WHAT DO WE MEAN BY BUILDING FUNCTION ?

Bill Hillier
Julienne Hanson
John Peponis

Bartlett School of Architecture and Planning, University College London

ABSTRACT

Scientific approaches to architecture usually avoid the issue of building form, preferring to focus on function. But how can there be a theory of function without a systematic analysis of the key architectural variable of form ? A theory of description is required. In this paper it is argued that such a theory can be built through the analysis of spatial form in buildings. Then once spatial form is describable in terms of a descriptive theory, a more powerfully scientific - and architectural - understanding of function is possible. The argument draws on several pieces of research carried out by the authors and their students, but focusses eventually on various types of medical building in order to illustrate certain general principles.

THE NEED FOR A THEORY OF DESCRIPTION

There are, it seems, now two distinct traditions in architectural discourse: a critical tradition, which is concerned first and foremost with the changing form of buildings, sometimes confining itself to the superfices of style, but at its best attending carefully to the systematics of spatial and morphological form (the work of Colin Rowe and Paul Frankl come immediately to mind (Rowe, 1976; Frankl,1914); and a research tradition which (as witness the preoccupations of this conference) studiously avoids the issue of form, and adresses itself almost exclusively to matters of function, in the belief, it would appear, that function is scientifically tractable whereas form is not.

But surely the schism is bizarre. How can there be a useful theory of building function unless it either incorporates or relates to a theory of the architectural malleability of form. Similarly, how can there be a theory of architectural form independent of a theory of the functional logistics of form. For both the scientist and the architect it is the essence of his discipline that the two issues are aspects of a single question: what is it about architectural form that works ?

For the architectural researcher, the question ought to be crystallised as his or her most pressing concern, since how can any investigation be truly systematic unless the architectural variable can be controlled ? That architectural variable is building form. Architecture decides form, and hopes for function. Controlling the architectural variable for research purposes means having a theory of description of building form. Without it we can do nothing. With it we can at least begin to find out what there might be to know about buildings.

The aim of this paper is to set out a theory for the description of building form, and through this to arrive at a reformulation of the problem of function. More precisely, it tries to answer two questions:
- what is it about what people do (function) that leaves its mark on building form ? and
- what is it about building form that leaves its mark on what people do ?

Both of these questions are, we believe, capable of exact answers given a theory of description.

ELEMENTS OF A SPATIAL DESCRIPTION

First, we must decide what it is about buildings that needs to be described. At first, this appears an impossibly complex question, not the least because buildings are so many different things: they are physical constructions, they are arrangements of space, and they are objects in a style culture. But provided we remember that only art deals with the whole of reality and science can only hope to understand its underlying dimensions, then we need not be pessimistic. Being a physical and a stylistic object the building shares with most other artefacts. It is the organisation of space that make the building unique. It is the distinguishing mark of space that sets the work of architecture apart from other artefacts, and it does so in a very important way.

A bridge, a vase or a surgical instrument is functionally useful insofar as it is a physical construction, and socially meaningful insofar as it participates in a style culture. Physical form and social meaning are, to some extent, separable questions. In architecture this is not so. Because buildings organise space for social purposes, social considerations are present in the very physical form of the building. Social meaning is not a gloss added to buildings: it is an intrinsic aspect of their physical form. Indeed it is their most important aspect. Buildings are for the organisation of space. This is the most general statement of function. In buildings, technology permits and style confirms that space has been created and organised for social purposes. This is the unavoidable basis of our discipline.

The task of a theory of description thus becomes one of a theory of spatial description. But not only that. If spatial description is to reveal functional dynamics, then it must somehow incorporate or lead to a theory of the social nature of space. We must learn to describe buildings spatially as social products, or we will not be able to arrive at a usable definition of building function

A descriptive theory for space has, in our view, to be built at three levels:

- the identification and representation of spatial elements;
- the categorisation and analysis of spatial relations; and
- the modelling of common, or 'genotypical' themes and patterns.

The first, identification and representation of elements must, we believe, be solved in three ways not one. The first identifies boundaries. A building is a more or less controlled domain, and this means a continuous boundary perforated by one or more entrances. possibly with other boundary-entrance pairs within:

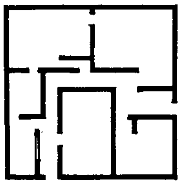


Fig.1a
Plan

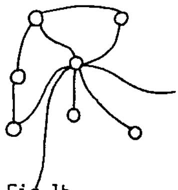


Fig.1b
Boundary graph

Fig.1.a. is a simple plan and Fig.1.b. represents its boundary structure as a graph in which circles represent bounded spaces and lines relations of direct permeability.

But bounded space by no means exhausts what a building is spatially. Some very complex buildings have very few bounded spaces- churches, some schools, offices and so on - yet achieve a highly complex and differentiated pattern of space. These can be analysed - at the level of the plan at least - by introducing two further kinds of spatial description. The first we call a 'convex', or two dimensional description. This identifies the fewest and fattest convex spaces that cover the system, by applying a rule which says that fat spaces always prevail over thin spaces. Fig.2.a and b show our example analysed convexly, together with its graph:

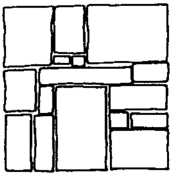


Fig.2a
Convex breakup

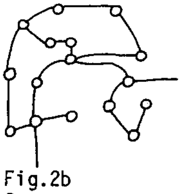


Fig.2b
Convex graph

The second we call an 'axial' or one dimensional description. This identifies the longest and fewest straight lines that cover all the convex spaces in the plan. Fig.3.a. and b. shows the circulation spaces of our example analysed axially, together with its graph:

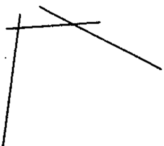


Fig.3a
Axial map of circulation

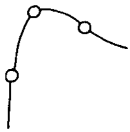


Fig.3b
Axial graph of circulation

Even before we begin relational analysis these simple representational devices can allow us to quantify and represent visually many of the spatial characteristics of plans, the more so as plans become complex and everyday language fails to provide a rigorous account. To begin with, we can simply count the numbers of bounded, convex and axial spaces and express them in relation to the total area of the plan - so many space of one or other type per so many square metres. These will tell us the rate at which a particular type of building or a particular architect adds each type of space to a growing structure.

More informatively, we can look at the ratios of each type of space to the others. Since the convex system must always have the maximum number of spaces, then it will serve as the best reference point. If we divide the number of bounded spaces by the number of convex spaces then we arrive at a figure between 0 and 1, where figures close to 1 indicate few convex spaces over and above the bounded, and close to 0 many more. Low figures thus indicate a higher degree of convex articulation of the plan.

If we then divide the number of axial lines by the number of convex spaces then we will have a similar measure of axial linking, since if few lines link many convex spaces we will have a low value, and if the number of lines approaches the number of convex spaces we will get a value close to 1, and little axial linking. We can quickly establish that, for a good sample of his villas and apartments at least, Corb convexly articulates more and axially integrates less than other major modern movement figures.

It should be noted by the way, that the measure of axial linking only works if the fewest axial lines to cover the convex spaces have been drawn. In cases where extra links between convex spaces can be made by additional axial lines, then the best measure of the axial organisation is one which compares the axial system to a perfect orthogonal grid by using the formula:

$$\frac{2\sqrt{R} + 2}{A} \qquad (1)$$

where R is the number of rings in the axial map and A the number of axial lines. High values indicate an axial organisation approximating a grid, very low values a high degree of axial breakup (Readers interested in the full account of these measures should consult Hillier and Hanson 1983, forthcoming).

RELATIONAL ANALYSIS

These representations and figures give useful data on the general characteristics of plans. Much greater precision in analysis can however be achieved through the numerical analysis of relations among spaces. The integration value of a space (code-named RRA in the house jargon of the space syntax research programme) will express how many spaces distant a particular space is from every other space in the plan, it being the case (though far from always obvious) that these values are different, from one space to another in the same system. Fig.4.a. and b. show this graphically and numerically by taking two convex spaces in our example and 'justifying' the graphs from those points that is, aligning all the other spaces above the selected space in layers according to how many spaces each is from it:

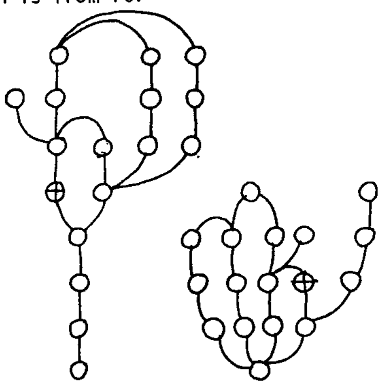


Fig.4a Fig.2 'justified' from its most segregating space. RRA= 2.152.

Fig.4b Fig.2 'justified' from its most integrating space. RRA= .717.

The numerical value is a number varying about 1, with low values indicating more integration (in effect, less distance to all the others) and high values more segregation, or more distance. It is calculated first by:

$$\frac{2(MD - 1)}{k - 2} \qquad (2)$$

where MD is the 'mean depth', or the mean number of spaces away, of all other spaces from the selected space, and k the total number of spaces in the system; and then applying a correcting factor to eliminate the empirical effects of size (see Hillier and Hanson for details) and permit cross comparisons of systems of different sizes. The mean of all the integration values from spaces in a system will the precisely express the overall degree of integration or segregation in that system. The 10% (or 2% or 25%) of most integrating spaces will then form the 'integration core' of that system, and likewise for segregation. In real systems, integration values correlate highly with patterns of global movement, to an average of about .75.

Integration is a 'global' measure since it takes into account every other space in the system. 'Control value', on the other hand is a 'local' value, since it takes into account only the neighbours of a space and the neighbours of those neighbours. It expresses, again with a value varying about 1 (but with high values indicating strong control) how much better or worse connected a space is than its neighbours. It is calculated simply by summing the reciprocals of the valencies (numbers of connections) of the neighbours of a particular space:

$$\sum \frac{1}{C_N} \qquad (3)$$

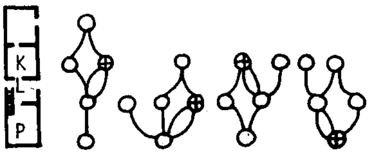
In effect, strong control spaces gain more than they give away, and vice versa for weak control spaces. In general, control values correlate very highly with local patterns of movement, to an average of about .87.

By applying integration and control measures to the subset of spaces that lie on at least one ring in the system - that is those which offer at least one alternative route to other spaces, in contrast to a ringless tree form where there is only one route from every space to every other space - one can then work out how a space relates to the route choices available in the system, that is, how far it integrates and controls those choices.

GENOTYPICAL THEMES AND PATTERNS

Differences in integration and control values from one space to another and from one part of the system to another, applied to the bounded, convex and axial patterns, are among the fundamental means by which social relations put their imprint on buildings. 'Differences' is the key word. Take the ordinary house for example, so often repressively interpreted in recent years as a set of relationally identical function-spaces linked to a central corridor. In every vernacular tradition of housing we have examined (see Hillier and Hanson 1982) the house is a set of spaces which relate differentially to the overall spatial pattern of the house. Each function is, in effect, recognisable from its unique set of integration and control characteristics. Function is what people do - but it is also a characteristic spatial pattern, and a characteristic part of an overall spatial pattern.

Fig.5 17th Century Banbury house



5a Plan | 5b From P RRA = 0.667 | 5c From L RRA= 0.167 | 5d From K RRA= 0.333 | 5e From ⊕ RRA= 0.333

Take an extremely simple example (from Wood-Jones 1963). Fig.5.a. shows a simple plan, then a - e show how different functions have different integration characteristics, both visually using the justified graph technique, and numerically using integration values. In fact this pattern in which, of all the daytime spaces, the parlour (P) integrates least, the main living area (L) most, and the kitchen (K) lies in between is a very common pattern, and underlies many different housing geometries. We have come to call this type of relational differentiation of the functions and categories located in different spaces within a plan an 'inequality genotype', because these important spatial differences can be expressed as numerical inequalities.

Now if such 'inequality genotypes' can be shown to exist across a sample of buildings within a particular cultural tradition - and this can obviously be simply by aggregating, taking means, testing the strength of differences across a sample as compared to individual cases, and so on - then it is reasonable to say not only that we have identified a cultural genotype by objective and numerical means, but also that we have done so by means of an analysis which is at once functional and spatial. We have simply looked at the relational spatialisation of those different functions. Once we had a theory of description, the problem in effect disappeared, and we were able to detect what we all know to exist anyway: the imprint of social relations and - dare we say it - even of the human mind on these extra-somatic organs we call buildings.

GLOBAL FUNCTIONS AND INTERFACES

But what about the overall, or 'global' pattern of space ? For this we need a new concept which we will call 'global function'. 'Global function' means something like the overall figure that characterises different 'types' of building: the 'churchness' of a church, 'schoolness' of a school, and so on - in effect the link between spatial form and building function that everyday language affirms by referring to both with the same word.

All such typical figures have, it seems, certain things in common. All building 'types' define two fundamentally different categories of people: a set of 'inhabitants' whose social identiy as individuals is durably recorded in the building form by control of space or a set of spaces; and a set of 'visitors' whose rights of presence in the building exist and distinguish them from the world of strangers, but not in a durable way as individuals and not through control of spaces. Family members in a dwelling, teachers in a school, medical professionals in a hospital are all inhabitants in this sense, while guests in a house, pupils in a school and patients in a hospital are all visitors. Note that length and constancy of occupation are not the criterion of membership. In many institutional buildings - asylums, prisons, and so on - visitors are much more permanently present than the inhabitants who escape at every opportunity.

In general, buildings can be defined in these terms as devices for making two kinds of interface: one between inhabitants and visitors, and the other between different categories of inhabitant. These interfaces are realised through some configuration of bounded, convex and axial space, and some set of relations of integration and control. These interfaces are what we mean by the global function of the building. It is this, we believe, that we name - at once socially and spatially - when we say 'school' or 'church' or 'hospital'.

It would seem then that if we could find a way to analyse and classify 'global functions' in spatial terms then we might be on a way to a theory of building form which linked the traditional notions of 'type' and 'function' is a new and systematic way. Unfortunately, although such an analysis can be very rewarding, we do not believe it can be completed in quite those terms. There are two main reasons for this. First, there are paradoxes and difficulties in the way of realising all but the simplest interfaces in perfect spatial form. For example, the preservation of statuses often requires segregation, while the parallel requirement to control others requires close proximity, and these contrary pulls can produce as great discontinuities within a type of building as between that type and others. Second, the effect of size is such as to in itself radically alter possibilities and even the desirability of realising social structures in spatial form. Buildings literally become quite different objects socially as they become large, and set us problems which are more the problems of large buildings in general, than the problems of a particular building type.

What follows, while sketchy, draws on several studies by ourselves, colleagues and students involving spatial analysis to identify genotypical patterns, field observations of space occupancy and movement, and computer simulation experiments on real buildings to see hpw space affects function theoretically. For the sake of continuity and clarity we will concentrate on 'medical interface' buildings, although our real purpose is to show general principles.

MEDICAL INTERFACES AND OTHERS

The simplest 'medical interface' is probably the common-or-garden doctor's surgery, often adapted from an ordinary house. It turns out that our initial example was such a building. Fig.6.a. shows the plan appropriately labelled, while 6.b. and 6.c. show how different the building is from the doctors' and the patients' spaces (the latter being the waiting room and the outside):

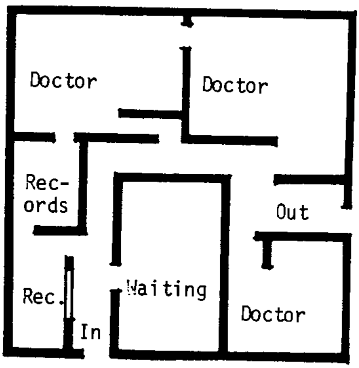


Fig.6a Plan of doctors' surgery

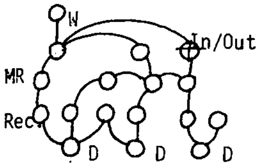


Fig.6b
Justified graph from the doctors' spaces, RRA= .475

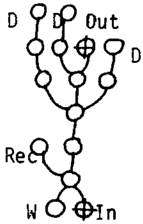


Fig.6c
Justified graph from patients' spaces,RRA=1.693

These differences arise largely from the very strong 'genotypical' spatial requirements that this type of interface commonly imposes. There are four main elements:

- the patients have to be held in an easily controlled space shallow in the building;
- the entrances to doctors' room have to be separated from the entrance to the patients' space by some kind of distancing devicing, usually an axial discontinuity of some kind;
- contact between doctors and between them and other inhabitants has to be possible without the possibility of accident contact with patients;
- for the same reason, doctors have to have a separate entrance and independent routes to their part of the building - a common device for preserving status which we call the 'stage door effect'.

All of these requirements can be stated and realised in terms of the spatial elements and relations we have defined. Collectively they have a very powerful effect that makes the building very unlike the house that it adapted. It becomes a building with a very strong programme, in the sense that virtually everything that occurs is specified and inscribed into the spatial pattern. Very little that is unprogrammed can occur. The strength of the genotypical model is such that if one or other element is missing from the spatial organisation, then in all likelihood it will be compensated for by some bevavioural practice or device.

The four elements of the model really come down to two general 'interface' requirements: the need to preserve the status and solidarity of the main 'inhabitants'; and the need to control the movement of patients. Insofar as the doctors' status is to be conserved by segregation and depth into the building, the patient must circulate into the building; but insofar as the patient is to be controlled then the circulation from his entrance is controlled by the receptionist. The same circulation space thus works in opposite ways for doctor and patient. For the doctor, it is his means of distancing; for the patient it is his means of control. For the doctor, the circulation is for the patient to travel to his segregated domain. For the patient, it is the means by which he is spatially restricted and confined to a room until his time comes. The relations of room to circulation are, as it were, reversed for patients and doctors. As we will see, under some conditions this reversal becomes a major theme in medical - as in other - interfaces, bringing about major discontinuities in the building form.

Fig. 7. is a larger medical interface building. selected, we admit, to illusrate a point. At first glance we might take it for just another building with a more extended circulation system appropriate for its size. In fact a more carefull examination will show that its global form exactly and only reproduces the four genotypical elements that we detected in the doctors' surgery. The geometry of the layout is quite different, but the 'syntactic' (all our spatial parameters) principles are identical. The only substantial effect of size would seem to be the duplication of the waiting area and the removal of its boundary. In this sense, the building increases the control of the patient and the 'reversal' effect, in that, locally at least, the function of the circulation system is to create an area where patients can be both static and surveyed:

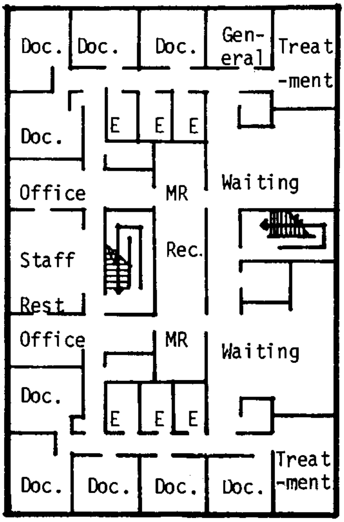


Fig.7 Plan of a 'category' type health centre.

Fig.8. is another kind of larger medical interface building: the out-patients department of a famous modern hospital, selected again, we admit, to illustrate a point. At first sight it seems that we must give up our theory of the genotypical elements of the medical interface, since none at all seem to be present to any degree in this plan. Patients are spread through the building from shallow to deep; there is no distancing of doctors' rooms - on the contrary, they interface directly with the patient areas; there is no separate circulation system, apart from links from one room to its neighbour; and there are no stage door effects to speak of.

But a moment's thought about the social ideas underlying the genotype suggests another possibility. Surely what has happened in this case is that the genotypical elements that have to do with the distancing and re-integration of doctors as professionals has been dropped, and those that have to do with the control of patients have taken over. The increase in axial organisation, so that the plan becomes very shallow from the outside and most points inside, the direct interface with doctors rooms, the splitting of patients into dispersed groups could all be thought to have this effect. If in the previous case the global pattern of circulation was defined by the issue of professional status, in this case it is dominated by the pragmatics of control. This is what we mean by a morphological discontinuity. The two types of solution cohere as part of a general theoretical model; but as actual spatial types the are utterly unlike each other.

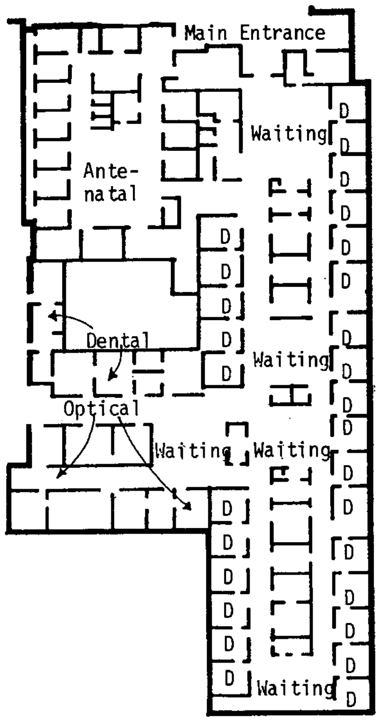


Fig.8 Hospital outpatients department

Fig.9. is then a more recent solution to the same problem: the OPD of one of London's most recent hospitals:

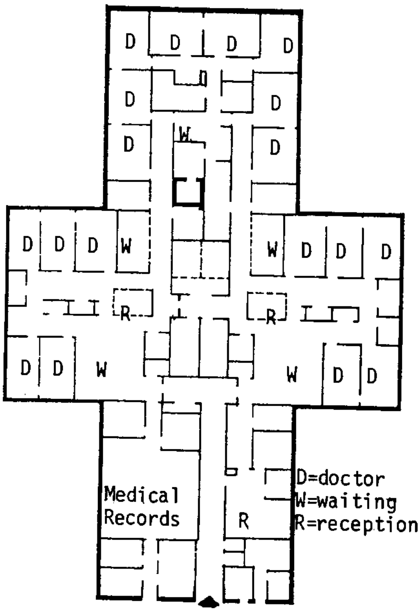


Fig.9 Hospital outpatients department

Once again we seem to have a morphological discontinuity. The original genotypical elements are not there to any degree, but nor are the principles of the first OPD. In fact, the very contrary principles seem to be at work. The main effect of this rather labyrinthine spatial planning seem to be to segregate the patients, rather than the doctors, and to separate them from the outside as strongly as they are separated from each other. A greater contrast both in working and in intuitive feeling with the first OPD can hardly be imagined. Do these feeling have an objective basis ? And does this have anything to with the relation between spatial form and function in buildings?

The answer to this question is difficult to formulate, but is nevertheless, we believe, vital to our understanding of large buildings complexes and why some don't seem to work as well as others. It runs something like this. As buildings grow larger, it becomes more and more difficult to maintain them as 'strong programme' buildings, that is buildings where most of what happens is specified by explicit or tacit rules, and built into the spatial structure of the building. There is a simple reason for this. As the numbers of people and the number of spaces necessary to accomodate them increase, so the amount of unprogrammed contact as the natural by-product of functionally defined movement is also likely to increase. The same effect is strong in settlements as they become larger, and indeed the liveliness that we often find in settlments is often largely a by-product of this process.

Our analyses of settlment form, together with our observations and simulations of them, have suggested in fact that settlements take the spatial form they do largely in order to generate, control and render predictable the unprogrammed by-products of complex patterns of spatial activity and movement. Settlements, being on the whole 'weak programme' spatial complexes use the natural laws that relate spatial parameters to movement probabilities to create a kind of order in what might otherwise be an unmanageably complex and unintelligible social and spatial environment. The laws of space literally become the basis for a certain kind of informal order to take over where social programmes are weak. In settlements, all that is predictable in the general spatial distribution of people, and that is predictable, intuitively as well as formally, from the spatial pattern.

We believe that large buildings also have this character and this problem. Those that solve it by using the global structure of the building to generate and control a pattern of movement and potential encounter over and above the specific and localised functional programmes acquire a kind of spatialised culture by which we begin to recognise them as institutions. UCL itself is an example of such a complex, and so is the AA. Those that pursue localised function at the expense of such a global patterning will not generate such a spatial culture, and will be experienced as fragmentary institutions.

This may sound like subjective evaluation, but we believe that our work on settlement forms and more recent work on buildings (especially an admirable study by a current MSc student of ours, Alan Penn) can begin to provide an objective basis for such views. Take for example the axial maps of the two OPD's: (Figs 10+11 over page). The integration analysis of the first shows an 'integration core' that links the outside to the inner recesses of the building, and gives the plan a strong global structure. The second, on the other hand, has a core which is both deep in the plan, and limited to its central areas. Bearing in mind that integration values are by far the most powerful determinants of global movement patterns, it is clear that all categories of user are going to be pressed towards relative immobility by the structure of space in the second case, whereas in the first they will be drawn naturally into global patterns of movement and encounter. The second may be better functionally at the local level. But it achieves this at the expense of building the global institution. The first may be functionally more difficult to manage, but it will contribute much more to the global institution.

These suggestions are confirmed powerfully by an examination of the whole floor hospital plans of which the two OPD's are a part. In the first case, it can be seen that the interior of the OPD itself forms a significant part of the integrating core, a core which forms a strong grid reaching deep into the building. In the second case, the core has a simple tree form based on a very powerful single main 'street', and in this case the OPD - a major source of moving visitors to the building - is in a highly segregated zone, with no penetration at all from the core (see next page for Figs.12 and 13). We have clue to believe that these powerful - and objective - differences in spatial structure, which are rendered visible by a theory of description, are powerfully implicated in the utterly different impressions that these two hospitals make both as functioning organisations and as spatial institutions.

What we are saying, in short, is something like this. In a small house, the role of space is not so much to reflect everyday functional patterns, but to add something to those patterns that would not otherwise be there. Because we have a best room, we do not need to go there, and we relax more. We have the space as a resource when we need it. Space adds a degree of extra culture to life, and life is the more pleasant because of it. At a rather larger scale, we find the 'strong programme' building, where everything is specified, and where typological-functional analysis is possible on the basis of interfaces and morphological discontinuities and social strategies. But with the large building we find a new phenomenon. Again space is adding something to function, as with the house, but here it is not adding the structure of a given culture, but something like the opposite. It is adding the generation of a social field that is unstructured, but which, like a settlement, acquires a predictability and a reproducibility - and therefore a social identity - through adapting its spatial organistion to nurture and organise this emergent phenomenon. The true function of large building complexes in our time is, we believe, to create these emergent social organisms.

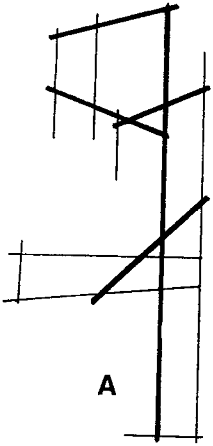


Fig.10 Axial map of Fig. 8 with 'integrating core' marked by thick lines

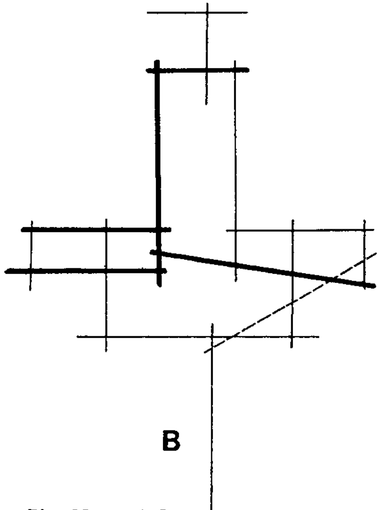


Fig.11 Axial map of Fig.9 with 'integrating core' marked by thick lines

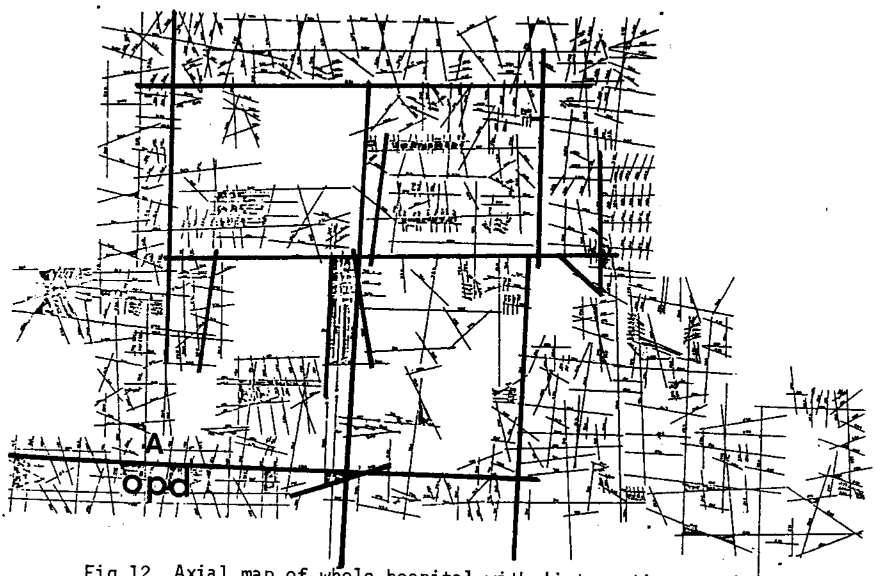


Fig.12 Axial map of whole hospital with 'integrating core' marked

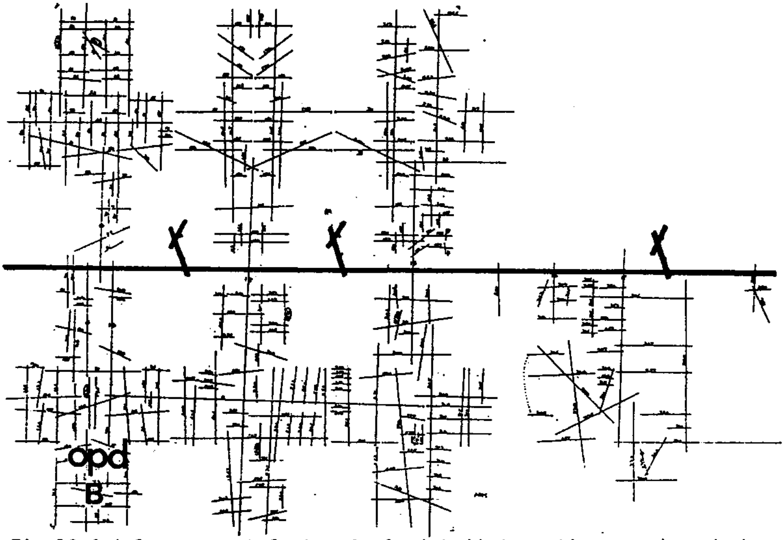


Fig.13 Axial map of whole hospital with 'integrating core' marked

REFERENCES

Frankl P. Principles of Architectural History : the four phases of architectural style, 1420-1900: MIT Press, 1973, translated from the German edition of 1914

Hillier B. and Hanson J. Discovering Housing Genotypes: working paper of the Unit for Architectural Studies, University College London, 1982

Hillier B. and Hanson J. The Social Logic of Space: Cambridge University Press, forthcoming 1983

Rowe C. The Mathematics of the Ideal Villa and other Essays: MIT Press, 1976

Wood Jones R. Traditional Domestic Architecture in the Banbury Region: Manchester University Press, 1963