, then to give her varying amounts of warning of journeys to be made within these path-networks. If, for a given subject and path-network, longer warning periods led to improved performance, this would indicate that planning was being undertaken during the warning period, and was affecting performance. The absence of such an improvement would not in itself show an absence of planning - it might have occurred without improving performance - so an attention-demanding distracter task should be given. Rewards for performance on the two tasks could be adjusted to discover how much faith the subjects put in advance planning to improve their subsequent performance under various conditions.

Because knowledge sufficient to enable a wayfinder to follow a route on the ground may be inadequate for mentally simulating the journey, it should be possible to produce subjects with near-optimal performance in actual travel through a path-network, but very different levels of competence in planning journeys, and related tasks such as giving travel directions. In fact, it should be possible to produce pairs of subjects of whom one excels during actual travel and the other in giving directions, simply by training the two differently: one on actual travel problems alone, the other on a mixture of these and problems requiring the subject to give directions to another person. A single individual's wayfinding performance on two path-networks (with similar training) should in many cases become more similar as they become more familiar. If the two networks have similar general appearance at a local level; equally distinctive junctions, exits, and paths; the same average number of exits per junction; and the same average number of paths in an optimal route containing a given number of links, then performance levels should grow increasingly similar - even if one network is a rectangular grid, and the other has no clear higher-level structure, or a misleading one. A theory of wayfinding in familiar path-networks relying on the continued use of information about the network's overall structure would predict continuing differences due to such higher-level structural differences.

This hypothesis, and others suggested below, might be tested experimentally, using the following general method.

Subjects could be trained to a high level of familiarity with one or more networks, and ability to travel around them; then be presented with a sequence of photographs showing views of a junction together with the name of a destination to be reached, and be asked to indicate the exit to be taken. Differences in error-rates (likely to be very low) and in response-times could be taken as a measure of the difficulty of recalling or constructing the information necessary to enable the wayfinder to choose the correct exit.

Nothing in the discussion here depends on whether the experimental environments are real or simulated. A major problem with empirical wayfinding research is the need to take subjects on trips through large-scale environments. This requires a lot of time, and reliable access to appropriate travel environments; the degree of control the experimenter can exert over those environments is generally limited. There is considerable promise in a novel approach: the use of computer-simulated travel environments to study route choice (Bonsall 1992).

Simulated environments allow a complete record to be kept of the subject's "travel" behaviour. This can be compared with other subjects, previous trips, and the results of running computer models of human wayfinding on the same simulated environments. Wayfinding problems could be embedded within a variety of computer games, allowing the introduction of distracter tasks where useful. Controlled changes in the simulated environment - for example, changes in signposting provision - can be tested for their effects.

The main methodological problem with computer-simulated travel experiments is that their validity, in relation to real travel, could be challenged: how can the investigator be sure that effects found in simulated environments will "scale-up" to the real world? Meeting this objection would necessitate the design of experiments that could be carried out in either real or simulated environments, with appropriate statistical analysis of the results. The more results prove to transfer from one to the other, the more confidence we can have in the simulation where direct comparisons are impossible. It should also be possible to test systematically which aspects of the real world it is most important to simulate correctly, by doing the same experiment on a range of simulators as well as a real travel environment.

A further central feature of the Rule-Follower model that could be tested using the approach outlined here is the predicted switch from egocentric to allocentric specification of target exits at a particular junction as the wayfinder acquires routes that <u>merge</u> at that junction (i.e. that use different links to reach the junction, but involve leaving by the same exit). Suppose subjects are trained extensively on a route that takes them into junction A (which must have at least four exits) via path-segment S1, and out via S2 to the destination B, and are not allowed to go to or through A in any other way (Fig.13a). This should maximize the chances of the exit-specification for S2 being encoded egocentrically. If the subjects are then split into two groups, and the members of group I are guided along a route that enters A via S3 but still leaves via S2 (Fig.13b), while members of group II are taken on a route that enters A via S3 and leaves by S1 (Fig. 13c), then members of group I are more likely to recode the exit-specification for S2 allocentrically than members of group II - if the arguments of 7.2 are correct. If members of both groups are then presented with a view of A as it appears when entered along yet another path-segment, S4, along with an instruction to get to B, the members of group I should show a marked advantage in their speed of response. A negative result would show that my account of the relationship between allocentric and egocentric exit-specification is wrong.

<u>8. Using the structure and properties of individual path-networks.</u>

<u>8.1. Section introduction</u>.

When an area is only partly familiar, the wayfinder's route-choice system may be considerably more complex and diverse in its problem-solving methods than that of a seasoned traveller in the same area. Subjectively difficult problems will occur more frequently, and a careful marshalling of fragments of information about the travel environment may be necessary to optimise route-choices. In these circumstances, many secondary structural characteristics of particular path-networks may be brought into use.

In 7.2, it was noted that there might be where-next rules with conditions specifying one destination, a set of destinations, or any destination but one of a limited set. There are additional possibilities: a where-next rule could specify that the wayfinder's destination must lie within or outside some specified region; or must meet some other condition concerning its location. Another important group of route-choice rules, intermediate destination rules, were not considered in section 7. Much of this section concerns these classes of rules, and certain aspects of the structure of path-networks to which they are related.

8.2. Network regions.

The most important of these aspects of path-network structure is the hierarchical: the part/whole relationships between places.

The hierarchical structure of travel environments means that either the number of junction-specific exit-selection rules a Rule-Follower Wayfinder requires, or the complexity of the rules' conditions on the wayfinder's destination, or both, can be reduced by the acquisition of appropriate route-choice rules of other types. These may be either where-next rules like:

> "If you are at the Brighton Clock Tower, and your destination is anywhere in Hove, take the Western Road exit",

or intermediate destination rules such as:

> "If you are in Brighton, and your destination is anywhere in Lewes, adopt the crossroads at Lewes prison as intermediate destination."

The first of these rules can be substituted for a collection of more place-specialized rules (one for each destination in Hove that the wayfinder may want to visit), and could enable the wayfinder to begin the trip from the Clock Tower to any Hove destination without planning, even if the particular journey has never been made before.

There are costs to the use of such less place-specialized rules. Given a specific destination, using the where-next rule requires the extra effort of testing whether that destination is in Hove; once this has been tested for a particular destination, it may be worth storing a more specialized (compiled) rule concerned with that destination alone.

The intermediate destination specified by an intermediate destination rule may be a place on the way to a final destination, as in the example above; or it may be a part of the path-network - a path or region - that contains the final destination.

A network region is a connected part of a path-network, other than a path, that the wayfinder treats as a unit for some purposes. Such regions must be distinguished from the mosaic regions of plane path-networks: a mosaic region includes everything contained within a cycle of links in a plane network, while a network region includes only a part of the network itself, excluding any non-network area it may surround. Network regions can be defined in any path-network, and one such region may be a part of another. The division of a path-network into network regions is a subjective matter, but people's decisions about the structure to assign to path-networks are strongly influenced by its objective physical and functional properties. Here, I want to consider particularly the influence that the path-network's node/link topology, and its perceived path structure, have on its network region structure.

Consider a path-network P1, and any connected part of that path - network, P2 (Fig.14). P2 is a set of nodes, links, and/or parts of links, such that any member of the set can be reached from any other without going outside P2. There will be a set of "boundary points", either at nodes or partway along links, where P2 meets the rest of P1 (i.e. the parts of P1 not in P2, which I shall call "P1-P2", and which may or may not be connected).

A network region as defined here must be a connected part of a path-network. Whether a connected part of a path-network is likely to be regarded as a network region by a wayfinder depends in part on the travel habits of the wayfinder, but also on the set of boundary-points that form its interface with the rest of the network. If P2 has only a small number of boundary-points with P1-P2, and both P2 and P1-P2 contain a number of the wayfinder's habitual destinations, then P2 is likely to be treated as a network region. The boundary-points will have a vital role in travel in the network: any journey beginning in P2 and ending outside it, or vice versa, must pass through one of these points. The fewer boundary points there are, and the more they are used by the wayfinder, the more likely P2 is to be thought of by the wayfinder as a significant structural element of P1 - that is, as a network region. In the extreme case, P2 meets P1-P2 at just one node, or partway along a single link. Either the boundary-point, or the network region itself, may reasonably be adopted as an intermediate destination when the starting point is within P2 and the destination outside it.

P2 is also likely to be treated as a network region if all its boundary-points lie on a small number of paths, particularly if these are major paths. If they all lie on a single path, then any journey into or out of P2 must cross that path (Fig.15) which therefore becomes a suitable intermediate destination for such journeys. Frequently, particularly in plane path-networks, the boundary points of the larger network regions (as perceived by many wayfinders) will lie along major paths, while those of smaller regions lie on minor paths.

The division of a path-network into either network or mosaic regions provides what can be called a spatial reference system, in terms of which location can be specified. The simplest form of spatial reference system is a straightforward dichotomy between locations within and outside some particular region. More complex systems can be produced by dividing an area into a set of minimally overlapping regions (sharing only their boundaries), that together cover the whole. For example, one can specify the location of a place in London by specifying the borough in which it lies. However, such patchwork systems often show no regularity in the way the elements of the patchwork relate to each other or the space they divide up; they may have very different shapes, sizes, and numbers and arrangements of neighbours.

Systematic or regular ways of dividing up space produce reference systems that are easier to comprehend and use. Many of the reference systems that we use in defining regions and specifying location make use of one of three elementary ways of dividing up a plane, which may be called "banded", "concentric", and "sectoral". These are diagrammed in Fig.16. A "banded" division divides a two-dimensional space into a series of bands, running side-by-side. A "concentric" division begins with a central region and divides such a space into a succession of rings, while a "sectoral" division cuts it into pie-slices that meet at a central point. All three types of spatial division tend to be associated with a pair of opposite "reference directions". These are shown by arrows in the figure.

These ways of dividing space are often useful in path-networks: banded divisions occur where a number of parallel paths divide a region into slices, allowing wayfinders to keep track of their location relative to this reference system by taking note whenever they cross one of these paths; the exits from a major junction often form a good basis for a division into sectoral regions. Concentric divisions of path-networks are less common, although there are, for example, towns with inner and outer ring roads. Although derived from ways of dividing up a plane, these ways of dividing space may be usefully applied to non-plane path-networks, so long as their path-structure produces a clear sequence of bands, sectors or rings.

In addition to place-specific route-choice rules such as those suggested above, part-whole relationships underlie a number of <u>non-place-specific</u> route-choice rules that can be used in conjunction with stored factual information to generate suitable intermediate destinations. For example, a wayfinder might use a non-place-specific rule such as the following:

> "If your destination lies within a region you are currently outside, and there is only one entrance to that region, adopt that entrance as an intermediate destination".

Where these conditions are met, there is no danger that following the rule without using planning techniques will lead the wayfinder into error. Less discriminating rules such as:

> "If your destination lies within a region you are currently outside, adopt that region as an intermediate destination",

could also be used without planning, but might lead to the choice of a sub-optimal route. Suppose, for example, that the region in question is the central part of a town, surrounded by a ring road, and that the situation is similar to that shown in Fig.17, in which the wayfinder is at A and wants to reach F. The quickest way into the region surrounded by the circular "ring road" (B,C,D,I,H,G) is evidently via B and E, and if the region has been adopted as intermediate destination this route will presumably be chosen if the wayfinder knows of it. Unless travel within the region is considerably faster (or less costly in some other way) than around the edge, however, a route to F via D would be preferable.

When you are learning your way around an area, it makes sense to store information acquired in the course of solving a given problem both in a specialized form that can be immediately applied if the same problem recurs, and in a more general form that will allow it to be used in solving other problems to which it may be relevant. For example, if you learn that in order to get from A to B, it is necessary to get into and remain in region R, this information could be encoded as a place-specialized intermediate destination rule, to be used when you wish to get from A to B in future. However, it is also worth storing the simple fact that B is part of R. Such a fact may, in conjunction with non-place-specific rules such as those above, help the wayfinder solve a wide range of problems concerning B, if the specialized rules required are not available: R becomes a possible intermediate destination if the wayfinder wants to get to B from anywhere outside R.

Before leaving the topic of part/whole relationships, one more point is worth noting: different parts of a path-network may repeat the same structure and layout. Examples commonly arise in large housing estates, and in cities with a regular grid of streets. Such symmetries can be useful if a wayfinder knows about them, as they can be used to predict what relatively unfamiliar parts of the network are like from what is known of the more familiar parts. Moreover, the existence of any such symmetry will tend to make the layout of the network easier to comprehend and remember. On the other hand, similarities of any sort between parts of a path-network may lead to confusions about where in the network the wayfinder is.

8.3. Path-network wayfinding and the local environment.

The relationship between the wayfinder's local environment and the destination of her journey can be used to classify wayfinding problems into three categories. In "local" wayfinding problems the destination is visible from the starting point, and so is a strip of territory covering the direct line from starting point to destination and containing at least one complete route between the two. In "destination-hidden" wayfinding, the destination is out of sight from the starting point (for a blind person, or in pitch darkness, all wayfinding problems are "destination-hidden"). In the intermediate case of "destination-visible" wayfinding, the destination is visible, but at least part of the area relevant to the choice of route is not: a typical example arises when the wayfinder is given the problem of reaching a tall building from a point whence its top is visible over the tops of nearer buildings on intervening streets, but its base is concealed. The destination-hidden and destination-visible cases can be grouped as "non-local" wayfinding.

Consider Fig.18, which shows a path-network layout. Three different path-networks with this same layout are to be compared. The problem in all three cases is to get from point S to point D, and movement is in all three cases confined to the network.

(a) Local wayfinding.

On a flat plain, you must pick your way along narrow, unfenced, but clearly marked paths running between areas of impassable terrain to reach a tall pillar (D) that is visible 100m away.

(b) Destination-visible wayfinding.

The destination is the same as in (a), but the paths run between 3m high brick walls, so the layout of the network cannot be seen. The pillar is tall enough to be visible above the walls.

(c) Destination-hidden wayfinding.

As for (b), except that the paths have been roofed over (and provided with artificial light), so that the pillar cannot be seen except when visible along a path.

Case (a) should be easy, while (b) and (c) are increasingly difficult, because a successively smaller proportion of the information relevant to the problem is available from the local environment.

In local wayfinding, there is no need to depend on the reliability of stored information to tell you which way to go. Provided the destination itself is readily recognizable, a local wayfinding problem can be solved using only locally available information. Because this is available during problem-solving, it is <u>relatively</u> unimportant how familiar the problem- solver is with the ground to be covered. The incentive to learn about the layout of a path-network is considerably less if most of the possible journeys within it are local; on the other hand, the <u>difficulty</u> of learning about the layout is reduced.

The difficulty of destination-hidden wayfinding compared to local wayfinding is clear. Unless stored non-local information is available to the wayfinder, such problems can only be solved by extensive search. If a mistake occurs during travel, it is more likely that the wayfinder will need to retrace her steps than in local wayfinding: in destination-hidden wayfinding, this may be necessary in order to get the wayfinder back to a place from which the way to the destination is known; in local wayfinding both the destination and the way to it will still be in sight.

Destination-visible wayfinding is intermediate in difficulty between the local and destination-hidden varieties. Destination-visible problems can be tackled without stored knowledge other than that needed to recognize the destination, but not with great confidence: the wayfinder has a reasonable basis for selecting which way to set out an which way to go at subsequently encountered junctions (taking the path that runs most directly towards the destination), but no surety that this path will reach the destination, unless stored information to that effect is available.

Knowing whether one place is "local" with respect to another (i.e., is in its local environment) is often in itself of considerable use in wayfinding: if a wayfinder reaches a location from which she knows her destination is normally visible, she can scan the local environment in order to find it and make sure she is going in the right direction.

Landmarks - which can be defined as objects that are part of the local environment over an unusually wide area - can be selected as intermediate destinations, and can also help the wayfinder keep track of her position, and the direction in which she is travelling. A city with a scattering of unusually tall buildings should be a relatively easy environment for wayfinding, compared to one where there are no buildings that are markedly taller than their neighbours.

Another type of circumstance in which directions of travel are particularly easy to keep track of occurs when a distant landmark, located outside the network (such as a mountain, or the sun), is visible at more or less the same bearing from any point within a wide area. This bearing can then be used as a standard direction.

It is interesting to compare separately the differences in the difficulty of planning, on the one hand, and of unplanned travel, on the other, according to whether the wayfinding problem concerned is local, destination-visible, or destination-hidden.

In local wayfinding, the fact that possible routes to the destination can be explored visually means that planning the journey to the destination in advance is much easier than in destination-hidden or even destination-visible wayfinding. The wayfinder can check possible ways to the destination without either recalling information from memory, or undertaking any actual travel. Comparatively little useful new information is likely to become available during the journey so there is less advantage to delaying decisions about later parts of the journey until they must be made, than in non-local wayfinding. Successful planning of a route is possible without any prior knowledge of the area.

In local wayfinding problems, the shortest possible course to the destination - that is, a straight line toward it from the current position - can readily be perceived. This line can be used as the basis of a solution, modified as necessary to go around obstacles. One possible strategy is just to undertake a depth-first visual search for a way to the destination, always looking first at those exits that keep the wayfinder nearest the direct line to the destination from the current position. However, more complex strategies are also available: a planned solution to a local wayfinding problem can be constructed in any order, working forward or backward as convenient. One possible approach is to pick out the most formidable obstacle lying across the direct line to the destination, check whether it extends further to the left or to the right of that line, and investigate the other end of the obstacle as a possible intermediate destination - checking both how to reach it, and how to get to the final destination from it. Further possible intermediate destinations could be chosen in a recursive fashion, if necessary.

For non-local wayfinding problems, planning a journey in any detail (without a map) is almost always more difficult than making that journey without prior planning - because of the information made available by the local environment when the journey is actually made. If you want to "run through" a route mentally from beginning to end, you must recall all the points at which decisions must be made, in the correct order. It is not necessary to be able to do this in order to make the journey in actual travel - the wayfinder must only be able to recognise each choice-point as it approaches, and make the correct decision.

So far as planning is concerned, the differences between destination-hidden and destination-visible problems are slight relative to the differences between local and non-local problems. The direction of the destination is perceptually available in destination-visible problems, and an approximate judgement of its straight-line distance will be possible, but these pieces of information in themselves do not allow the wayfinder to plan a route.

An <u>unplanned</u> wayfinding approach changes the situation. Here, the largest difference appears to lie between destination-hidden problems on the one hand, and the two types in which the destination is visible on the other. In the latter cases, the wayfinder has available a criterion for exit choice during unplanned wayfinding: taking the exit pointing most directly toward the destination. Although this is not an infallible strategy, it is better than choosing exits arbitrarily. Furthermore, the visibility of the destination provides the wayfinder with a means of checking that she has not gone badly wrong: if the destination is known to remain visible throughout the journey from a given starting point, then its disappearance behind something else implies that a mistake has been made. The destination-visible and local cases are indistinguishable if no planning is done.

The likely effect of these differences is to make planning much more common for local than for destination-hidden wayfinding (because it is so much easier), and more common for the destination-hidden than for the destination-visible variety - because the local environment in destination-visible wayfinding makes available an exit-selection criterion for unplanned wayfinding that is not available in the destination-hidden case.

8.4. Using special path-network properties.

There is a well-known method of solving certain classes of wayfinding problems in bidirectional plane path-networks (or two-dimensional mazes): always keep your hand on the right-hand (or left-hand) wall - or equivalently, always take the rightmost (or leftmost) exit at a junction. This method works if the maze is one with an entrance and an exit, or if it is one with a central goal, and no loop, or cycle of path-segments, surrounding that goal. It would also permit a wayfinder to make an exhaustive search of any bidirectional plane path-network that is also a tree. The possibilities and limitations of this maze- solving method can be explained by thinking of a plane path-network that is not a tree as dividing the plane in which it is embedded into a collection of minimal mosaic regions, as described in section 2.

A mosaic region, minimal or otherwise, can be defined by the sequence of nodes encountered in a tour of the region boundary. Fig.19 has the minimal mosaic regions CDFH, DEPONMHF, and MNOPR. Notice that there can also be junctions that lie within a minimal mosaic region - such as junctions G and L in Fig.19.

A path-network that lies in a plane has a perimeter: those nodes and path-segments that are not surrounded by any loop or cycle of links belonging to the network. (If the path-network is a tree, or if none of its loops or cycles surround any part of the network, the whole network is the perimeter.) In Fig.19, the perimeter is the shaded area. The "right-hand" algorithm's properties can be best understood if the area outside the perimeter is considered as one more minimal mosaic region. (If the network were laid out on the surface of a sphere, this region would be a bounded part of the surface undivided by path-segments, just like all the others.) The effect of starting at an arbitrary point between nodes, setting out in either direction, and then following a "rightmost exit" (or "leftmost exit") exit-selection rule at each node encountered, is to make a tour of the network region corresponding to a single minimal mosaic region - i.e, going through all nodes and path-segments of the network that are either within the mosaic region or on its boundary. For a starting point on the boundary between two minimal mosaic regions (like point X in Fig.19), reversing direction or changing from leftmost to rightmost exit will alter which of the two regions is toured. For a starting point like Y, which is within a minimal mosaic region, such changes just alter the direction in which the tour is made. Fig.20 shows the four possible tours in the network shown in Fig.19 - unarrowed, as they may be made in either direction.

If, therefore, you know that your destination is in, or on the boundary of, the same minimal mosaic region as your current location, you can reach the destination by a "leftmost" or "rightmost" tour.

If all the path-segments in a path-network are <u>straight</u>, it can be called a <u>rectilinear</u> path-network (e.g., Fig. 21 - but a rectilinear network need not be a plane network). The main advantage of rectilinear networks compared to non-rectilinear ones is the greater ease of keeping track of the direction of travel in such networks: mentally integrating the change due to curves in the route followed is far more difficult than keeping track of a sequence of discrete changes, although this would not be easy for a network like that of Fig.21, with links running in many different directions and few long straight paths.

If a rectilinear network is also a plane network and contains a number of long straight paths, like that shown in Fig.22, these can be very useful as the basis for dividing the path-network into mosaic regions, and subjecting it to a systematic search. Two completely straight paths, of course, cannot meet in more than one junction: if they do not meet (like P1 and P2) they divide the path-network into three "bands", while if they cross (like P2 and P3), they divide it into four sectors.

Fig.23 shows a <u>rectangular</u> path-network, in which all paths run in one of two orientations, which are perpendicular to each other. Because there are only four possible directions of travel in a rectangular network, it is both particularly easy and particularly useful to keep track of your changes of direction. This makes worthwhile some unplanned wayfinding techniques employing episodic knowledge about the current journey, but not dependent on recognizing places visited before.

One simple but useful technique is simply to avoid turning in the same direction at four successive junctions. A sequence of four turns in the same direction in a rectangular network is likely to take the traveller along a path-segment she has already been along (Fig.24a), although not certain to do so (Fig.24b).

Unless you have just turned through 180 degrees and retraced your steps, you must have moved in all four directions in a rectangular network in order to be back at a point previously visited - a fact that can allow the wayfinder to avoid misrecognition. It can also be used in carrying out a search of such a network: so long as no U-turns are made, and movement in one of the four directions is avoided, the wayfinder can be certain that the links she is on have not been previously visited during the search.

Fig.25 shows a path-network of a type that I call a "grid". A grid consists of two ordered sets or "bundles" of paths, each consisting of a continuous sequence of alternating nodes and links. Every junction in the grid is formed by the intersection of two paths, one from each bundle. It is not necessary for every member of one bundle of paths to intersect every member of the other. Grids for which this is the case I shall call "complete grids".

There are a number of advantages to grid networks from the point of view of the wayfinder. First, the structure of a grid network is particularly easy to comprehend and remember: if the two orderings of paths can be committed to memory, the sequence of junctions along any one path should be easy to recall. Second, grids can easily be provided with a systematic naming system, as in the grid-based cities of North America with their parallel bundles of numbered streets and avenues, cutting across each other and providing a way of specifying any junction (by the paths that meet there) and any part of a path (by the junctions it lies between).

Third, the techniques described above for rectangular path--networks also work for grids. Indeed, the technique of avoiding four successive turns in the same direction is useful in a rectangular network precisely because such networks frequently include what might be called "minimal grids": cycles of four path-segments, like that which the traveller is shown traversing in Fig.24a. making four such successive turns in a grid will always bring you back to the path-segment you started from. Also, as with a rectangular network, a journey that returns to its starting point must either have been a trip up and down a single path, or must have included moves in all four directions.

The fourth advantage of a grid from a wayfinder's viewpoint is that a systematic search through all the nodes in a grid is comparatively easy. The wayfinder can check each member of one of the two path-bundles in turn from end to end, moving a single link along an intersecting path to get to the next member of the bundle. As Fig.26 shows, this avoids the need to traverse some links at all, while most of the rest are traversed only once; in a complete grid (Fig.27) every node can be visited without covering any link twice.

Kuipers (1977, pp.80-81) notes that people think of considerably distorted grids of paths as consisting of two sets of parallel paths, and finds it paradoxical that this inaccurate representation of the world is more useful than an accurate one would be. Probably, grids are assumed to be metrically regular unless anything is known to the contrary, and it is the relative unimportance of accurate metrical information that allows such cognitive "distortions" to survive uncorrected. However, the assumption can certainly lead to inefficient travel: in Fig.28, the route suggested by the grid structure is certainly not the shortest, and for travel by foot will not be the quickest, unless there are considerable differences in the ease of travel along different paths.

9. Conclusions.

This paper has indicated some of the complexities that arise when we try to build theories of human wayfinding, and of the route-choice behaviour of road-users. Yet it has only scratched the surface even of the purely cognitive factors in route-choice. To fully understand route-choice, we need to take into account attitudinal factors as well: two people with the same information may still make different route-choices because one trusts the signposting system more than the other, or attaches greater importance to avoiding congestion and less to arriving quickly.

Replacing the mythical driver with perfect knowledge of the network and a simple criterion for choosing routes with something more psychologically realistic in models of route-choice is not going to be easy. Yet it must be attempted: a community of road-users will be highly variable in the amount and types of information they possess about the travel environment, and the ways they process it. Human beings are opportunistic in their problems-solving, and real-world path-networks offer complex challenges and multiple resources to the wayfinder. Both the nature of a particular city's road-network, and the mix of different levels of familiarity with that network among road-users, will affect the public knowledge-base and travel habits into which network alterations or new travel information systems have to be slotted: this is going to make it difficult to come up with a single design of a VMS or IVRGI system which will be useful in any town. What is more, once significant numbers of drivers make use of any such system, their behaviour changes will in turn alter the distribution of traffic in the network and hence the behaviour of other drivers, whether or not these other drivers are using the route-guidance system. Of course, changes in travel behaviour are precisely what is being sought; but if we are to have any idea in advance whether these changes will be desirable ones or not, we must have models which will tell us what they are likely to be.

Acknowledgements.

This paper was written under funding from the Science and Engineering Research Council of Great Britain, within the research programme "Fundamental Requirements of Full-Scale Dynamic Route-Guidance Systems", which involves the universities of Leeds, Southampton and York.

I would like to thank Peter Bonsall for most helpful comments and criticisms, and Ferri Ghahri-Seremi and Sally Mortimer for essential computing and secretarial help respectively. I would also like to thank my successive thesis supervisors, Aaron Sloman and Maggie Boden. The views expressed, however, remain my sole responsibility.

References.

Appleyard, D. (1969)
"City Designers and the Pluralistic City"
Chapter 23 in L. Rodwin and Associates (eds)
Planning Urban Growth and Regional Development
MIT Press.

Bartram, D. and Smith, P. (1984)
"Everyday Memory for Everyday Places"
Chapter 3, pp.35-52 in Harris, J.E. and Morris, P.E. (eds)
Everyday Memory, Actions and Absent-Mindedness
Academic Press.

Bonsall, P. (1992)
"The Influence of Route Guidance Advice on Route Choice in Urban Networks"
Transportation 19(1), pp.51-73.

Briggs, R. (1973)
"Urban cognitive distance"
in Downs, R.M. and Stea, D. (eds)
Image and Environment: Cognitive Mapping and Spatial Behaviour
Aldine Publishing.

Byrne, R. (1979)
"Memory for urban geography"
Quarterly Journal of Experimental Psychology 31, pp.147-154.

Cadwallader, M.T. (1979)
"Problems in cognitive distance"
Environment and Behavior 11, pp.559-576.

Chandrasekaran, B. and Mittal, S. (1982)
"Deep versus Compiled Knowledge Approaches to Diagnostic Problem Solving"
Proceedings of the Conference of the American Association for Artificial Intelligence 1982 pp.349-354.

Chase, W.G. (1983)
"Spatial Representations of Taxi Drivers"
in Rogers, R. and Sloboda, J.A. (eds)
The Acquisition of Symbolic Skills
Plenum.

Chase, W.G. and Chi, M.T.H. (1981)
"Cognitive Skill: Implications for Spatial Skill in Large-Scale Environments"
in Harvey, J.H. (ed)
Cognition, Social Behavior and the Environment
Lawrence Erlbaum Associates.

Cousins, J.H., Siegel, A.W. and Maxwell, S.E. (1983)
"Wayfinding and Cognitive Mapping in Large-Scale Environments: A Test of a Developmental Model"
Journal of Experimental Child Psychology 35, pp.1-20.

Dijkstra, E.W. (1959)
"A Note on Two Problems in Connexion with Graphs"
Numerische Mathematik 1, pp.269-271.

Elliott, R.J. and Lesk, M.E. (1982)
"Route Finding in Street Maps by Computers and People"
Proceedings of the Conference of the American Association for Artificial Intelligence 1982 pp.258-261.

Gardner, M. (1959)
"About mazes and how they can be traversed"
Scientific American 200, 1, pp.132-137.

Gärling, T., Böök, A., Lindberg, E. and Nilsson, T. (1981)
"Memory for Spatial Layout of the Everyday Physical Environment: Factors Affecting Rate of Acquisition"
Journal of Environmental Psychology 1, pp.263-277.

Gärling, T., Böök, A., Lindberg, E. and Säisä, J. (1983)
The Spatiotemporal Sequencing of Everyday Activities: How People Manage to Find the Shortest Path to Travel in their Home Town
Environmental Psychology Research Unit, University of Umeå: unpublished paper.

Gärling, T., Böök, A., Lindberg, E. and Säisä, J. (1984)
"Everyday Activities in Large-Scale Environments: A Conceptual Framework and a Model"
in Gärling, T. (ed)
Human Action and Cognitive Mapping of Large-Scale Environments: Some New Directions in Recent Research
Environmental Psychology Research Unit, University of Umeå.

Golledge, R., Briggs, R. and Demko, D. (1969)
"The Configuration of Distances in Intra-Urban Space"
Proceedings of the Association of American Geographers 1, pp.60-65.

Gotts, N.M. (1988)
Unplanned Wayfinding in Path-Networks: A Theoretical Study of Human Problem-Solving
Unpublished D.Phil Thesis, University of Sussex

Harary, F. (1969)
Graph Theory
Addison-Wesley.

Hart, P.E. (1982)
"Directions for AI in the Eighties"
SIGART Newsletter 79, pp.11-16.

Hayes-Roth, B. and Hayes-Roth, F. (1979)
"A Cognitive Model Of Planning"
Cognitive Science 3, pp.275-310.

Heft, H. (1979)
"The Role of Environmental Features in Route-Learning: Two Exploratory Studies of Way-finding"
Environmental Psychology and Nonverbal Behavior 3, pp.172-185.

Kuipers, B.J. (1977)
Representing Knowledge Of Large-Scale Space
Massachusetts Institute of Technology Technical Report AI-TR-418.

Kuipers, B.J. (1979)
"Commonsense Knowledge Of Space: Learning From Experience"
Proceedings of the International Joint Conference on Artificial Intelligence 6 Vol.1, pp.499-501.

Kuipers, B.J. (1982)
"The Map in the Head Metaphor"
Environment and Behavior 14, pp.202-220.

Kuipers, B.J. (1983)
"Modeling Human Knowledge of Routes: Partial Knowledge and Individual Variation"
Proceedings of the Conference of the American Association for Artificial Intelligence 1983 pp.216-219.

Lee, T.R. (1970)
"Perceived Distance as a Function of Direction in the City"
Environment and Behavior 2, pp.40-51.

Lindberg, E., Gärling, T. and Böök, A. (1984)
"Wayfinding and Spatial Orientation in the Designed Environment"
in Gärling, T. (ed)
Human Action and Cognitive Mapping of Large-Scale Environments: Some New Directions in Recent Research
Environmental Psychology Research Unit, University of Umeå.

MacEachren, A.M. (1980)
"Travel Time as the Basis of Cognitive Distance"
The Professional Geographer 32, pp.30-36.

Neves, D.M. and Anderson, J.R. (1981)
"Knowledge Compilation: Mechanisms for the Automization of Cognitive Skills"
in Anderson, J.R. (ed)
Cognitive Skills And Their Acquisition
Lawrence Erlbaum Associates.

O'Keefe, J. and Nadel, L. (1978)
The Hippocampus as a Cognitive Map
Clarendon Press.

Pailhous, J. (1970)
Representation de l'Espace Urbain: l'Exemple du Chauffeur de Taxi
R.U.P., Paris.

Pailhous, J. (1980)
"Representation of Urban Space"
in Farr, R. and Moscovici, S. (eds)
Social Representation
Cambridge University Press.

Passini, R.E. (1980a)
"Wayfinding: a Conceptual Framework"
Man-Environment Systems 10, pp.22-30.

Passini, R.E. (1980b)
"Wayfinding in Complex Buildings: an Environmental Analysis"
Man-Environment Systems 10, pp.31-40.

Sadalla, E.K. and Staplin, L.J. (1980)
"The Perception of Traversed Distance: Intersections"
Environment and Behavior 12, pp.167-182.

Säisä, J., Gärling, T. and Lindberg, E. (1984)
"The Spatiotemporal Sequencing of Everyday Activities in Large-Scale Environments: Test of a Model of Consumers' Sequential Spatial Choices"
in Gärling, T. (ed)
Human Action and Cognitive Mapping of Large-scale Environments: Some New Directions in Recent Research
Environmental Psychology Research Unit, University of Umeå.

Siegel, A.W. and White, S.H. (1975)
"The Development of Spatial Representations of Large-Scale Environments"
Advances in Child Development and Behavior 10, pp.10-55.

Sugie, M., Menziliciogiu, O. and Kung, H.T. (1984)
CARGuide - On-board Computer for Automobile Route Guidance
Carnegie-Mellon University Report CMU-CS-84-144.

Thorndyke, P.W. (1981)
"Spatial Cognition and Reasoning"
Chapter 7 in Harvey, J.H. (ed)
Cognition, Social Behavior and the Environment
Erlbaum.

Thorndyke, P.W. and Goldin, S.E. (1983)
"Spatial Learning and Reasoning Skill"
Chapter 9 in Pick, H.L. and Acredolo, L.P.
Spatial Orientation: Theory, Research, and Application
Plenum.

Thorndyke, P.W. and Hayes-Roth, B. (1982)
"Differences in Spatial Knowledge Acquired from Maps and Navigation"
Cognitive Psychology 14, pp.560-589.

Tolman, E.C. (1948)
"Cognitive Maps in Rats and Men"
Psychological Review 55, pp.189-208.

Fig.1

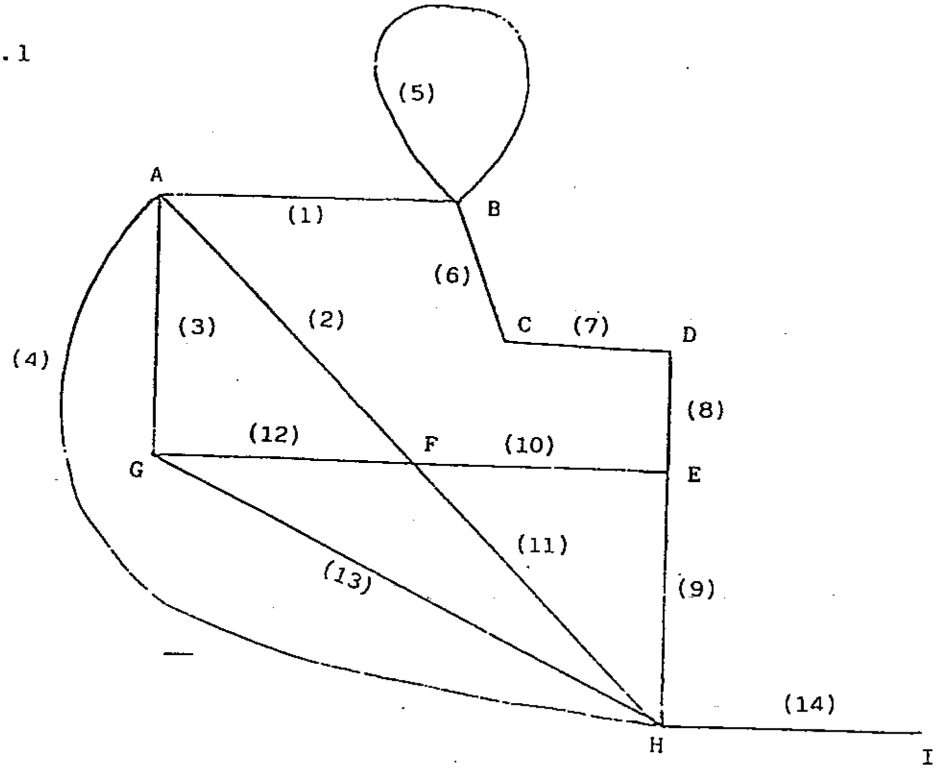


Fig.2

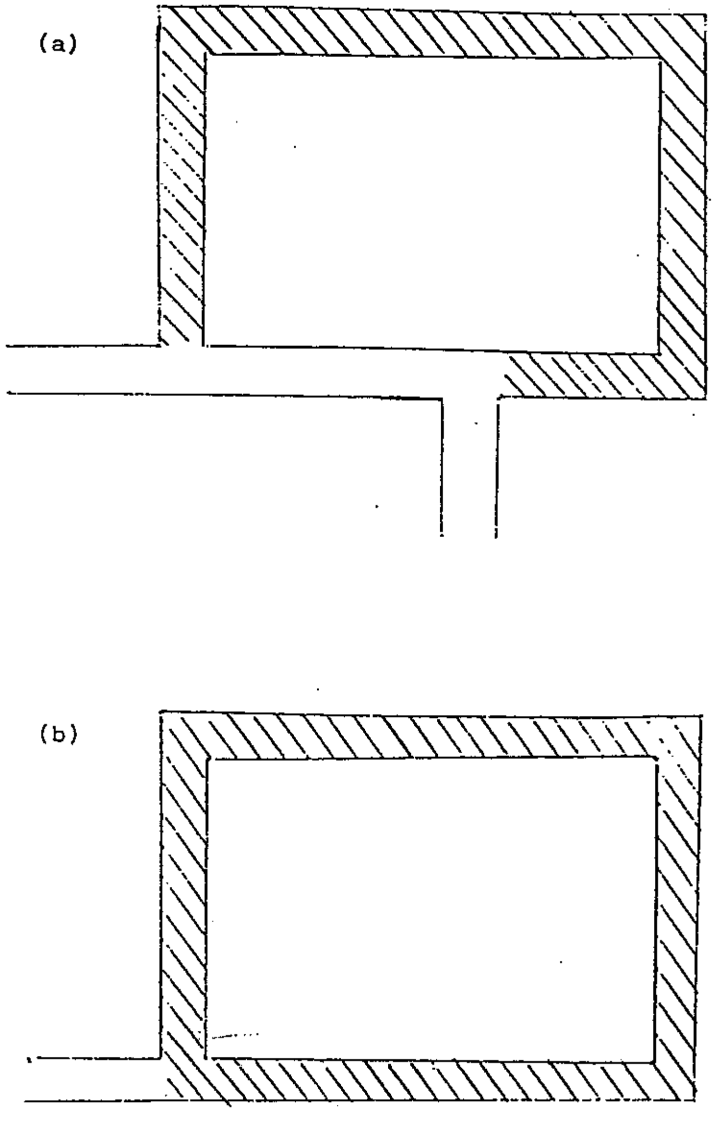


Fig.3

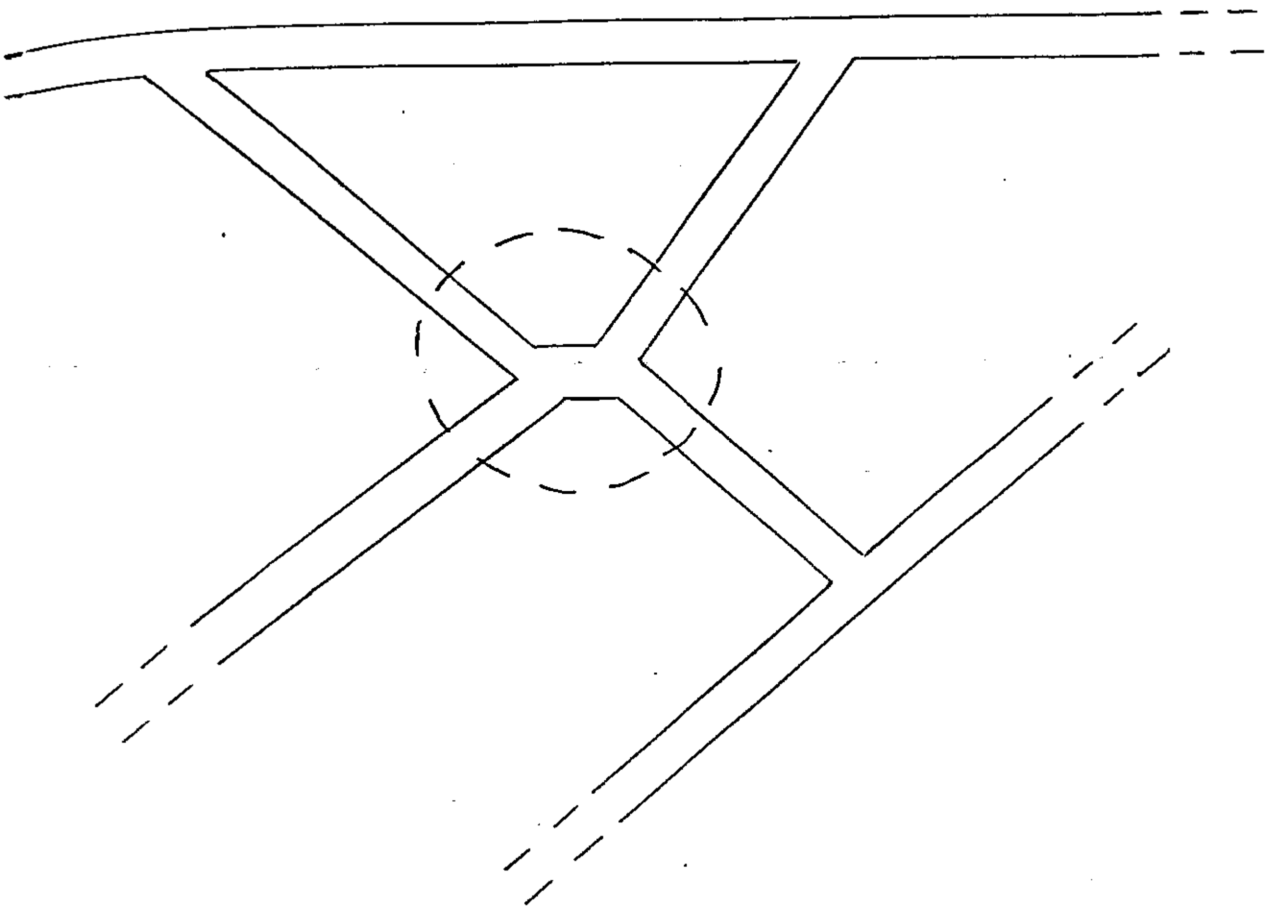


Fig.4

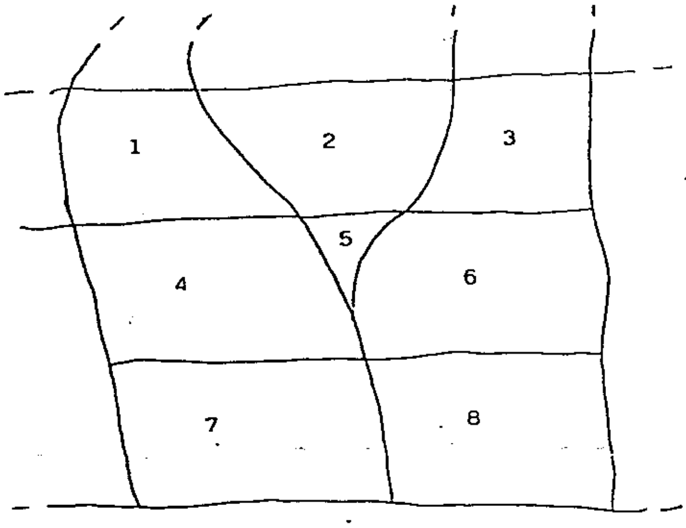


Fig.5

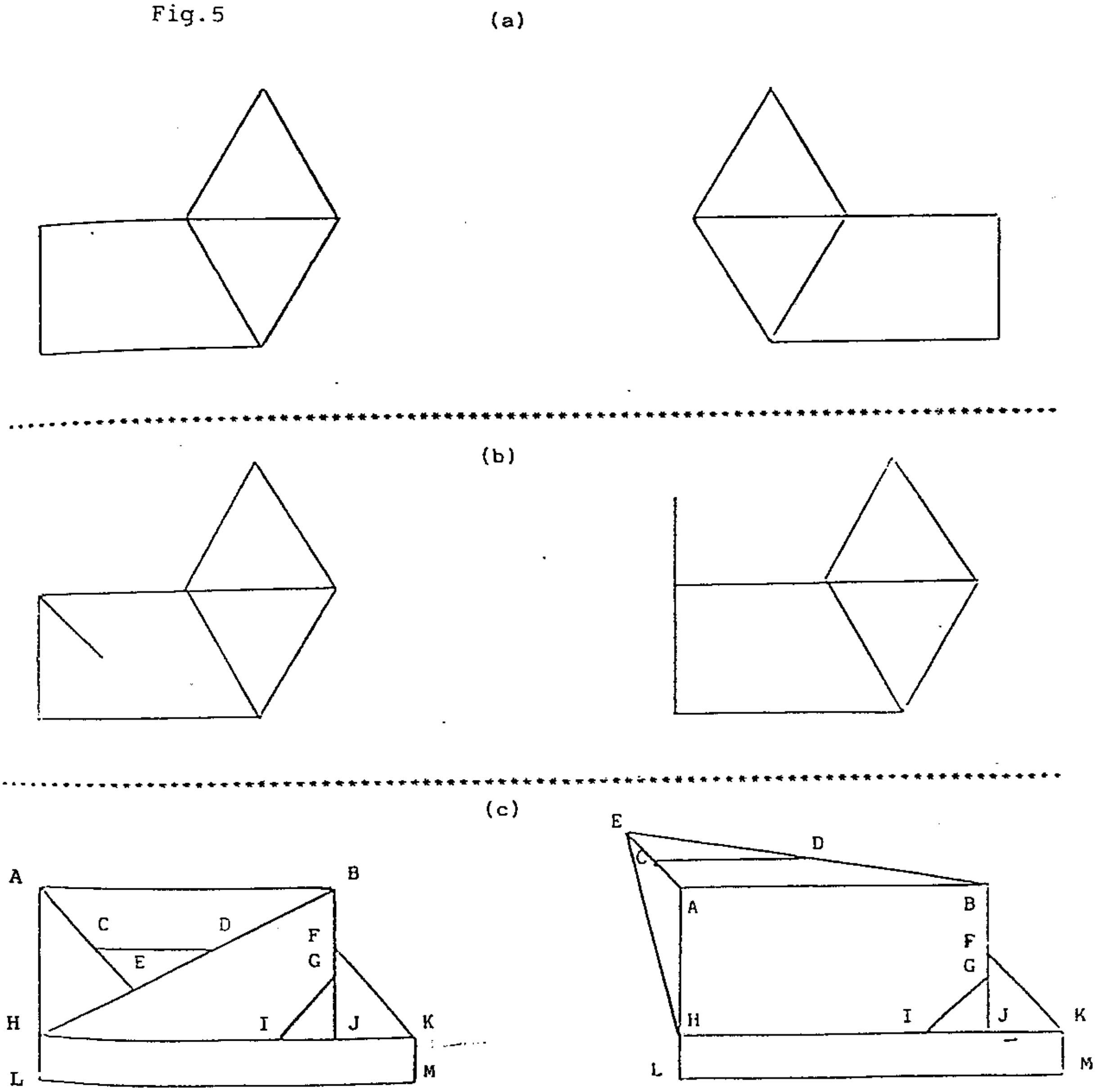
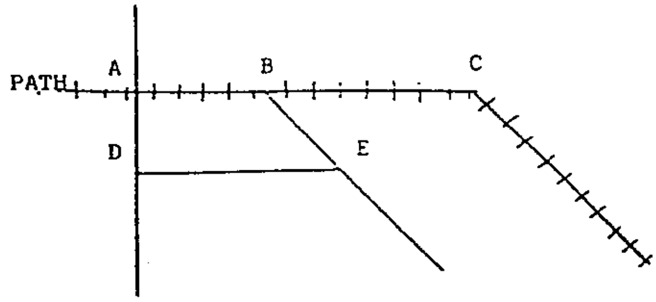


Fig.6

(a)



(b)

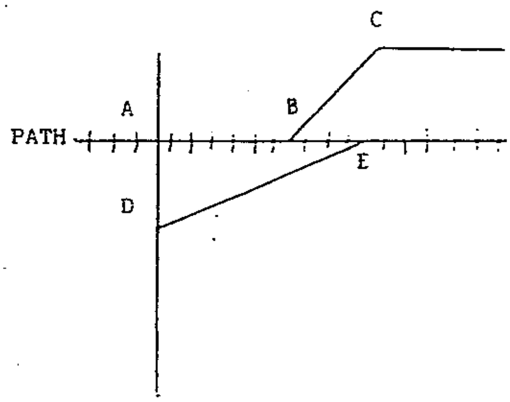


Fig.7

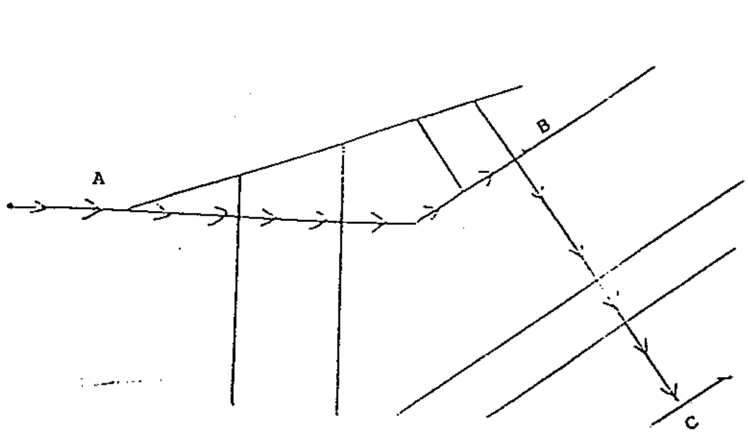


Fig.8

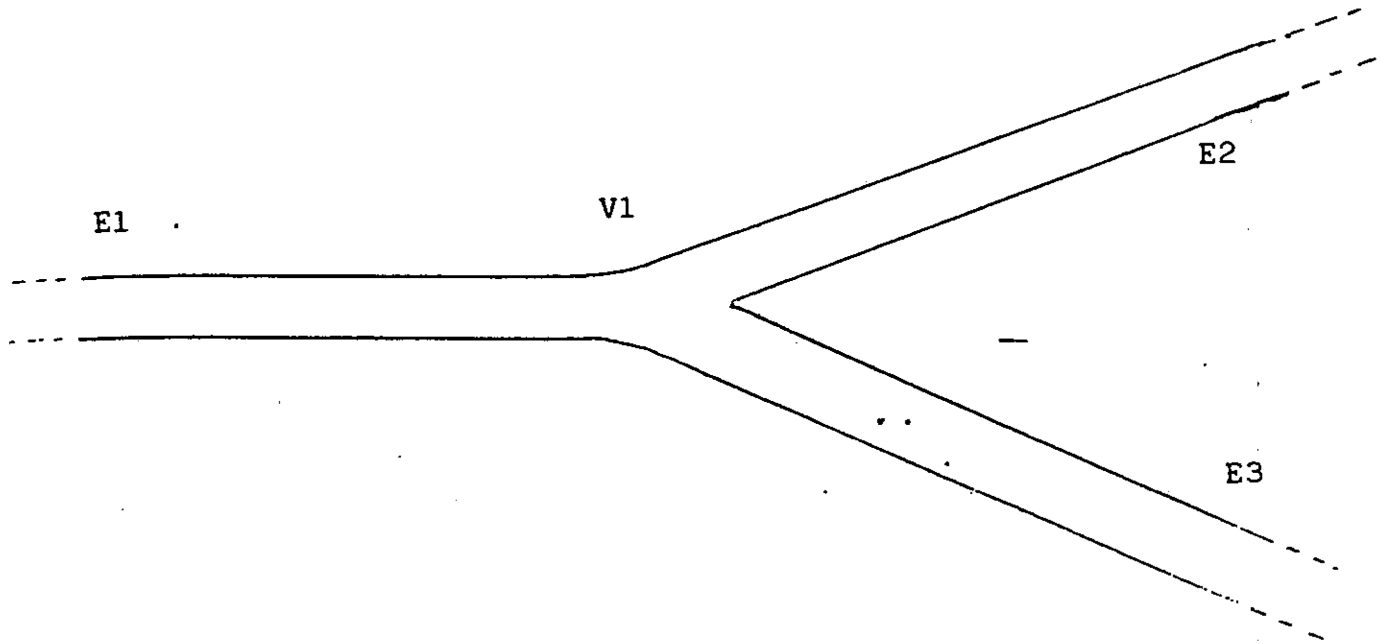
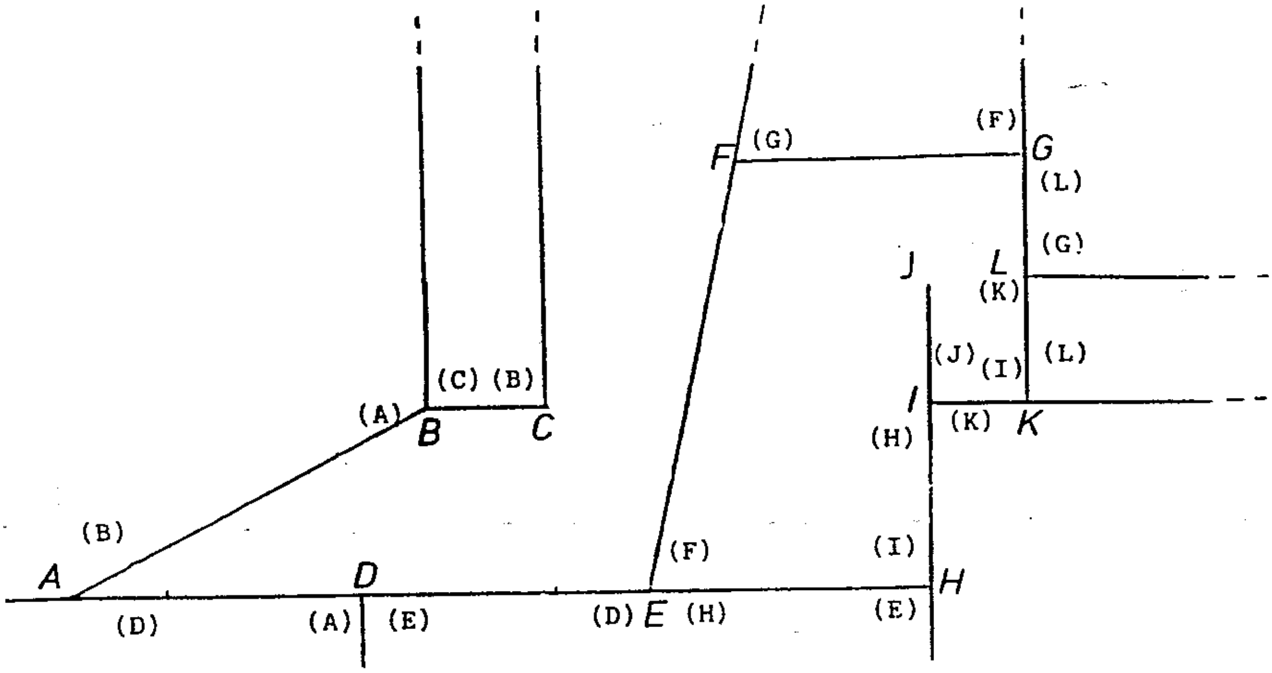
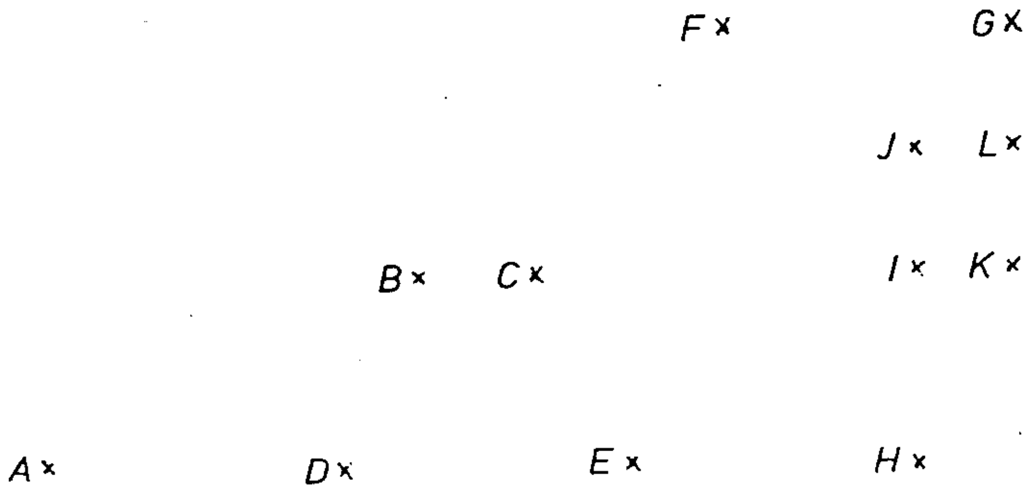


Fig.9

(a)



(b)

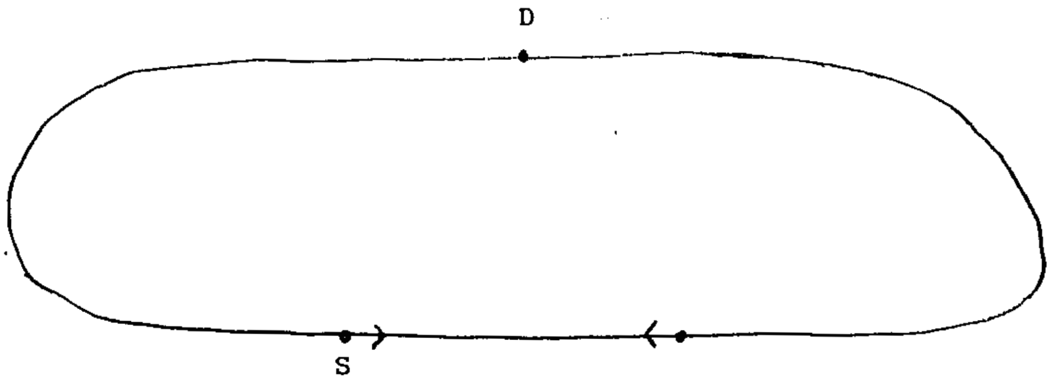


(c)

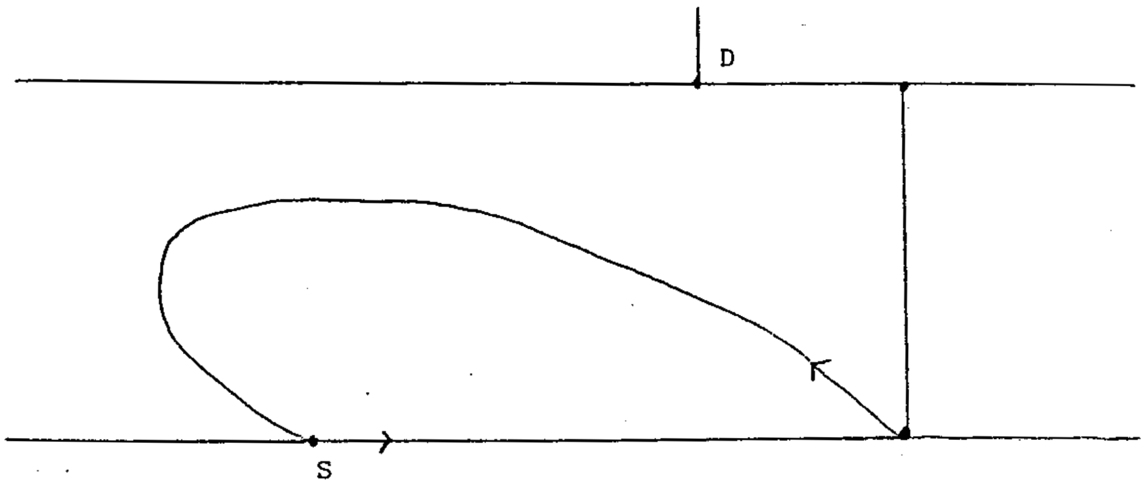
AB	AD	BC	DE
EF	EH	FG	GL
HI	IJ	IK	KL

Fig.10

(a)



(b)



(c)

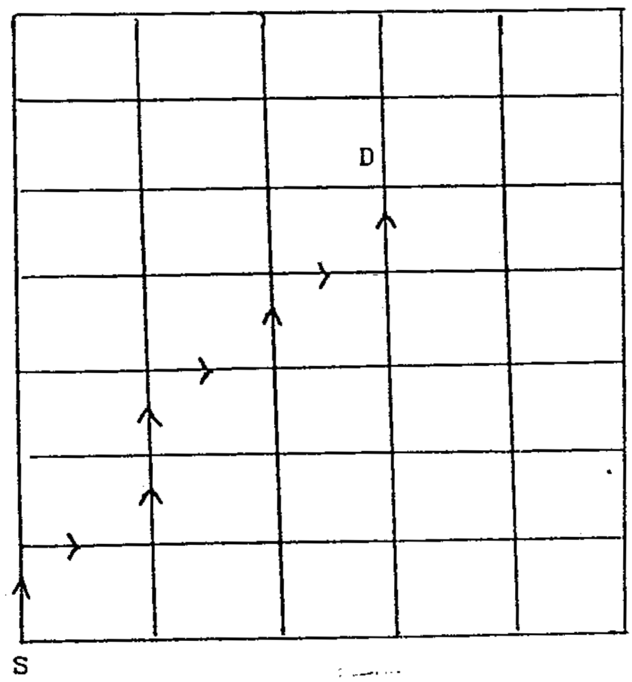


Fig.11

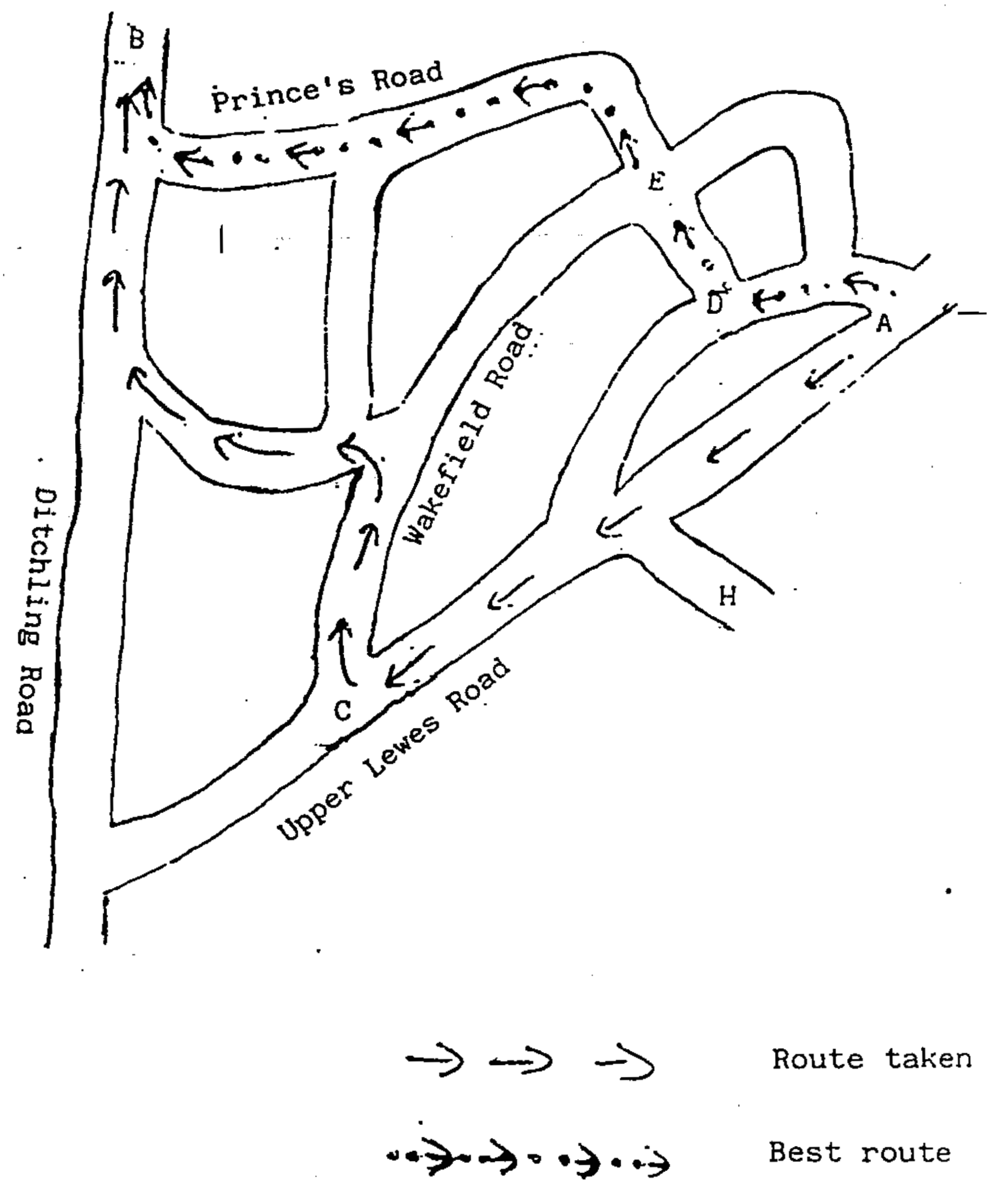

B
Prince's Road
E
D
A
Wakefield Road
Ditchling Road
H
C
Upper Lewes Road
Route taken
Best route

Fig.12

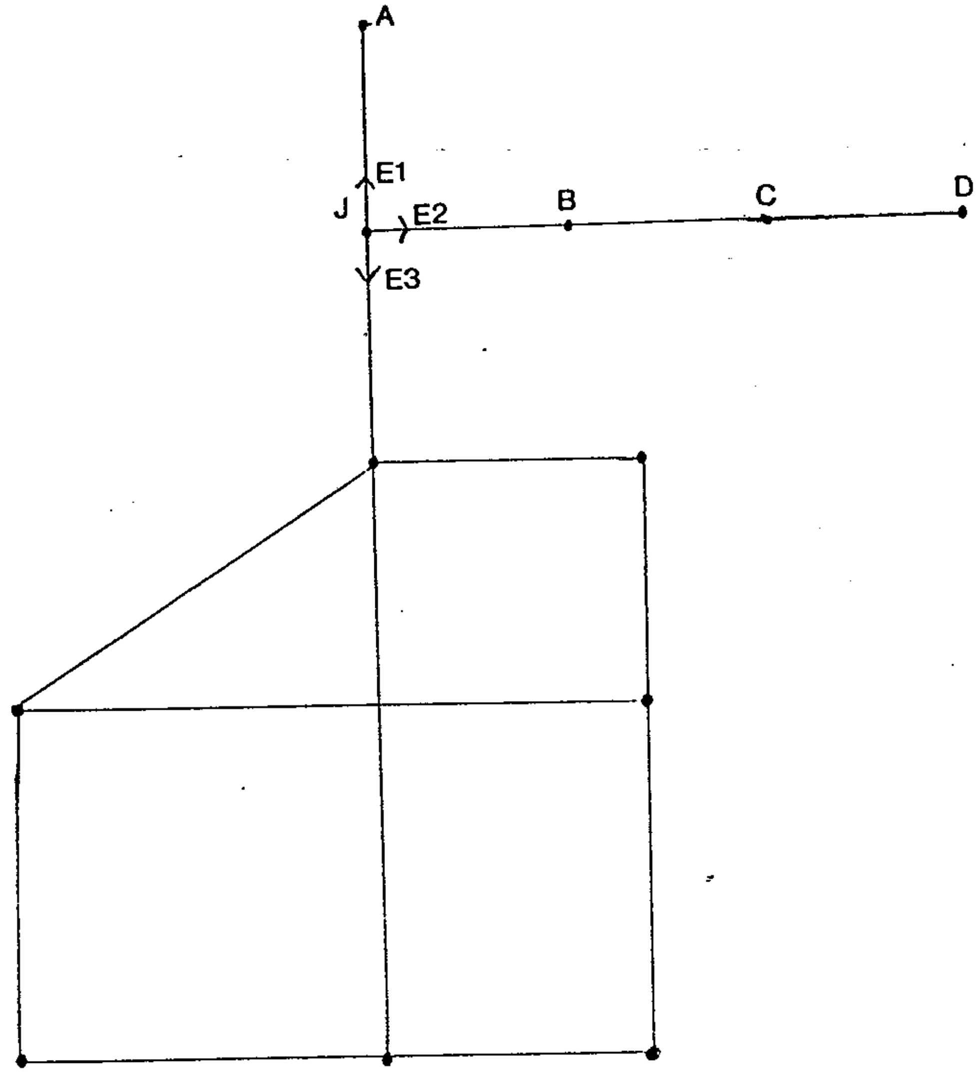


Fig.13

Groups I and II

Group I

Group II

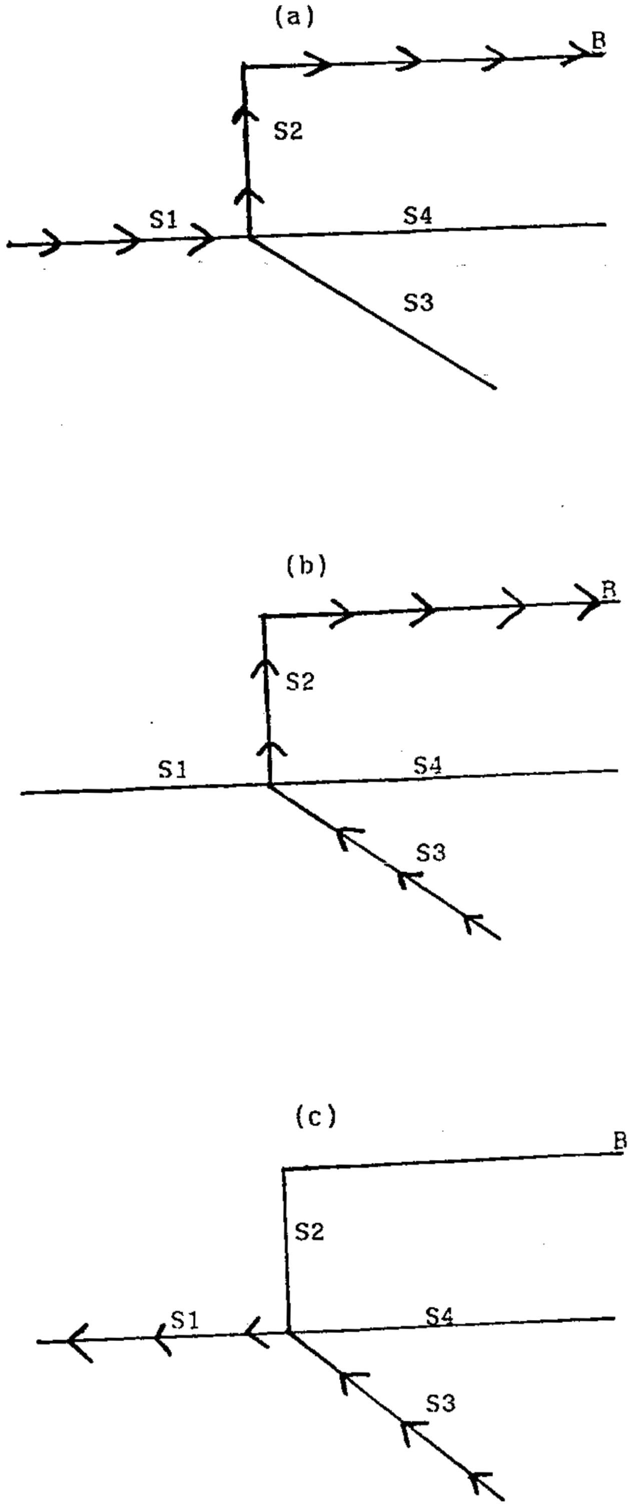
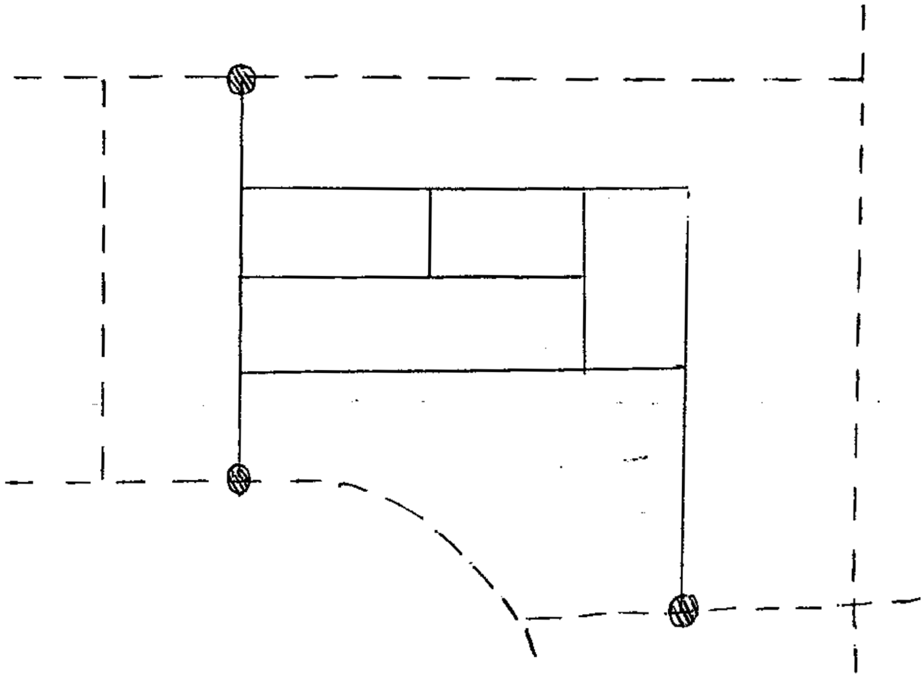


Fig.14



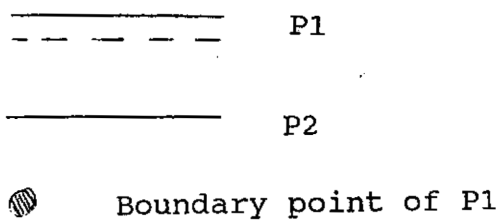


Fig.15

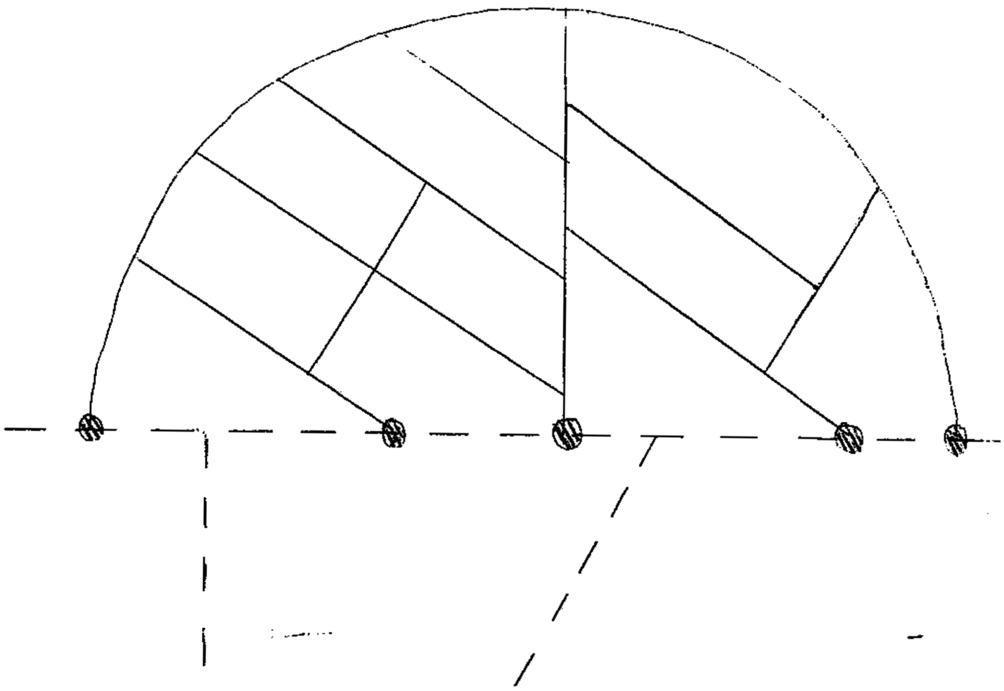
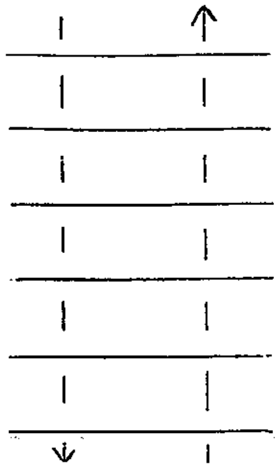


Fig.16

Banded division of space

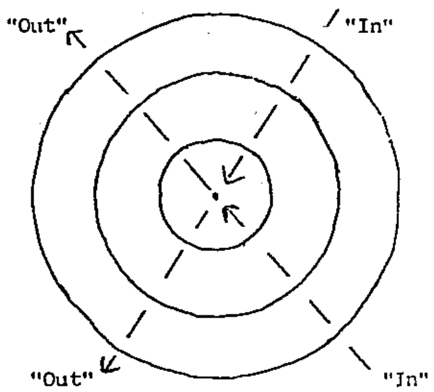
(a)



(b)

(c)

Concentric division of space



Sectorial division of space

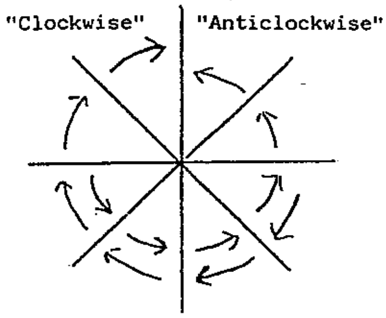


Fig.17

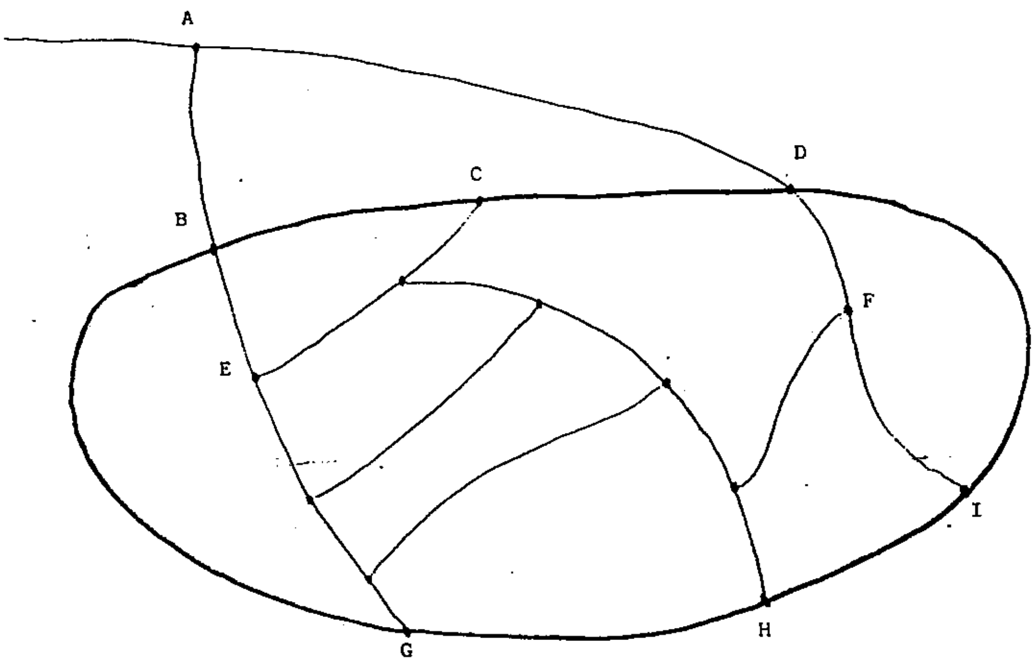


Fig.18

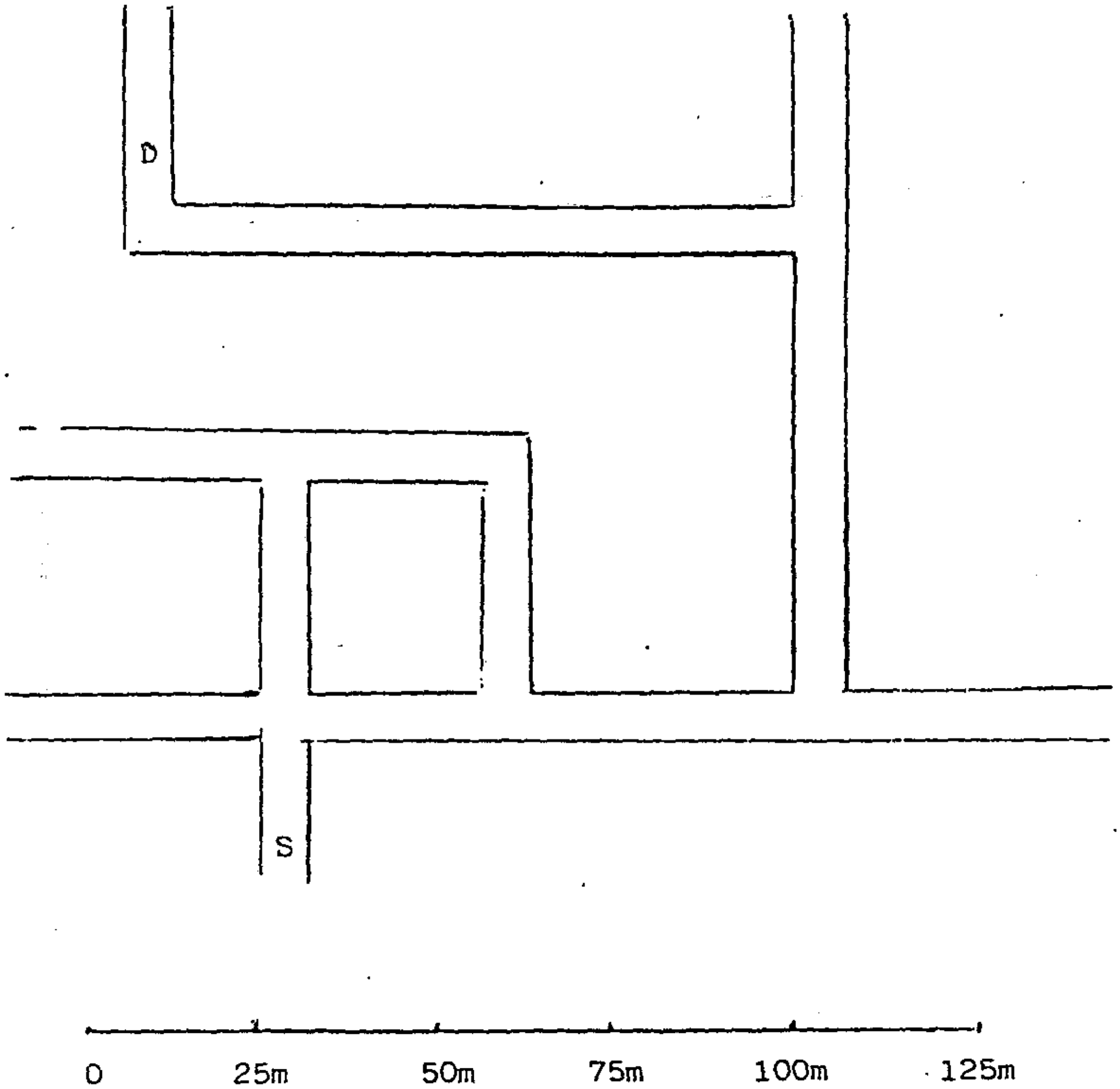


Fig.19

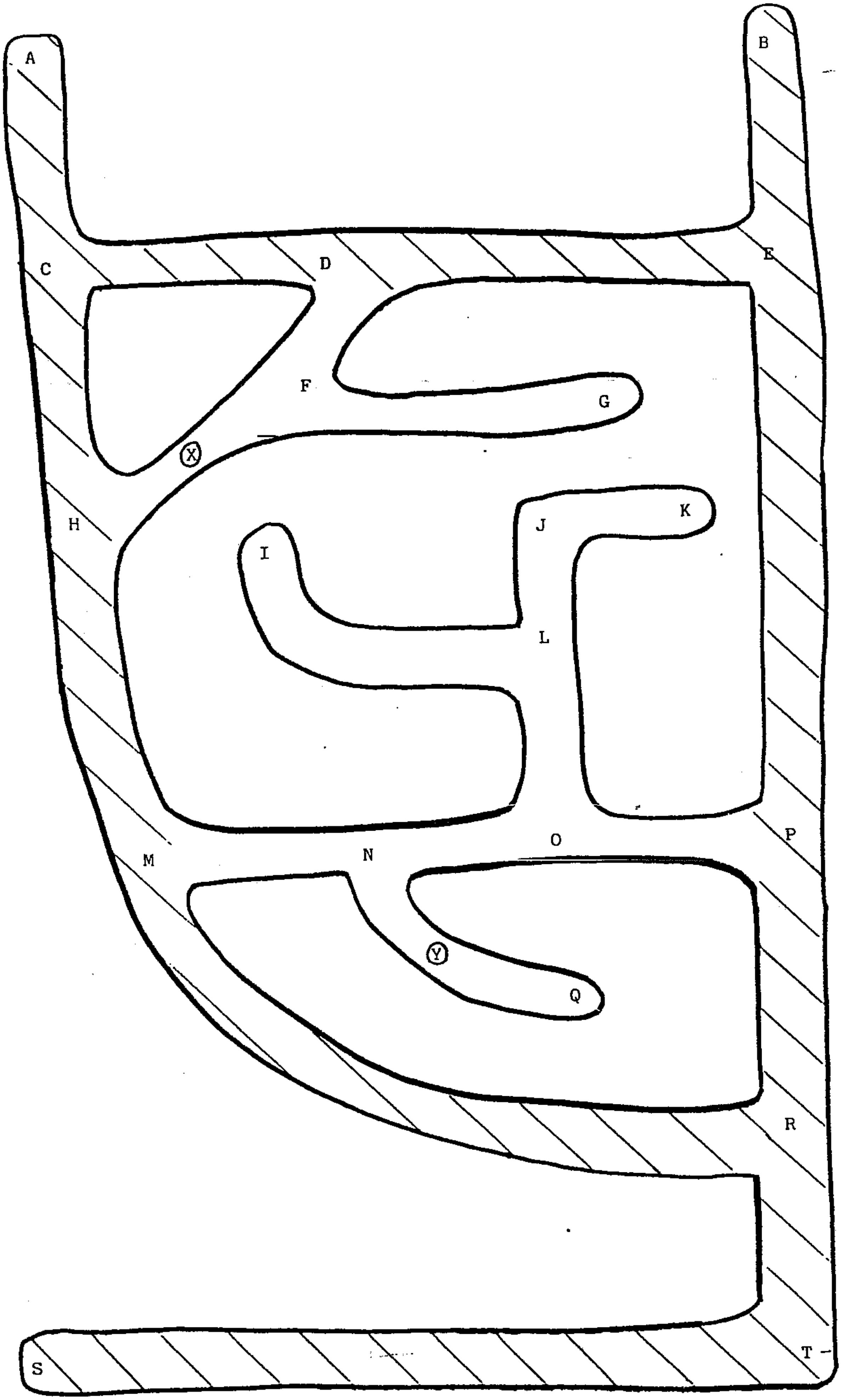


Fig.20

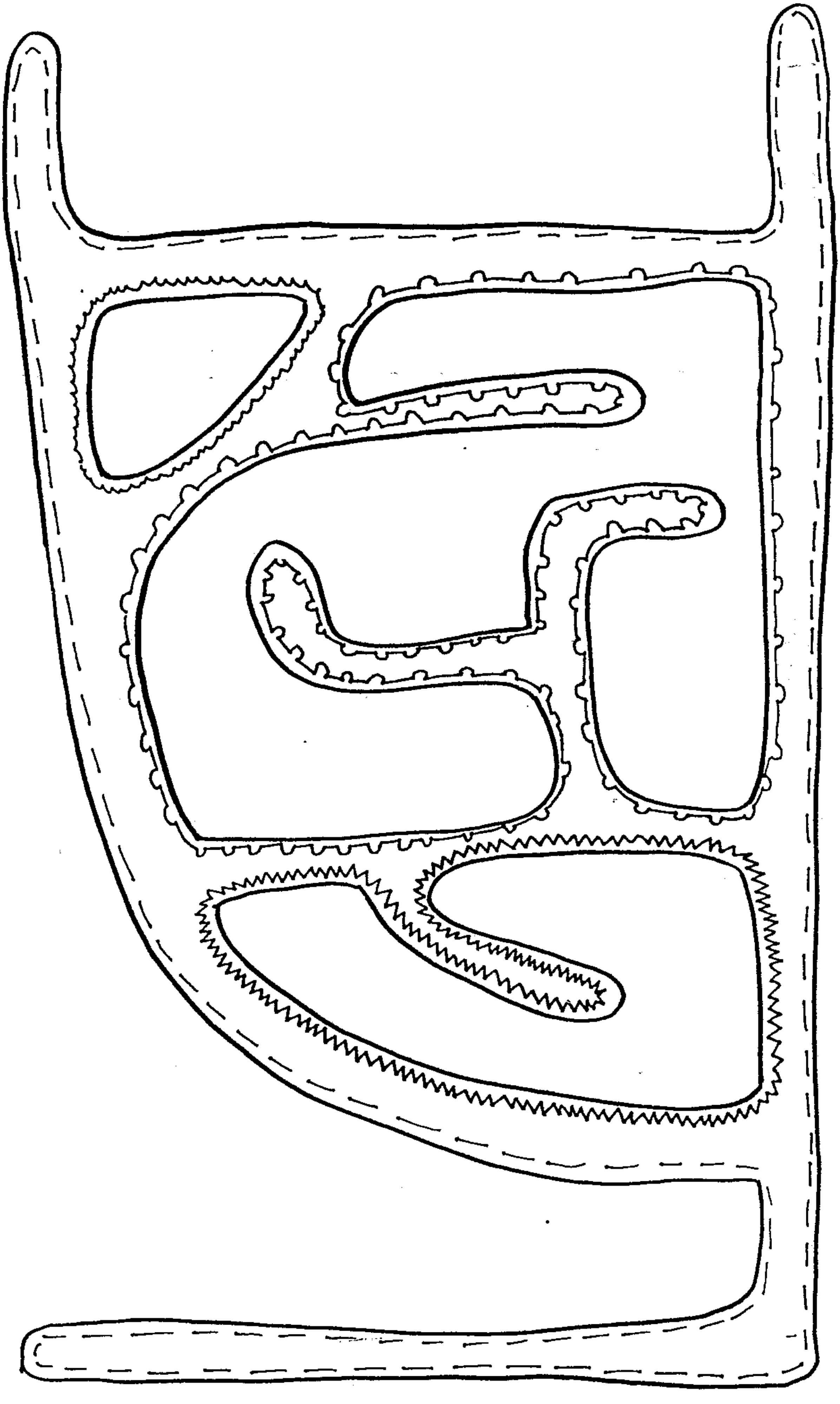


Fig.21

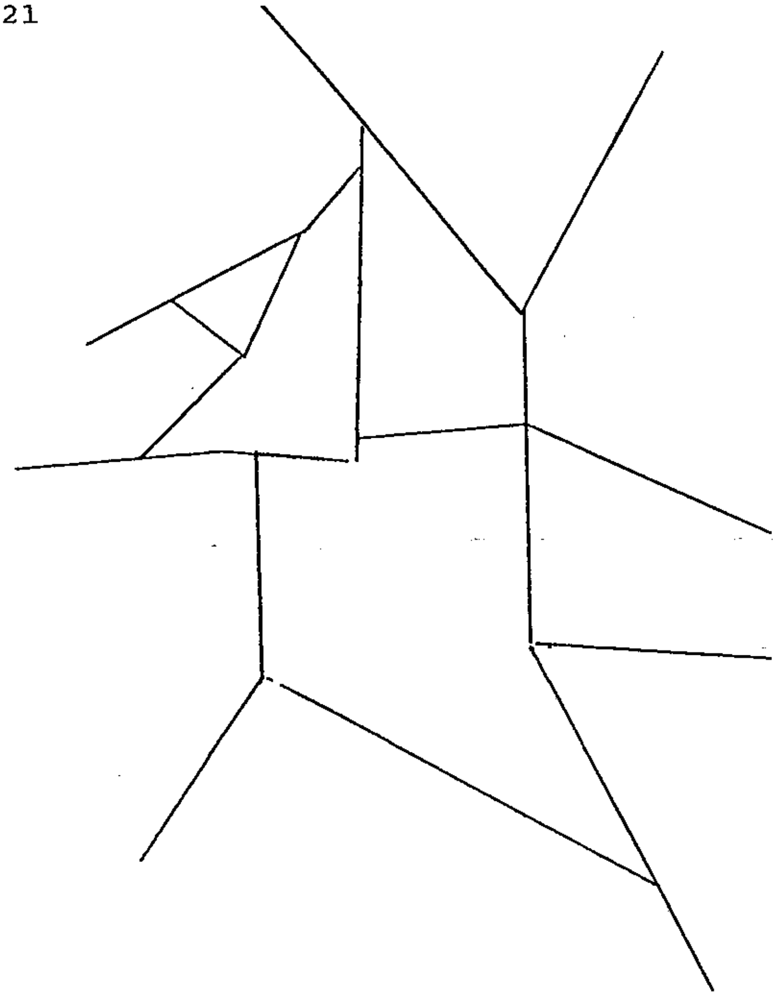


Fig.22

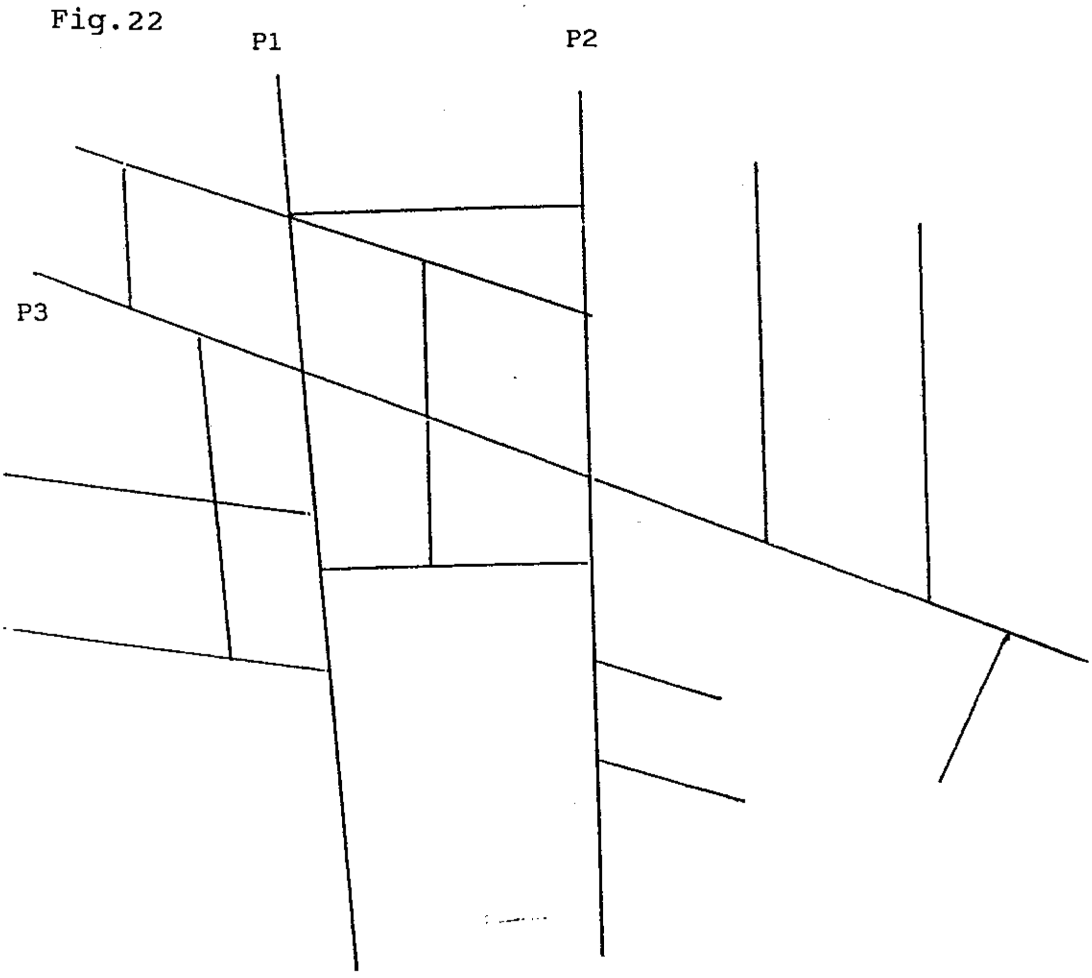


Fig.23

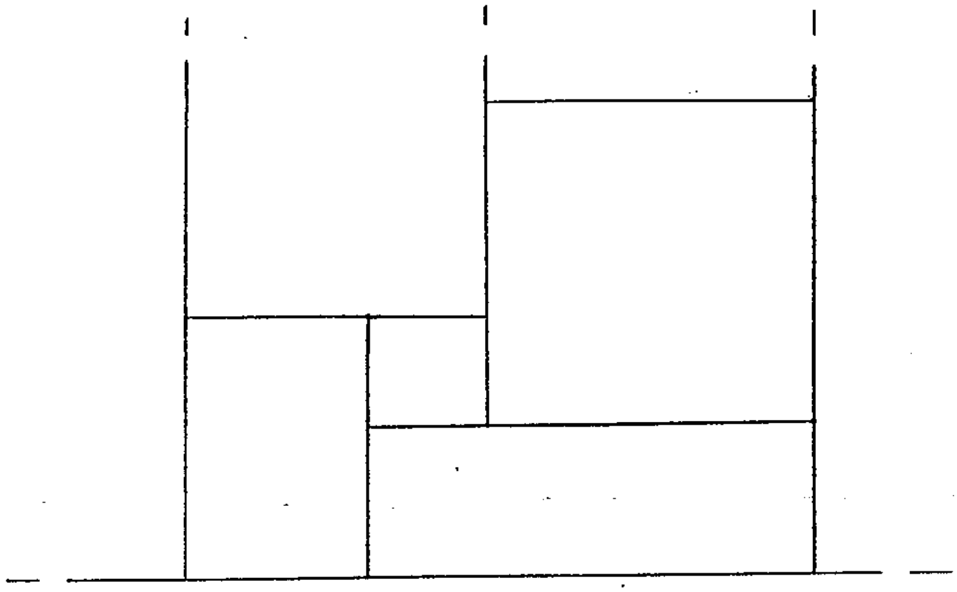


Fig.24

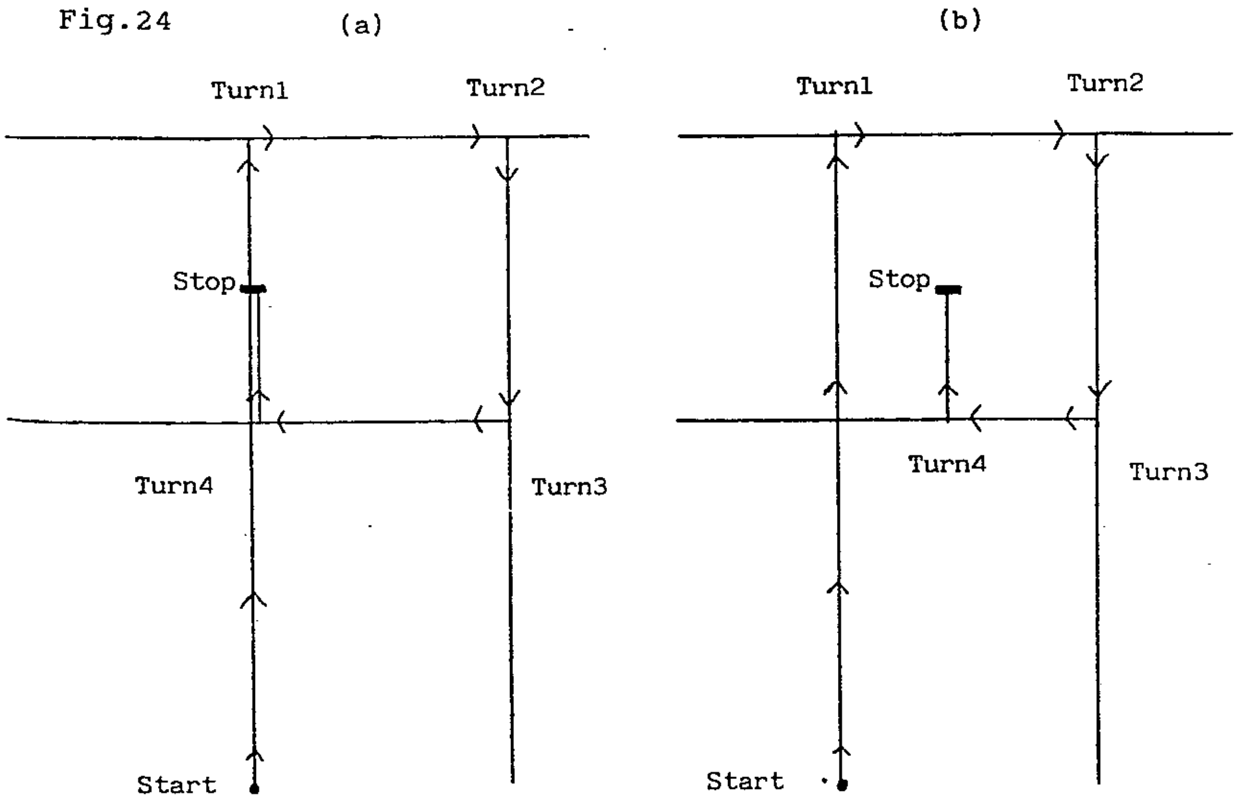

(a)
Turn1
Turn2
Stop
Turn4
Turn3
Start
(b)
Turn1
Turn2
Stop
Turn4
Turn3
Start

Fig.25

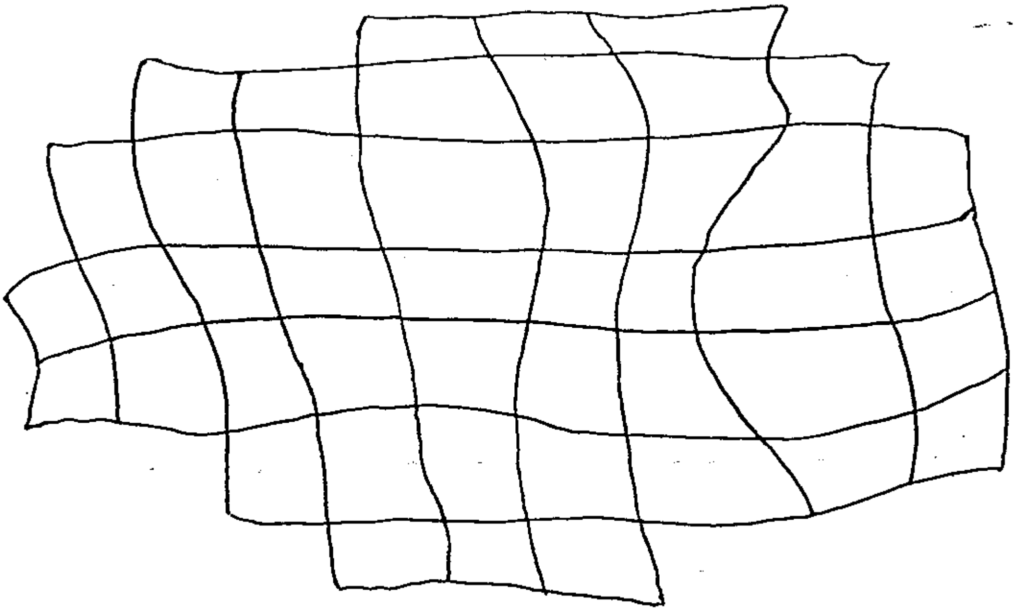


Fig.26

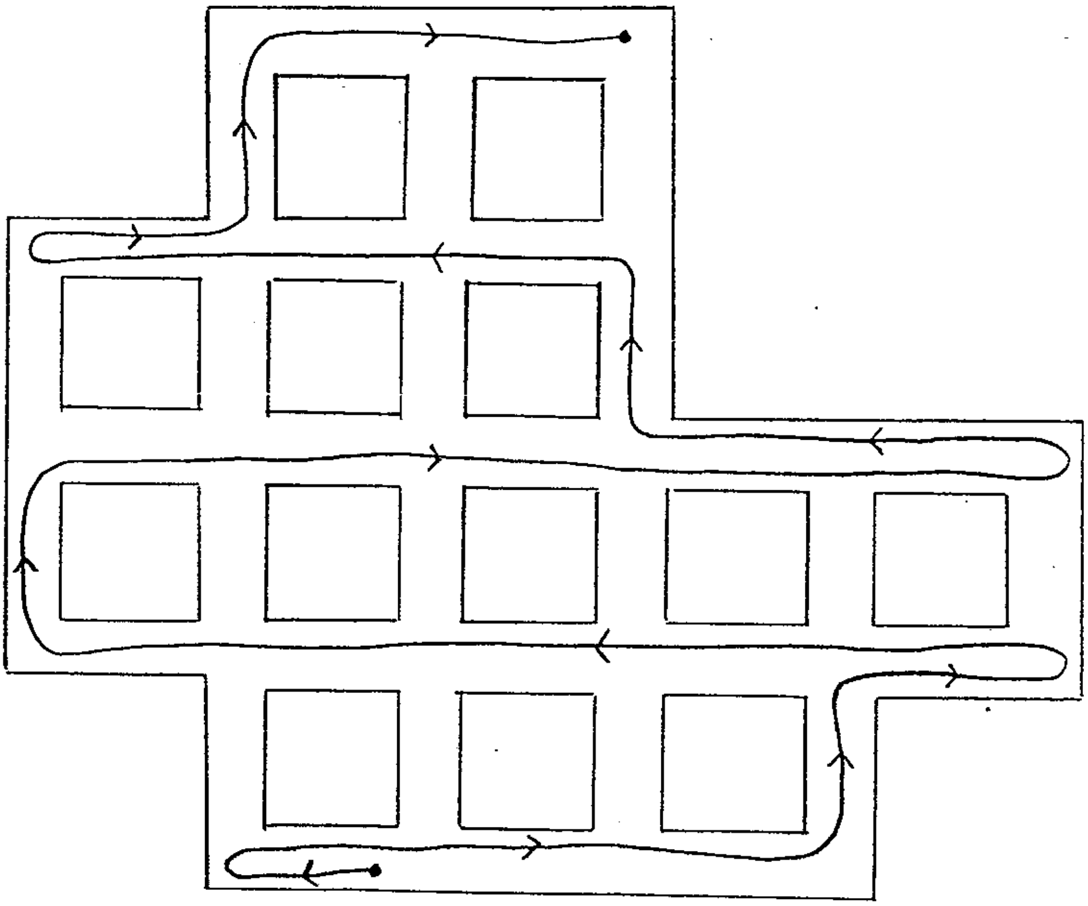


Fig.27

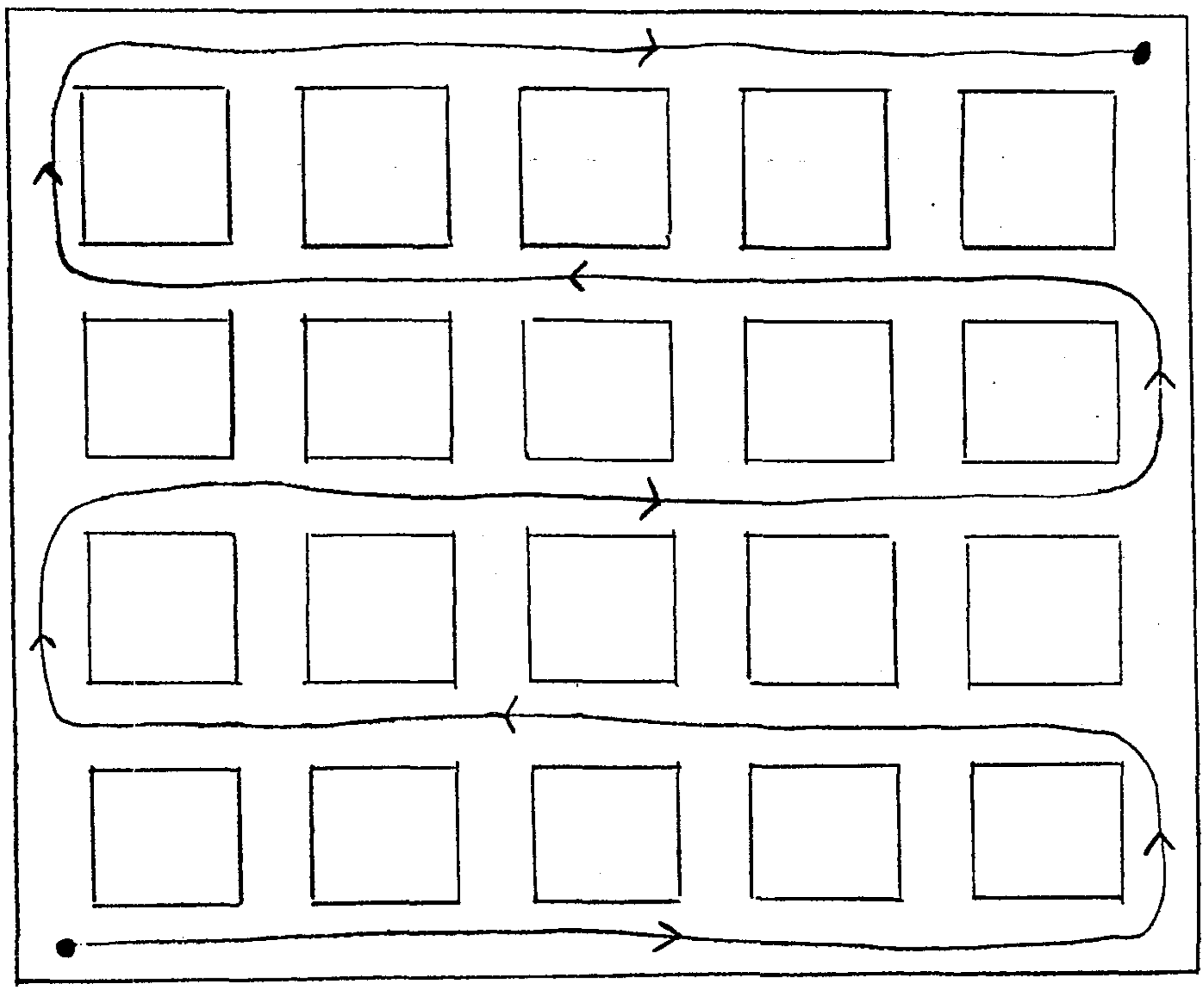


Fig.28