Space is the machine

Bill Hillier

Space Syntax

Since *The social logic of space* was published in 1984, Bill Hillier and his colleagues at University College London have been conducting research on how space features in the form and functioning of buildings and cities. A key outcome is the concept of 'spatial configuration' – meaning relations which take account of other relations in a complex. New techniques have been developed and applied to a wide range of architectural and urban problems. The aim of this book is to assemble some of this work and show how it leads the way to a new type of theory of architecture: an 'analytic' theory in which understanding and design advance together. The success of configurational ideas in bringing to light the spatial logic of buildings and cities suggests that it might be possible to extend these ideas to other areas of the human sciences where problems of configuration and pattern are critical.

Space is the machine
Bill Hillier

A configurational theory of architecture

Space Syntax

‘A house is a machine for living in...’
Le Corbusier (1923)

‘But I thought that all that functional stuff had been refuted. Buildings aren’t machines.’ *Student*

‘You haven’t understood. The building isn’t the machine. Space is the machine.’ *Nick Dalton, Computer Programmer at University College London (1994)*

Space Syntax

UCL

Hardback and paperback editions first published in the United Kingdom in 1996 and 1999, respectively, by the Press Syndicate of the University of Cambridge.

This electronic edition published in 2007 by:

Space Syntax
4 Huguenot Place, Heneage Street
London E1 5LN
United Kingdom

www.spacesyntax.com

ISBN 978-0-9556224-0-3

Layout and design
by Christian Altmann
Set in Haas Unica

Contents

Preface to the e-edition

Space is the Machine was first published in 1996 by Cambridge University Press. The book built on the theory of society and space set out in *The Social Logic of Space* (Cambridge University Press 1984), to outline a configurational theory of architecture and urbanism. Unfortunately, although *The Social Logic of Space* is still in print after 23 years, when the initial print-run of *Space is the Machine* was exhausted, the number of colour plates forbad the use of the cheap reprinting technology that would have made a succession of reprints economically viable. So, although the book was selling well at the time, it fell out of print. As demand for the book has continued, for several years copies of the book have either been impossible to find or prohibitively expensive.

I am now immensely pleased that Space Syntax Limited, with support from University College London (UCL), have decided to rectify this situation by creating a new e-edition of the book and making it available for free on the web. I am particularly grateful to Tim Stonor for the initial decision to fund the project, to Tim, Chris Stutz and Shinichi Iida for organizing and managing the project and acting as effective editors of the new edition, and to Laura Vaughan and Suzanne Tonkin of UCL for their encouragement throughout the process. Thanks and appreciation are also due to Christian Altmann for the new design of the publication; to Rodrigo Mora for preparing electronic images from the original artworks; to Marco Gandini, Joseph Laycock, Sacha Tan, and Saussan Khalil for proofreading; to Molly Hall for creating a new index; and to Christian Beros for image manipulation and creation of the web distribution pages for the e-edition.

Looking back on *Space is the Machine*, as on *The Social Logic of Space*, I find myself pleasantly surprised that the foundations set out there for the 'space syntax' approach to human spatial phenomena still seem robust. At the same time, the developments in the subject since 1996 have been substantial, not least through the inauguration of the bi-annual space syntax symposia in 1997 (originally the brainchild of Mark David Major). These have created a resource of several hundred papers on developing the theory, methodology and applications of the space syntax approach and now constitute one of its most important resources. To me personally, it is most gratifying that a set of ideas created by a small group of people working together at UCL in the nineteen seventies has now flowered into a large and coherent body of work belonging to a world-wide research community.

At the risk of being unfair to others, however, it does seem to me that certain contributions to the theory and method of space syntax have been so significant as to deserve review in this preface to what is now an eleven-year-old text. For example, on the theoretical foundations of space syntax, the three papers published by John Peponis and his colleagues of the Georgia Institute of Technology in Atlanta in *Environment and Planning B* in 1997 and 1998 (Peponis et al 1997, Peponis et al 1998a, b) on the geometrical foundations seems of permanent significance, as do the two papers of Mike Batty of CASA at UCL (Batty 2004a, b) on the graph theoretic

Preface to the e-edition

foundations. I would also hope that my own attempts to show that the effects on ambient space of the placing and shaping of physical objects are systematic and can be mathematically expressed will prove similarly robust. The importance of these effects both for the understanding of urban form (Hillier 2002), and human spatial cognition (Hillier 2007) will, I hope, lead to a more unified understanding of the link between these two realms.

On the methodological side, there has been a remarkable flourishing of new syntactic methods from many sources and locations. From UCL, the most significant of these have been the 'syntacticising' of visibility graph analysis by Alasdair Turner in his Depthmap software (Turner & Penn 1999, Turner et al 2001) and the development of segment based axial analysis with angular, metric and topological weightings, initially through the pioneering work of Shinichi Iida and his Segmen software with subsequent implementation in Depthmap. It was these more complex and disaggregated forms of line analysis that allowed us to show not only that human movement was spatially guided by geometrical and topological rather than metric factors but also to clarify why a powerful impact of space structure on movement was to be mathematically expected (Hillier & Iida 2005). Other key methodological developments include the pioneering work of Dalton on angular analysis (Dalton 2001), now available in the WebMap and WebMapatHome software; the work of Figueiredo and Amorim of the University of Pernambuco in Brazil on 'continuity lines' in the Mindwalk software (Figueiredo & Amorim 2005), which extend lines by discounting angular changes below a certain threshold; and the Spatialist software development by Peponis and his colleagues in connection with the three papers referred to above. Other significant software developments focus on linking space to other urban factors such as land use patterns and densities, notably the Place Syntax software from Marcus and his colleagues at the Royal College of Technology in Stockholm, Sequence software developed by Stegen at ARSIS in Brussels, and the Confeego software pioneered by Stutz, Gil, Friedrich and Klaasmeyer for Space Syntax Limited.

In the more substantive areas of theory, my own research has explored the inter-relations of space, movement at different scales and land use patterns, and it can now arguably be seen to be pointing in the direction of a design-level (meaning precise enough for the ideas to be usable in design) theory of cities as self-organising systems. The theory is in two parts: on the one hand, a theory of how the spatial form of cities is shaped by spatial laws linking the emergence of characteristically urban space patterns to cognitive as well as to social and economic factors; on the other, a theory of how the emergent patterns of space shape movement, and through this shape land use patterns, leading through feedback and multiplier effects, to the generic form of the city as a foreground network of linked centres at all scales set into a background network of largely residential space. Critical to the emergence of this theory was the paper "Centrality

Preface to the e-edition

as a process" (Hillier 1999) which showed how local processes with an essentially metric nature combined with the larger scale geometric and topological properties of the spatial network to create the processes by which centres and sub-centres emerge in the network through the logic of the network itself - though each is of course also affected by its relation to others.

Taken together these developments in space syntax suggest that it offers a powerful complement to traditional methods for modelling cities, not least transport modelling methods. These have conceptual foundations quite different from syntactic models and seek to explain different things, but they could be brought into a symbiotic relation with syntactic models to the benefit of both. A key research priority in the immediate future will be to explore their inter-relations. In fact, following the pioneering work of Penn on the configurational analysis of vehicular movement (Penn et all 1998) work by Chiaradia, Raford and others in Space Syntax Limited has already suggested that configuational factors can contribute insights into other kinds of movement networks, including cycles, buses, and overground and underground rail networks.

One aspect of the deepening relation between space syntax and the wider spatial research community has been the debate as to how far space syntax's basic tenets, such as the representation of cities as line networks and the setting aside of Euclidean metric factors at the larger spatial scale in favour of topological and/or geometric ones, are theoretically valid and methodologically viable. From the syntactic point, certain points of criticism, such as that axial maps are 'subjective' and measures should be metricised, seem to have been answered. Turner et al (2005) have showed that least line graphs (allowing random selection among syntactically equivalent lines) are rigorously defined and indeed are objects of great theoretical interest in themselves, as is shown by recent work suggesting they have fractal properties (Carvalho & Penn 2004). Likewise the criticism that syntax disregards metric information has been answered by showing clearly that in terms of functionality this is a scale issue. As shown in (Hillier 1999) referred to above, at a sufficiently localised scale space works in a metric way, perhaps reflecting the scale up to which people can make reasonably accurate judgement about distance in complex spaces, so an account of the metric properties of space is necessary to a functionally sensitive and predictive analysis of space at this level. But at the non-local level, it seems that the functionality of space reflects people's use of a geometrical picture of the network connectivity rather than a metric picture in navigating the urban grid, and at this scale introducing metric weighting into the measures is positively misleading (Hillier et al 2007).

The study of space within buildings using space syntax methods has also much advanced since 1996, not least of course through the publication of Julienne Hanson's *Decoding Homes and Houses* (1999), the third of the syntax books from Cambridge University Press. Also notable has been the work of Penn

Preface to the e-edition

and his colleagues on spatial form and function in complex buildings, in particular the influential work on spatial design and innovation in work environments. Although not strictly within the syntax context, the highly original work of Steadman (Steadman 1998, 2001) on the enumeration of built forms through a clarification of geometric, constructional and environmental constraints both answers questions about enumerability raised in *Space is the Machine*, and offers a platform for a new approach to spatial enumerability which could and should be taken up within the syntax community.

Against the background of these theoretical and methodological developments, and cross-disciplinary exchanges, space syntax research is now becoming much more interdisciplinary. Following a special issue of *Environment and Behaviour* in 2003 edited by Ruth Conroy Dalton and Craig Zimring bringing together papers on space syntax and cognition from the 2001 Atlanta Symposium, the 2006 conference on Spatial Cognition at the University of Bremen organised a well-attended all day workshop on space syntax. The link between space syntax and cognitive studies is now becoming a well-established branch of syntax research. At the same time the pioneering work of Laura Vaughan and her colleagues is taking syntax in the direction of a greater engagement with social studies, and a special issue of *Progress in Planning* will shortly appear on the use of space syntax in the study of space as a dimension of social segregation and exclusion (Vaughan (ed.) 2007).

Overall, space syntax is becoming a flourishing paradigm for spatial studies, increasingly well integrated with other approaches and increasingly expanding its scope and scale of investigation. But the real test of theory and method is application in the real world of projects and development. Here the contribution of Space Syntax Limited cannot be overestimated. Since its foundation as an active company offering spatial design and spatial planning consultancy under the leadership of Tim Stonor, it has tested the theory and technology on a wide range of projects, many of them high profile. There are now a significant number of projects in which Space Syntax has exerted a key spatial design influence, including of course in the redesign of Trafalgar Square (with Norman Foster) and Nottingham's Old Market Square (with Gustafson Porter), arguably the two most famous squares in the UK, both now functioning in a new and highly successful way following their respective re-designs. Other up and running projects include the Brindley Place development in Birmingham, Exchange Square and Fleet Place in London, and of course the Millennium Bridge, where Space Syntax showed not only how well the bridge would be used but also how strong and beneficial its long term effects would be on the areas on both sides of the river. Equally interesting to space syntax are cases where aspects of space syntax advice was not followed, since in each case problems have appeared that were clearly foreseen by syntax at the design stage.

Preface to the e-edition

Carefully and responsibly used, it is clear that syntax works as a design and planning tool. One consequence of its success in relatively small-scale design and planning problems is that syntax is now increasingly being used as the foundation for the space-based master-planning of whole parts of cities or even of whole cities, and so in effect as a new way of modelling cities. It is increasingly well understood that a syntactic model of a city has two great advantages as a complement to an orthodox model. First, a syntactic model allows the designer or planner to work across all urban scales using the same model, so that one form of analysis will identify the large scale movement networks and its land use effects, while another will similarly identify micro-scale features and land use potentials of the local urban grid. Second, exactly the same model that is used in research mode to investigate and understand how the city is working now can be used in design and planning mode to simulate the likely effects of different design and planning strategies and schemes, allowing the rapid exploration of the long term consequences of different strategies.

Space Syntax Limited also constitutes an experiment in how the relations between a university and a spin-out company can be organised. Although Space Syntax Limited carries out its own research, it maintains a very close relation to the university research department, feeding problems into it and testing new ideas and new technologies. Collaboration is both at the strategic research level, but also reaches down to the level of individual projects where necessary. The experience of a working collaboration between the university and the company has convinced us all that in this field even the most basic research cannot be separated from the demands and questions raised by practice. Many theoretical developments have been sparked by questions raised by projects, and at the same time projects have provided a superb early testing ground for turning research ideas into workable and proven technologies. The fact that it is Space Syntax Limited which is now re-publishing one of the basic theoretical texts of space syntax is an emblem of the closeness with which theory and practice, and the university and the commercial world, have developed collaboratively over the past decade.

bh
June 6th 2007

Preface to the e-edition

References

Batty, M. (2004) A New Theory of Space Syntax, Working Paper 75, Centre for Advanced Spatial Analysis, UCL, London: available from WWW at*http://www.casa.ucl.ac.uk/working_papers/paper75.pdf*

Batty, M. (2004) Distance in Space Syntax, Working Paper 80, Centre for Advanced Spatial Analysis, UCL, London: available from WWW at *http://www.casa.ucl.ac.uk/working_papers/paper80.pdf*

Carvalho, R., Penn, A. (2004) "Scaling and universality in the micro-structure of urban space." Physica A **332** 539-547.
Dalton, N. (2001) "Fractional configuration analysis and a solution to the Manhattan Problem." Proceedings of the 3rd International Space Syntax Symposium, Georgia Institute of Technology, Atlanta, GA, 7-11 May 2001.

Figueiredo, L., Amorim, L. (2005) "Continuity lines in the axial system." Proceedings of the 5th International Space Syntax Symposium, Technische Universiteit Delft, the Netherlands, 13-17 June 2005.

Hanson J (1999) *Decoding Homes and Houses* Cambridge University Press, Cambridge.

Hillier, B. (1999) "Centrality as a process: accounting for attraction inequalities in deformed grids." *Urban Design International* **4** 107-127.

Hillier, B. (2002) "A theory of the city as object." *Urban Design International* **7** 153-179.

Hillier, B., Iida, S. (2005) "Network and psychological effects in urban movement." In Cohn, A. G., Mark, D. M. (eds) *Spatial Information Theory: COSIT 2005*, Lecture Notes in Computer Science number 3693, 475-490, Springer-Verlag, Berlin.

Hillier, B. (2007) "Studying cities to learn about minds: how geometric intuitions shape urban space and make it work." *Environment and Planning B: Planning and Design* (forthcoming).

Hillier, B., Turner, A., Yang, T., Park, H. T. (2007) "Metric and topo-geometric properties of urban street networks: some convergences, divergences and new results" Proceedings of the 6th International Space Syntax Symposium, ITU, Istanbul, Turkey, 12-15 June 2007.

Preface to the e-edition

Penn, A., Hillier, B., Banister, D., Xu, J. (1998) "Configurational modelling of urban movement networks" *Environment and Planning B: Planning and Design* **25** 59-84.

Penn, A. and Desyllas, J. and Vaughan, L. (1999) "The space of innovation: interaction and communication in the work environment" *Environment and Planning B: Planning and Design*, 26 (2) 193-218

Peponis, J., Wineman, J., Rashid, M., Kim, S. H., Bafna, S. (1997) "On the description of shape and spatial configuration inside buildings: convex partitions and their local properties." *Environment and Planning B: Planning and Design* **24**(5) 761-781.

Peponis, J., Wineman, J., Bafna, S., Rashid, M., Kim, S. H. (1998) "On the generation of linear representations of spatial configuration." Environment and Planning B: Planning and Design 25(4) 559-576.

Peponis, J., Wineman, J., Rashid, M., Bafna, S., Kim, S. H. (1998) "Describing plan configuration according to the covisibility of surfaces." *Environment and Planning B: Planning and Design* **25**(5) 693-708.

Steadman, J. P. (2001) "Every built form has a number." Proceedings of the 3rd International Space Syntax Symposium, Georgia Institute of Technology, Atlanta, GA, 7-11 May 2001.

Steadman, J. P. (1998) "Sketch for an archetypal building." *Environment and Planning B: Planning and Design, 25th Anniversary Issue*, 92-105.

Turner, A., Penn, A. (1999) "Making isovists syntactic: isovist integration analysis." Proceedings of the 2nd International Space Syntax Symposium, Universidade de Brasília, Brasilia, Brazil, 29 March – 2 April 1999.

Turner, A., Doxa, M., O'Sullivan, D., Penn, A. (2001) "From isovists to visibility graphs: a methodology for the analysis of architectural space." *Environment and Planning B: Planning and Design* **28**(1) 103-121.

Turner, A., Penn, A., Hillier, B. (2005) "An algorithmic definition of the axial map." *Environment and Planning B: Planning and Design* **32**(3) 425-444.
Vaughan, L. (ed.) (2007) Progress in Planning (The Spatial Syntax of Urban Segregation) **67**(4) (in press).

Acknowlegdements

Acknowledgements and thanks are due first to the many friends and colleagues who have, over the years, made an enormous contribution to the ideas and research set out in this book, most notably Dr. Julienne Hanson, Alan Penn and Dr. John Peponis, each of whose contributions has been too large and diverse to acknowledge in detail; to Nick 'Sheep' Dalton for the title of the book, and also for the brilliant software on which much of the research is founded, the outward and visible sign of which is in the Plates in the book; to Mark David Major for masterminding the gradual and painful evolution of the text and illustrations; to Myrto-Gabriella (Petunia) Exacoustou for reading, criticising and helping me substantially improve the final draft; to Professor Pat O'Sullivan, Head of the Bartlett, for a six months sabbatical in 1992, when I said I would finish but he knew I wouldn't; to the Engineering and Physical Sciences Research Council for continued research funding; to the many contributors to the research, and especially Tim Stonor, Kayvan Karimi, Beatriz de Campos, Xu Jianming, Gordon Brown, John Miller, Tad Grajewski, Lena Tsoskounoglou, Laura Vaughan, Martine de Maesseneer, Guido Stegen and Chang Hua Yoo (who drew the first version of the map of London); to the MSc and doctoral students of the Bartlett School of Graduate Studies who continue to give so much intellectual buzz to the department; to Professors Philip Steadman, Tom Markus and Mike Batty for sustained intellectual support over the years; to Stuart Lipton and his team at Stanhope Properties Plc for providing us with so many opportunities to apply our research on real development and design projects, and to learn so much from them, and also to Gordon Graham of the London Regeneration Consortium, the South Bank Employers Group, Chesterfield Properties Plc, Ove Arup and Partners, and Peter Palumbo; to the many public bodies who have invited us to contribute to their work through applied research, including National Health Service Estates, British Railways, British Airways, Powergen, the Department of Education and Science, Technical Aid for Nottingham Communities, the London Boroughs of Croydon and Camden, and the Tate Gallery; to the many architectural practices who have invited us to work with them on their projects, but most especially Sir Norman Foster and Partners, the Richard Rogers Partnership, Terry Farrell and Company, Skidmore, Owings and Merrill, Nicholas Grimshaw and Partners, Bennetts Associates, SW Architects, and Avanti Architects; to Professor Sheila Hillier and Martha Hillier for tolerating an obsessive lap-topper in the house for longer than was reasonable; to Kate, Charlotte and Ben Hillier for continuing to be the good friends and supporters of a preoccupied and inconsiderate father; to the thief who took all copies of the draft of the first four chapters from my home when stealing my computer, thus saving me from premature publication; to Rose Shawe-Taylor, Karl Howe, Emma Smith, Susan Beer and Josie Dixon of Cambridge University Press; and finally to UCL for continuing to be the most tolerant and supportive of Universities.

Introduction

In 1984, in *The Social Logic of Space*, written in collaboration with Julienne Hanson and published by Cambridge University Press, I set out a new theory of space as an aspect of social life. Since then the theory has developed into an extensive research programme into the spatial nature and functioning of buildings and cities, into computer software linking 'space syntax' analytic tools with graphical representation and output for researchers and designers, and into an expanding range of applications in architectural and urban design. During this time, a large number of articles, reports and features have appeared, theses have been written in many universities using the theory and methods of 'space syntax', and research has been initiated in many parts of the world into areas as diverse as the analysis of archaeological remains and the design of hospitals.

During this time, many theoretical advances have also been made, often in symbiosis with the development of new techniques for the computer representation and analysis of space. One key outcome of these advances is that the concept of 'configuration' has moved to centre stage. Configuration means, put simply, relations taking into account other relations. The techniques of 'configurational analysis' - of which the various 'space syntax' techniques are exemplars - that have been built from this idea have made it possible to bring the elusive 'pattern aspect' of things in architecture and urban design into the light of day, and to give quantitative expression to the age-old idea that it is 'how things are put together' that matters.

This has in turn led to a clear articulation of a philosophy of design. Architectural and urban design, both in their formal and spatial aspects, are seen as fundamentally configurational in that the way the parts are put together to form the whole is more important than any of the parts taken in isolation. The configurational techniques developed for research can, in fact, just as easily be turned round and used to support experimentation and simulation in design. In linking theoretical research to design in this way, we are following a historical tradition in architectural theory which has both attempted to subject the pattern aspect of things in architecture to rational analysis, and to test these analyses by embodying them in real designs. The difference now is only that the advent of computers allows us to bring a much great degree of rigour and testing to theoretical ideas.

The aim of this book is to bring together some of these recent developments in applying configurational analysis to issues of architectural and urban theory into a single volume. The surprising success of configurational ideas in capturing the inner logic of at least some aspects of the form and functioning of built environments, suggests that it might in due course be useful to extend these ideas to other areas where similar problems of describing and quantifying configuration seem to be central, including some aspects of cognitive psychology, but also perhaps sociology itself. At present we are encouraged by the current interest in these ideas across a range of disciplines and, just as the last decade has been devoted to the development and testing of techniques of configurational

Introduction

analysis within architecture and urban design, so we hope that the coming decade will see collaborations amongst disciplines where configuration is identified as a significant problem, and where some development of the configurational methodology could conceivably play a useful role.

The immediate context of the book is the changing theoretical debate within and around architecture. Looking back, it is easy to see that in spite of the attention paid to theory in architecture in the twentieth century, and in spite of the great influence that theories have had on our built environment, architectural theories in the last decades have in general suffered from two debilitating weaknesses. First, most have been strongly normative, and weakly analytic, in that they have been too much concerned to tell designers how buildings and environments should be, and too little concerned with how they actually are. As a result, theories of architecture have influenced our built environment enormously, sometimes for good, sometimes for ill, but they have done little to advance our understanding of architecture.

Second, there has been an explosion of the historic tendency to form architectural theories out of ideas and concepts borrowed from other disciplines. As a result, architectural discourse has been dominated by a series of borrowings, first from engineering and biology, then from psychology and the social sciences, then from linguistics and semiology, and most recently of all from literary theory. Each of these has had the merit that it allowed architecture to become part of wider intellectual debate. But there has been a price, in that very little attention has been given to the internal development of architecture as a discipline. Through this turning away, architecture has increasingly ignored the lessons waiting to be learned from the intensive study of experimental twentieth-century architecture, and acquired what now amounts to a hidden history in which key aspects of recent architectural reality have been suppressed as though they were too painful to talk about.

The aim of this book is to begin the process of remedying this bias towards overly normative theories based on concept borrowing from other disciplines, by initiating the search for a genuinely analytic and internal theory of architecture, that is, one based on the direct study of buildings and built environments, and guided by concepts formed out of the necessities of this study. The guiding belief is that what we need at the end of the twentieth century is a better and deeper understanding of the phenomenon of architecture and how it affects people's lives, and how this relates to innovative possibility in architecture, and the central role of the architectural imagination.

This book is therefore concerned with what buildings and cities are like, why they are as they are, how they work, how they come about through design, and how they might be different. The word 'theory' is used not in the common architectural sense of seeking some set of rules which, if followed, will guarantee architectural success, but in the philosophical and scientific sense that theories are the abstractions through which we understand the world. An architectural theory, as we see it, should deepen our grasp of architectural phenomena, and only subsequently and with great modesty, suggest possible principles on which

Introduction

to base speculation and innovation in design. Such a theory is analytic before it is normative. Its primary role is to enquire into the puzzle that we see and experience architecture, but we do not understand what we see and experience. However strongly we may feel that architecture may be wrong or right, we rarely understand the architectural grounds on which such judgments are made. This book therefore seeks an understanding of the theoretical content of architecture.

The book is in four parts. The first, *'Theoretical Preliminaries'*, deals with the most basic of all questions which architectural theory tries to answer: what is architecture, and what are theories, that they can be needed in architecture? In the first chapter, 'What architecture adds to building', the key concepts of the book are set out on the way to a definition of architecture. The argument is that in addition to functioning as bodily protection, buildings operate socially in two ways: they constitute the social organisation of everyday life as the spatial configurations of space in which we live and move, and represent social organisation as physical configurations of forms and elements that we see. Both social dimensions of building are therefore configurational in nature, and it is the habit of the human mind to handle configuration unconsciously and intuitively, in much the same way as we handle the grammatical and semantic structures of a language intuitively. Our minds are very effective in handling configuration in this way, but because we do work this way, we find it very difficult to analyse and talk rationally about the configurational aspects of things. Configuration is in general 'non-discursive', meaning that we do not know how to talk about it and do not in general talk about it even when we are most actively using it. In vernacular buildings, the configurational, or non-discursive, aspects of space and form are handled exactly like the grammar of language, that is, as an implication of the manipulation of the surface elements, or words and groups of words in the language case, building elements and geometrical coordinations in building. In the vernacular the act of building reproduces cultural given spatial and formal patterns. This is why it seldom seems 'wrong'. Architecture, in contrast, is the taking into conscious, reflective thought of these non-discursive and configurational aspects of space and form, leading to the exercise of choice within a wide field of possibility, rather than the reduplication of the patterns specific to a culture. Architecture is, in essence, the application of speculative and abstract thought to the non-discursive aspects of building, and because it is so, it is also its application to the social and cultural contents of building.

Chapter 2, 'The need for an analytic theory of architecture', then takes this argument into architectural theory. Architectural theories are essentially attempts to subject the non-discursive aspects of space and form to rational analysis, and to establish principles to guide design in the field of choice, principles which are now needed as cultural guidance is no longer automatic as it is in a vernacular tradition. Architectural theories are both analytic in that they always depend on conjectures about what human beings are like, but they are also normative, and say how the world should be rather more strongly than they say how it is. This means that architecture can be innovative and experimental through the agency of theories, but

Introduction

it can also be wrong. Because theories can be wrong, architects need to be able to evaluate how good their theories are in practice, since the repetition of theoretical error - as in much of the modernist housing programme - will inevitably lead to the curtailment of architectural freedom. The consequence of this is the need for a truly analytic theory of architecture, that is, one which permits the investigation of the non-discursive without bias towards one or other specific non-discursive style.

Chapter 3, 'Non-discursive technique', outlines the prime requirement for permitting architects to begin this theoretical learning: the need for neutral techniques for the description and analysis of the non-discursive aspects of space and form, that is, techniques that are not simply expressions of partisanship for a particular type of configuration, as most architectural theories have been in the past. The chapter notes a critical difference between regularities and theories. Regularities are repeated phenomena, either in the form of apparent typing or apparent consistencies in the time order in which events occur. Regularities are patterns in surface phenomena. Theories are attempts to model the underlying processes that produce regularities. Every science theorises on the basis of its regularities. Social sciences tend to be weak not because they lack theories but because they lack regularities which theories can seek to explain and which therefore offer the prime test of theories. The first task in the quest for an analytic theory of architecture is therefore to seek regularities. The first purpose of 'non-discursive technique' is to pursue this task.

Part II of the book, *'Non-discursive Regularities'*, then sets out a number of studies in which regularities in the relation between spatial configuration and the observed functioning of built environments have been established using 'non-discursive techniques' of analysis to control the architectural variables.

Chapter 4, 'Cities as movement economies' reports a fundamental research finding: that movement in the urban grid is, other things being equal, generated by the configuration of the grid itself. This finding allows completely new insights into the structure of urban grids, and the way these structures relate to urban functioning. The relation between grid and movement in fact underlies many other aspects of urban form: the distribution of land uses, such as retail and residence, the spatial patterning of crime, the evolution of different densities and even the part-whole structure of cities. The influence of the fundamental grid-movement relation is so pervasive that cities are conceptualised in the chapter as 'movement economies', in which the structuring of movement by the grid leads, through multiplier effects, to dense patterns of mixed use encounter that characterise the spatially successful city.

Chapter 5, 'Can architecture cause social malaise?' then discusses how this can go wrong. Focussing on specific studies of housing estates using configurational analysis coupled to intensive observation as well as social data it is shown how the overly complex and poorly structured internal space of many housing estates, including low-rise estates, leads to impoverishment of the 'virtual community' – that is, the system of natural co-presence and co-awareness created by spatial design and realised through movement - and this in turn leads to anti-

Introduction

social uses of space, which are the first stage in decline towards the 'sink estate'. Because the role of space in this process is to create a disorderly and unsafe pattern of space use, and this is then perceived and experienced, it is possible to conceptualise how architecture works alongside social processes to create social decline. In a sense, the creation of disorderly space use through maladroit space design creates the first symptoms of decline, even before any real decline has occurred. In a sense then, it is argued, we find that the symptoms help to bring about the disease.

Chapter 6, 'Time as an aspect of space' then considers another fundamental difference between urban forms: that between cities which serve the needs of production, distribution and trade, and those which serve the needs of social reproduction, that is of government, major social institutions and bureaucracies. A series of 'strange towns' are examined, and it is shown how in their spatial properties, they are in many senses the opposite to the 'normal' towns considered in Chapter 5. The detailed spatial mechanisms of these towns are examined, and a 'genotype' proposed. An explanation is then suggested as to why 'cities of social reproduction' tend to construct these distinctive types of spatial patterns.

Chapter 7, 'Visible Colleges', then turns to the interiors of buildings. It begins by setting out a general theory of space in buildings, taking into account the results of settlement analysis, and then highlights a series of studies of buildings. A key distinction is made between 'long and short models', that is, between cases where space is strongly governed by rules, and therefore acts to conserve given social statuses and relationships and cases where space acts to generate relations over and above those given by the social situation. The concept of long and short models permits social relations and spatial configuration to be conceptualised in an analogous way. A ritual is a long model social event, since all that happens is governed by rules, and a ritual typically generates a precise system of spatial relationships and movements through time, that is, a spatial 'long model'. A party is a short model event, since its object is to generate new relationships by shuffling them in space, and this means that rules must be minimised by using a spatial 'short model'. In a long model situation space is adapted to support the rules, and behavioural rules must also support it. In a short model situation, space evolves to structure, and often to maximise, encounter density.

Part III of the book, *'The Laws of the Field'*, then uses these noted regularities to reconsider the most fundamental question of all in architectural theory: how is the vast field of possible spatial complexes constrained to create those that are actually found as buildings? First, in Chapter Eight, 'Is architecture an ars combinatoria?', a general theory of 'partitioning' is proposed, in which it is shown that local physical changes in a spatial system always have more or less global configurational effects. It is the laws governing this passage form local physical moves to global spatial effects that are the spatial laws that underlie building. These local-to-global spatial laws are linked to the evolution of real buildings through what will be called 'generic function', by which is meant the

Introduction

spatial implications of the most fundamental aspects of human use of space, that is, the fact of occupation and the fact of movement. At this generic level, function imposes restraints on what is spatially viable, and this is responsible for what all buildings have in common as spatial designs. Generic function is the 'first filter' between the field of possibility and architectural actuality. The second filter is then the cultural or programmatic requirement of that type of building. The third filter is the idiosyncrasies of structure and expression that then distinguish that building from all others. The passage from the possible to the real passes through these three filters, and without an understanding of each we cannot decipher the form-function relation. Most of all, without a knowledge of generic function and its spatial implications we cannot understand that what all buildings have in common in their spatial structures is already profoundly influenced by human functioning in space.

In Chapter 9, 'The fundamental city', the theory of generic function and the three filters is applied to cities to show how much of the growth of settlements is governed by these basic laws. A new computer modelling technique of 'all line analysis', which begins by conceptualising vacant space as an infinitely dense matrix of lines, containing all possible structures, is used to show how the observable regularities in urban forms from the most local to the most global can be seen to be products of the same underlying processes. A fundamental settlement process is proposed, of which particular cultural types are parameterisations. Finally, it is shown how the fundamental settlement process is essentially realised through a small number of spatial ideas which have an essentially geometrical nature.

Part IV of the book, *'Theoretical Syntheses'*, then begins to draw together some of the questions raised in Part I, the regularities shown in Part II and the laws proposed in Part III, to suggest how the two central problems in architectural theory, namely the form-function problem and the form-meaning problem, can be reconceptualised. Chapter 10, 'Space is the machine', reviews the form-function theory in architecture and attempts to establish a pathology of its formulation: how it came to be set up in such a way that it could not be solved. It then proposes how the configuration paradigm permits a reformulation, through which we can not only make sense of the relation between form and function in buildings, but also we can make sense of how and why buildings, in a powerful sense are 'social objects' and in fact play a powerful role in the realisation and sustaining of human society. Finally, in Chapter 11, 'The reasoning art', the notion of configuration is applied to the study of what architects do, that is, design. Previous models of the design process are reviewed, and it is shown that without knowledge of configuration and the concept of the non-discursive, we cannot understand the internalities of the design process. A new knowledge-based model of design is proposed, with configuration at its centre. It is argued from this that because design is a configurational process, and because it is the characteristic of configuration that local changes make global differences, design is necessarily a top down process. This does not mean that it cannot be analysed, or supported by research. It shows however that only configurationally biased knowledge can really support the design

Introduction

process, and this, essentially, is theoretical knowledge. It follows from this that attempts to support designers by building methods and systems for bottom up construction of designs must eventually fail as explanatory systems. They can serve to create specific architectural identities, but not to advance general architectural understanding.

In pursuing an analytic rather than a normative theory of architecture, the book might be thought by some to have pretensions to make the art of architecture into a science. This is not what is intended. One effect of a better scientific understanding of architecture is to show that although architecture as a phenomenon is capable of considerable scientific understanding, this does not mean that as a practice architecture is not an art. On the contrary, it shows quite clearly why it is an art and what the nature and limits of that art are. Architecture is an art because, although in key respects its forms can be analysed and understood by scientific means, its forms can only be prescribed by scientific means in a very restricted sense. Architecture is law governed but it is not determinate. What is governed by the laws is not the form of individual buildings but the field of possibility within which the choice of form is made. This means that the impact of these laws on the passage from problem statement to solution is not direct but indirect. It lies deep in the spatial and physical forms of buildings, in their genotypes, not their phenotypes.

Architecture is therefore not part art, and part science, in the sense that it has both technical and aesthetic aspects, but is both art and science in the sense that it requires both the processes of abstraction by which we know science and the processes of concretion by which we know art. The architect as scientist and as theorist seeks to establish the laws of the spatial and formal materials with which the architect as artist then composes. The greater scientific content of architecture over art is simply a function of the far greater complexity of the raw materials of space and form, and their far greater reverberations for other aspects of life, than any materials that an artist uses. It is the fact that the architect designs with the spatial stuff of living that builds the science of architecture into the art of architecture.

It may seem curious to argue that the quest for a scientific understanding of architecture does not lead to the conclusion that architecture is a science, but nevertheless it is the case. In the last analysis, architectural theory is a matter of understanding architecture as a system of possibilities, and how these are restricted by laws which link this system of possibilities to the spatial potentialities of human life. At this level, and perhaps only at this level, architecture is analogous to language. Language is often naïvely conceptualised as a set of words and meanings, set out in a dictionary, and syntactic rules by which they may be combined into meaningful sentences, set out in grammars. This is not what language is, and the laws that govern language are not of this kind. This can be seen from the simple fact that if we take the words of the dictionary and combine them in grammatically correct sentences, virtually all are utterly meaningless and do not count as legitimate sentences. The structures of language are the laws

Introduction

which restrict the combinatorial possibilities of words, and through these restrictions construct the sayable and the meaningful. The laws of language do not therefore tell us what to say, but prescribe the structure and limits of the sayable. It is within these limits that we use language as the prime means to our individuality and creativity.

In this sense architecture does resemble language. The laws of the field of architecture do not tell designers what to do. By restricting and structuring the field of combinatorial possibility, they prescribe the limits within which architecture is possible. As with language, what is left from this restrictive structuring is rich beyond imagination. Even so, without these laws buildings would not be human products, any more than meaningless but syntactically correct concatenations of words are human sentences.

The case for a theoretical understanding of architecture then rests eventually not on aspiration to philosophical or scientific status, but on the nature of architecture itself. The foundational proposition of the book is that architecture is an inherently theoretical subject. The very act of building raises issues about the relations of the form of the material world and the way in which we live in it which (as any archaeologist knows who has tried to puzzle out a culture from material remains) are unavoidably both philosophical and scientific. Architecture is the most everyday, the most enveloping, the largest and the most culturally determined human artefact. The act of building implies the transmission of cultural conventions answering these questions through custom and habit. Architecture is their rendering explicit, and their transmutation into a realm of innovation and, at its best, of art. In a sense, architecture is abstract thought applied to building, even therefore in a sense theory applied to building. This is why, in the end, architecture must have analytic theories.

Part one
Theoretical preliminaries

1

Chapter one

What architecture adds to building

The visual impression, the image produced by differences of light and colour, is primary in our perception of a building. We empirically reinterpret this image into a conception of corporeality, and this defines the form of the space within...Once we have reinterpreted the optical image into a conception of space enclosed by mass, we read its purpose from its spatial form. We thus grasp...its content, its meaning.
Paul Frankl

Defining architecture

What is architecture? One thing is clear: if the word is to serve a useful purpose we must be able to distinguish architecture from building. Since building is the more basic term, it follows that we must say in what sense architecture is more than building. The essence of our definition must say what architecture adds to building.

The commonest 'additive' theory is that architecture adds art to building. In this analysis, building is an essentially practical and functional activity on to which architecture superimposes an artistic preoccupation which, while respecting the practical and functional, is restricted by neither. The extreme version of this view is that architecture is the addition to building of the practically useless and functionally unnecessary.[1] The more common is that builders make buildings while architects add style.

From the point of view of finding what people 'really mean' when they say 'architecture', there are serious problems with these views. The most obvious is that it defines architecture in terms of what is normally thought of as its degeneration, that is, that architecture is no more than the addition of a surface appearance to building. Even if we take the view that this is what architecture has become, it is surely unacceptable as a definition of what it should be. Architects believe, and clients on the whole buy, the idea that architecture is a way of being concerned with the whole building, and a means of engaging the deepest aspects of what a building is. If architecture is defined as an add-on which ignores the main substance of building, then architecture would be an addition to building, but would not be more than building. On the contrary, it would be considerably less. If we accuse architecture of being no more than this, we imply that architecture ought to be much more. We are therefore back to the beginning in our pursuit of a definition.

An equally difficult problem with this view is that it is very hard to find examples of building with a purely practical and functional aim. Wherever we find building, we tend to find a preoccupation with style and expression, however modest. Some of the most striking instances of this have come from our growing awareness of building by technologically simple societies, where we do not find that simplicity of technique is associated with simplicity of cultural intent or the elimination of the preoccupation with style. On the contrary, we find that through the idiosyncrasies of style, building and settlement form becomes one of the primary – though most puzzling and variable – expressions of culture.[2] The term that expresses this discovery. 'architecture without architects' confirms the existence of architecture as something over and above building, even though at the same time it affirms the absence of architects.[3]

It is the awareness of the cultural richness of everyday building that lead Roger Scruton, in his *The Aesthetics of Architecture* to try to solve the definition problem for architecture by arguing that since all building shares a preoccupation with the aesthetic and the meaningful, all building should be seen as architecture.[4] Scruton seeks to reintegrate architecture with the whole of building. In his view, all that we ever find in architecture is found, at least in embryonic form, in the everyday vernacular in which most of us participate through our everyday lives. Thus: 'Even when architects have a definite "aesthetic" purpose, it may not be more than the desire that their work should "look right" in just the way that tables and chairs, the lay of places at a table, the folds in a napkin, an arrangement of books, may "look right" to a casual observer.' This leads him to a definition: 'Architecture is primarily a vernacular art: it exists first and foremost as a process of arrangement in which every normal man [sic] may participate.'[5]

The difficulty with this definition is that it leads to exactly the wrong kind of distinction between, for example, the careful formal and spatial rules that governed the English suburban house as built endlessly and repetitiously between the wars by speculative builders, and the works of, say, Palladio or Le Corbusier. The work of both of these architects is characterised by radical innovation in exactly those areas of formal and spatial organisation where according to Scruton's definition, there should be a preoccupation with cultural continuity and reduplication. It would seem to follow that Scruton's definition of architecture would cover the familiar English spec builders' vernacular more easily than it would the works of major architectural innovators.

While it may be reasonable, then, to *prefer* the English inter-war vernacular to the works of Palladio and Le Corbusier, it does not seem likely that a *definition* of the ordinary use of the word architecture lies in this direction. On the contrary, Scruton's definition seems to lead us exactly the wrong way. Architecture seems to be exactly not this preoccupation with cultural continuity, but a preference for innovation. Far from using this as a basis for a definition then Scruton's preoccupation with the vernacular seems to accomplish the opposite. It tells us more how to distinguish everyday building from the more ambitious aspirations of what we call architecture.

Is architecture a thing or an activity?

In what direction should we look then for a definition of architecture as more than building? Reflecting on the common meanings of the word, we find little help and more difficulties. The word 'architecture' seems to mean both a *thing* and an *activity*. On the one hand it seems to imply buildings with certain 'architectural' attributes imposed on them. On the other, it seems to describe what architects do, a certain way of going about the process of making buildings. This double meaning raises serious problems for a definition of architecture. If 'architecture' means both attributes of things and attributes of activities, then which 'really is' architecture'? The definition surely cannot encompass both. Properties of things seem to exist regardless of the activity that creates them, and activities are what they are regardless of their product. Is architecture, then, 'essentially' a thing or an activity? It must, it seems, be one or the other.

However, when we try each definition in isolation we quickly run into paradoxes. Let us experiment first with the idea that architecture is essentially a thing; that is, certain attributes found in some, but not all, buildings. If that is what architecture 'essentially' is, then it would follow that a copy of a building which possesses the architectural attributes will also be architecture, to exactly the same degree and in the same way as the original building. But we baulk at this idea. Copies of architectural buildings seem not themselves to be architecture, but what we have named them as, that is, copies of architecture. Certainly we would not normally expect to win an architectural prize with a deliberate copy. On the contrary, we would expect to be disqualified, or at least ridiculed.

What then is missing in the copy? By definition, it cannot be properties of the building since these are identical in both cases. The disqualifying factor must lie in the act of copying. The act of copying somehow makes a building with architectural attributes no longer, in itself, architecture. This means that what is missing in the copy is not to do with the building but to do with the process that created the building. Copying is therefore in some crucial sense not 'architectural'. Even if we start from the proposition that architecture is attributes of building, and therefore in some sense, 'in the object', the problem of the copy shows that after all architecture implies a certain kind of activity, one which is missing in the act of copying.

What then is missing in the act of copying? It can only be that which copying denies, that is, the intention to *create*, rather than simply to reproduce, architecture. Without this intention, it seems, a building cannot be architecture. So let us call this the 'creative intention' and try to make it the focus of a definition of architecture. We may experiment with the idea as before. This time, let there be an ambitious but talentless architect who intends as hard as possible to make architecture. Is the product of this intention automatically architecture? Whether it is or not depends on whether it is possible to approve the intention as architectural but disqualify the result. In fact this is a very common form for architectural judgments to take. The products of aspiring architects are often judged by their peers to have failed in exactly this way. A jury may legitimately say: 'We understand your intention but do not think you have succeeded.' How are such judgments made? Clearly there is only one answer: by reference to the objective attributes of the proposed buildings that our would-be architect has designed.

It seems then the normal use of words and common practice has led us in a circle. Creative intention fails as a definition of architecture by reference to positive attributes of things, just as positive attributes of things previously failed by reference to intentions. Yet architecture seems at the same time to mean both. It seems it can only be that the idea of architecture is at once a thing and an activity, certain attributes of buildings and a certain way of arriving at them. Product and process are not, it seems, independent. In judging architecture we note both the attributes of the thing and the intellectual process by which the thing is arrived at.

This may seem at first sight rather odd. It violates the common conception that attributes of things are independent of the processes that put them there. But it does reflect how people talk about architecture. Architectural talk, whether by lay people or by critics, typically mixes comment on product with comment on process. For example, we hear: 'This is an ingenious solution to the problem of...', or 'This is a clever detail', or 'This spatial organisation is boldly conceived', 'I like the way the architect has...', and so on. Each of these is at one and the same time a comment on the objective attributes of the building and a comment on the creative intellectual process that gave rise to it. In spite of the unlikelihood of product and process somehow being interdependent in the idea of architecture, this does seem to be exactly the case. In describing our experience of architecture we describe not only the attributes of things, but also the intellectual processes of which the thing is a manifestation. Only with the simultaneous presence of both do we acknowledge architecture.

There is, it seems, some inconsistency between our normal way of reasoning about things and the way we talk, reasonably and reasoningly, about architecture. We might even say that the idea of architecture exhibits some confusion between subjects and objects, since the judgment that a building is architecture seems at one and the same time to depend on the attributes of the 'objective' thing and on attributes of the 'subjective' process that gives rise to the thing. It might be reasonable to expect, then, that further analysis would show that this strangeness in the idea of architecture was pathological and that, with a more careful definition, product and process, and object and subject, could and should be separated.

In fact, we will find the contrary. As we proceed with our exploration of what architecture is and what it adds to building we will find that the inseparability of products and processes and of subject and objects is the essence of what architecture is. It is our intellectual expectations that it should be otherwise which are at fault. Architecture is at once product and process, at once attribute of things and attribute of activity, so that we actually see, or think we see, both when we see and name architecture.

How does this apparent interdependence of product and process then arise as architecture from the act of making a building? To understand this we must first know what building, the allegedly lesser activity, is, and we must understand it both as product and as process. Only this will allow us to see what is distinctive about architecture, and how this distinctiveness involves both product and process. To allow this to become fully clear, the argument that follows will be taken in two stages. first we will look at building as a product, in order to ask what it is about the building as product that architecture takes hold of and adds something to. Then we will look at building as a process, in order to ask how the process of architecture, as adding something to building, is different.

So what is a building?

The question 'what is a building?' tends to provoke two kinds of simplification. The first is that because buildings are purposeful objects we can say what they are by saying what their purpose is. The second is that there must be some simple primordial purpose which was the original reason for buildings and therefore constitutes a kind of continuing essence of building. The first simplification is a logical error, the second a historical one. Both find their commonest, but not only, expression in such ideas as that buildings are essentially 'shelter'.

Both simplifications arise because purposes are seen to be anterior to objects and therefore in some sense explanatory of them. But logically, functional definitions are absurd. In defining building in terms of a function, rather than an object, no distinction is made between buildings as objects and other entities which also can or do provide that function, as for example trees, tents, caves and parasols also provide shelter. Functional definitions are also dishonest. One who defines a building as a shelter has a picture of a building in mind, but one which is implicit rather than explicit, so that the imprecision of the definition is never revealed to the definer. If we say 'a building is a shelter' we mentally see a building and conceive of it functioning as a shelter, so that the function seems to 'explain' the object. Functional definitions only appear to work because they conceal an implicit idea of the object. This prevents the imprecision of the definition from being apparent to the definer. Even if the function were thought to be unique to the object, the definition of an object through its function would never be satisfactory since we could never be sure either that this function is necessarily unique to this object, or that this is the only 'essential' function of this object.

Historically in fact all the evidence is that neither is the case. If we consider the phenomenon of building even in the earliest and simplest societies, one of the most striking things that we find is that buildings are normally multifunctional: they provide shelter from the elements, they provide some kind of spatial scheme for ordering social relations and activities, they provide a framework for the arrangement of objects, they provide a diversity of internal and external opportunities for aesthetic and cultural expression, and so on. On the evidence we have, it is difficult to find historical or anthropological grounds for believing that buildings are not in their very nature multifunctional.

Nor is there any reason why we should expect them to be. In spite of the persistence of the absurd belief that humankind lived in caves until neolithic times (beginning about 10–12,000 years ago), and then used the cave as the model for the building,[6] there is evidence that human beings have created recognisable buildings for a very long time, perhaps as long as at least three hundred thousand years.[7] We do not know how the antiquity of building compares with that of language, but it is clear that the evolutionary history of each is very long, and that conjectural historical ontologies are equally irrelevant to both in trying to understand the complex nature of either as social and cultural phenomenon. The speculation that buildings are somehow 'explained' by being defined as shelters, because we imagine that there must have been a time when this was *all* that building was, is about as useful in understanding the social and cultural complexities of building as the idea that language began with pointing and grunting is to theories of the structure and functioning of language.

But it is not only time that has given buildings their variety of cultural expression. The nature of the building as an object itself has complexities which in themselves naturally tend to multifunctionality and diversity of cultural expression. It is only by understanding the complex nature of the building as object that we can begin to understand its natural tendency to multifunctionality. At the most elementary level, a building is a construction of physical elements or materials into a more or less stable form, as a result of which a space is created which is distinct from the ambient space. At the very least then, a building is both a physical and a spatial transformation of the situation that existed before the building was built. Each aspect of this transformation, the physical and the spatial, already has, as we shall see, a social value, and provides opportunity for the further elaboration of this value, in that the physical form of the building may be given further cultural significance by the shaping and decoration of elements, and the spatial form may be made more complex, by conceptual or physical distinctions, to provide a spatial patterning of activities and relationships.

However, even in the most primitive, unelaborated state, the effect of this elementary transformation of material and space on human beings – that is, its 'functional' effect – is complex. Part, but only part, of this complexity is the functional effect that the 'shelter' theorists have noted, namely the physical effect that bodies are protected from ambient elements that in the absence of the building might be experienced as hostile. These elements include inclement weather conditions, hostile species or unwelcome conspecifics. When we say that a building is a 'shelter', we mean that it is a kind of protection for the body. To be a protective shelter a building must create a protected space through a stable construction. What is protective is the physical form of the building. What is protected is the space. Buildings have a bodily function, broad and non-specific, but classifiable as bodily, as a result of which the building has space able to contain bodies, and certain physical properties through which bodies are protected.

However, even the simplest bodily act of making a shelter is more complex than might appear at first sight. To enclose a space by a construction creates not only a physical distinction on the surface of the earth, but also a logical, or categoric distinction. We acknowledge this through terms like 'inside' and 'outside'. These are relational notions with an essentially logical nature, not simple physical facts. They arise as a kind of 'logical emergence' from the more elementary physical fact of making a boundary. The relationality of these 'logical emergents' can be demonstrated by simply pointing to the interdependence of 'inside' and 'outside'. One implies the other, and we cannot create a space inside without also making a space outside. Logicality can be demonstrated by direct analogy. The physical process of drawing a boundary is analogous to naming a category, since when we do so we also by implication name all that is not that category, that is, we imply the complement of that category, in the same sense that when we name the space inside we also imply all the space that is outside. In that sense the space outside is the complement of the space inside. Logicians confirm this analogy by drawing Venn diagrams, that represent concepts as all that falls within the space of a circle, an exactly analogous logical gesture to the creation of a boundary in real space.

As Russell has pointed out,[8] relations, especially spatial relations, are very puzzling entities. They seem to exist 'objectively', in the sense (to use the example given by Russell) that 'Edinburgh is to the north of London', but we cannot point directly to the relation in the way that we can to other entities which seem to 'really exist'. We must accept, Russell argues, that 'the relation, like the terms it relates, is not dependent on thought, but belongs to the independent world which thought apprehends, but does not create'. We must then accept, he continues, that a relation 'is neither in space nor in time, neither material nor mental, yet it is something'.

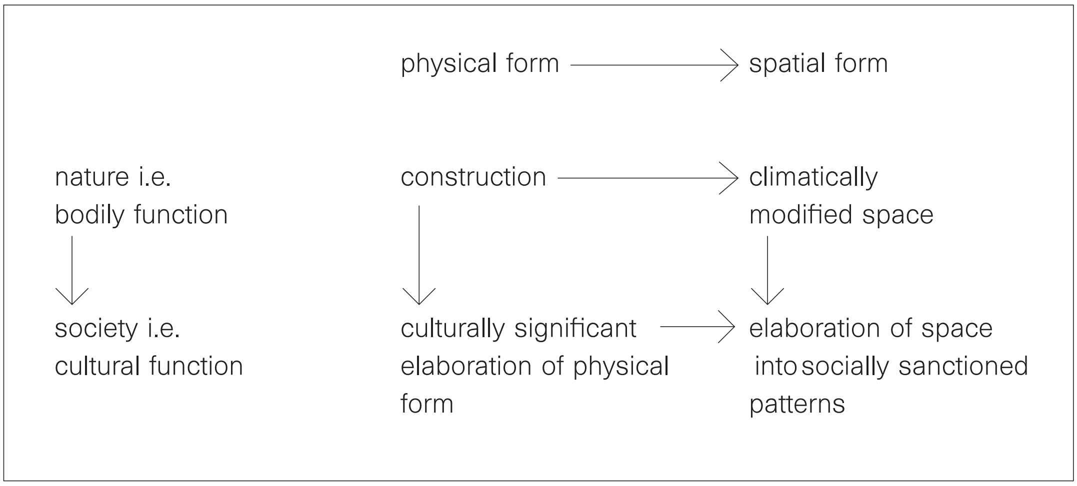
The 'objectivity' of relations, and of the more complex relational schemes we call 'configurations', will be a continuing theme in this book. However, even at the simplest level of the creation of a boundary by the simplest act of building, matters are yet more complex. The logical distinctions made by drawing boundaries are also sociological distinctions, in that the distinction between inside and outside is made by a social being, whose power to make this distinction becomes recognised not only in the physical making of the boundary and the creation of the protected space but also in the logical consequences that arise from that distinction. This is best expressed as a right. The drawing of a boundary establishes not only a physical separateness, but also the social separateness of a domain – the protected space – identified with an individual or collectivity, which creates and claims special rights in that domain. The logical distinction and the sociological distinction in that sense emerge from the act of making a shelter even if they are not intended. The primary act of building, we might say, is already complex in that minds, and even social relations, are engaged by bodily transformations.

As is the case with the logical complexity, the sociological complexity implied by the boundary is in its very nature relational. Indeed, it is the logic of the relational complex that gives rise to the sociological distinctions through which building first begins to reflect and intervene in social relations. It is this essential relationality of form and of space which is appropriated in the processes by which buildings are transformed from bodily objects to social and cultural objects. The fundamental relational complex of form and space created by the act of making the simplest built object is the seed of all future relational properties of spaces through which buildings become fully social objects.

A building then becomes socially significant over and above its bodily functions in two ways: first by elaborating spaces into socially workable patterns to generate and constrain some socially sanctioned – and therefore normative – pattern of encounter and avoidance; and second by elaborating physical forms and surfaces into patterns through which culturally or aesthetically sanctioned identities are expressed. The fundamental duality of form and space that we noted in the most elementary forms of the building thus continues into its complex forms. By the elaboration of space, a social domain is constituted as a lived milieu. By the elaboration of form a social domain is represented as significant identities and encounters. In both senses, buildings create more complex patterns from the basic bodily stuff of form and space. It is through these patterns that buildings acquire their potential at once to constitute and represent – and thus in time to appear as the very foundation of – our social and cultural existence.

What architecture adds to building

Figure 1.1



We may summarise what we have said about the nature of buildings as objects in a diagram which we will use from now on as a kind of fundamental diagram of the building as object, (see fig. 1.1). The essence of the diagram is that a building even at the most basic level embodies two dualities, one between physical form and spatial form and the other between bodily function and socio-cultural function. The link between the two is that the socio-cultural function arises from the ways in which forms and spaces are elaborated into patterns, or, as we will in due course describe them, into configurations. We must now look more carefully at what we mean by the elaboration of form and space into configuration, since this will be the key to our argument not only about the nature of buildings, but also, in due course, to how architecture arises from building.

Let us begin with a simple and familiar case of the elaboration of the physical form of the building: the doric column. When we look at a doric column, we see a plinth, a pedestal, a shaft, a capital, and so on, that is, we see a construction. The elements rest one upon the other, and their relation to each other takes advantage of and depends on the natural law of gravity. But this is not all that we see. The relations of the elements of a column governed by the law of gravity would hold regardless of the 'doricness' of the elements. If, for example, we were to replace the doric capital with an ionian capital, the effect on the construction would be negligible, but the effect on the 'doricness' of the ensemble would be devastating.

So what is doricness? Clearly it is not a type of construction, since we may substitute non-doric elements in the ensemble without constructional penalty. We must acknowledge that doricness is not then in itself a set of physical relations, although it depends on them. Doricness is a scheme in which elements with certain kinds of elaboration are 'above' and 'below' others in a certain relational sequence which emerges from construction but is not given by construction. On the contrary, the notion of 'above' and below' as we find them in doricness seem to be 'logical emergents' from the act of construction in exactly the same sense that 'inside' and 'outside' were logical emergents from the physical construction of a boundary. Doricness is then a logical construction, one built on the back of a physical construction but a logical construction nonetheless. Through the logical doricness of the ensemble, we may say that we move from the simple visuality of the physically interdependent system, to enter the realm of the intelligible. Doricness is a configuration of properties that we understand, over and above what we see as physical interdependencies, a form of relational elaboration to something which exists in physical form, but which through this elaboration stands clear of its physicality. This process of moving from the visible to the intelligible is, we will see in due course, very basic to our experience both of building and of architecture, and, even more so, to the difference between one and the other.

Spatial patterns in buildings also arise as elaborations on primitive logical emergents from the physical act of building. As with doricness, they depend on but cannot be explained by natural law (as many have tried to do by appeal to biological 'imperatives' such as 'territoriality'). The origins of relational schemes of space lie somewhere between the ordering capacities of the mind and the spatial ordering inherent in the ways in which social relationships are realised in space. With space, as with form, we therefore find a split in building between a bodily nature, albeit with a rudimentary relational nature, and a more elaborated configurational nature which relates to minds and social experience rather than to bodies and individual experience. The passage from the simple space to a configuration of space is also the passage from the visible to the intelligible.

Space is, however, a more inherently difficult topic than physical form, for two reasons. First, space is vacancy rather than thing, so even its bodily nature is not obvious, and cannot be taken for granted in the way that we think we can take objects for granted. (See Chapter 10 for a further discussion of this assumption.) Second, related spaces, almost by definition, cannot be seen all at once, but require movement from one to other to experience the whole. This is to say that relationality in space is rarely accessible to us as a single experience. We must therefore digress for a moment to talk about space as a phenomenon, and how we can overcome the difficulties that exist in talking about it. We will take this in two stages. First, we will talk about the problem of how far space can be seen as an objective, independent 'thing-in-itself'. We must do this because there is great confusion about the status of space and how far it can be regarded as an independent entity rather than simply as a by-product of, say, the arrangement of physical things. Second, we will talk about space as configuration, since it is as configuration that it has its most powerful and independent effects on the way buildings and built environments are formed and how they function for their purposes.

What architecture adds to building

About space

It is far from obvious that space is, in some important sense, an objective property of buildings, describable independently of the building as a physical thing. Most of our common notions of space do not deal with space as an entity in itself but tie it in some way to entities that are not space. For example, even amongst those with a interest in the field, the idea of 'space' will usually be transcribed as the 'use of space', the 'perception of space', the 'production of space' or as 'concepts of space'. In all these common expressions, the idea of space is given significance by linking it directly to human behaviour or intentionality. Common spatial concepts from the social sciences such as 'personal space' and 'human territoriality' also tie space to the human agent, and do not acknowledge its existence independently of the human agent. In architecture, where concepts of space are sometimes unlinked from direct human agency, through notions such as 'spatial hierarchy' and 'spatial scale' we still find that space is rarely described in a fully independent way. The concept of 'spatial enclosure' for example, which describes space by reference to the physical forms that define it rather than as a thing in itself, is the commonest architectural way of describing space.

All these concepts confirm the difficulty of conceptualising space as a thing in itself. On occasion, this difficulty finds an extreme expression. For example, Roger Scruton believes that the idea of space is a category mistake made by pretentious architects, who have failed to understand that space is not a thing in itself, but merely the obverse side of the physical object, the vacancy left over by the building. For Scruton, it is self-evident that space in a field and in a cathedral are the same thing except insofar as the interior built surfaces of the cathedral make it appear that the interior space has distinctive properties of its own. All talk about space is error, he argues, because it can be reduced to talk about buildings as physical things.[9]

In fact, even at a practical level, this is a bizarre view. Space is, quite simply, what we use in buildings. It is also what we sell. No developer offers to rent walls. Walls make the space, and cost the money, but space is the rentable commodity. Why then is Scruton embarrassed by the concept of space? Let me suggest that Scruton is making an educated error, one that he would not have made if he had not been so deeply imbued with the western philosophical tradition in which he has earned his living – and to which, incidentally, he has written an outstanding introduction.[10]

The dominant view of space in western culture has been one we might loosely call the 'Galilean-Cartesian'. This view arises from a scheme of reasoning first set out in full clarity by Descartes.[11] The primary properties of physical objects are, he argued, their 'extension', that is, their measurable properties like length, breadth and width. Because extension can be quantified by measuring devices which do not depend on human agency, extensions can be seen as the indubitably objective properties of things, unlike 'secondary' properties like 'green' or 'nice' which seem to depend in some way on interaction with observers.

Now if extension is the primary property of objects, then it is a short step to see it also as the primary property of the space within which objects sit. As Descartes says: ‘After examination we shall find that there is nothing remaining in the idea of body excepting that it is extended in length, breadth and depth; and this is comprised in our idea of space, not only of that which is full of body, but also that which is called a vacuum.’[12] In other words, when we take the object away from its space its extension is still present as an attribute of space. Space is therefore generalised extension, or extension without objects. Descartes again: ‘In space… we attribute to extension a generic unity, so that after having removed from a certain space the body which occupied it, we do not suppose we have also removed the extension of that space.’[13]

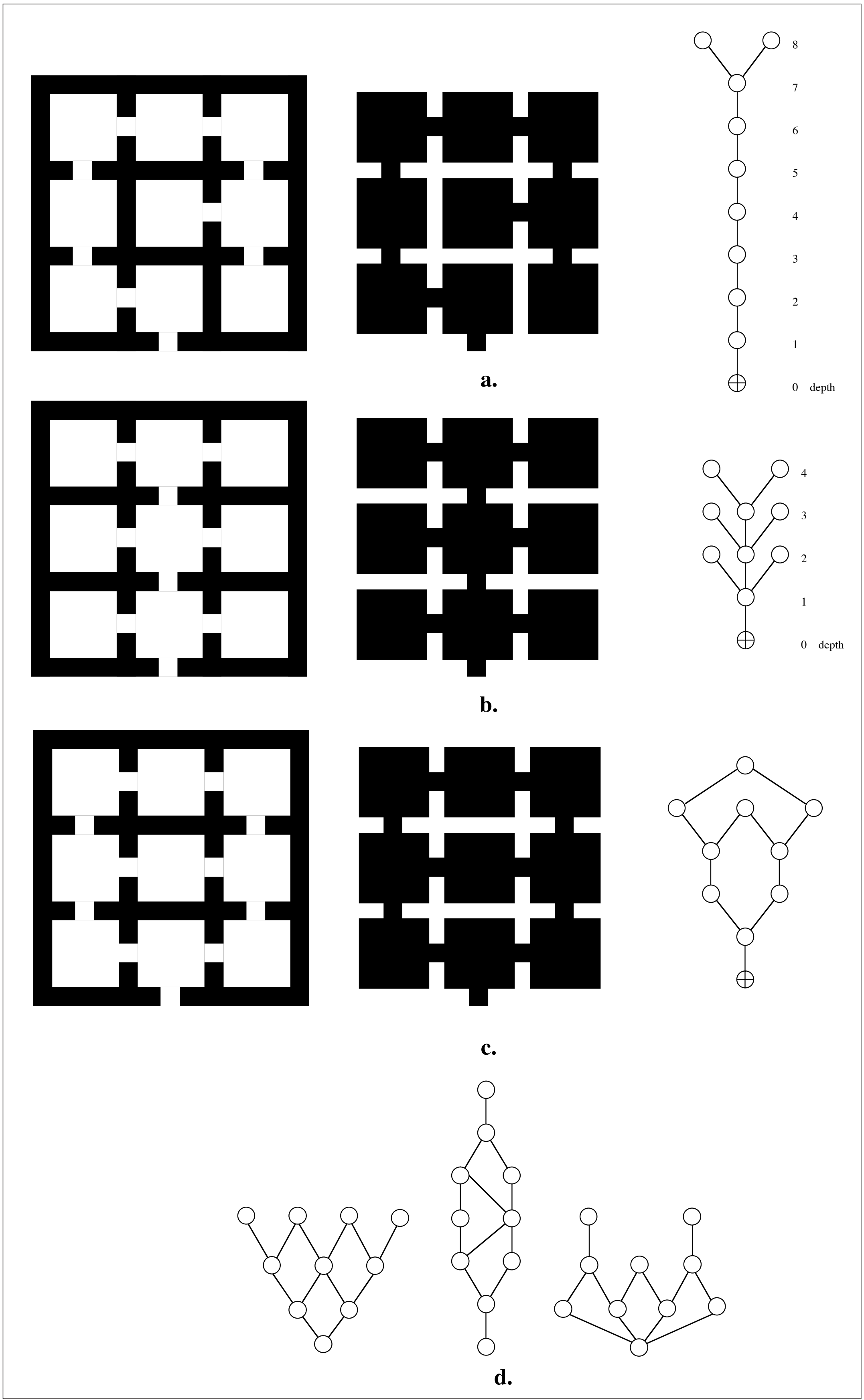
Following this reasoning, space comes to be seen as the general abstract framework of extension against which the properties of objects are defined, a metric background to the material objects that occupy space. This view of space seems to most of us quite natural, no more than an extrapolation of commonsense. Unfortunately, once we see space in this way, we are doomed not to understand how it plays a role in human affairs. Culturally and socially, space is never simply the inert background of our material existence. It is a key aspect of how societies and cultures are constituted in the real world, and, through this constitution, structured for us as ‘objective’ realities. Space is more than a neutral framework for social and cultural forms. It is built into those very forms. Human behaviour does not simply happen in space. It has its own spatial forms. Encountering, congregating, avoiding, interacting, dwelling, teaching, eating, conferring are not just activities that happen in space. In themselves they constitute spatial patterns.

It is because this is so that spatial organisation through buildings and built environments becomes one of the principle ways in which culture is made real for us in the material world, and it is because this is so that buildings can, and normally do, carry social ideas within their spatial forms. To say this does not imply determinism between space to society, simply that space is always likely to be structured in the spatial image of a social process of some kind. The question is: how exactly does this happen, and what are these structures like?

Space as configuration

One thing is clear. Encountering, congregating, avoiding, interacting, dwelling, conferring are not attributes of individuals, but patterns, or configurations, formed by groups or collections of people. They depend on an engineered pattern of co-presence, and indeed co-absence. Very few of the purposes for which we build buildings and environments are not ‘people configurations’ in this sense. We should therefore in principle expect that the relation between people and space, if there is one, will be found at the level of the configuration of space rather than the individual space. This is confirmed by commonsense. Individual spaces place little limit on human activity, except for those of size and perhaps shape. In most reasonable spaces, most human activities can be carried out. But the relation between space and social existence does not lie at the level of the individual space, or individual activity.

Figure 1.2



What architecture adds to building

It lies in the relations between configurations of people and configurations of space.

To take the first steps towards understanding how this happens, we must understand how, in principle, a configuration of space can be influenced by, or influence, a configuration of people. Let us therefore consider some simple hypothetical examples. The two notional 'courtyard' buildings of figure 1.2a and b show in the first column in black, in the normal way, the pattern of physical elements of the buildings. The corresponding figures in the second column then show in black the corresponding pattern of spatial elements. The basic physical structures and cell divisions of the two 'buildings' are the same, and each has the same pattern of adjacencies between cells and the same number of internal and external openings. All that differs is the location of cell entrances. But this is enough to ensure that from the point of view of how a collection of individuals could use the space, the spatial patterns, or 'configurations', are about as different as they could be. The pattern of permeability created by the disposition of entrances is the critical thing. Seen this way, one layout is a near perfect single sequence, with a minimal branch at the end. The other is branched everywhere about the strong central spaces.

Now the pattern of permeability would make relatively little difference to the building structurally or climatically, that is, to the bodily aspect of buildings, especially if we assume similar patterns of external fenestration, and insert windows wherever the other had entrances onto the courtyard. But it would make a dramatic difference to how the layout would work as, say, a domestic interior. For example, it is very difficult for more than one person to use a single sequence of spaces. It offers little in the way of community or privacy, but much in the way of potential intrusion. The branched pattern, on the other hand, offers a definite set of potential relations between community and privacy, and many more resources against intrusion.

These differences are inherent in the space patterns, and would apply to whole classes of human activity patterns. In themselves the spatial layouts offer a range of limitations and potentialities. They suggest the possibility that architectural space might be subject to limiting laws, not of a deterministic kind, but such as to set morphological bounds within which the relations between form and function in buildings are worked out.

We will see from Chapter 3 onwards that it is by expressing these pattern properties in a numerical way that we can find clear relations between space patterns and how collections of people use them. However, before we embark on numbers, there is a visually useful way of capturing some of the key differences between the two spatial patterns. This is a device we call a justified graph, or j-graph. In this we imagine that we are in a space which we call the root or base of the graph, and represent this as a circle with a cross inscribed. Then, representing spaces as circles, and relations of access as lines connecting them, we align immediately above the root all spaces which are directly connected to the root, and draw in the connections. These are the spaces at 'depth one' from the root. Then an equal distance above the 'depth one' row we align the spaces that connect directly to first row spaces, forming the line of 'depth two' spaces, and connect these to the depth one spaces, and so on. Sometimes we will have to draw rather long and circuitous lines to link spaces at different levels, but this does not matter. It is the fact of connection that matters. The laws of graphs guarantee that if the layout is all at one level then we can make all the required connections by drawing lines connecting the spaces without crossing other lines.[14]

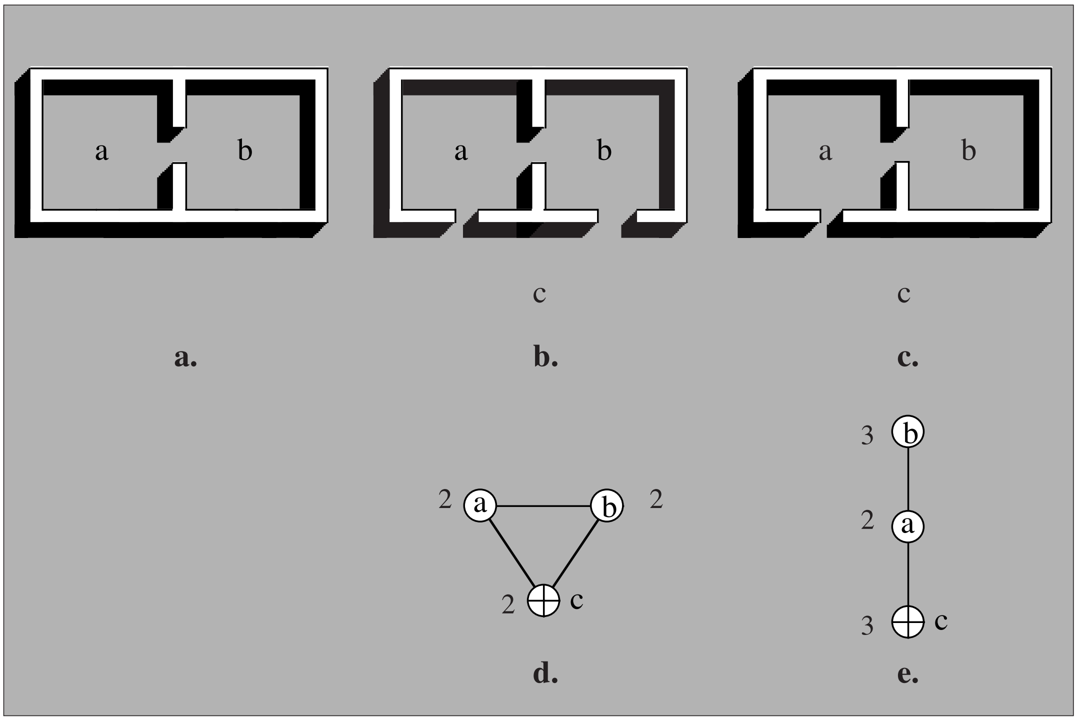
The resulting j-graph is a picture of the 'depth' of all spaces in a pattern from a particular point in it. The third column in figure 1.2a and b shows j-graphs for the corresponding spatial structures, drawn using the exterior space as root. We can immediately see that the first is a 'deep tree' form, and the second a 'shallow tree' form. By 'tree' we mean that there is one link less than the number of cells linked, and that there are therefore no rings of circulation in the graph. All trees, even two as different as in the two in the figures, share the characteristic that there is only one route from each space to each other space – a property that is highly relevant to how building layouts function. However, where 'rings' are found, the justified graph makes them as clear as the 'depth' properties, showing them in a very simple and clear way as what they are, that is, alternative route choices from one part of the pattern to another. The series of figures in figure 1.2c shows a hypothetical case, based on the same basic 'building' as the previous figures.

We do not have to justify the graph using the outside space as root. This is only one way – though a singularly useful way – of looking at a building. We can of course justify the graph from any space within it, and this will tell us what layout is like from the point of view of that space, taking into account both depth and ring properties. When we do this we discover a fact about the spatial layouts of buildings and settlements that is so fundamental that it is probably in itself the key to most aspects of human spatial organisation. This is the simple fact that a pattern of space not only looks different but actually is different when justified from the point of view of its different constituent elements. The three notional j-graphs shown in figure 1.2d appear very different from each other, but all three are in fact the same graph justified from the point of view of different constituent spaces. The depth and ring properties could hardly appear more different if they were different configurations. It is through the creation and distribution of such differences that space becomes such a powerful raw material for the transmission of culture through buildings and settlement forms, and also a potent means of architectural discovery and creation. Let us see how.

Formally defining configuration

First we need to bring a little more formality into the definition of 'configuration'. Like the word 'pattern' (which we do not use because it implies more regularity than we will find in most spatial arrangements), configuration seems to be a concept addressed to the whole of a complex rather than to its parts. Intuitively, it seems to mean a set of relationships among things all of which interdepend in an overall structure of some kind. There is a way of formalising this idea that is as simple as it is necessary. If we define spatial relations as existing when there is any type of link – say adjacency or permeability – between two spaces, then

Figure 1.3



configuration exists when relations between two spaces are changed according to how we relate one or other, or both, to at least one other space.

This rather odd sounding definition can be explained through a simple graphic example. Figure 1.3a shows a cell divided by a partition into two, sub-cell *a* and sub-cell *b*, with a door creating a relation of permeability between the two. It is clear that the relation is formally 'symmetrical' in the sense that cell *a* is to cell *b* as *b* is to *a*. The same would be true of two cells which were adjacent and therefore in the relation of neighbour to each other. If *a* is *b*'s neighbour, then *b* must also be *a*'s neighbour. This 'symmetry', which follows the algebraic rather than the geometrical definition, is clearly an objective property of the relation of *a* and *b* and does not depend on how we choose to see the relation.

Now consider figures 1.3b and c in which we have added relations to a third space, *c* (which is in fact the outside space), but in a different way so that in 1.3b both *a* and *b* are directly permeable to *c*, whereas in 1.3c, only *a* is directly permeable to *c*. This means that in 1.3c we must pass through *a* to get to *b* from *c*, whereas in 1.3b we can go either way. In 1.3c therefore, *a* and *b* are different with respect to *c*. We must pass through *a* to get to *b* from *c*, but we do not need to pass through *b* to get to *a* from *c*. With respect to *c*, the relation has become asymmetrical. In other words, the relation between *a* and *b* has been redefined by the relation each has to a third space. This is a configurational difference. Configuration is a set of interdependent relations in which each is determined by its relation to all the others.

We can show such configurational differences rather neatly, and clarify their nature, by using the j-graph, as in figure 1.3d and e, corresponding to 1.3b and 1.3c respectively. Compared to 1.3a, spaces *b* and *c* in 1.3e have acquired 'depth' with respect to each other, in that their relation is now indirect and only exists by virtue of *a*. The numbers adjacent to each space in the j-graph index this by showing the total depth of each space from the other two. In contrast, 1.3d has acquired a 'ring' that links all three spaces, meaning that each has a choice of route to each of the others. The graph of 1.3d is identical when seen from each of its spaces, while in 1.3e, *b* and *c* are identical, but *a* is different.

Society in the form of the object

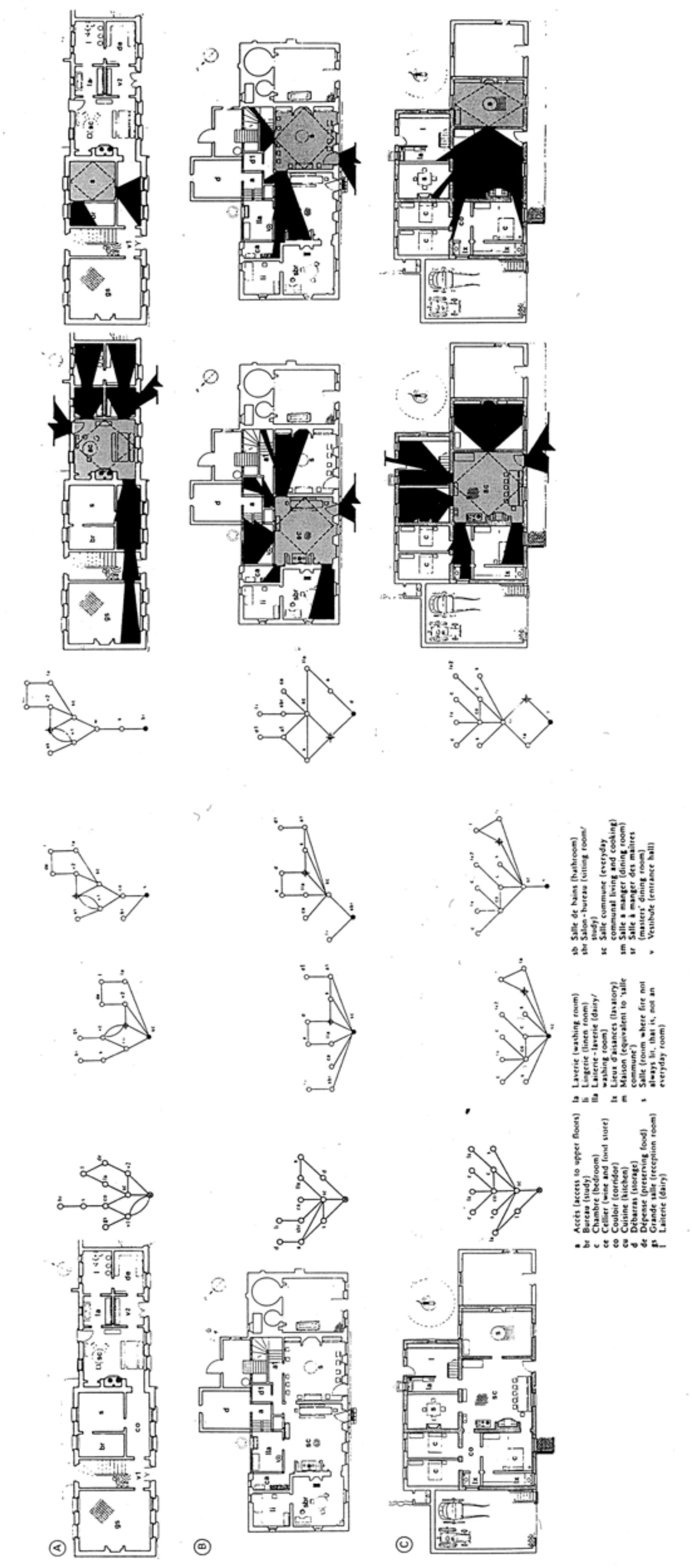
Now let us use this concept of configuration, and its key spatial dimensions of depth and rings, to try to detect the presence of cultural and social ideas in the spatial forms of buildings. Figures 1.4a, b and c show, on the left, the ground-floor plans of three French houses, and to their immediate right, their j-graphs drawn initially from the outside, treating it as a single space, then to the right again three further j-graphs justified from three different internal spaces.[15] Looking at the j-graphs drawn from the outside, we can see that in spite of the geometrical differences in the houses there are strong similarities in the configurations. This can be seen most easily by concentrating on the space marked sc, or salle commune, which is the main everyday living space, in which cooking also occurs and everyday visitors are received. In each case, we can see that the salle commune lies on all non-trivial rings (a trivial ring is one which links the same pair of spaces twice), links directly to an exterior space – that is, it is at depth one in the complex – and acts as a link between the living spaces and various spaces associated with domestic work carried out by women.

The salle commune also has a more fundamental property, one which arises from its relation to the spatial configuration of the house as a whole. If we count the number of spaces we must pass through to go from the salle commune to all other spaces, we find that it comes to a total which is less than for any other space – that is, it has less depth than any other space in the complex. The general form of this measure[16] is called *integration*, and can be applied to any space in any configuration: the less depth from the complex as a whole, the more integrating the space, and vice versa. This means that every space in the three complexes can be assigned an 'integration value'.[17]

Now once we have done this we can ask questions about how the different functions in the house are 'spatialised', that is, how they are embedded in the overall spatial configuration. When we do this, we find that it is very common that different functions are spatialised in different ways, and that this can often be expressed clearly through 'integration' analysis. In the three French houses, for example, we find that there is a certain order of integration among the spaces where different functions are carried out, always with the salle commune as the most integrated, as can be seen in the j-graphs beside each plan. If all the functions of the three houses are set out in order of the integration values of the spaces in which they occur, beginning with the most integrated space, we can read this, from left to right, as: the salle commune is more integrated (i.e. has less depth to all other spaces) than

What architecture adds to building

Figure 1.4



the corridor, which is more integrated than the exterior, and so on. To the extent that there are commonalities in the sequence of inequalities, then we can say that there is a common pattern to the way in which different functions are spatialised in the house. We call such common patterns 'inequality genotypes', because they refer not to the surface appearances of forms but to deep structures underlying spatial configurations and their relation to living patterns.[18]

These results flow from an analysis of space-to-space permeability. But what about the relation of visibility, which passes through spaces? The three rows of figures on the right in figure 1.4 (lower panel) show all the space that can be seen with the doors open from a diamond-shaped space within each salle commune and one other space, drawn by joining the centre points of each wall of a room, and thus covering half of the space in the room. The idea of the diamond shape is that space use (in most western cultures) is normally concentrated within this diamond shape, the corners commonly being reserved for objects. The diagrams show that in each case the salle commune has a far more powerful visual field than the salle. In other words, the spatial and functional differences between spaces that we find through the analysis of permeability in the houses also appear in the analysis of visibility. These visibility differences can also form the basis for quantitative and statistical analysis.

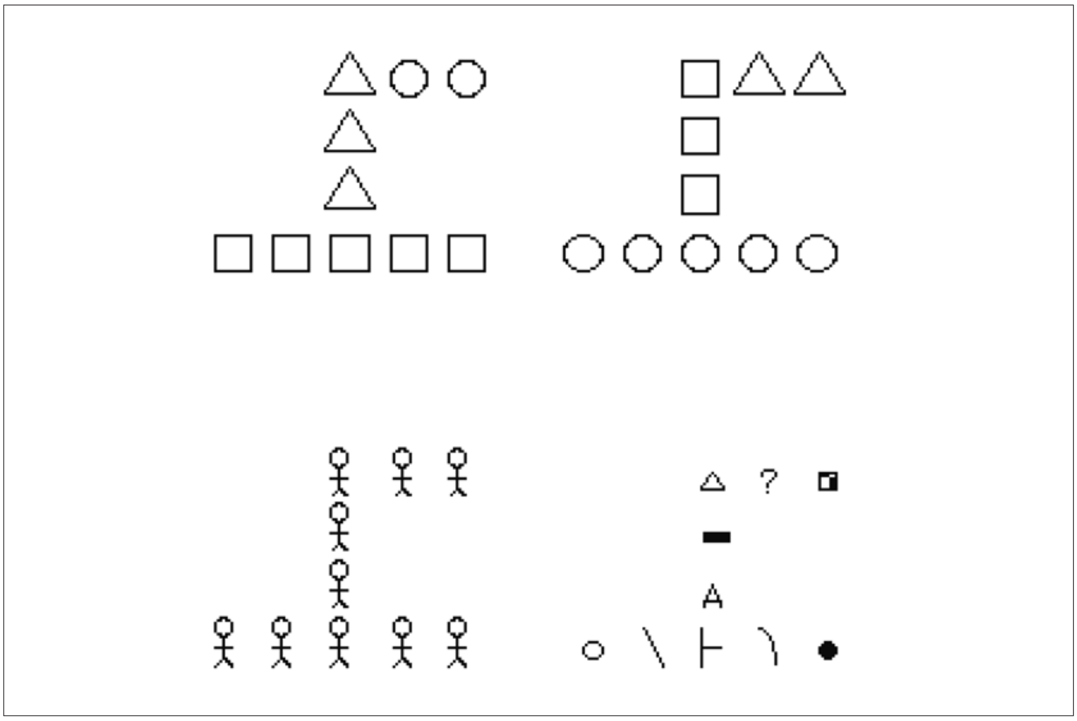
This type of method allows us to retrieve from house plans configurational properties that relate directly to the social and cultural functioning of the house. In other words, through spatial configuration culturally determined patterns are embedded in the material and spatial 'objectivity' of buildings. By the analysis of spaces and functions in terms of their configurational relations within the house, and the search for common patterns across samples, we can see how buildings can transmit common cultural tendencies through spatial form. We must now ask how and why this is the case, and what follows from it?

The non-discursivity of configuration: ideas we think of and ideas we think with

The answer will take us to the centre of our argument: the non-discursivity of configuration. Non-discursivity means that we do not know how to talk about it. The difficulty of talking about spatial or formal configurations in architecture has always seemed a rather peripheral problem to architectural theory. I suggest it is the central problem, and part of a much more general problem in human affairs.

Let us begin to explore the intuitive aspects of the idea of configuration a little further. Consider the four groups of elements in figure 1.5. Each group is a different set of 'things', but placed in more or less the same overall 'configuration'. The human mind has no difficulty in seeing that the configurations are the same, in spite of the differences in the constituent 'things', and this shows that we easily recognise a configuration, even where we have no way of giving it a name and thus assigning it to a category – although we might try to do so by making analogies with configurations for which names are already at hand, such as 'L-shaped', or 'star-shaped'. However, the fact that our minds recognised configurations as being the same even when there is no name at hand to link them shows that our ability

Figure 1.5



to recognise and understand configuration is prior to the assignment of names.

Configuration seems in fact to be what the human mind is good at intuitively, but bad at analytically. We easily recognise configuration without conscious thought, and just as easily use configurations in everyday life without thinking of them, but we do not know what it is we recognise and we are not conscious of what it is we use and how we use it. We have no language for describing configurations, that is, we have no means of saying what it is we know. This problem is particularly salient in buildings and architecture, because both have the effect of imposing spatial and formal configuration on the world in which we live. But the problem is not confined to architecture. On the contrary it appears to be present to some degree in most cultural and social behaviours. In using language, for example, we are aware of words and believe that in speaking and hearing we are handling words. However, language only works because we are able to use the configurational aspects of language, that is, the syntactic and semantic rules which govern how words are to be assembled into meaningful complexes, in a way which makes their operation automatic and unconscious. In language we can therefore distinguish ideas we think *of*, that is, the words and what they represent, and ideas we think *with*, that is, syntactic and semantic rules which govern how we deploy words to create meaning. The words we think *of* seem to us like things, and are at the level of conscious thought. The hidden structures we think *with* have the nature of configurational rules, in that they tell us how things are to be assembled, and work below the level of consciousness. This 'unconscious configurationality' seems to prevail in all areas where we use rule systems to behave in ways which are recognisable as social. Behaviour at table, or the playing of games, appear to us as spatio-temporal events, but they are given order and purpose by the underlying configurational 'ideas-to-think-with' through which these events are generated. We acknowledge the importance of this unseen configurationality labelling it as a form of knowledge. We talk about 'knowing how to behave', or 'knowing a language'.

We can call this kind of knowledge 'social knowledge', and note that its purpose is to create, order and make intelligible the spatio-temporal events through which we recognise the presence of culture in everyday life. We must of course take care to distinguish social knowledge from forms of knowledge which we learn in schools and universities whose purpose is to understand the world rather than to behave in it, and which we might therefore call analytic, or scientific knowledge. In itself (though not necessarily in its consequences) analytic knowledge leaves the world as it is, since its purpose is to understand. Analytic knowledge is knowledge where we learn the abstract principles through which spatio-temporal phenomena are related – we might say the 'configurationality' – consciously. We are aware of the principles both when we acquire and when we use the knowledge. As a result, through the intermediary of the abstract, we grasp the concrete. In social knowledge, in contrast, knowledge of abstract configurationality is acquired through the process of creating and experiencing spatio-temporal events. Social knowledge works precisely because the abstract principles through which spatio-temporal phenomena are brought together into meaningful patterns are buried beneath habits of doing, and never need be brought to conscious attention.[19]

In spite of these functional differences, social knowledge and analytic knowledge are made up of the same elements: on the one hand, there is knowledge of spatio-temporal phenomena, on the other, there are abstract 'configurational' structures that link them together. But whereas in social knowledge the abstract ideas are held steady as ideas to think with in order to create spatio-temporal events in the real world, so that the abstract ideas become the normative bases of behaviour, in scientific knowledge, an attempt is made to hold spatio-temporal phenomena steady in order to bring the abstract structures through which we interpret them to the surface in order to examine them critically and, if necessary, to reconstitute them.

This can be usefully clarified by a diagram, see figure 1.6. The difference between the two forms of knowledge lies essentially in the degree to which abstract ideas are at the level of conscious thought and therefore at risk. The whole purpose of science is to put the abstract 'ideas we think with' in making sense of spatio-temporal events at risk. In social knowledge, the whole purpose of the 'knowledge' would be put at risk by bringing them to conscious thought since their function is to be used normatively to create society. However, it is clearly a possibility that the abstract structures of social knowledge could, as with science, themselves become the object of conscious thought. This, in a nutshell, is the programme of 'structuralism'. The essence of the structuralist method is to ask: can we build a model of the abstract principles of a system (e.g. language) that 'generates' all and only the spatio-temporal events that can legitimately happen? Such a model would be a theory of the system. It would, for example, 'explain' our intuitive sense

Figure 1.6

	abstract principles	spacial-temporal events
social knowledge	codes, rules ideas to think	speech, social behaviour, spaces, ideas to think
analytic knowledge	theories, hypotheses paradigms	'facts', phenomena

that some strings of words are meaningful sentences and others – most – are not. Structuralism is rather like taking the output of a computer as the phenomena to be explained, and trying to find out what programme could generate all and only these phenomena. Structuralism is an enquiry into the unconscious configurational bases of social knowledge, that is, it is an inquiry into the non-discursive dimensions of social and cultural behaviour.

Building as the transmission of culture through artefacts

The spatial and formal patterns that are created through buildings and settlements are classic instances of the problem of non-discursivity, both in the sense of the configurational nature of ideas we think with in creating and using space, and in the sense of the role these play in social knowledge. As has already been indicated, one of the most pervasive examples of this is the dwelling. Domestic space varies in the degree to which it is subject to social knowledge, but it is not uncommon for it to be patterned according to codes of considerable intricacy which govern what spaces there are, how they are labelled, how bounded they are, how they are connected and sequenced, which activities go together in them and which are separated, what individuals or categories of persons have what kinds of rights in them, how they are decorated, what kinds of objects should be displayed in them and how, and so on. These patterns vary from one cultural group to another, but invariably we handle domestic space patterns without thinking of them and even without being aware of them until they are challenged. In general, we only become aware of the degree of patterning in our own culture when we encounter another form of patterning in another culture.

But domestic space is only the most intensive and complex instance of a more generalised phenomenon. Buildings and settlements of all kinds, and at all levels, are significantly underpinned by configurational non-discursivity. It is through this that buildings – and indeed built environments of all kinds – become part of what Margaret Mead called 'the transmission of culture through artefacts'.[20] This transmission occurs largely through the configurational aspects of space and form in those environments. For example, we think consciously of buildings as physical or spatial objects and we think of their parts as physical or spatial parts, like columns or rooms. But we think of 'buildings' as whole entities through the unconscious intermediary of configuration, in that when we think of a particular kind of building, we are conscious not only of an image of an object, but at the same time of the complex of spatial relations that such a building entails. As space – and also as meaningful forms – buildings are configurational, and because they are configurational their most important social and cultural properties are non-discursive. It is through non-discursivity that the social nature of buildings is transmitted, because it is through configuration that the raw materials of space and form are given social meaning. The social stuff of buildings, we may say, is the configurational stuff, both in the sense that buildings are configurations of space designed to order in space at least some aspects of social relationships, and in the sense that it is through the creation of some kind of configuration in the form of the building that something like a cultural 'meaning' is transmitted.

Building as process

How then can this help us make the distinction between architecture and building? We note of course that we now begin not from the notion that buildings prior to architecture are only practical and functional objects, but from the proposition that prior to architecture buildings are already complex instances of the transmission of culture through artefacts. This does not mean of course that buildings of the same type and culture will be identical with each other. On the contrary, it is common for vernacular architectures to exhibit prodigious variety at the level of individual cases, so much so that the grounds for believing that the cases constitute instances of a common vernacular style, either in form or space, can be quite hard to pin down.

The crucial step in arriving at our definition of architecture is to understand first how the vernacular builder succeeds in making a building as a complex relational structure through which culture is transmitted, while at the same time creating what will often be a unique individual building. We do not have to look far for the answer. This combination of common structure and surface variety is exactly what we find where social knowledge is in operation in the form in which we have just described it: complex configurational ideas at the non-discursive level guide the ways in which we handle spatio-temporal things at the surface level, and as a result configurational ideas are realised in the real world. In building terms, the manipulation of the spatial and formal elements which make up the building will, if carried out within the scope of non-discursive configurational ideas to think with, which govern key aspects of their formal and spatial arrangement, lead to exactly the combination of underlying common structure and surface variety that characterises vernacular architectures in general.

To understand how this happens in particular cases, we can draw on the remarkable work of Henry Glassie.[21] Glassie proposes that we adapt from Noam Chomsky's studies of language a concept which he calls 'architectural competence.' 'Architectural competence' is a set of technological, geometrical and manipulative skills relating form to use, which constitute 'an account not of how a house is made, but of how a house is thought...set out like a programme...a scheme analogous to a grammar, that will consist of an outline of rule sets interrupted by prosy exegesis'.

The analogy with language is apposite. It suggests that the rule sets the vernacular designer uses are tacit and taken for granted in the same way as the rule sets that govern the use of language. They are ideas the designer thinks with rather than of. They therefore have a certain degree of abstraction from the material reality they help to create. They specify not the specific but the generic, so that the vernacular designer may use the rules as the basis of a certain restrained creativity in interpreting the rules in novel ways.

Now the implication of Glassie's idea is that 'architectural competence' provides a set of normative rules about how building should be done, so that a vernacular building reproduces a known and socially accepted pattern. The house built by a builder sharing the culture of a community comes out right because it draws on the normative rules that define the architectural competence of the community. In this way buildings become a natural part of 'the transmission of culture by artefacts'. Through distinctive ways of building, aspects of the social knowledge distinctive of a community are reproduced. Thus the physical act of building, through a system of well defined instrumentalities, becomes the means by which the non-discursive patterns we call culture are transmitted into and through the material and spatial forms of buildings. The non-discursive aspects of building are transmitted exactly as we would expect them to be: as unconscious pattern implications of the manipulation of things.

So what is architecture?

To understand building, then, we must understand it both as a product and as a process. Having done this, we can return to our original question: what is it that architecture adds to building? By unpacking the cultural and cognitive complexity of building, it will turn out that we are at last in sight of an answer. Whatever architecture is, it must in some sense go beyond the process by which the culturally sanctioned non-discursivities are embedded in the spatial and physical forms of buildings. In what sense, then, is it possible to 'go beyond' such a process?

The answer is now virtually implied in the form of the question. Architecture begins when the configurational aspects of form and space, through which buildings become cultural and social objects, are treated not as unconscious rules to be followed, but are raised to the level of conscious, comparative thought, and in this way made part of the object of creative attention. Architecture comes into existence, we may say, as a result of a kind of intellectual *prise de conscience*: we build, but not as cultural automata, reproducing the spatial and physical forms of our culture, but as conscious human beings critically aware of the cultural relativity of built forms and spatial forms. We build, that is, aware of intellectual choice, and we therefore build with reason, giving reasons for these choices. Whereas in the vernacular the non-discursive aspects of architecture are normative and handled autonomically, in architecture these contents become the object of reflective and creative thought. The designer is in effect a configurational thinker. The object of architectural attention is precisely the configurational ideas to think with that in the vernacular govern configurational outcomes. This does not mean that the designer does not think of objects. It means that at the same time the designer thinks of configuration.

The essence of architecture lies therefore in building not by reference to culturally bound competences, and the way in which they guide the non-discursive contents of buildings through programmes of social knowledge specific to one culture, but by reference to a would-be universalistic competence arrived at through the general comparative study of forms aimed at principle rather than cultural idiosyncrasy, and, through this, at innovation rather than cultural reduplication. It is when we see in the non-discursive contents of buildings evidence of this concern for the abstract comparability of forms and functions that building is transcended and architecture is named. This is why the notion of architecture seems to contain within itself aspects of both the product which is created and of the intellectual process through which this creation occurs.

Architecture exists, we might say, where we note as a property of things evidence not only of a certain kind of systematic intent – to borrow an excellent phrase proposed by a colleague in reviewing the archaeological record for the beginning of architecture[22] – in the domain of non-discursivity, but of something like theoretical intent in that domain. In a key sense architecture transcends building in the same way that science transcends the practical crafts of making and doing. It introduces into the creation of buildings an abstract concern for architectural possibility through the principled understanding of form and function. The innovative imperative in architecture is therefore in the nature of the subject. We should no more criticise architects for their penchant towards innovation than we should scientists. In both cases it follows from the social legitimations which give each its name and identity. Both architecture and science use the ground of theoretical understanding to move from past solutions to future possibility, the latter in the direction of new theoretical constructs, the former in the direction of new realities. The judgment we make that a building is architecture arises when the evidence of systematic intent is evidence of intellectual choice and decision exercised in a field of knowledge of possibility that goes beyond culture into principle. In this sense, architecture is a form of practice recognisable in its product. The judgment we make that a building is architecture comes when we see evidence in the building both of systematic intent which requires the abstract and comparative manipulation of form within the general realm of architectural possibility, and that this exploration and this exercise of intellectual choice has been successfully accomplished.

Architecture is thus both a thing and an activity. In the form of the thing we detect evidence of a systematic intent of the architectural kind. From the built evidence we can judge both that a building is intended to be architecture and, if we are so inclined, that it is architecture. We see now why the definition of architecture is so difficult. Because it is the taking hold of the non-discursive contents of building by abstract, universalistic thought, it is at once an intentional mental act and a property we see in things. It is because we see in things the objectivised record of such thought that we name the result architecture.

What architecture adds to building

It is clear from this analysis that architecture does not depend on architects, but can exist within the context of what we would normally call the vernacular. To the extent that the vernacular shows evidence of reflective thought and innovation at the level of the genotype, then that is evidence of the kind of thought we call architectural within the vernacular. This does not mean that the innovative production of buildings which are phenotypically individual within a vernacular should be thought of as architecture. Such phenotypical variety is normal as the product of culturally constrained non-discursive codes. It is only when the innovation, and therefore the reflective thought, changes the code that underlies the production of phenotypes that we detect the presence of abstract and comparative – and therefore architectural – thought within the confines of vernacular tradition. It is therefore perhaps at times of the greatest change that we become aware of this type of thought in vernacular traditions, that is, when a new vernacular is coming into existence. This is why the demarcation between the vernacular and architecture constantly shifts. The reproduction of existing forms, vernacular or otherwise, is not architecture because that requires no exercise of abstract comparative thought, but the exploitation of vernacular forms in the creation of new forms can be architecture.

Architecture exists then to the degree that there is genotypical invention in the non-discursive, that is, invention with the rules that govern the variability that is possible within a style. The precondition for such invention is an awareness of possibilities which are not contained in contemporary cultural knowing but which are at the same time within the laws of what is architecturally possible. Architecture is characterised therefore by a preoccupation with non-discursive means rather than non-discursive ends. This is not the outcome of a perverse refusal to understand the cultural nature of building, but a taking hold of this very fact as a potentiality to explore the interface between human life and its spatial and physical milieu. In the act of architectural creation, the configurational potentialities of space and form are the raw materials with which the creator works.

Like any creative artist, therefore, the architect must seek to learn, through intellectual inquiry, the limits and potentialities of these raw materials. In the absence of such inquiry, there are manifest and immediate dangers. In the vernacular the pattern of form and the pattern of space which give the building its social character are recreated through the manipulation and assembling of objects. We can say then that the form, the spatial pattern and the functional pattern – the form-function relation, in short – are known in advance and need only be recreated. Because architecture of its nature unlinks the pattern aspects of the building from their dependence on social knowledge, these aspects of the building – and above all their relation to social outcomes – become uncertain.

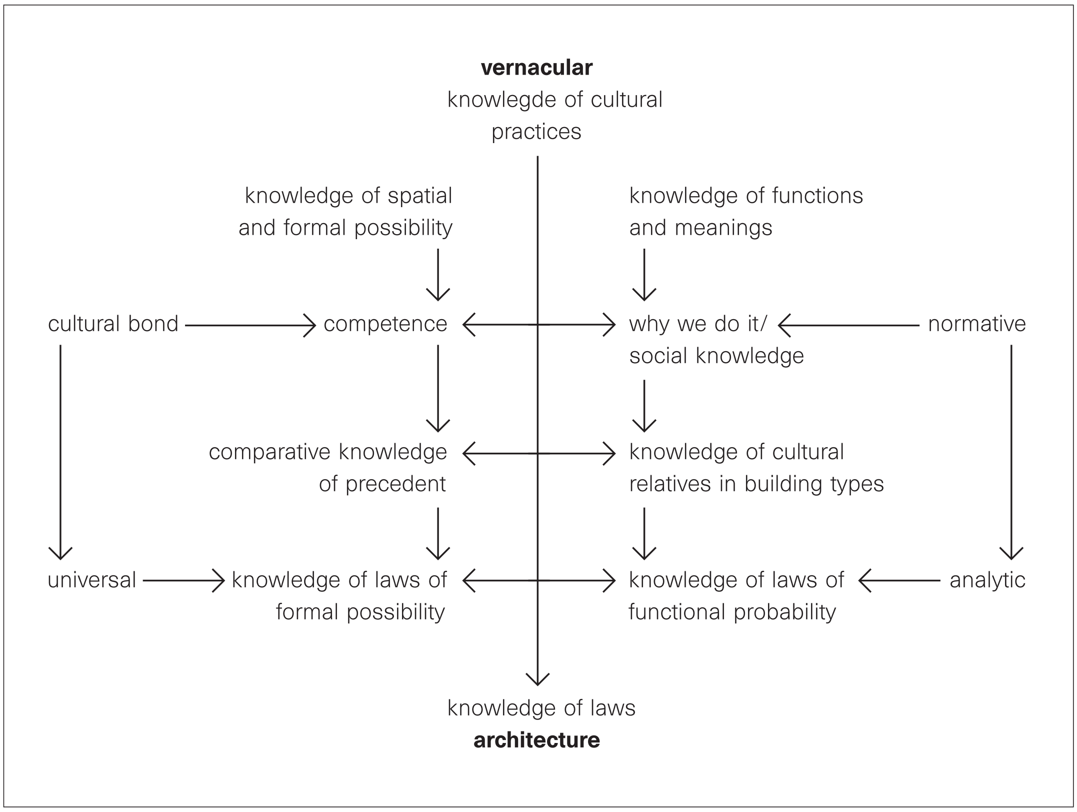
In architecture then, because these crucial relations between non-discursive forms and outcomes are not known in advance, architecture has to recreate in a new, more generalised form, the knowledge conditions that prevail in the vernacular. Because architecture is a creative act, there must be something in the place of the social knowledge structure as ideas to think with. Since architecture is based on the general comparability of possible forms, this knowledge cannot simply encompass particular cases. It must encompass the range of possible cases and if possible cases in general. There is only one term for such knowledge. It is theoretical knowledge. We will see in the next chapter that all architectural theories are attempts to supply principled knowledge of the non-discursive, that is, to render the non-discursive discursive in a way that makes it accessible to reason. In the absence of such knowledge, architecture can be, as the twentieth century has seen, a dangerous art.

The passage from building to architecture is summarised in figure 1.7. The implication of this is that, although we know the difference between architecture and building, there is no hard and fast line to be drawn. Either can become the other at any moment. Taking a broader view which encompasses both, we can say that in the evolution of building we note two ways in which things are done: in obedience to a tradition, or in pursuit of innovation. Building contains architecture to the degree that there is non-discursive invention, and architecture becomes building to the degree that there is not. Vernacular innovation is therefore included within architecture, but the reduplication of vernacular forms is not. Architecture is therefore not simply what is done but how it is done.

The bringing of the non-discursive, configuration dimension of built form from cultural reproduction to reflective awareness and abstract exploration of possibility is at once a passage from the normative to the analytic and from the culture-bound to the universal, the latter meaning that all possibilities are open rather than simply the permutations and phenotypical innovations that are sanctioned by the vernacular. The passage is also one which transforms the idea of knowledge from cultural principle to theoretical abstraction.

In a strong sense, then, architecture requires theory. If it does not have theoretical knowledge, then it will continue to depend on social knowledge. Worse, there is every possibility that architecture can come to be based on social knowledge masquerading as theoretical knowledge, which will be all the more dangerous because architecture operates in the realms of the non-discursive through which society is transmitted through building.[23] Architecture is therefore permanently enjoined to theoretical debate. It is in its nature that it should be so. In that it is the application of reflective abstract thought to the non-discursive dimensions of building, and in that it is through these dimensions that our social and cultural natures are inevitably engaged, architecture is theory applied to building. In the next chapter we will therefore consider what we mean by theory in architecture.

Figure 1.7



Notes

1 J. Ruskin: *Seven Lamps of Architecture*, London 1849, chap. 1.

2 The literature on vernacular architecture as culture is now extensive, and growing rapidly. Among the seminal texts offering wide coverage are Rudovsky's *Architecture Without Architects*, 1964; Paul Oliver's *Shelter and Society*, Barrie & Rockliff, The Cresset Press, 1969 and its follow-up *Shelter in Africa*, Barrie & Jenkins, London, 1971; Amos Rapoport's *House Form and Culture*, Prentice Hall, 1969; Labelle Prussin's classic review of the contrasting vernaculars within a region, *Architecture in Northern Ghana*, University of California Press, 1969; Susan Denyer's African Traditional Architecture, Heineman, 1978; and Kaj Andersen's African Traditional Architecture, Oxford University Press, 1977; in addition to earlier anthropological classics such as C. Daryll Forde's *Habitat, Economy and Society*, Methuen, 1934. Studies of specific cultures are now too numerous to mention, as are the much large-number of texts which have now dealt with the architecture of particular cultures and regions, but which are not yet available in English. Among recent studies of the vernacular, the most important to my mind – and by far the most influential in this text – has been the work of Henry Glassie, and in particular his *Folk Housing in Middle Virginia*, University of Tennessee Press, 1975, references to which, explicit and implicit, recur throughout this text.

3 The same has often been said of 'industrial' architecture. J. M. Richards, for example, in his *An Introduction to Modern Architecture*, Penguin, 1940, describes

Thomas Telford's St Katharine's Dock as 'Typical of the simple but noble engineer's architecture of his time'.

4 Roger Scruton, *The Aesthetics of Architecture*,. Methuen, 1977.

5 Ibid., p. 16.

6 For a recent restatement of this belief see S. Gardiner, *The Evolution of the House*, Paladin, 1976.

7 See for example prehistorical sections of the most recent (nineteenth) edition of Sir Banister fletcher's *A History of Architecture* (edited by Professor John Musgrove) written by my colleague Dr Julienne Hanson. It is a comment on architectural history that it is only very recently that the true antiquity of building has been reflected in the histories of world architecture. Some of Dr Hanson's sources are in themselves remarkable texts which if better known would entirely change popular conception of the history not only of building but also of human society. The key texts are given in Dr Hanson's bibliography, but I would suggest the remarkable R. G. Klein, *Ice Age Hunters of the Ukraine*, Chicago and London, 1973 as a good starting point.

8 B. Russell, *The Problems of Philosophy*, Home University Library, 1912, Oxford University Press paperback, 1959; Chapter 9 'The world of universals'.

9 R. A. Scruton, *The Aesthetics of Architecture*, p. 43 et seq.

10 R. A. Scruton, *A Short History of Modern Philosophy: from Descartes to Wittgenstein*, ARK Paperbacks,1984.

11 R. Descartes, *The Principles of Philosophy*, Part 2, Principle X in *The Philosophical Works of Descartes*, Cambridge University Press, vol. 1, p. 259.

12 Descartes, Principle XI, p. 259.

13 Descartes, Principle X ,p. 259.

14 Graphs which have this property are called 'planar' graphs. Any spatial layout on one level, considered as a graph of the permeability relations, is bound to be planar.

15 These examples are taken from a study of seventeen houses in Normandy carried out for the Centre Nationale de Recherche Scientifique, and published as 'Ideas are in things' in *Environment and Planning B, Planning and Design* 1987, vol. 14, pp. 363–85. This article then formed one of the basic sources for a much more extended treatment in J. Cuisenier, *La Maison Rustique: logique social et composition architecturale*, Presses Universitaires de France, 1991.

16 The 'normalisation' formula for taking the effect of the number elements in the graph out of the total depth calculation from an element is 2(md–1)/k–2, where md is the mean depth of other elements from the root element, and k is the number of elements. There is a discussion of this measure in P. Steadman, *Architectural Morphology*, Pion, 1983, p. 217. The measure was first published in Hillier et al., 'Space Syntax: a new urban perspective' in the *Architect's Journal*, no 48, vol. 178, 30.11.83. There is an extensive discussion of its theoretical foundations and why it is so important in space in Hillier and Hanson, *The Social Logic of Space*, Cambridge University Press, 1984. The measure theoretically eliminates the effect of the numbers of elements in the system. However, in architectural and urban reality there is an additional problem: both buildings and settlements, for practical and empirical

reasons (as will be fully discussed in Chapter 9) tend to become relatively less deep as they grow. A second, 'empirical' normalisation formula is therefore required to take account of this. Such a formula is set out in *The Social Logic of Space*, which has proved robust in use, but has been extensively discussed, for example in J. Teklenberg, H. Timmermans & A. van Wagenberg, 'Space syntax: standardised integration measures and some simulations', *Environment & Planning B: Planning & Design*, vol. 20, 1993, pp. 347–57. See also M. Kruger, 'On node and axial grid maps: distance measures and related topics', paper for the European Conference on the Representation and Management of Urban Change, Cambridge, September 1989, Unit for Architectural Studies, University College London.

17 There is a further measure called 'difference factor', which expresses how strong these differences are, set out in 'Ideas are in things', cited in note 15 above.

18 It should be noted that the argument in the paper from which these examples are taken, 'Ideas are in things' is a great deal more complex than that presented here to illustrate the technique. In fact, it was proposed that two fundamental typological tendencies would be identified within the sample, which were more to do with differ-ences in the relations of the sexes than anything else. A new version of this paper will be published in J. Hanson, *The Social Logic of Houses*, forthcoming from Cambridge University Press.

19 These issues are dealt with at greater length and for a slightly different purpose in Chapter 7.

20 Margaret Mead, *Continuities in Cultural Evolution*, Yale University Press, 1964, Chapter 5.

21 For example, Henry Glassie, *Folk Housing in Middle Virginia*.

22 J. Hanson, written for the intended *Encyclopaedia of Architecture*, McGraw-Hill, New York, but not yet published.

23 We will see in later chapters, and particularly in Chapters 6 and 11, exactly how this can occur and what its consequences are.

Chapter two

The need for an analytic theory of architecture

Do architects need theories?

In the previous chapter, architecture was defined as the taking into reflective thought of the non-discursive, or configurational, aspects of space and form in buildings. In vernacular traditions, these aspects are governed by the taken for granted *ideas to think with* of a culture. In architecture, *ideas to think with* become *ideas to think of*. Spatial and formal configuration in buildings ceases to be a matter of cultural reproduction and becomes a matter of speculative and imaginative enquiry.

It follows from this definition that architecture is an aspiration, not a given. To bring to conscious thought the principles that underlie the spatial and formal patterns that transmit culture through buildings, and to formulate possible alternatives that work as *though they were culture* – since architecture must be an addition to culture not simply a removal of it – is an intellectual as well as a creative task. It requires not only the conceptualisation of pattern and configuration *in vacuo*, but also comparative knowledge and reflective thought. This is why architecture is a reflective as well as an imaginative project, one which seeks to replace – or at least to add to – the social knowledge content of building with an enquiry into principle and possibility.

Architectural theory is the ultimate aim of this reflection. An architectural theory is an attempt to render one or other of the non-discursive dimensions of architecture discursive, by describing in concepts, words or numbers what the configurational aspects of form or space in buildings are like, and how they contribute to the purposes of building. In a sense, theory begins at the moment architecture begins, that is, when spatial and formal configuration in buildings, and their experiential and functional implications, are no longer given through a tradition of social knowledge transmitted through the act of building itself. As soon as building moves free from the safe confines of cultural programming, something like a theory of architecture is needed to support the creative act by proposing a more general understanding of the spatial and formal organisation of buildings than is available within the limits of a single culture.

This is not to say that creative architecture depends on theory. It does not. But in that architecture is the application of speculative abstract thought to the material world in which we live, the reflective aspects of architectural enquiry lead to the formulation if not of theory then at least of theory-like ideas. The need for theory becomes greater as architecture advances. Theory is most required when architecture becomes truly itself, that is, when it becomes the free exploration of formal and spatial possibility in the satisfaction of the human need for buildings.

However, the fact that theory is an inevitable aspect of architecture does not mean that all theories will have a positive effect on architecture. On the contrary, the dependence of architecture on theoretical ideas creates a new type of risk: that theories may be wrong, maybe disastrously wrong. The much discussed 'failure' of modernism in architecture is seen as at least the failure of a theory – the most ambitious and comprehensive ever proposed – and even by some as the failure of the very idea of a theory of architecture.

The need for an analytic theory of architecture

As a result, in the late twentieth century a number of new questions are posed about theories of architecture which are also questions about architecture itself. Does architecture really need theories, or are they just a pretentious adjunct to an essentially practical activity? If architecture does need theories, then what are they like? Are they like scientific theories? Or are they a special kind of theory adapted for architectural purposes? If architectural theories can be wrong, and have apparently adverse consequences, then can they also be right? How can we set about making architectural theories better? And most difficult of all: how can architecture as a creative art be reconciled to the disciplines of theory? Are the two not opposed to each other, in that better theories must lead inevitably to the elimination of architectural freedom.

The answer proposed in this chapter is that once we accept that the object of architectural theory is the non-discursive – that is, the configurational – content of space and form in buildings and built environments, then theories can only be developed by learning to study buildings and built environments as *non-discursive objects*. To have a theory of non-discursivity in architecture in general we must first build a corpus of knowledge about the non-discursive contents of architecture as a phenomenon. This of course runs counter to most current efforts in architectural theory, which seek to build theory either through the borrowing of concepts from other fields, or through introspection and speculation.

However, the product of the first-hand study of non-discursivity in buildings and built environments will lead to a new kind of theory: an analytic theory of architecture, that is, one which seeks to understand architecture as a phenomenon, before it seeks to guide the designer. An analytic theory of architecture is, it will be argued, the necessary corollary of architectural autonomy. Without the protection of an analytic theory, architecture is inevitably subject to more and more externally imposed restrictions that substitute social ideology for architectural creativity. Analytic theory is necessary in order to retain the autonomy of creative innovation on which the advance of architecture depends.

Are architectural theories just precepts for builders?

Before we can embark on the task of building an analytic theory of architecture, however, we must first explore the idea of theory in architecture a little to prevent our enquiry being obscured by some of the more common misconceptions. Architectural theories do take a very distinctive form, but all is not as it seems at first sight, and it is important that we do not allow appearances to disguise their true nature and purposes.

We may usefully begin by examining the views of a well-known critic of architectural theories. In his 1977 polemic against architectural modernism and its intellectual fashions, *The Aesthetics of Architecture*[1] Roger Scruton is dismissive of the very idea of a theory of architecture: 'Architectural theory', he says in a footnote, is 'usually the gesture of a practical man, unused to words'. Elsewhere he goes further. There is not and cannot be a theory of architecture. What has been called

The need for an analytic theory of architecture

architectural theory are merely '...precepts...which...guide the builder'. While such precepts can be useful canons, they can never amount to a real theory, because they cannot be universal, and it is only with the claim to universality that theory arises.[2]

At first sight, Scruton seems to be right. For the most part – modernism is one of the few exceptions – we associate theories of architecture with individual architects. When we think of Palladio's or Le Corbusier's theory of architecture we take it to mean something like the intellectual ground of a style, the generic principles underlying an approach to design. It seems self evident that no such principles could ever be universal. The idea even leads to paradox. A universal formula for architecture would, if followed, render architecture the same and unchanging, and therefore ultimately dull.

But does theory in architecture really only mean a formula for architectural success? A scientist would find this a strange use of the word 'theory'. For a scientist a theory is a rational construct intended to capture the lawfulness of how the world *is*, not a set of guidelines as to how it *should be*. Scientific theories help us act on the world, but only because they have first described the world independently of any view of how it should be. The essence of science is that its theories are *analytic*, not normative in intent. They describe how the world is, not prescribe how it ought to be.

Might we then suggest that this is exactly the difference between architectural and scientific theories, namely that scientific theories are analytic, and about understanding how things are, whereas architectural theories are normative, and about telling us what to do? There seems to be some truth in this. It is reasonable to say that architecture is about how the world should be rather than how it is, and that its theories should therefore tend to express aspirations rather than realities. In fact, on closer examination, it turns out that this is not and can never be the case. Admittedly, architectural theories are normally presented in normative form, but at a deeper level they are no less analytic than scientific theories.

Take for example, two theories which are about as far apart as they could be in focus and content, Alberti's theory of proportion,[3] and Oscar Newman's theory of 'defensible space'.[4] Both are presented as precepts for successful design, in that both authors' books are aimed primarily at guiding the architectural practitioner in design, rather than explaining the nature of architectural experience as experienced, as Scruton's book is. But if we read the texts carefully, we find that this is not all they are. In each case, the normative content of the work rests on clear, if broad, analytic foundations. Alberti's theory of proportion rests on the Pythagorean notion of mathematical form in nature,[5] and the coincidence it asserts between the principles of natural form and the powers of the mind, as evidenced by the relationship between our sense of harmony in music and the simple numerical ratios on which those harmonies are based. If architecture follows the mathematical principles found in nature, Alberti argues, then it cannot help reproducing the intelligibility and harmony that we find in natural forms. Similarly, Newman's 'defensible space' theory rests on the theory of 'human territoriality', by which genetic tendencies in certain species to defend territory against others of the species, are generalised to human beings, both

The need for an analytic theory of architecture

as individuals and – mistakenly in my view – as groups. If, Newman argues, architects design space in conformity with 'territorial' principles, then it will be following biological drives built into us by nature.[6]

It is notable that in both of these theories, the principles for design are said to be based on principles to be found in nature. In a very strong sense, then, in both cases the normative content of the theory depends on the analytic. On reflection, it must be so to some degree in all cases. Any theory about how we should act to produce a certain outcome in the world must logically depend on some prior conception of how the world is and how it will respond to our manipulations. Careful examination will show that this is always the case with architectural theories. We invariably find that the precepts about what designers should do are set in a prior framework which describes how the world is. Sometimes this framework is explicitly set out, and rests on a specific scientific or quasi-scientific foundation, as in the two cases we have instanced. Sometimes it is much more implicit, reflecting no more than a currently fashionable way of looking at the world, as for example many recent theories have rested on the fashionable assumption that 'everything is a language' so that designers can and should design following the principles of linguistic theories in making their buildings 'meaningful'.

Although presented normatively, then, architectural theories must have a great deal of analytic content, whether this is explicit or implicit. In point of fact, faced with an architectural theory, our first reaction would usually be to treat it exactly as we would a scientific theory. Offered a general proposition on which to base architectural precepts for design – say a proposition about the psychological impact of a certain proportional systems or the behavioural effects of a certain kind of spatial organisation – our first reaction would be to question the general proposition, or at least to subject it to test by a review of cases. We usually find quite quickly that would-be general propositions run foul of cases known to us, which we then instance as counter-examples to the theory. In other words, we treat an architectural theory very much in the same way as we would treat a scientific theory: that is, we treat it as an analytic theory by trying to find counter-examples which would refute its generality. Even when it survives, we would be inclined to treat it with continuing scepticism as at best a provisional generalisation, which we can make use of until a better one comes along.

It is a mistake, then, to treat architectural theories simply as normative precepts, as Scruton does. Architectural theories are not and cannot be simply normative, but are at least analytic-normative complexes, in which the normative is constructed on the basis of the analytic. It follows that properly theoretical content of architectural theories is specified by the analytic. If the analytic theory is wrong, then the likelihood is that the building will not realise its intention. Architectural theories, we might say, are about how the world should be, but only in the light of how it is believed to be.

The need for an analytic theory of architecture

Theories in design

Why should architectural theories take this distinctive form of combining propositions about how the world should be with propositions about how it is believed to be? The answer is to be found in the nature of what architects do, that is, design. Through its nature as an activity, design raises issues to which architectural theorists propose solutions in the form of analytic-normative complexes of theoretical ideas. To understand why this is so, we must understand a little about design.

Design is of course only a part of the protracted processes by which buildings come into existence. The 'building process' involves formulating a need for a possible building, conceptualising what it might be like, initiating a process of resourcing, negotiating and organising, creating some kind of representation, or series of representations of increasing refinement, of what the building will be like, then constructing, fitting, operationalising, and finally occupying the completed building. Vernacular building is of course a less complex process. But if the circumstances exist in which 'design' is a function, then the corollary is that this more complex building process, or something approximating to it, also exists. Design does not exist as a function independent of this larger process. On the contrary, it implies it.

How then do we define design within this process? First, we note that it is only at the end of the process that the object of the process – an occupied building – exists. For most of its duration, the process is organised around a surrogate for the building in the form of an abstract idea or representation which continually changes its form. It begins as an idea for the building, then becomes an idea of the building, then a more formalised concept, then a series of more and more refined representations, then a set of instructions and finally a building. For the most part, the complex process of building takes place around this shifting, clarifying, gradually materialising idea.

The process of seeking, fixing, and representing a realisable concept of a building from an idea for a building is design. Design is what architects do, though it is not all they do, and not only architects do it. But it is design that keeps what architects do – whether or not it is architects that do it – fixed in the process of creating buildings. There has to be a control of the process of searching out, conceptualising, and representing the surrogate building through the process. Let us call this the 'design function', so that we can see that it is independent of who actually carries it out.

The design function exists within the building process for one fundamental reason: because at all stages of the process – though with differing degrees of accuracy – the properties and performance of the building as it will be when built must be foreseen in advance, that is, they must be knowable from the surrogate. Without this foresight, the commitments of resources necessary at each stage of the process cannot be made with confidence. The design function is essentially a matter of stage-managing a constantly changing representation of what will eventually be a building, so that at all stages of the process there is in view a proposal for an object

The need for an analytic theory of architecture

that does not yet exist, and which is probably unique – since if it were a copy there would be no need for design – but whose technical, spatial, functional and aesthetic properties if and when built are, as far as possible, predictable in advance.

The design function in the building process therefore involves on the one hand searching out and creating a representation of a possible solution for the design problem in hand, and on the other the prediction of the performance of the building when built from the representation. The activities that make up the design process reflect this duality. Design essentially is a cyclic process of generating possible design proposals, then selecting and refining them by testing them against the objectives the building must satisfy – to be beautiful, to be cheap, to be ostentatious, to represent an idea, to repay investment, to function for an organisation by providing adequate and well-ordered accommodation, and so on.[7] These two basic aspects to the design process can be called the creative phases and the predictive phases. In the creative phases the object is to create possible design proposals. In the predictive phases, the object is to foresee how proposals will work to satisfy the objectives.

Once we understand the creative-predictive nature of the design process, then it is easy to see how the normative and analytic aspects of theories can usefully contribute to the process. Theories can be used, and often are used, tacitly or explicitly, in two quite distinct modes in the design process: as aids to the creative process of arriving at a design; and as aids to the analytic process of predicting how a particular design will work and be experienced. Often of course these two aspects will be conflated in a undifferentiated thought process. The normative aspects of a theory tells the designer where to search for candidate solutions in the creative phases, the analytic aspects how the solution will work. For example, if you are a Palladian, then in the creative phases of design you search for a formal and spatial solution with Palladian properties – a certain range of envelope geometries, certain symmetries of plan and façade, certain kinds of detailing, and so on – confident that if you proceed in a Palladian manner then you can predict a Palladian outcome. If you are a Newmanite, then you search for formal and spatial solutions with a certain layering of spatial hierarchies, certain possibilities of surveillance, the avoidance of certain formal themes and so on, again confident that by proceeding this way a safe environment will result. Theory thus structures the search for a possible design in a solution space that might otherwise be both vast and unstructured, and it does so in a way that gives the designer confidence – which may of course be quite misplaced – that the nature and properties of the eventual building can be known from the theory.

The use of theory is of course only one way of structuring the design process. In fact few designers claim to create designs from theory, and many would go out of their way to deny it. But this does not mean that they do not design under the influence of theory. Much use of theoretical ideas in architecture is tacit rather than explicit. This is not due to malign intent on the part of designers, but much more to do with the need for theory in design, however little this is recognised.

The need for an analytic theory of architecture

Consider, for example, the problem of prediction. Having created a candidate design, the designer now has the task of foreseeing how the 'unknown non-discursivities' of form and space that will be created by the design will work and be experienced when built. Logically there are only two possible bases for such prediction: known precedent and theoretical principle. Prediction by precedent means prediction by reference to known cases that already exist. Prediction by principle means prediction by reference to the generality of known cases. Both are essentially claims based on experience, but the former is specific, and the latter general.

Prediction from precedent raises two problems. The idea of architecture includes the idea that the building to be created will not simply be a copy of one which exists. This means that precedent cannot be used lock stock and barrel for the whole building. Precedent can therefore only be used piecemeal for aspects or parts of the building. Since formally and spatially buildings are complex configurations, and not simply assemblages of parts, it can never be clear that the new embedding of a precedent attribute or part will not work differently in the context of the new whole. The use of precedent in design is necessary, since it brings in concrete evidence in support of prediction, but it is never sufficient, because each new synthesis recontextualises each aspect of precedent. The use of precedent therefore necessarily involves interpretation.

The pressure on designers to work at least in part from knowledge of theoretical principle is therefore intense. The apparent advantage to the architect of working within a particular theory becomes the solution to the prediction problem appears already to be contained within the theory. The normative theoretical concepts that guide the generation of a candidate design also take the form of analytic concepts which indicate that if the designer follows the precepts of the theory, then it is to be expected that the design will work in the way the architect intends. The analytic foundations of the normative theory return at the predictive stages to appear to guarantee architectural success. This is why architectural theories take the form of normative-analytic complexes. They fulfil the two primary needs of the design process with a single set of propositions.

However, it is clear that these advantages will only exist to the extent that the theory's analytic foundations are not illusory. If they do not offer a realistic picture of how the world works, then it is likely that the designer's predictions will refer only to an illusory reality. A poorly founded analytic theory will not inhibit the designer in the creative phases of design, but it would lead him or her to look in the wrong place. It would also mean that the designer's predictions would be unlikely to be supported by events when the building is built. This is why bad theories are so dangerous in architecture.[8] They make design appear to be much easier, while at the same time making it much less likely to be successful. This, in the last analysis, is why architects need analytically well founded theories.

However, this is not the same as to say that architects simply need scientific theories to guide them in design. The dual use of theory in architecture both to generate designs and to predict their performance permits us to introduce a very

The need for an analytic theory of architecture

important comparison: between theories in art and theories in science, and to argue that architecture needs theories both in the sense that the word is used in art and in the sense that the word is used in science.

Theories in art are not analytic-normative complexes of the kind we typically find in architecture. They are primarily about supporting the creative process, that is, they are in essence about *possibility*. Theories in art expand the realm of the possible, by defining a new way to art or even by defining a new form of art. There need in principle be no constraints on what type of theories are used. The role of a theory in art is not to claim a universal art, or to set up one form of art as superior to another, but to open up one more possible kind of art. Theory in art is then essentially generative. It does not have to take much account of functional or experiential consequences. It uses abstract thought only to generate new possibilities in art that had not been seen before.

If architecture were simply an art, it would need theories only in the sense that painters or sculptors have theories: that is, as speculative extensions of the realm of the artistically possible. It is clear that architecture as art has and needs this kind of theory. But this is not all it has and needs. The difference between architecture and art is that when an artist works, he or she works directly with the material that will eventually form the art object – the stone, the paint and so on. What the artist makes is the work of art. Architecture is different. An architect does not work on a building, but a representation of a building we call a design. A design is not simply a picture of a building, but a picture of a potential object and of a potential *social* object – that is, an object that is to be experienced, understood and used by people. A design is therefore not only a prediction of an object, rather than an object itself, but, however functionally non-specific it claims to be, a prediction of people in relation to building. This is where analytic theories are needed, and analytic theories are analogous to scientific theories. Theories in science are sets of general, abstract ideas through which we understand and interpret the material phenomena the world offers to our experience. They deal with how the world is, not how it might be. Because architecture is creative it requires theories of possibility in the sense that they exist in art. But because architecture is also predictive, it needs analytic theories of actuality as well as theories of possibility.

It is this double nature that makes architectural theories unique. They require at once to have the generative power of theories in art and at the same time the analytic power of theories in science. The first deals with the world as it might be, the second with the world as it is. The question then is: how may there be theories of architecture which are at once creative and analytic. One aspect of the answer turns out to be simple: good analytic theories are already likely to be also good theories of possibility. The entire usefulness of scientific theories in their applications in science and technology is in fact founded on the simple but unobvious fact: that analytic theories do not simply describe the world as it is, but also describe the limits of how it can be. Scientific theories are arrived at through the examination of the world as it is. But it is exactly the theoretical understanding of the world as it is that opens up

The need for an analytic theory of architecture

whole realms of new possibility that do not yet exist.

It is this fundamental link between actuality and possibility that opens the way to an analytic theory of architecture. But before we explore it, we must first look a little more carefully at architectural theories to see how they are structured, and why, and how they might eventually move in the direction of becoming more analytic.

The problem of architectural theory

The most common problem with architectural theories is that they have too often been strongly normative and weakly analytic, that is, it has been too easy to use them to generate designs, but they are too weak in predicting what these designs will be like when built. The theories of modernism were, for example, quite easy to follow in generating designs to satisfy normatively stated objectives. The problem was that the architectural means proposed were not the means required to achieve those objectives. The theories were weakly analytic. They did not deal with the world as it actually is. The normative dominated the analytic.

Exactly how normatively strong but analytically weak architectural theories are held in place can be seen by taking one more step in disaggregating what architectural theories are like and how they work. For example, looking a little more closely at our two exemplars of architectural theories – the Albertian and the Newmanite – we find both have two quite distinct components: one in the realm of broad intention, telling architects what they should aim to achieve through architecture, and one in the realm of what we might call architectural technique, telling architects how to realise that intention. Alberti's theory, for example, tells architects that in order to design buildings that people will experience as harmonious, they should aim to reflect in their buildings the mathematical order found in nature. He then goes on to offer a method for calculating proportions to serve as a technique for realising this aim in architectural terms.[9] Newman tells architects they should aim to design spaces beyond the dwelling so that inhabitants may identify with them and control them, then specifies hierarchical techniques of space organisation in order to realise this. We might call these the broad and narrow propositions about architecture contained in a typical architectural theory. The broad proposition, or intention, sets a goal while the narrow proposition, or architectural technique, proposes a way of designing through which the intended effect will be realised.

One difference between the broad and narrow propositions lies in what they engage. The broad proposition engages a world of ideas which may be very large in its scope and may contain much that is poorly defined and little understood. The narrow proposition, on the other hand, engages the realities of architectural design and experience. If in general theories are abstract propositions which engage the real world of experience, then the broad and narrow propositions of architectural theories occupy opposite ends of the spectrum covered by theories. The broad propositions are in the realm of philosophical abstraction, where the theory engages the vast world of ideas and presuppositions, implicit and explicit, which eventually rests nowhere but in the evolution of human minds. The narrow propositions are in

The need for an analytic theory of architecture

the realm of direct experience of the world where theories engage the minutiae of everyday experience.

Broad proposition and narrow proposition also differ in their intended universality. Broad propositions are intended to be universalistic in that they attempt to say things about architecture which are held to be generally true, and to say it in such a broad way as to allow it to be true in quite different architectural circumstances. But it is clear that we should not regard the narrow propositions as universalistic.[10] For the most part the narrow propositions are offered as possible techniques for realising an abstractly stated aim, not the only such techniques. On reflection, again this must be so. The narrow propositions of an architectural theory are techniques for bridging between the abstract and the concrete. Only an abstraction can be general. We should not mistake a technique for realising an analytic abstraction for the abstraction itself.

Now consider these broad and narrow propositions in relation to what is required of theory in the two phases of design, that is, in the first phase, ideas about possible forms and, in the second phase, ideas about the relations between forms and performance outcomes. Both of the theories we have been considering appear to supply both needs. Ideas of possible forms are contained in the narrow propositions, that is, the constructive techniques through which the theorist advises the designer to go about design to ensure success. In the case of Alberti's theory, this means the systems of worked out proportions which guide the designer in setting up the building as a physical form. In Newman's case, this means the diagrams of spatial hierarchy which the designer can follow in setting up the spatial design. Ideas of the relation between form and functional outcome are then expressed at the more philosophical level of the broad propositions. In Alberti's case, this means the broad propositions, based on the analogy with music, about the human experience of visual harmony.[11] In Newman's case, it means the broad propositions about 'human territoriality' and its spatial implications.[12] In other words, in both cases, it is the highly specific narrow propositions which guide the creative process of design, and the very generalised broad propositions which guide the designer in predicting functional effect from formal configuration.

Now the problem with most architectural theories is that this is exactly the opposite of what is required for architecture which is creatively innovative and functionally successful. In the generative phase of design, what is needed if architectural creativity is to be maximised is ideas about formal and spatial configuration which are as unspecific as possible about specific solutions, in order to leave the solution space as open as possible to creative invention. In the predictive phases, what is needed is precision about specific forms since what is at issue is the prediction of the functional outcome of this or that real design. In the generative phases, where what is required are abstract or genotypical ideas which open up realms of possibility just as theories do in art, architectural theories of this type offer a rather narrow range of solution types which are essentially no more than a set of abstract exemplars to follow – particular systems of numerical proportions in one case, particular diagrams of hierarchical spatial relations in the other. Then when in the predictive phases of design the designer needs a much greater degree of analytic precision in order to foresee how this or that innovative form will work functionally or experientially, all that the theories offer is the vague analytic generalisations of the broad propositions.

In other words, architectural theories of this type are over-specific where they should be permissive and vague where they should be precise. The designer is given concrete models to follow when he or she needs constructive creative ideas to search the solution space, and vacuous abstractions when he or she ought to be given techniques to predict the performance of particular designs. This is, in a nutshell, the problem with most architectural theories, and this is how, in real design, the normative aspects of theory come to dominate the analytic. What is needed are theories with the reverse properties, that is, theories that are as non-specific as possible to particular solutions in the generative phases of design in order to leave the solution field as large and dense as possible, and as specific and rigorous as possible in the predictive phases in order to be able to deal predictively with unknown forms where the need for effective prediction is greatest. The implication of this is that we need a fully fledged analytic theory which would offer abstract understanding rather than specific models in the creative phases of design, and phenotypical precision rather than vague generalisations at the testing stages.

What exactly, are theories?

How should we go about setting up such a theory? The first step must be to make sure we understand exactly what an analytic theory is. This turns out to be not as easy as looking the word up in a dictionary. Few words are in fact more ambiguous in their origins than 'theory'. In its ancient Greek origins, the verb *theoreein* means to be a spectator, and the products of this speculative activity, *theoremata*, were, not surprisingly, speculations. For Bacon theories were simply errors, the 'received systems of philosophy and doctrine', to be replaced in due course by something altogether better.[13] This meaning is still reflected in everyday use. In common usage, theories are speculations, of lesser status than facts, at best a temporary fix until the facts are known. A fictional detective with a premature 'theory' about a case will almost certainly be shown to be wrong. The expression 'only a theory' clearly expects theory not to be eventually supported by 'facts', but to be replaced by facts. In these senses, theories embody irremediable uncertainties, and appear to constitute a form of thought whose object is to replace itself with a-theoretical, and therefore secure, knowledge. In complete contrast, in modern science the word 'theory' today stands for the deepest level of understanding of phenomena. Successful theories in areas where none had previously prevailed, like evolution theory, are the most epoch making of intellectual events. Conflicts between rival theories of, say, the origins of the universe or the nature of matter, conducted on the obscure battlefields of macro and micro phenomena, are among the epics of the late twentieth-century thought.

So what then is 'theory', that it can be subject to such a range of interpretations and ambiguities? The source of this ambiguity lies of course not in the vagaries of etymological history but in the nature of theories themselves. Theories are found in the realm of speculative thought, because they are at root, speculations. They are not in themselves, for example, statements about observable phenomena, nor even statements about the regularities that are to be found in observable phenomena. They are propositions about hypothetical processes which might be responsible for the regularities we see in phenomena. As such they have a necessarily abstract nature, and are purely conceptual entities. You cannot see a theory, only its consequences, so you cannot verify a theory, only phenomena that are consistent with it. When we test a theory we do not simply look at the theory to see if all the parts are in working order and properly related, though we do also do this. We check the theory by seeing how far the phenomena available in the real world are consistent with the theory, and preferably with no other. To check a theory, in effect, we look away from the theory. Theories are in themselves unobservable and unexperiencable, and this is why in the end even the best and the most durable remain in some sense speculative.

But even when we accept the abstract and speculative nature of theories, we have not yet exhausted the apparent indeterminacy of the idea. No set of concepts which become part of a theory can exist in isolation. On the contrary, concepts can only exist as part of conceptual schemes through which we interpret our experience of the world and turn information into knowledge. No concept or set of concepts can exist in a vacuum. Each must be embedded in a broader range of propositions or assumptions about what the world is like and how it works. These broader frameworks have been known as paradigms since Thomas Kuhn first drew attention to their existence.[14]

With all this indeterminacy in what we mean by theory, how is it that they can be so important and so useful. To answer this we must understand the circumstances in which theories arise and what purposes they serve. Theorisation begins when we note a certain type of phenomenon and make a certain type of presupposition. The phenomenon we note is that of surface regularity in the world as we experience it. The presupposition we make is that *surface regularity* implies *underlying invariance* in the processes that give rise to the phenomena we see.

The first of these – the noting of regularities – theorisation shares with language. The fact that language has words for classes of things rather than simply for individual things assumes that we know the difference between order and chaos, that is, that we can discern in the objective world ‘structural stabilities’ [15] which are sufficiently well defined and repetitious to support the assignment of names. These names are, as philosophers have endlessly noted, abstract terms for classes in the guise of names for things, with the consequence that even such a simple apparently concrete act of pointing at a thing and naming it depends on the prior existence not only of the abstract universal constituted by that class name, but also of the scheme of such abstractions of which that particular abstraction forms a part. These schemes, as we have known since de Saussure,[16] differ from one language to another so that we are compelled to acknowledge that names are not

The need for an analytic theory of architecture

neutral, simple handles on things, but conceptual instruments by which we create an organised picture of the world. Names create understanding, and it is against the background of the organised picture of the world already given to us by language and culture that theorisation begins.

Theory begins in the same place as language where we note, in the flux of experience, regularities, but adds a further presupposition: that since regularity is unlikely to be the product of chance, there must be some kind of order not only in the regular phenomena that we observe but also in the processes that give rise to the phenomena. Why we should make this presupposition is not clear. But it seems plausible that just as language seems intimately bound up with how we cognise the world so theorisation is bound up with how we act in the world. When, for example, we strike stones to make sparks and then fire, the sequence of events from one to the other is not inscribed on the surface of things but implies some interior process which is set in motion by our actions. Just as the world responds to our actions on it by producing regularities, so we presuppose that the existence of regularities which do not result from our actions must be the result of invariant processes analogous to our actions. If then language arises from our being in the world and needing to know its objective persistences, so theorisation seems to arise from our acting in the world and on the world and needing to know the interior processes by which outcome reliably follows from action.

We thus see that regularities are the starting point of theory, but they are not the theory itself. Regularities initiate the process of theorisation since we infer from the existence of regularities that there must be some invariant structure in whatever process it is that produces these surface regularities. Theories are concerned with the nature of that process, more precisely they are attempts to model the invariant structure of processes which are thought to exist for there to be surface regularities. A theory, then, is not a list of regularities. Regularities are what theory seeks to explain, but are not in themselves theory. They initiate the search for theory but are not and cannot be its end point. A theory which seeks to 'explain' regularities is an entity of an altogether different kind from a list of regularities.

Moreover, although theorisation moves on from language by seeking to identify the hidden processes that give rise to surface regularities, it does not begin in a conceptual or linguistic void. It begins in the only place it can, in the evolution of thought and language, and their relation to the space-time phenomena that we experience 'without trying'. Because thought and language already give us a picture of the world which, at some level at least, seems to reflect its order and therefore to explain it, we are compelled to acknowledge that when we begin the process of theorisation we are already in possession of a view of the world which in many ways is very like a theory, in that it makes the world seem a more or less coherent and organised place. The difference is that the theory-like understanding we acquire from culture and language reflects not an interior order which gives rise to the surface regularities but an order in those surface regularities themselves. When for example language tells us that 'the sun rises', it reflects the regularities that we note on the surface of things, not a hidden process which gives rise to this surface regularity. We might usefully then think of such everyday constructions as 'theory in the weak sense'. Analytic, or scientific theories, are 'theories in the strong sense'. They aim at a greater truth because they seek not to bring order to surface regularities but to show how those surface regularities arise from invariant necessities buried deep in the nature of things.

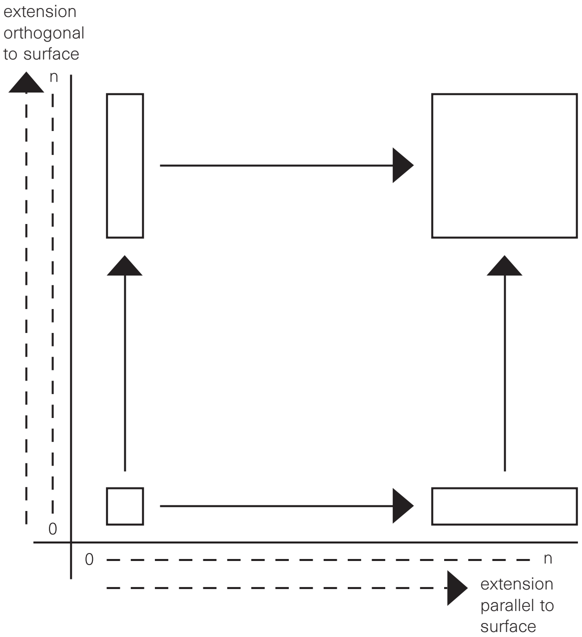
Formally defining simple regularities

Because surface regularities are the object of theory, the first step in theorisation is to formalise the idea. In fact, there is a beautifully simple way to extract the idea of regularity from phenomena and represent it as pure regularity, independent of the overall qualitative nature of things. The idea is that of translating the properties of objects in the world as we see them in real space into an abstract space which allows us to be quite clear about what these properties are. This is done by the familiar technique of replacing the space within which the object exists with an abstract co-ordinate system in which the axes represent those properties of the object that seem to be of interest as regularities. Thus one co-ordinate might represent the height of the object, another the length and another the breadth. We may then represent any object which has these properties as a single point in the 'property space'.

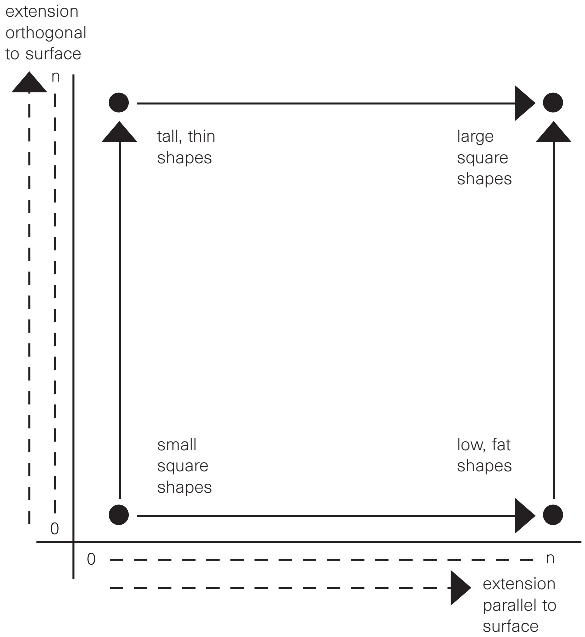
Once we can represent the properties of an object as a point in a property space rather than as that set of actual properties in real space, we can easily represent exactly what we mean by a regularity as far as these properties are concerned. For example, to the extent that things are comparable to each other in more that one property in the property space, the points representing them in the property space will cluster in a particular region of the space. Clusters in the property space give a formal meaning to the idea of a type or class of things, in so far as those properties are concerned. If things when represented as points in the property space are randomly distributed throughout the space, that is, if there are no clusters, then we would say either that there were no types, but only individuals, or that we had selected the wrong properties for analysis. If on the other hand we see clusters, we infer that things tend to fall into types, by which we mean that variation on one property tends to be associated with variation on at least one other, or perhaps many others. This is shown graphically in the top two diagrams of figure 2.1. We may equally use the property space to formalise the idea that the regularity we see lies not in apparent classes or types of things but sequences of states of things. In this case we ask: when an entity changes on one dimension, does it change in any other? It it does, then the regularity will show itself as a regular pattern in the distribution of entities in the property space. This is shown in the bottom two diagrams in figure 2.1. When we see such a pattern, we would infer that some process if not of cause and effect then at least of regular co-variation was in operation, since each time one variable was changed a change in another variable regularly appeared.

The need for an analytic theory of architecture

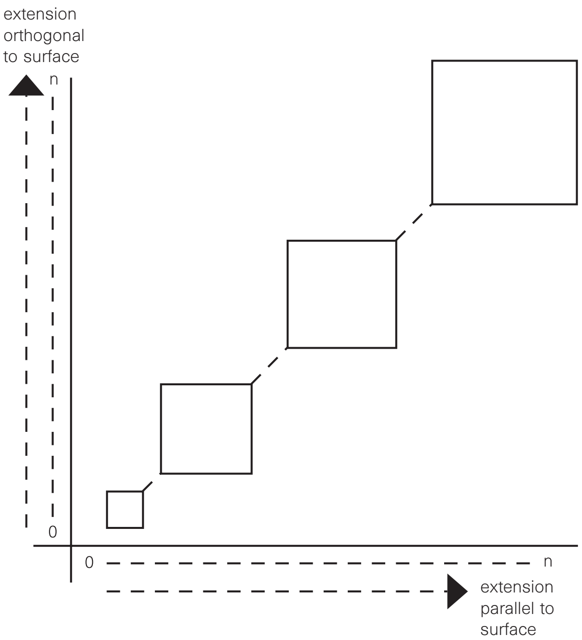
Figure 2.1



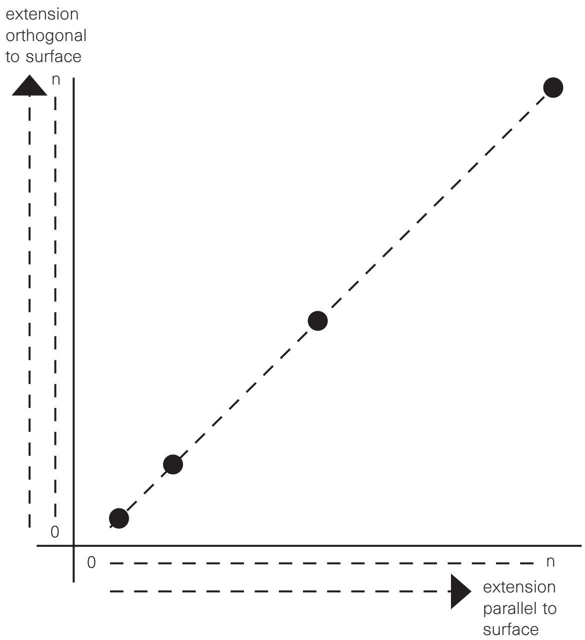
We first create a coordinate system in which vertical axis represents increasing size orthogonal to the surface and the horizontal axis increasing size parallel to the surface. We then draw in the space shapes conforming to the coordinates. In the bottom left corner we start with an initial square shape. This becomes extended vertically as we move upwards, laterally as we move horizontally and in both dimensions if we move in both directions.



Because the properties of lateral and vertical expansion are represented inthe axes, then the points in the property space represent the changes in spape in the left figure no less than the actual shapes themselves. Other properties not in the coordinate system would of course be omitted. The points represent only the properties selected for representation in the property space. The idea of types or classes of shapes can thus be shownby point clusters.



In the above case, the shapes drawn in the coordinate space show that an increase on one dimension is always related to an increase in the other. This type of regular association between changes is the surface phenomenon we associate with 'cause and effect'-one change seems to bring aboiut another. But what is shown in the figure is not 'cause andeffect', merely the regular association of changes. To 'explain' these regularities we would need to show why they are necessary.



Associated changes can also be represented as points in the property space. The position of points indicates that a change in one dimension is always associated with a change in the other. This representation is also known as a 'scatergram'. The degree to which points fomr a line from bottom left to top right can be indexes by a 'correlation coefficient', a value between 0 and 1 which indicates how strongly one change implies the other.

The need for an analytic theory of architecture

We might then reasonably say that questions about types, that is, about similarities and differences, are questions of the form: do entities cluster in particular regions of the property space?; while questions about cause and effect are of the form: when entities move in one dimension of the property space do they move in another?[17] Both of these describe the apparent regularities of surface phenomena, that is, the appearance of types and the appearance of cause and effect, in an abstract way. The property space is a means of controlling the attributes that are to be accounted for in the pattern of similarities and differences. Where the real object is present, all its properties are manifest. In the property space, only selected properties are present. Of course, everything depends on our selecting the right properties for the property space in the first place. For this reason we can never be sure from the absence of a regularity that no regularities are present in these phenomena.

But even if we go through a long process of experimenting with different properties until we eventually find the clusters or covariations that indicate the presence of regularities, it will always still be the surface phenomena that are represented regardless of the degree of abstraction. We are still seeing the surface of things, that is, apparent regularities of things as presented to our experience. We are not seeing the theory that purports to account for those regularities, that is, we are not seeing the model of the structures of the process which might account for these regularities. What we are doing is recording phenomena in such a way as to be able to see clearly what we mean by regularities, by translating properties into the dimensions of a coordinate space and locating objects as points within this space so that only the regular properties are represented in what we see. This both seems to be and is a fundamental way – maybe the fundamental way – of rigorously recording similarities and differences, and constant associations between things, within an objective and independent framework.

The meaning of the word theory can then be made precise. As we have said, just as the *a priori* given for the noting of regularities is that we know the difference between order and randomness, the *a priori* given for taking this into theorisation is that regularity on the surface implies some systemic process below the surface, such that the structure of that system is in some sense invariant. A theory is an attempt to model these invariants in a system of interdependent concepts. A theory is a model because it deals with the way in which things must be interrelated in order to produce the surface phenomena, and abstract because it represents the system by some means other than that of the system itself. A theory is a model, but not in the sense that a physical model is a model, that is, a small copy of the thing itself, but in the contrary sense of a model taking as abstract a form as possible, uncommitted to any particular kind of representation or embodiment. In its purest form, a theory is a kind of abstract machine, since it is an attempt to create an abstract representation of the working of processes which give rise to what we see.

Now the enormous power of theories arises from one very specific property of such 'abstract machines', a property we have already touched upon. Because theories are abstract working models of processes which give rise to the actual,

The need for an analytic theory of architecture

they also give a basis for conjecturing about the possible. Theories in effect allow us to go beyond the accumulated experience of reality and conjecture possible states of reality that are compatible with the model. It is this link between the actual and the possible that makes theories so useful for prediction. To 'apply' a theory is essentially to pose the question: is what is proposed a possible case?

It is too limiting then to call theories 'explanations' of how the world is. A theory defines the invariants that underlie many different states of reality. It is in principle unlikely that all possible states of a particular set of phenomena already exist or are already known. It is likely then that the theory will also predict possible states that do not exist but could according to the model. It is this property above all others that imparts to theory its immense power as a tool of thought and as an agent of human creativity, and also its practical usefulness. However, it is clear that these virtues will arise only to the degree that the theory captures invariants that really are 'out there'. But how can this be? How can an abstraction capture what is really 'out there'. To take this next step, we must know a little more about how theories are put together, how they work, and what they are made of.

What are theories made of?
The first thing we must note is that theories are made of concepts, usually in the form of a system of interdependent concepts with two forms of expression: words, and formal expression, usually mathematical. Since everyday life and language is also run on concepts we must know the difference between a scientific concept and an unscientific one. What then is the difference? We can do no better than to discuss the concepts on which both language and science seem to be founded, that is, the difference between order and randomness.

Order and randomness are both concepts which have a powerful intuitive meaning. Both are very broad indeed in their application, so much so that it is very hard to pin down what the two terms mean with any real clarity. Both terms, and even more the way they are related, express complex intuitions about the way the world is. Each term can be used in a wide range of situations, and the meaning only becomes clear enough to feel understood in the spoken or written context. This is common enough. The intuitive concepts that pervade and give sense to our languages have this richness and imprecision, so they can be used in a great variety of situations, and indeed it is only in the context in which a concept is used that its meaning becomes unambiguous.

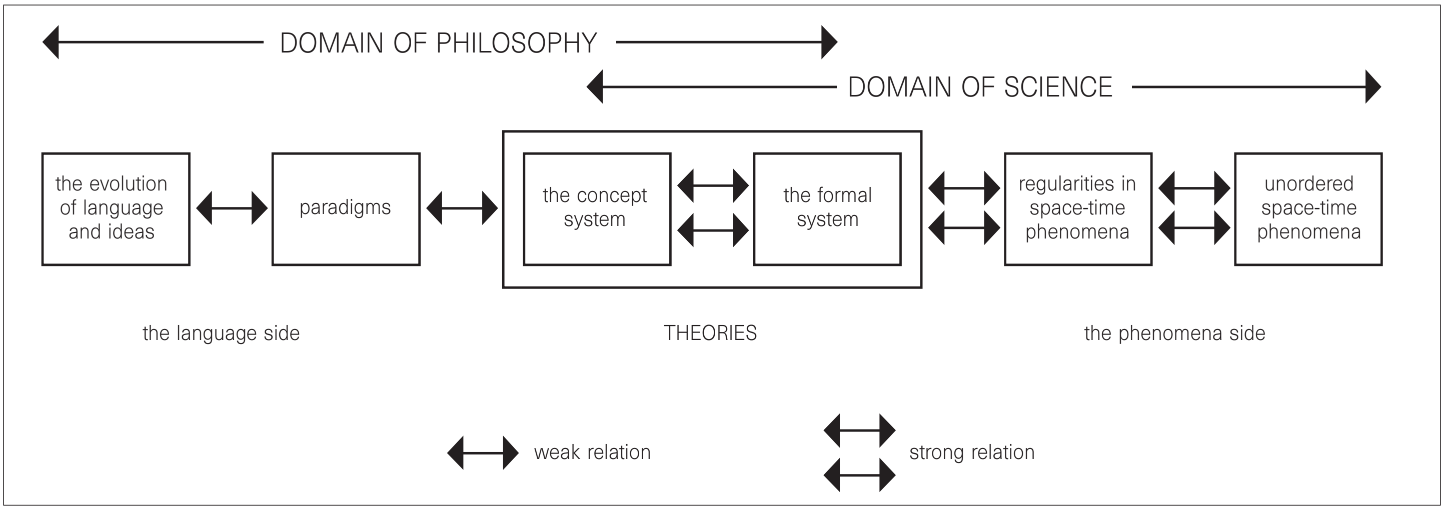
In science, it is exactly this richness and imprecision that is restricted. Scientific concepts, although expressed in language, are much narrower in their potential application than normal linguistic concepts. But they are also more systemic, in that they compress and express more interrelationships between concepts. They express more connection between things, but at the cost of a narrowing of the range of application. The concept of 'entropy' is a good example of this because it relates both order and chaos in a systemic way, and in doing so restricts the range of application of the new synthetic concept to those situations where precisely these systemic relations hold. The degree of entropy in a system describes that system's position in a continuum from order to chaos. Like many scientific concepts of great profundity and generality it can be explained simply, though not through words but through a simple model.[18] Imagine two jars, *a* and *b*, with *a* containing 100 balls numbered 1–100, and *b* empty, and some system for selecting a random number between 1 and 100 – say, a pointer on a spindle which can be spun so that it lands with equal likelihood on the numbers set out in a circle. Spin the pointer and when the point rests on a number, find the ball with that number and transfer it from whichever jar it is in to the other one. Then repeat this operation as many times as necessary. What happens? Intuitively – and correctly – we say that the process will settle down to about half the balls in each jar. Why? The answer tells us what entropy is and how it can be measured. The first time the pointer selects a number, the probability that the ball selected will go from *a* to *b* is 1, that is, it is certain, because all the balls are in *a*. The second time, there is one chance in 100 that the single ball in *b* will return to *a*, but 99 chances out of 100, that is, a probability of .99, that another ball will go from *a* to *b*. The next time there is one chance in 50 that one of the balls in *b* will return to *a*, but 98 chances out of 100 that another ball will go from *a* to *b*. Clearly, as the process goes on, the chances of balls going back from *b* to *a* gradually increase and the chances of balls going from *a* to *b* diminish correspondingly.

When about half the balls are in each jar, the probabilities are about equal, so the system tends to settle down to small variations about this state. To see why this happens let us define a microstate of the system as a particular distribution of individual balls in jars and a macrostate as a particular number of balls in each jar. There are, clearly, only 200 possible microstates of the system for the macrostate in which one ball is in one jar and the rest are in the other, that is, one for each of the hundred balls in each jar. For the macrostate with two balls in one jar and 98 in the other there are all possible combination of two balls for each jar, that is, 200 × 200. For the macrostate with three in one and 97 in the other there are all combinations of three balls. In other words, the number of microstates for the macrostate is maximised when the largest possible number are in the least full jar – that is, when half the balls are in each – because beyond that point there will be fewer balls in the other jar and everything happens in reverse.

This is why the system tends to the half and half state. There are far more microstates corresponding to the half and half (or near half and half) macrostates than for those in which a few balls are in one jar and many in the other. In other words all the system does is to tend to its most probable state. This is also the definition of the state of maximum entropy. Entropy is maximal in a system when the system is in one of the macrostates for which there are the largest number of microstates. An example of this, is where two gases are each randomly distributed in a container, without regions where one or other gas predominates. There are far more microstates with random distribution than microstates with concentrations of one or other gas in a certain region. Our model of jars and balls is then a statistical

The need for an analytic theory of architecture

Figure 2.2



representation of the mixing of two gases in a closed compartment – or for the gradual heat death of the universe as the universe tends from its current improbable state to its most probable state, that is, one in which heat is more or less evenly dispersed throughout the universe.[19]

In other words, entropy relates the notions of order and chaos into a single concept, but at the same time gives it a much more precise and limited reference to the world. However, it also does something else of no less importance. It permits the concept to be captured in a formal mathematical expression as well as through words. It is through this formal expression that the link between the concept and the observable world is made. This two-way emancipation of concepts, on the one hand reorganising concepts into more precise systems of interdependence and on the other relating them to the real world by associating them with formal expressions is the essence of what theories are.

Theories are therefore made of two things: words and formal expressions. But both represent concepts. A theory is a system of concepts with one type of expression, the verbal, which links the concepts back into our understanding, necessarily with some imprecision; and another, mathematical form which links the concepts forward into phenomena, necessarily with great exactness. Theories thus link our understanding to the world, connected to our understanding by linguistic concepts and connected to phenomena by formal expressions corresponding to the concepts.

This two-way relation using language and formalism to link concepts to our understanding on the one hand and to the real world on the other is the heart of what theories are. We may clarify all these complex relations in a diagram, see figure 2.2. This figure shows not only how theories intervene between language and the world, but also how science relates to philosophy, which overlaps with science in part of this overall scheme. The overall form of the diagram sets the evolution of language and ideas on the left and the phenomena of space-time on the right. Theories are in the centre, defined as a relation between a system of concepts and a system of formal expressions which looks two ways: through the concepts it looks

The need for an analytic theory of architecture

back first into the broader conceptual schemes we call paradigms, then into the evolving structure of language and ideas which are both an inevitable context and an inevitable constraint on theorisation; and through the formalism it looks forward towards the regularities in space-time phenomena which theories seek to account for, and then onwards into the general foreground of space-time phenomena which do not form part of the regularities but which may at any stage arbitrarily engage the theory by offering phenomena which are inconsistent with the 'abstract machine' for generating phenomena proposed by the theory.

The earliest ancestors of what we would recognise as 'scientific' theories, such as those of the Pythagoreans who are said to have first noted the relation between numerical ratios and forms occurring in nature, are probably best seen as paradigms rather than as fully fledged theories, although in their preoccupation with the relation between space-time regularities and formal expression they certainly prefigure theories in the modern sense.[20] Pythagoreanism (as we earlier noted as influencing Alberti) is a generalisation of a single concept which generated a way of looking at the world on the basis of a few results. This is legitimately a precursor to theory but not in itself what we ought to be calling a theory. However the attraction of such over-generalisation remains, as is seen in the prevalence of variants on Pythagoreanism in the mystical substitutes for theory which have continued to occupy the fringes of architectural thought throughout the twentieth century.[21]

Theories in the scientific sense are one step in from both paradigms on the one side and regularities on the other in that they are composed of concepts which are focused and related to each other to form a system, with precise relations between each concept and formal techniques or expressions which are used to check how far the regularities implied by the system of concepts are detectable in space time phenomena. Scientific theories thus require three relations to be particularly strong: the relations among concepts which form the conceptual system; the relations between concepts and formal techniques of measurement; and the relations between these formal techniques and space-time phenomena. In terms of the diagram, we may say then that science needs to be strong from the 'concept system' in the direction of phenomena.

Science is, and must expect to be, weaker in the other direction, that is, in the passage back through paradigms into the more general evolution of ideas. This tends to be ground occupied by philosophy. Philosophy overlaps with science in being interested in theories, and relating them back to broader families of concepts[22] right through to those that prevail in everyday life and social practices,[23] but does not normally preoccupy itself with the rigorous testing of theories against real space-time phenomena. Science and philosophy are rivals in the realm of theory, but only because their preoccupations reach out from theory in contrary directions with the effect that between them science and philosophy cover the ground that needs to be occupied by theoretical thought. However, it is because science moves from concepts to phenomena that its theories eventually come to have a puzzling status, because the intuitive sense that they 'explain' things comes

The need for an analytic theory of architecture

from the relation between the concepts that make up the theory and the sense we have from everyday language that our ideas 'explain' the world. Scientific theories are in this sense psychologically strongest where they are in fact weakest, that is, where the concepts that form the theory relate back into the broader conceptual systems which inform everyday life.[24]

Towards an analytic theory of architecture

Given these definitions, how then can there be an analytic theory of architecture? first, let us be completely clear about one thing. If there are no objective regularities in the real world of architectural form and space, linking the configurational aspects of form and space with behavioural and experiential outcomes, then there are no grounds whatsoever for seeking to build an analytic theory. The need for and the possibility of an analytic theory both stand or fall with the existence of such 'non-discursive regularities'.

This means that to build an analytic theory, non-discursive regularities must first be investigated and, if they exist, brought to light. How can this be done? We may first recall that an architectural theory is an attempt to render one or other of the non-discursive aspects of architecture discursive, by describing non-discursivity in concepts, words and numbers. We may say that an architectural theory seeks to create a 'non-discursive technique', that is, a technique for handling those matters of pattern and configuration of form and space that we find it hard to talk about. In research terms we could say that an architectural theory, at least in the 'narrow' aspects through which it describes and prescribes design decisions, is an attempt to control the architectural variable.

Now, as we have seen, architectural theories in the past have tended to be strongly normative and weakly analytic, because the non-discursive techniques proposed are only able to describe certain kinds of configuration. This is why in application they are partisan for that kind of configuration. For example, if a non-discursive technique describes systems of proportion in terms of numerical or geometric ratios, it is unlikely to be able to deal with configurations which lack such proportionality. It will only describe those cases where these proportions hold. In any attempt to apply such partisan techniques generally, they are more likely therefore to act as distorting mirrors than a discovery of new regularities. Likewise, if our non-discursive technique is a system of diagrams expressing spatial hierarchy, it is unlikely that those techniques can be usefully applied to the vast range of cases where such clear hierarchisation is not found. It follows again that such a technique will be useless for investigating spatial patterns in general.

We can say then that a non-discursive technique which is partisan for – usually because it is a product of a preference for – one particular kind of non-discursivity, will not be usable as an analytic tool, and cannot therefore be used for the discovery of non-discursive regularities. This deficiency, however, does point us in the direction of what is needed. To bring to light non-discursive regularities, we need non-discursive techniques for the description of either spatial patterns or formal patterns (or conceivably both) which are uncommitted to any particular type of spatial or formal configuration or pattern, and which are capable of general application to describe all possible types of pattern. For example, it ought to be able to handle spatial patterns or built form patterns which lack geometric regularity as well as those which have it. Unless this can be done with rigour there there is little hope that theoretical propositions in architecture can ever be analytic in the sense that we require them to be.

The next chapter of this book will introduce such a set of non-discursive techniques for the analysis of configuration, first developed in spatial form as 'space syntax', but now being broadened to cover other aspects of configuration. These techniques have been used over several years for two principle purposes, first to discover how far it was possible to bring to light and subject to rigorous comparative analyses the configurational aspects of space and form in building through which culture is transmitted, and second, through these comparative studies to develop a corpus of material which would permit the gradual development of a general theory of architectural possibility. The remainder of this book is essentially an account of the progress that has so far been made in this project.

As we will see, what we discover through applying these techniques to the analysis of spatial and formal patterns in architecture, wherever they are found and whatever their embodiment in either buildings or urban systems, are invariants in patterns which lie not on the surface of things but which are buried in the nature of configurations themselves. These invariants we can think of as deep structures or genotypes. Each cultural manifestation through building, whether as a building 'type' for a particular purpose, or a particular architectural ethos or imprinting of culture on building, does so through such genotypes. For example, seen as systems of organised space, it turns out that towns and cities have deep structures which vary with culture. Likewise, seen as organised spaces, buildings for different function purposes also have deep structures or genotypes. These genotypes are – or embody – cultural or typological invariants. These are not of course general laws. They are at best the 'covering laws' of cultures. There are the genotypical invariants by which each society and each function in society seeks to express itself through architecture.

However, as we build our corpus of genotypes we gradually begin to see that there is another level of invariance: there are genotypes of the genotypes. Below the level of cultural variation in architecture there exist invariants across cultures and types. These 'genotypes of genotypes' are not the covering laws of cultures but the invariant laws that bind humankind in general to its artificial material world. They are the abstract raw material out of which all configurational possibility in space and form in the built world are constructed. It is at this level of invariance – and only at this level – that we can build a genuine analytic theory. These possibilities will be dealt with in Chapters 8 and 9.

The need for an analytic theory of architecture

Architecture as art and as science

If this theoretical project is eventually to succeed – and it is beyond the scope of any single book to do more than take a few faltering steps towards such a theory – then it is clear that such a theory would liberate rather than constrain design. At root, the need for architectural theory arises from the need to formulate principles from the experience of having built to inform and guide us on how we might build. This dynamic between the actual and possible is the essence of architectural theorising. Architectural theory arises from the fact that architects can neither forget the architectural tradition, nor repeat it. In architecture, theory is not simply a means to fix a picture of the world in a certain form. It is also the means by which form is destabilised and a new future is conceived. Architecture progresses by incorporating its reflection on the past into an abstract frame of possibility. This frame is theory. Without it, historical thought is sterile, and can only lead to imitation of the past. Through the intermediary of theory, reflection on the past becomes possible future. History constrains, but theory liberates, and the more general the theory, the greater the liberation.

Does this mean then that the line between architecture as science and architecture as art needs to be redrawn closer to science? I do not believe so. We can call on the beautiful ideas of Ernst Cassirer on the relation between art and science.[25] 'Language and science', he writes, 'are the two main processes by which we ascertain and determine our concepts of the external world. We must classify our sense perceptions and bring them under general notions and general rules in order to give them an objective meaning. Such classification is the result of a persistent effort towards simplification. The work of art in like manner implies such an act of condensation and concentration...But in the two cases there is a difference of stress. Language and science are abbreviations of reality; art is an intensification of reality. Language and science depend on one and the same process of abstraction; art may be described as a continuous process of concretion... art does not admit of...conceptual simplification and deductive generalisation. It does not inquire into the qualities or causes of things; it gives the intuition of the form of things...The artist is just as much the discoverer of the forms of nature as the scientist is the discoverer of facts or natural laws.'

Those of us who believe that science is on the whole a good thing, accept that science is in one sense an impoverishment – though in others an enhancement – of our experience of the world in that it cannot cope with the density of situational experience. It has to be so. It is not in the nature of science to seek to explain the richness of particular realities, since these are, as wholes, invariably so diverse as to be beyond the useful grasp of theoretical simplifications. What science is about is the dimensions of structure and order that underlie complexity. Here the abstract simplifications of science can be the most powerful source of greater insight. Every moment of our experience is dense and, as such, unanalysable as a complete experience. But this does not mean to say that some of its constituent dimensions are not analysable, and that deeper insight may not be gained from such analysis.

The need for an analytic theory of architecture

This distinction is crucial to our understanding of architecture. That architectural realities are dense and, as wholes, unanalysable does not mean to say that the role of spatial configuration (for example) in architectural realities cannot be analysed and even generalised. The idea that science is to be rejected because it does not give an account of the richness of experience is a persistent but elementary error. Science gives us quite a different kind of experience of reality, one that is partial and analytic rather than whole and intuitive. As such it is in itself that it is valuable. It needs to be accepted or rejected on its own terms, not in terms of its failure to be like life or like art.

It is in any case clear that the dependence of architecture on theories, covert or explicit, does not diminish its participation in Cassirer's definition of art. This is true both in the sense that architecture is, like art, a continuous process of concretion, and also in the sense that, like art, 'its aspects are innumerable'. But there are also differences. The thing 'whose aspects are innumerable' is not a representation but a reality, and a very special kind of reality, one through which our forms of social being are transformed and put at risk. The pervasive involvement of theory in architecture, and the fact that architecture's 'continuous concretion' involves our social existence, defines the peculiar status and nature of 'systematic intent of the architectural kind': architecture is theoretical concretion. Architects are enjoined both to create the new, since that is the nature of their task, but also to render the theories that tie their creation to our social existence better and clearer. It is this that makes architecture distinct and unique. It is as impossible to reduce architecture to theory as it is to eliminate theory from it.

Architecture is thus both art and science not in that it has both technical and aesthetic aspects but in that it requires both the processes of abstraction by which we know science and the processes of concretion by which we know art. The difficulty and the glory of architecture lie in the realisation of both: in the creation of a theoretical realm through building, and in the creation of an experienced reality 'whose aspects are innumerable'. This is the difficulty of architecture and this is why we acclaim it.

Notes

1 Roger Scruton, *The Aesthetics of Architecture*, Methuen, 1977.
2 Ibid., p. 4.
3 L. B. Alberti, *De Re Aedificatoria*, 1486; translation referred to: Rykwert et al. (1988), MIT Press, 1991.
4 O. Newman, *Defensible Space*; Architectural Press, 1972.
5 Alberti, Chapter 9.
6 Newman, pp. 3–9.
7 How this happens as a cognitive process is the subject of Chapter 11: 'The reasoning art'.
8 Chapter 11 includes a case study in the dangers of bad theory.
9 Alberti, for example, Book 9.

The need for an analytic theory of architecture

10 Scruton's fundamental error is to confuse these two aspects, and in effect to believe that the narrow propositions of architectural theory are intended to be universalistic. See Scruton, *The Aesthetics of Architecture*, p. 4.
11 Alberti, Book 9.
12 Newman, pp. 3–9.
13 F. Bacon, *The New Organon* (1620), Bobbs Merrill, 1960, Aphorisms Book 1, Aphorism cxv, p. 105.
14 T. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, 1962.
15 To use Rene Thom's admirable expression for what we observe – see *Structural Stability and Morphogenesis*, Benjamin, New York, 1975 – originally in French, 1972, as *Stabilite Structurelle et Morphogenese*. See for example p. 320.
16 F. De Saussure, (originally in French 1915) version used *Course in General Linguistics*, McGraw Hill, 1966, translated by C. Bally and A. Sechahaye with A. Riedlinger - see for example pp. 103–12.
17 These examples of course deal with linear variation, but the basic arguments also apply to non-linear variation.
18 This model, the 'Ehrenfest game', is taken from M. Kac and S. Ulam, *Mathematics and Logic*, Pelican Books, 1971, p. 168. Originally Praeger, 1968.
19 For a further discussion see H. Reichenbach, *The Direction of Time*, University of California Press, 1971, particularly Chapter 4.
20 See K. Popper K, *Conjectures and Refutations*, Routledge and Kegan Paul, 1963, Chapter 5: 'Back to the presocratics'.
21 See for example M. Ghyka M, *Geometrical Composition and Design*, Tiranti, London, 1956.
22 For example in the work of Alexander Koyre, e.g. *Metaphysics and Measurement*, Chapman and Hall, 1968 (originally in French) and *Newtonian Studies*, Chapman and Hall, 1965 or Georges Canguilhem e.g. *La Connaissance de la Vie*, Librairie Philosophique J.Vrin, Paris, 1971.
23 As pioneered in the work of Michel Foucault.
24 In the past, this has led to a quite rapid permeation by new scientific concepts of the conceptual schemes of everyday life, bringing changes in consciousness which may seem entirely progressive, as for example with the theories of Newton or Darwin. It may indeed be the loss of this illusory strength that has bought about much of the alienation from science in the late twentieth century. As science has progressed farther into micro and macro phenomena and discovered patterns which are utterly remote from everyday intuition the concepts that make up scientific theories become so strange that they cannot even be formulated so as to interface effectively with the established conceptual system of linguistic normality. This has happened with quantum theory. But what has happened with quantum theory confirms our model as set out in the diagram: science intervenes through formalisms between concepts and phenomena. It is no part of its function or its morality that these concepts should 'fall within the lighted circle of intuition' (to use Herman Weyl's admirable phrase – see his *Philosophy of Mathematics and Natural*

Science, Atheneum, New York, 1963, p. 66) and so be translatable into the available concepts of everyday life and language. There is no greater arrogance than that we should expect them to be, except perhaps the belief that the world itself in its deepest operations should conform itself to the apparatus of our intuitions.
Ernst Cassirer, *An Essay on Man*, Yale University Press, 1944.
25 Edition used: Bantam Matrix, 1970. Chapter 9, 'On art', pp. 152–88.

Chapter three

Non-discursive technique

'Environments are invisible. Their...ground rules...evade easy perception.' **Marshall McLuhan**

Object artefacts and abstract artefacts

One of the durable intellectual achievements of the twentieth century has been to initiate the scientific study of human artefacts. At first sight, such a study might seem paradoxical. Most artefacts are physical objects that adapt natural laws to human purposes. To make an object for a purpose surely presupposes that we understand it. But twenty-five years ago, Herbert Simon, in his *The Sciences of the Artificial*, showed that this was far from the whole story.[1] Even if the objects we make are not puzzling in themselves, they are so when seen in the context of the ramifying effects of their dispersion throughout our socio-technical ecosystem. He was thinking, amongst other things, of computers. It would be as enlightening, he argued, to have a natural history of computers in our increasingly artificial world, as of any natural phenomenon. Empirical sciences of artefacts were therefore not only a possibility, but a necessity.

But object artefacts are only the lesser aspect of the puzzle of the artificial. There also exists a class of artefacts which are no less dramatic in their impact on human life, but which are also puzzling in themselves precisely because they are not objects, but, on the contrary, seem to take a primarily abstract form. Language is the paradigm case. Language seems to exist in an objective sense, since it lies outside individuals and belongs to a community. But we cannot find language in any region of space-time. Language seems real, but it lacks location. It thus seems both real and abstract at the same time. Other artefacts which share some of the attributes of language, such as cultures, social institutions, and even, some would argue, society itself, all seem to raise this central puzzle of being, it seems, 'abstract artefacts'.

It cannot of course be said that 'abstract artefacts' are not manifested in space-time. They appear in the form of linguistic acts, social behaviours, cultural practices, and so on. But these space-time appearances are not the artefact itself, only its momentary and fragmentary realisations. We apprehend speech, as de Saussure would say, but not language.[2] In the same way, we see social behaviours, but we never see social institutions, and we see cultural events but we never see cultures. Yet in all these cases, the space-time events that we witness seem to be governed in their form by the abstract, unrealisable artefacts that we give a name to. The material world provides the milieu within which the abstract artefact is realised, but these realisations are dispersed and incomplete. The existence of languages, social institutions and cultures can be inferred from space-time events but not seen in them.

In spite of this strange mode of existence, abstract artefacts seem to be the stuff of which society is made. We cannot conceive what a society would be like if deprived of its languages, its characteristic social behaviours, its cultural forms and its institutions. It is not clear that anything would be left which we could reasonably call 'society'. We may conjecture, perhaps, that abstract artefacts are the way they are precisely because their purpose is to generate and govern dispersed events, and through this to convert a dispersed collectivity of speakers, behaviours or social actors into some semblance of a system. The multipositionality of the space-time realisation of abstract artefacts seems to be an essential part of how they work.

However, to say this is to restate the problem, not to solve it. In fact, in spite of their apparent oddity, abstract artefacts pose many of the puzzles which science seeks to explain for natural systems. For example, they seem able both to reproduce themselves over time, and also to undergo morphogenesis, though whether this is by a constant or sudden process is entirely obscure. If abstract artefacts have such properties, then it would seem to follow that they must therefore have some kind of internal principles or laws which give rise to stability and change, as do natural systems.[3] Yet whatever these laws are like, they must also pass through the human mind, since it is only through human mental activity that the self reproduction and morphogenesis of these systems occurs. It seems inconceivable, therefore, that the laws which govern the forms of abstract artefacts are similar to, or even commensurable with, the laws that govern natural systems. At the same time, such laws must be part of nature, since they cannot be otherwise. They must reflect some potentialities within nature.

In view of all these apparent paradoxes, it was the great merit of Lévi-Strauss and other pioneers of the study of abstract artefacts to have both identified the key insight necessary for their study, and to have pointed to a possible methodology for research.[4] The insight was to have seen the dependence of the concrete on the abstract in systems like language and culture, as clearly as Plato once noted it for the natural world.[5] Now, as then, this fundamental insight provides the starting point and initial stance for the setting up of sciences. The methodology was that, as with natural systems, we would expect to find clues to the nature of these organising laws by studying the regularities that abstract artefacts generate in space-time, that is, in speech, behaviour, cultural practices and institutional forms. Accordingly, the movement called structuralism aimed to assign abstract formal models with the structure and variety manifested in the space-time output of such systems - observed speech, social behaviour, organisational dynamics and so on - and through this to account not only for the internal systemness of such phenomena, but also to show how the human mind was capable of holding and creatively transforming such powerfully structured information. In this sense, structuralism was no more or less than orthodox science rewritten for the study of abstract artefacts.[6]

This research strategy reflects the fundamental fact that abstract artefacts manifest themselves to us in two ways: through the space-time events they generate; and through the configurational patterns which seem to support them and which enable us both to generate and interpret them. These two ways in which we experience abstract artefacts are bound together by the fact that in using configurational structures to generate space-time events we also project these configurational structures into space-time and in doing so help to transmit them into the future. This double take between the conscious manipulation of space-time events and the transmission of configurational structure is the defining characteristic of the abstract artefact and the reason it is able to be the stuff of society. By deploying objects and creating space-time events we necessarily transmit structures, and through them the abstract artefacts which hold society together as a communicative system. The object of structuralism is to capture the dynamics of these processes.

Formal methods were therefore critical to structuralism. However, as Heisenberg once remarked: 'Our scientific work in physics consists in asking questions about nature in the language that we possess and trying to get an answer from experiment by the means that are at our disposal.'[7] This is surely true of all scientific enquiry. Unfortunately, it seems to point directly to the failure of structuralism to deliver on its promises. Examining the space-time regularities of the phenomena generated by abstract artefacts, we cannot fail to note one overwhelming consistency; that they seem to be governed by pattern laws of some kind. The words that make up speech and the behaviours that seem social are all manifested in space-time as sequences or dispositions of apparent elements whose interdependencies seem to be multiplex, and irreducible to simple rules of combination. For example, to say, as Chomsky did,[8] that sentences, which appear to be sequences of words, cannot be generated by a left-right grammar, is a configurational proposition. Some degree of syncretic co-presence of many relations is involved whose nature cannot be reduced to an additive list of pairwise relations. This is to say that the laws governing abstract artefacts seem to be *configurational* in something like the sense we have defined it in the previous chapters.

It is in this respect that structuralism seems to have lacked methodology. Its formal techniques did not try to drive straight to the problem of configuration, but confined themselves to the more elementary aspects of logic and set theory, those branches of mathematics, that is, that sought to axiomatise the thinking processes of minds, rather than to model real world complexity.[9] Consequently, just as the 'languages' available for Plato in his time were inadequate for his vision of nature,[10] so the tools picked up in the mid-twentieth century by structuralism were too frail for the vision of artificial phenomena that had initiated their search. The phenomena that structuralist analysis sought to explain were in the main configurational, but the formal techniques through which investigators sought to demonstrate this rarely were.

Built environments as artefacts

The purposes of this digression into abstract artefacts are twofold: first, to draw attention to certain properties of built environments that might otherwise be missed; second, to point to certain advantages of the built environment in providing a platform for taking on the problem of configuration in a new way. First, however, we must understand the very peculiar status of built environments as artefacts.

Built environments appear to us as collections of object artefacts, that is, of buildings, and as such subject to ordinary physical laws, and deserving of Simonian enquiry. But that is not all that they are. As we noted in Chapter 1, in terms of spatial and formal organisation, built environments are also configurational entities, whose forms are not given by natural laws. If we wish to consider built environments as organised systems, then their primary nature is configurational, principally because it is through spatial configuration that the social purposes for which the built environment is created are expressed. The collections of object artefacts in space-time that we see, are then a means through which socially meaningful configurational entities are realised. In other words, in spite of appearances, built environments possess a key property of abstract artefacts. Its objects are more durable than, say, the spoken words of a language, or the rule-influenced individual behaviours that make up a social event, but they are of the same kind. They are space-time manifestations of configurational ideas which also have an abstract form. The built environment is only the most durable of the space-time manifestations of the human predilection for configuration. This has an epistemological consequence. We should not expect the built environment merely to be the material backdrop to individual and social behaviour, as it is often taken to be. It is a social behaviour, just as the use of language is a social behaviour and not just a means to social behaviour. We cannot therefore regard the built environment as merely an inert thing, and seek to understand it without understanding the 'social logic' of its generation.

But just as we cannot treat a built environment as a thing, we can no more treat it as though it were no more than a language. The built environment is, apart from society itself, the largest and most complex artefact that human beings make. Its complexity and its scale emerge together, because, like society, a built environment is not so much a thing as a process of spatio-temporal aggregation subject to continual change and carried out by innumerable agencies over a long period of time. Although these processes of aggregation may be locally characterised by the same kind of autonomic rule following as we find for individual acts of building, there are other no less fundamental attributes that make the built environment a special case.

The most obvious, and the most important, is that the spatio-temporal outputs of built environment processes are not ephemeral like those of language or social behaviour. They are long-lasting, and they aggregate by occupying a particular region of space for a long time. This means that over and above thinking of built environments as the products of abstract rule systems, we must also recognise that they have an aggregative dynamic which is to some extent independent of these rule systems, although, as we will see, it is rarely quite out of their control. These aggregative processes have quite distinctive properties. Spatio-temporal additions to a system usually occur locally, but the dynamics of the system tend to work at the more global aggregative levels.[11] Complexity arises in part from the recursive application, in increasingly complex aggregations, of rules which may initially be simple, but themselves may be transformed by the evolving context in which they are applied. A locally driven aggregative process often produces a global state which is not understood[12] but which needs to be understood in order for the locally driven process to be effective. This is the essential nature of the large aggregates of buildings which form most built environments.

Non-discursive technique

This complex, processual aetiology is the main reason why built environments have proved so resistant to orthodox attempts to model their structure mathematically. Buildings and cities are not crystalline objects, unfolding under the influence only of laws of growth. The elementary spatial gestuaries of humankind and its cultures may construct local elemental configurations, but these then operate as local orderings within growth processes and act as constraints on the ‘natural’ evolution of global patterns. Architectural and, even more so, urban forms occur at the interface between natural processes and human interventions. Human actions restrict and structure the natural growth processes, so that they cannot be understood without insight into both individually, and into the relations between the two. The intervention of the mind in the evolving complexity must be understood, but so must its limitations.

The built environment may then be the most obvious of objects, and the one that forms our familiar milieu, but at the same time its inner logic and structure is as inaccessible to us as anything in nature. However, it has one great advantage as an object of study. Its very scale, manifestness and slow rate of change offer it up as the paradigm case for configurational investigation. The essence of the problem is to capture the local-to-global dynamics of architectural and urban systems, that is, to show how the elementary generators, which express the human ability to cognise and structure an immediate spatial reality, unfold into the ramified complexities of large-scale systems.

In this, methodological difficulties are central. The aim of a method must be to capture the local or elemental ordering, the emergence of global complexity, and how both relate to the human mind. For any of these, the manifest problem of configuration must be tackled head on, and must be approached first and foremost as an empirical problem. If the space-time products of abstract artefacts are held together by configuration, then configuration can be found by examining them. The corpus of configurations that can be built through the study of real cases must be some indicator of where we might seek for the configurational invariants of built environment processes. For this task, the very scale, relative stability and availability of built environments make them the ideal vehicle for an enquiry. All we need are techniques that permit the extraction of configuration from its space-time embodiments - that is, non-discursive technique.

Simplicity as a means to complexity

The configurational formalisms proposed here as the basis for non-discursive technique are in some ways much simpler than others proposed for the similar classes of phenomena over the last twenty years.[13] Yet they have proved the most powerful in detecting formal and functional regularities in real systems. There are probably three reasons for this. First, the quantitative methods proposed are directed straight at the problem of *configuration*, that is, the problem of understanding the simultaneous effects of a whole complex of entities on each other through their pattern of relationships. Lack of attention to this central problem is the prime reason why past formalism often seemed to offer mathematical sophistication out of proportion to the empirical results achieved. With configurational analysis it is the other way round. Exceedingly simple quantitative techniques have led to a disproportionate success in finding significant formal and form-functional regularities. Configuration, as defined below, seems to be at least one of the things that architectural and urban patterns are about.

Second, in configurational analysis, as much theoretical attention has been given to the *representation* of the spatial or formal system that is to be analysed as to the method of quantification. As we will see, this quite normally gives rise to a whole family of representations of the same spatial system, each one relevant to some aspect of its functioning. It is also normal to combine representations, literally by laying one representation on top of the other and treating the resulting connections as real connections in the system. Through this, we find that pairs or even triples of representations taken together yield formally or functionally informative results. In terms of research strategy, this means trying to represent space in terms of the type of function in which we are interested. For example, simple line structures drawn through spaces, temporarily discounting other properties, have proved sufficient (as we will see in the next chapter) to account for many aspects of movement within buildings and urban areas.

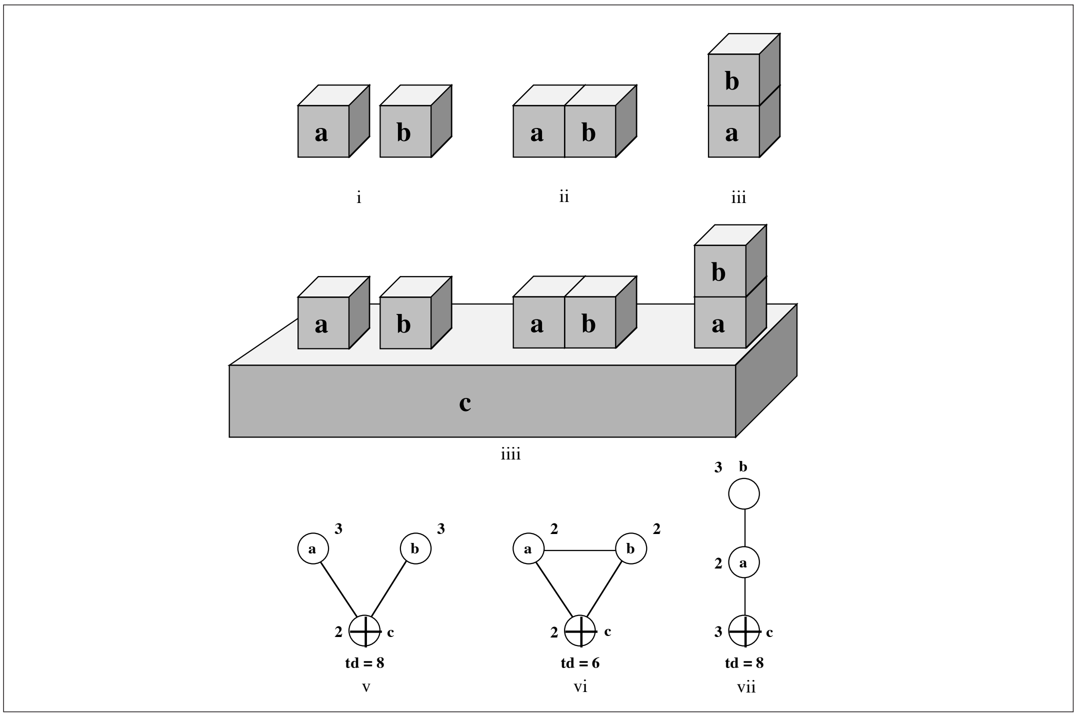
Third, and synthesising the previous two, much attention has been given to the graphic representation of the results of mathematical analysis, so that the formal structures identified in spatial or formal complexes can be intuitively seen and understood without the intermediary of mathematical formalism. This means that much can be understood by those whose temperaments lead them to prefer a graphical rather than a mathematical understanding. By representing mathematical results graphically, a level of communication is possible that permits large numbers of people to be interested and knowledgeable who would otherwise fall at the first fence of mathematical analysis. In parallel to this graphical representation of results, usually drawn by computer, there is a parallel emphasis in the initial stages of investigation to the drawing of spatial or formal ideas by investigators and by students as a constant adjunct to, and check on, formal analysis.

No apology is then offered for the simplicity of some of the notions presented here. Others have discussed some of these properties but have not been minded to explore their full empirical or theoretical relevance, or how they might be fitted into the overall form-function picture. Perhaps one reason for researchers to miss key relations while 'going close', has been what we would see as an overarching and in some ways premature concern with design at the expense of the empirical investigation of buildings. The 'space syntax' research at UCL has been driven by a remark of Lionel March's: 'The only thing you can apply is a good theory.'[14] Another possible reason why formal exploration has missed theoretical insight has been the frequent lack of a close enough relation between mathematical and empirical aspects of the problems posed by real buildings and cities. In contrast, the techniques of spatial representation and quantification proposed here are essentially survivors of an intensive programme of empirical investigation spread over the best part of two decades in which formal questions have been explored in parallel to the empirical puzzles posed by architectural and urban realities. We have already discussed the idea of configuration at some length in Chapter 1. Now we need to define it formally, and to show some of its power to say simple things about space and form. It should be noted that what follows is not a methodological cook book, but a theoretical exploration of the idea of configuration. At this stage, the examples given are illustrations of ideas, not worked examples of analysis. Case studies will come in ensuing chapters. The relation of this chapter to those that follow is that of a quarry, which future chapters return to to pick up one of the possibilities set out here, and refine it for the purposes of that chapter. This chapter shows the bases and connection of the whole family of methods.

Defining configuration

Let us begin by defining exactly what we mean by configuration, using an example directly analogous to figure 1.3 in Chapter 1, but taking a slightly different form. We may recall that in Chapter 1, a *simple* relation was defined as a relation - say, adjacency or permeability - between any pair of elements in a complex. A *configurational* relation was then defined as a relation insofar as it is affected

Figure 3.1

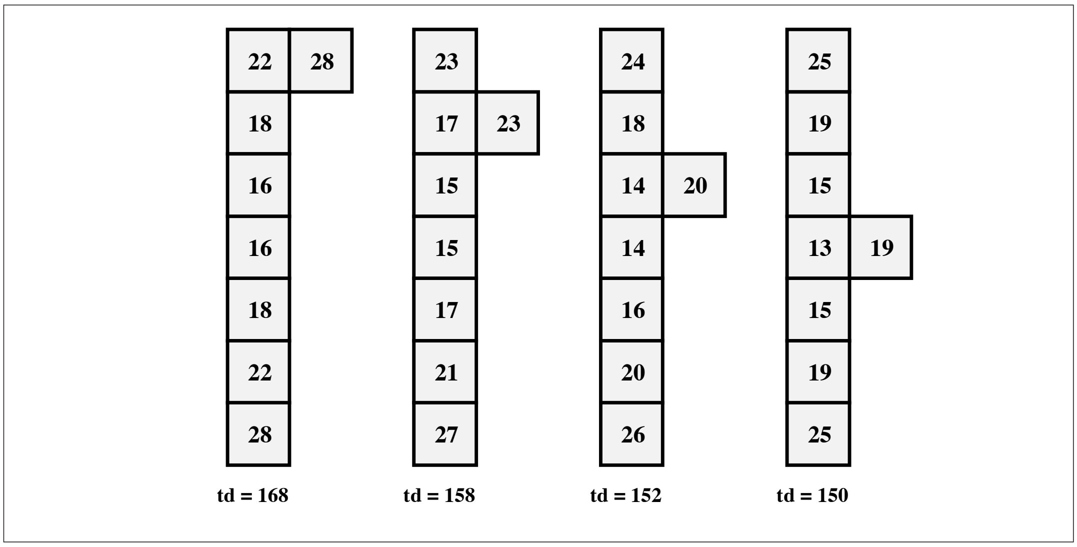


by the simultaneous co-presence of at least a third element, and possibly all other elements, in a complex. In figure 3.1 i, for example, *a* and *b* are two cubes standing on a surface. In 3.1 ii, the cubes are brought together full facewise to make a conjoint object. The relation of *a* and *b* is symmetrical in that *a* being the (contiguous) neighbour of *b* implies that *b* is the (contiguous) neighbour of *a*. One could equally say, though with less obviousness, that in 3.1 i *a* and *b* were non-contiguous neighbours, and were therefore symmetrical in this sense. Either way, the relation of the two remains symmetrical, and in fact this is implicit in the ‘neighbour’ relation. In 3.1 iii, the conjoint object formed by *a* and *b* in 3.1 ii is taken and rested on one of its ends, without changing the relation of *a* to *b*. But *b* now appears to be ‘above’ *a*, and the relation of ‘being above’, unlike that of ‘being the neighbour of’ is not symmetrical but asymmetrical: *b* being above *a* implies that *a* is not above *b*.

How has this happened? The temptation is to say that relations like ‘above’ and ‘below’ depend on an exogenous frame of reference, like ‘east’ and ‘west’, or ‘up’ and ‘down’. In fact, what has happened can be said more simply, as shown in 3.1 iiii. The surface on which the cubes stand - say, the surface of the earth - was not referred to in describing the relation between *a* and *b* in 3.1 i and ii. It should have been, had we wanted to foresee the effects of standing the conjoint object on its end. Let us call it *c*. In 3.1 ii, the relation of both *a* and *b*, taken separately, to the third object, *c*, is also symmetrical, as is their relation to each other. So, incidentally, is the relation of the conjoint object formed by *a* and *b* to the third object. These are all simple relations. But we can also say something more complex: that in 3.1 ii, *a* and *b* are symmetrical with respect to *c*, as well as with respect to each other. This is a configurational statement, since it describes a simple spatial relation in terms of at least a third. What happens in 3.1 iii is now clear. Although *a* and *b* remain symmetrical with respect to each other, they are no longer symmetrical with respect to *c*. On the contrary, they are asymmetrical with respect to *c*. The difference between 3.1 ii and iii is then a configurational difference. The relation of *a* and *b* to each other is changed if we add the ‘with respect to’ clause which embeds the two cubes in a larger complex which includes *c*.

The situation is clarified by the justified graphs (or j-graphs: graphs in which nodes are aligned above a root according to their ‘depth’ from the root – see

Figure 3.2



Chapter 1) of the configurations shown in 3.1 v, vi and vii. In each, the bottom node is the earth, and is inscribed with a cross to indicate that it is the root. In 3.1 v, *a* and *b* are each independently connected as neighbours to the earth. In 3.1 vi, the relation of neighbour between *a* and *b* is added. In 3.1 vii, the relation between *b* and *c*, the earth, is broken creating a 'two deep' relation between *b* and *c*. One may note that this set-up already exists in 3.1 v between the two non-contiguous cubes with respect to the earth. In this sense, 3.1 vii recreates a graph which already exists in v. This is also shown in the numbers attached to each of the nodes of the graph, which indicate the sum of 'depth' from that node to the other nodes in the system. The total depth of 3.1 v and vii is therefore 8, while that of vi is 6. We might say, then, that the distributions of total depths and their overall sum describe at least some configurational characteristics of these composite objects.

Now let us explore this simple technique a little further by examining figure 3.2, a series of simple figures composed of square cells joined together through their faces (but not their corners) with 'total depths' for each cell to all others inscribed in each cell, and the sums of these total depths for each figure below the figure. The figures are all composed of seven identically related cells, plus an eighth which is joined to the original block of seven initially at the top end in the leftmost figure, then progressively more centrally from left to right. There are two principal effects from changing the position of this single element. First, the total depth values and their distributions all change. Second, the sums of total depth for each figure change, reducing from left to right as the eighth element moves to a more central location. The effects, however, are quite complex. This is not of course surprising, but it illustrates two key principles of configurational analysis. First, changing one element in a configuration can change the configurational properties of many others, and perhaps all others in a complex. Second, the overall characteristics of a complex can be changed by changing a single element, that is, changes do not somehow cancel out their relations to different elements and leave the overall properties invariant. On the contrary, virtually any change to elements that is not simply a symmetrical change, will alter the overall properties of the configuration. We will see in due course that configurational changes of this kind, even small ones, play a vital role in the form and functioning of buildings and built environments.

Shapes as configurations

Another way of saying this, is that different arrangements of the same numbers of elements will have different configurational properties. For example, figure 3.3 is a set of rearrangements of the same eight square cells that we considered in figure 3.2, again with 'total depths' inscribed in each cell, but also with a number of other simple properties, including the total depth, set out close to the figure: *td* is total depth, *d* bar is the average for each cell, *sd* is the standard deviation, *df* is the 'difference factor' indicating the degree of difference between the minimum, maximum and mean depth in each complex (Hillier et al. 1987a), and *t/t* is the number of different depth values over the number of cells.

Non-discursive technique

In treating shapes as configurations in this sense, that is, as composites made up of standardised elements, we are in effect treating a shape as a graph, that is, as a purely relational complex of some kind in which we temporarily ignore other attributes of the elements and their relations. It is clear that such descriptions are very much less than a full description of the shape. For many shape properties, and for many of the purposes for which we might seek to understand shape, a configurational description of this kind would be quite inadequate or inappropriate. But there is one sense in which the configurational structure of the shape is a uniquely powerful property, and gives insights into properties of spatial and formal shapes which are increasingly manifesting themselves as the most fundamental, especially in studies of architectural and urban objects. This property is that graphs of shapes and spatial layouts are significantly different when seen from different points of view within the graph. This can be demonstrated visually by using the j-graph. By drawing j-graphs from all nodes in a shape, then, we can picture some quite deep properties of shapes.

For example, a highly interesting property of shapes is the number of different j-graphs they have, and how strong the differences are. For example, figure 3.4 shows all different j-graphs for a selection of the shapes in figure 3.3. The number varies from 3 to 6. The reason for this is that if we find that the j-graphs from two nodes are identical, then this means that from these two points of view, the shape has a structural identity, which we intuitively call symmetry. This is why in the shapes in figure 3.4, the smaller the number of different j-graphs as a proportion of the total number of j-graphs (that is, the number of elements in the graph) then the more the shapes appear regular because there are more symmetries in the shape. This is the ratio given as t/t (types over total) in figure 3.3. This aspect of the structure of the graph thus seems to reflect our sense that shapes can be regular or irregular to different degrees.

This analogy can be made more precise. In fact, the symmetry properties of shapes can be exactly translated as configurational properties. Mathematically, symmetry is defined in terms of invariance under transformation. In their book *Fearful Symmetry*, Ian Stewart and Martin Golubitsky illustrates this with singular clarity. ‘To a mathematician’ they argue, ‘an object possesses symmetry if it retains its form after some transformation.’[15] They illustrate this with a diagram showing the symmetries of the square, as in figure 3.5, in which ‘a typical point in the plane is mapped into eight different images by the...eight rigid motions that leave the square invariant’. Thinking of symmetries in terms of points in a shape is useful configurationally, since we may immediately ask what will be the characteristics of j-graphs drawn from each of the points. It is immediately clear that the j-graphs drawn from each of Stewart’s points will be identical, and that this would also be the case for any other comparable set of points which Stewart had selected. It is also clear that once a point has been selected there will only be seven other points in the shape from which j-graphs will be identical. The principle is in fact very simple: in a shape, every symmetry will create exactly one point from which the j-graph is

Figure 3.3

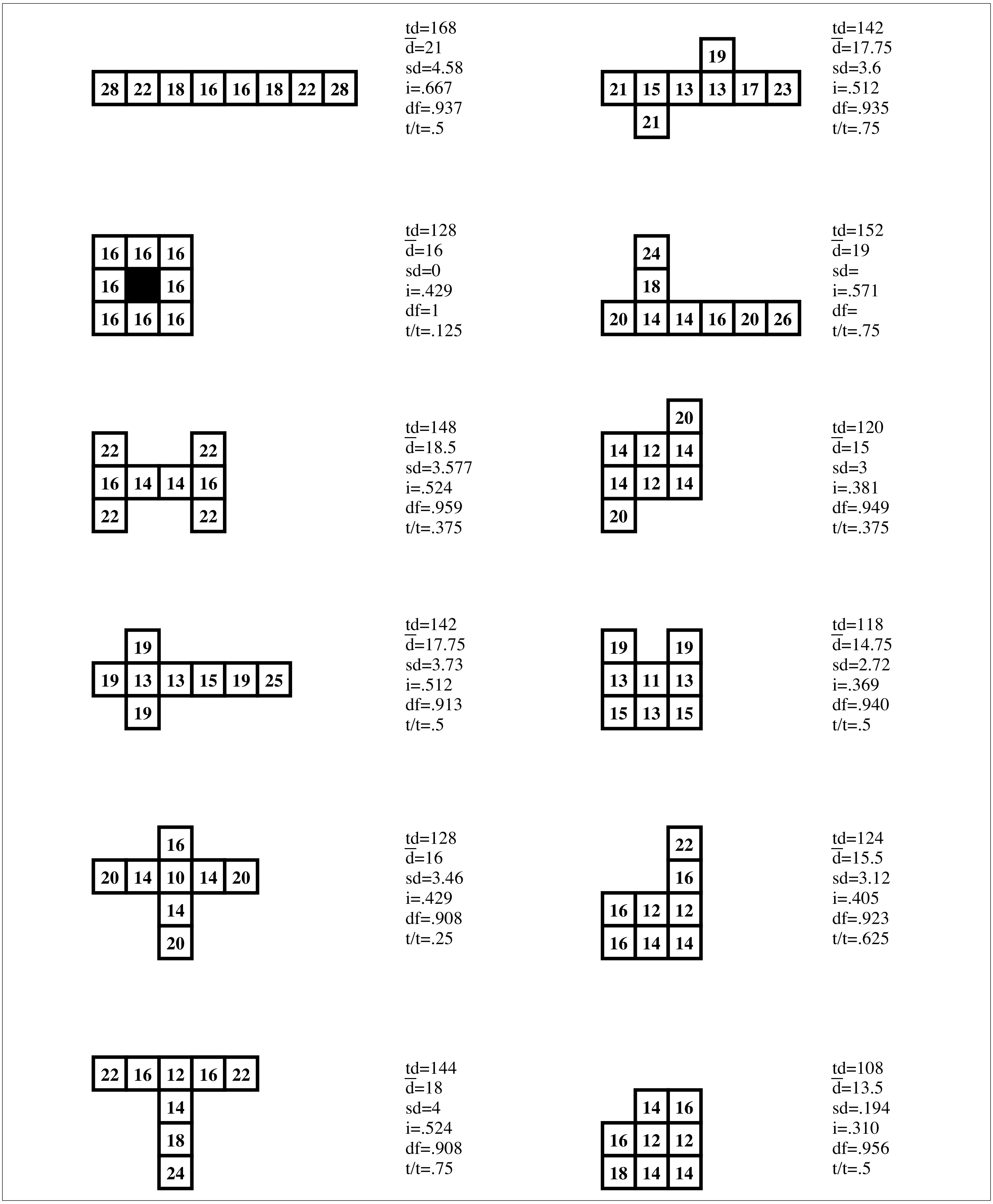

28 22 18 16 16 18 22 28
td=168
d=21
sd=4.58
i=.667
df=.937
t/t=.5
19
21 15 13 13 17 23
21
td=142
d=17.75
sd=3.6
i=.512
df=.935
t/t=.75
16 16 16
16 16
16 16 16
td=128
d=16
sd=0
i=.429
df=1
t/t=.125
24
18
20 14 14 16 20 26
td=152
d=19
sd=
i=.571
df=
t/t=.75
22 22
16 14 14 16
22 22
td=148
d=18.5
sd=3.577
i=.524
df=.959
t/t=.375
20
14 12 14
14 12 14
20
td=120
d=15
sd=3
i=.381
df=.949
t/t=.375
19
19 13 13 15 19 25
19
td=142
d=17.75
sd=3.73
i=.512
df=.913
t/t=.5
19 19
13 11 13
15 13 15
td=118
d=14.75
sd=2.72
i=.369
df=.940
t/t=.5
16
20 14 10 14 20
14
20
td=128
d=16
sd=3.46
i=.429
df=.908
t/t=.25
22
16
16 12 12
16 14 14
td=124
d=15.5
sd=3.12
i=.405
df=.923
t/t=.625
22 16 12 16 22
14
18
24
td=144
d=18
sd=4
i=.524
df=.908
t/t=.75
14 16
16 12 12
18 14 14
td=108
d=13.5
sd=.194
i=.310
df=.956
t/t=.5

Figure 3.4

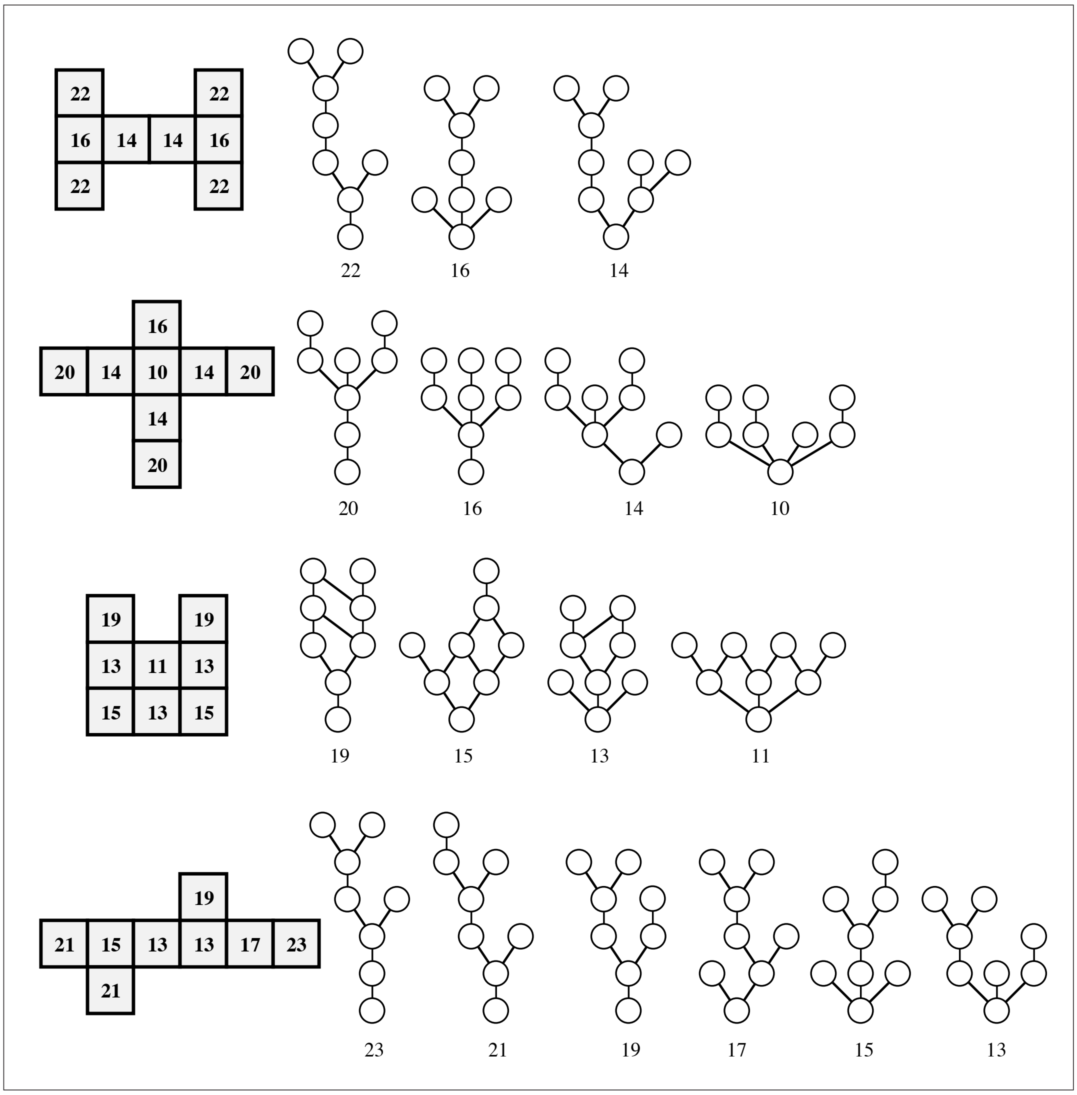
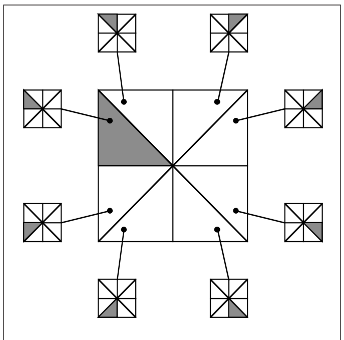


Figure 3.5



isomorphic. In effect, j-graph isomorphism is a test for symmetry. The j-graph allows us to look at symmetry as an internal property, in contrast to the more external view presupposed by the 'invariance under motion' definition. In a sense, the invariance under motion exists because there are different points within the shape from which the shape is identical. We might say that in a shape with symmetry there are points within the shape with identity of positional information in relation to the object as a whole, and this is demonstrated by j-graph isomorphism.

Universal distances

The distributions of depths that are shown through the j-graphs, and which underlie both architectural and geometrical effects - are in fact the most fundamental idea in quantifying the configuration properties of spatial or formal complexes. The idea first made its appearance in the literature of applied graph theory in 1959 when Harary applied it to sociometry under the name of 'status'. 'Status' is defined by Buckley and Harary[16] thus: 'The status s(v) of a node v in

G (a graph) is the sum of distances from v to each other node in G', distance meaning the fewest number of nodes intervening between one node and another. The problem with status defined in this way as 'total depth' is that the value will be very substantially affected by the number of nodes in the graph. Accordingly, as discussed in Chapter 1, a normalisation formula was proposed in *The Social Logic of Space*[17] which eliminates the bias due to the number of nodes in the graph. With this normalisation, numerical values can be assigned expressing 'total depth' independently of the size of the system. This normalisation formula was discussed and clarified by Steadman in *Architectural Morphology*[18] We will call these normalised values i-values, to express the idea of the degree of 'integration' of an element in a complex, which we believe these values express.

The need for the normalisation formula and the intuition of the form it might take in fact came from using the justified representation of the graph, or j-graph. Simply as a consistently used representation, the j-graph makes the structure of graphs, and more importantly the differences in their structures, extraordinarily clear. However, by representing them in a standard format, it also makes clear the need for comparative numerical analysis and how it might be done. For example, it is immediately clear what graph will be maximally and what minimally deep. It is a simple matter from there to find the normalisation. The fact that no one found this useful expression before, when it opens up whole new vistas for the empirical analysis and comparison of forms, is presumably because no one saw either its necessity or possibility.

However, although the i-value formula allows the theoretical elimination of the effects of the size of the system, it does not deal with the fact that, empirically, architectural and urban spatial complexes use only a small proportion of those theoretically possible, and this proportion shrinks as the size of the system grows. These effects are discussed in full in Chapter 9, and in fact become the basis of a full theory of urban spatial form. A second, empirical normalisation formula was therefore introduced to cope with this empirical fact.[19] The second formula is an empirical approximation with some theoretical justification (that it approximates a normal distribution of depth values from any node in a graph) and as such it lacks elegance. However its robustness has been demonstrated in large numbers of empirical studies over the years, during which time no need has arisen to call it into question.[20] No doubt, as studies advance, it will be possible to eliminate this second normalisation formula and replace it with an expression with more theoretical elegance. In the meantime, 'integration' will refer to the outcomes of both normalisations, unless 'total' depth' (status, with no normalisation) or 'i-value' (status with the first, theoretical normalisation for size) are specified. All these terms are different ways of referring to the same quantity.

Why has this quantity proved so fundamental in the empirical study of spatial and formal configurations? It is possible that its simplicity conceals a very fundamental theoretical property: that it is essentially a generalisation of the idea of distance. Our common concept of distance is that of a specific number of metric

Figure 3.6

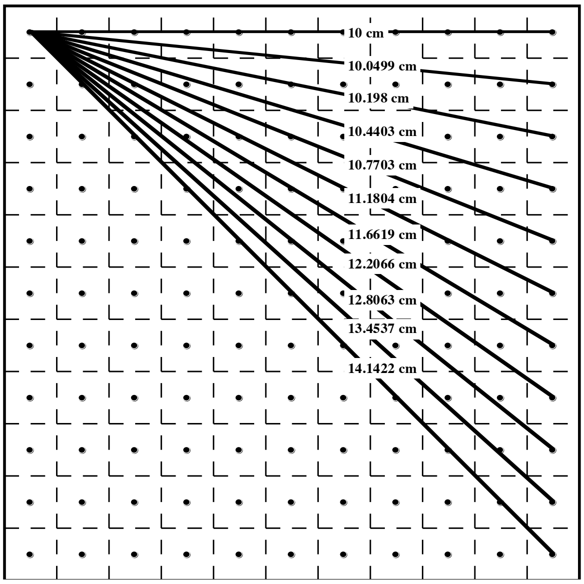
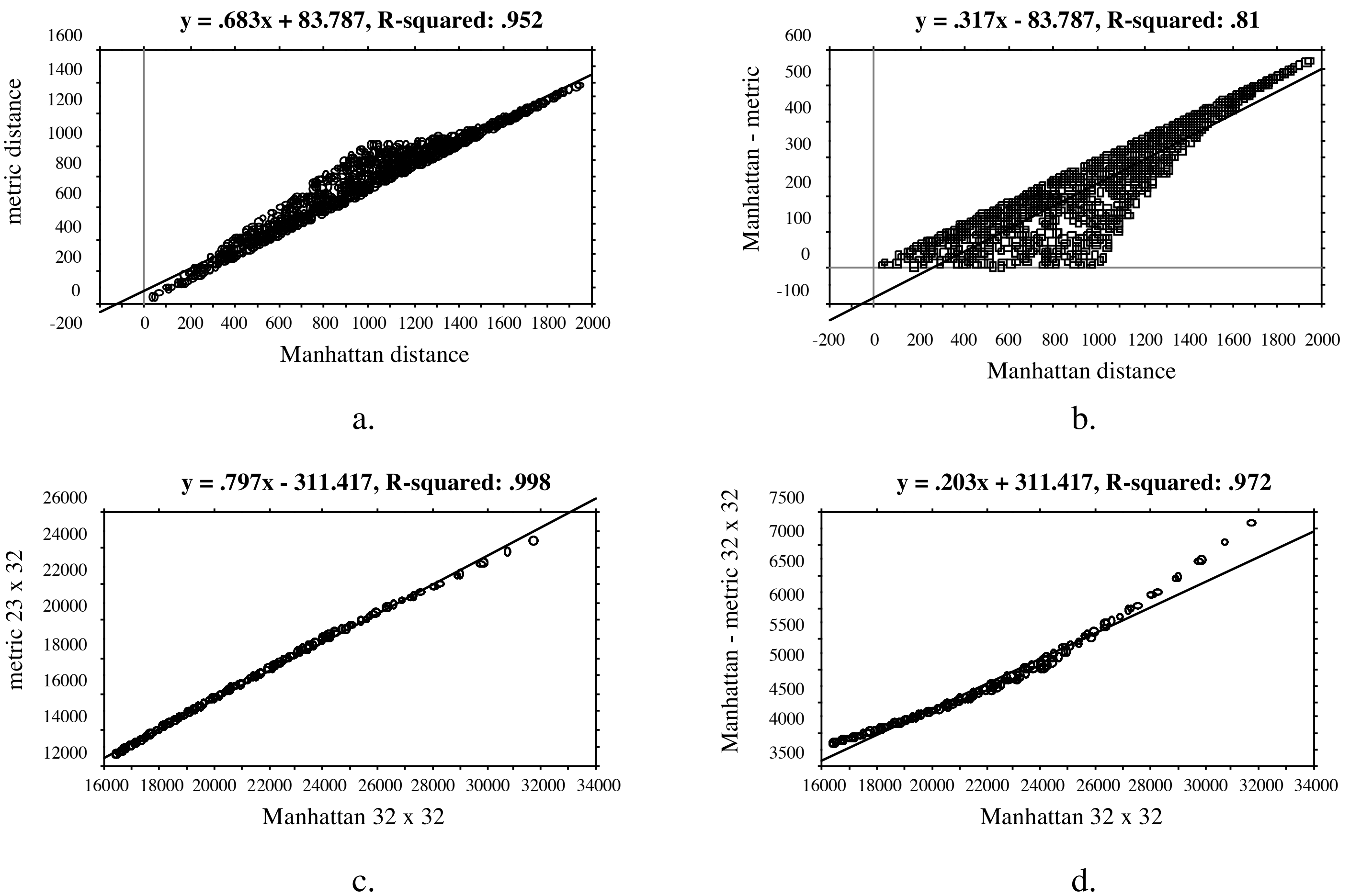


Figure 3.7



units between one point and another within some system of spatial reference. We can call this a specific distance. Total depth sums all specific distances from a node to all others. We may therefore think of it as a 'universal distance' from that node. If specific distance is about the metric properties of shapes and complexes, universal distances seem to be the key to configurational properties. Universal distance seems to be a generalisation of the idea of depth that permits configuration to become the central focus of analysis.

It may be objected that such a concept of universal distance has only been made possible through an unacceptable simplification of the idea of a shape to that of a graph, rather than an infinite set of points. This is a difficulty, but it seems that it might not be as great as it might at first appear. If we consider a square shape made up of square cells, and therefore representable as a graph, as in figure 3.6, and measure distances from and to the centroid of each cell, it is clear that graph distances will approximate metric distances only when they are orthogonally related. On the diagonal, metric distances will be either shorter or longer than graph distances, depending on whether or not we connect the graph diagonally across cell corners, or only allow joins through the faces. If corner links are not allowed, then graph distances will be $n + m$ (or 'Manhattan' distances, by analogy with the Manhattan grid) where m is the horizontal distance and n the vertical distance, while the metric (or 'as the crow flies') distance will be the square root of m squared + n squared. This will be maximal between opposite top and bottom corners. If diagonal links to adjacent nodes are allowed, then the distance between opposite top and bottom corners will be m or n, whichever is the greater, which equally misrepresents the metric distance. If we plot graph distance against metric specific distances in such a system we will find that not only are the differences substantial, but also that they vary in different parts of the system. In other words, graph and metric specific distances are not linearly related, so we cannot use one as a proxy for the other. Figure 3.7a is a plot of metric specific distance against graph (Manhattan) specific distance for 1000 randomly selected pair of points in a 100×100 square cell arrangement of the type shown in the previous figure, and figure 3.7b plots the difference between metric and graph specific distance on the vertical axis for increasing graph distance on the horizontal axis.

However, if we substitute universal for specific distances, and carry out the same analysis, this problem is significantly diminished. Figure 3.7c shows graph (Manhattan) against metric universal distances for all nodes in a 32×32 (i.e. 1024 cells) square cell complex, and figure 3.7d plots graph distance against the difference between metric and graph distances. Although the values are still exactly as different overall, they are now more or less linearly related, so that it is much more reasonable to use one as a proxy for the other. This fortunate fact permits a far more flexible use of graph based measure of configuration than would otherwise be the case. As we will see, such matters as shape and scale, area and distance can all be brought, as approximations at least, within the scope of the configurational method. All will be in some sense the outcome of seeing a

Non-discursive technique

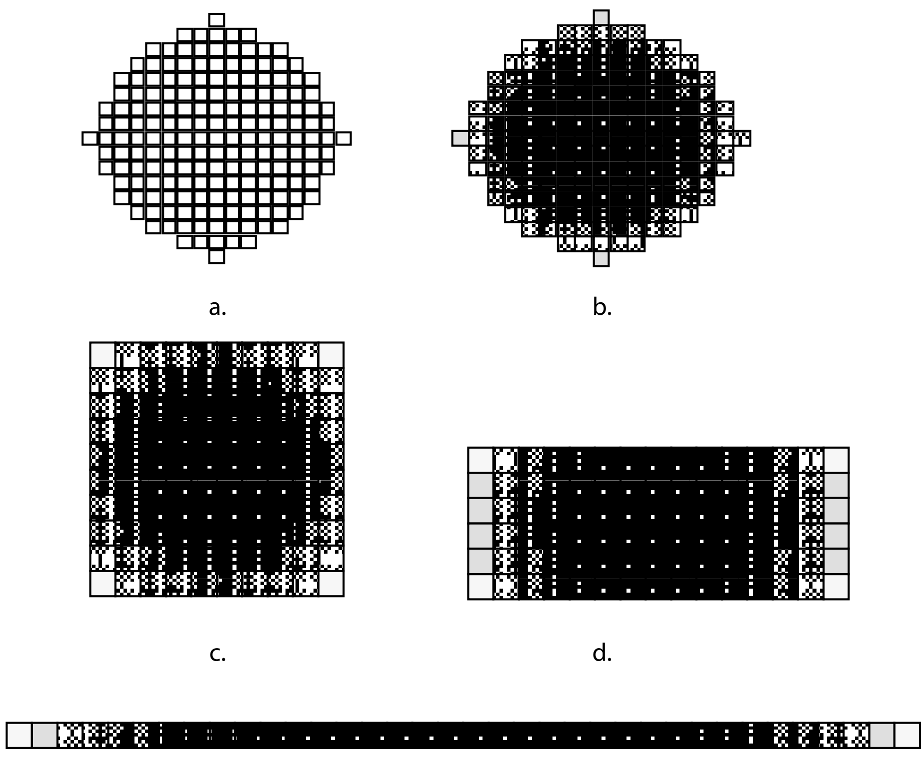
complex of related elements as a set of j-graphs. The j-graph in effect redefines the element of a complex in terms of its relation to all other elements in the complex. Summing the properties of j-graphs to express properties of the whole complex means summing the different points of view from which the complex can be seen internally. The eventual justification of this formalism is that architectural and urban systems are exactly this kind of complex. They are global systems whose structure, functioning and growth dynamics are manufactured out of the innumerable different points of view from which they can be seen.

Regular shapes as configurations

Now let us take the idea a little further, and closer to everyday experience. It is clear that any shape can be represented as a regularly constructed mesh of cellular elements, or tessellation, provided we can scale the mesh as finely as we need. This can then be treated as a graph, and thus expressed as a pattern of universal graph distances. By describing simple everyday shapes in this way, it turns out that we can capture important aspects of how they fit into everyday living patterns.

Suppose, for example, we create an (approximately) circular tessellation of arbitrarily small square cells, as in figure 3.8a. We may calculate the mean depth of each cell from all others, and express the results in a distribution of dot densities for the square elements in which the higher densities, or darker colours, stand for greater integration - that is, less depth – graded through to lightest colours for the least

Figure 3.8



integration, or greatest depth. It is clear that the centre has the highest integration, and that integration reduces evenly in concentric rings around the centre. In a perfect circle, all edge locations will have an identical degree of integration.

If we then consider the square tessellation in figure 3.8c we find that the pattern of integration not only runs from centre to edge, but also from the centre of the edge to the corners. The square form is thus more complex than the circular form in a simple, but critical way. We may say that in the square form, the 'central integration' effect occurs twice: once in the global structure from centre to edge, and once more locally on each side of the form. We can also easily calculate that the square form is less integrated - that is, has greater average universal distances per tessellation element - than the circular form.

As we elongate the square into a rectangle, as in figure 3.8d, the overall form is even less integrated, and the properties first found in the square become more exaggerated. The global structure of the form is now a group of integrated central squares, which includes some on or near the periphery of the object, with the two 'ends' substantially less integrated than other parts. Each side has a central distribution of integration, but one in which the long sides have much greater differentiation than the short sides, and correspond increasingly to the global structure of the tessellation as we elongate it. In the limiting rectangular tessellation, the single sequence of squares, then the local and the global structures are all identical, as in figure 3.8e.

We may summarise this by saying that while all these forms are globally structured from centre to edge, in the circular form the local or lateral structure is uniform, in the square form the lateral structure is maximally different from the global structure, while in the rectangular form the local lateral structure tends to become the global structure as we elongate it, until the limiting form of the single sequence is reached when the two structures become identical. The correspondence between these 'structures' of shapes and the ways in which shape is exploited for social purposes in everyday life is intriguing. For example, on square dining tables the centre side is more advantageous than corner locations, because it is a more integrated location. Similarly, the English prime minister sits in the centre of the long side of a broad rectangular table, maximising this advantage in integration. In contrast, where status rather than interaction is the issue, caricature dukes and duchesses sit at opposite ends of a long table, maximising proxemic segregation but also surveillance, while students and monks classically sit on the sides of a long thin 'refectory' table with no one at the ends, thus making all but localised conversations difficult. The politics of landholding knights with a peripatetic king sitting at a round table are equally manifest, as are the endless political debates over the shapes of conference tables and parliament chambers. The ways in which shapes are exploited and used all follow the pattern of integration in some way, though with opposite tendencies depending on whether interactive status or symbolic status is more critical.

Non-discursive technique

Plans as shaped space

Now let us consider the more complex case of the house plan. In the sequence of plans in figure 3.9i is a slightly simplified version of the plan of one of the farm houses in rural France that were considered in Chapter 1. The salle commune is the everyday space where cooking, eating and the reception of everyday visitors take place. The grande salle is a space for more formal reception of guests. The workspaces to the right are a dairy, washing room and storage, all associated with the female role in the house, the bureau is the office of the principal male occupant, and the salle is an indeterminate space, perhaps functionally associated with the bureau. What does it mean to analyse this plan as a shape?

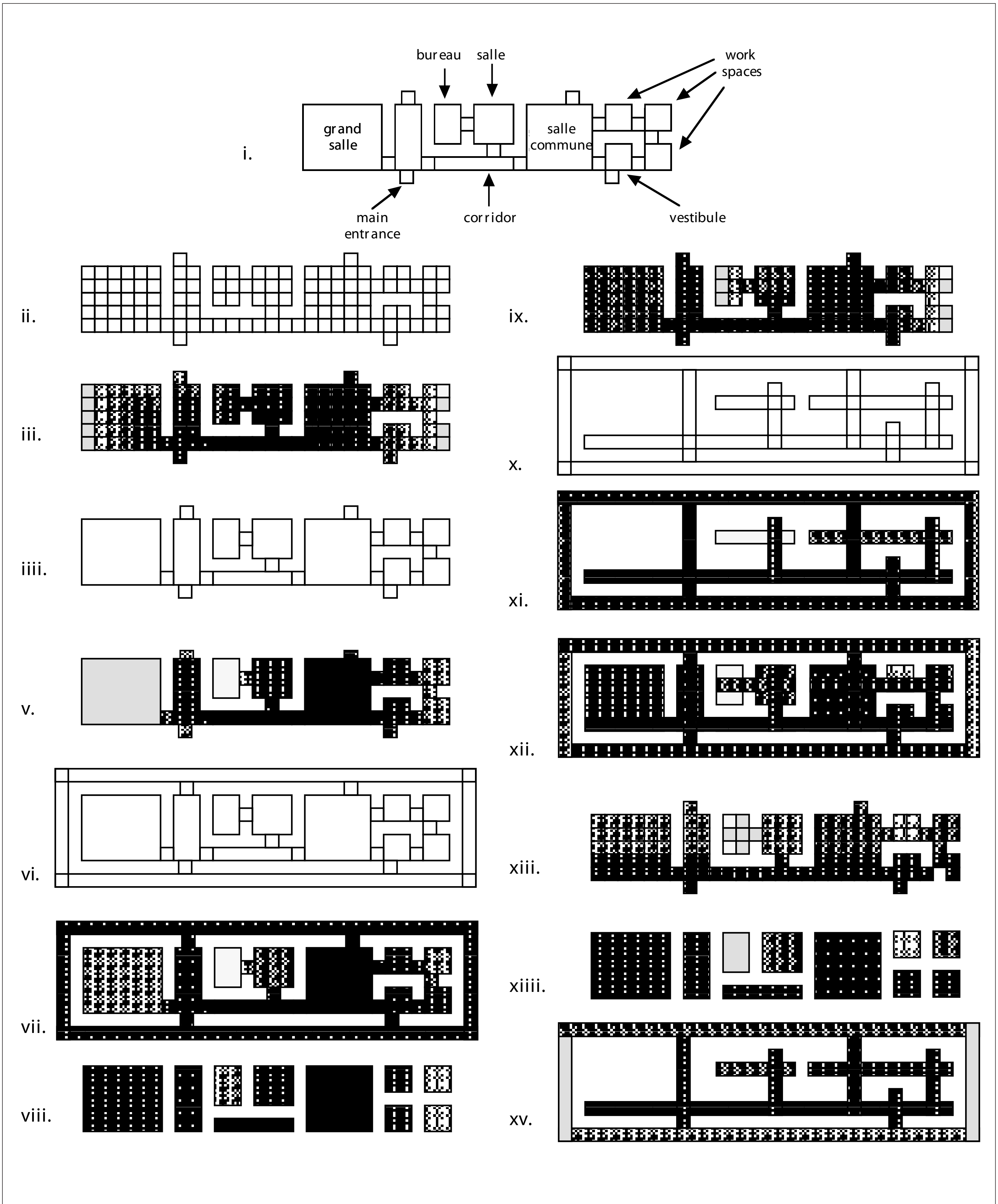
A plan is, first, a shape, which can be represented as a tessellation, see 3.9ii. For convenience and speed of analysis we use a rather large element, and treat thresholds as single elements. This leads to some unrealism in wall thickness, but this does not affect the analysis. The tessellation may be analysed into a pattern of universal distances. Since this reflects the distribution of centrality in the shape, in this elongated plan the least universal distances - shown darkest - are found in the front corridor between the large space mid-right - the salle commune - and the main entrance mid-left as in 3.9iii.

The metric distribution of universal distances represents the degree to which physical effort must be made to move from one part of the shape to another. If we compare the plan shape to a square shape with the same number of elements we have a simple index of the overall metric integration of the shape. In this case, the mean universal distance of cells in the shape is 10.3 whereas for an equivalent square it would be 4.9. Dividing the former into the latter, we find that our shape has 2.1 times the universal distance of an equivalently sized square, indicating that about twice as much effort must be made to move around this plan as in an equivalent square. We may think of the reciprocal of this number as indexing the degree to which a shape gets towards being a square. In this case the value is .462. The degree and distribution of universal distances thus indexes something like the physical economy of the shape, the human counterpart to which is the amount of physiological effort needed to overcome universal distances. We may perhaps think of this way of looking at the plan as its bodily or physiological structure. It represents the inertia a particular shape offers to the human body occupying it.

However, as we saw in Chapter 1, the plan is also an arrangement of convex elements, that is, rooms, corridors, halls, and so on. We can represent it as such, again, by using single element thresholds, as in 3.9iiii. Again we analyse this for its pattern of integration, this time treating the convex elements as elements, and therefore ignoring actual distances and sizes, giving 3.9v. Now of course, as was shown in Chapter 1, the strongest integrator is the salle commune. Though the colour coding makes it look the same as the corridor, the integration value of the space (.197, using the i-value formula) is a little stronger (that is, has lower universal distance) than the corridor (.205). This means that in terms of convex as opposed to metric organisation, the focus of integration has been displaced from the geometric

Non-discursive technique

Figure 3.9



centre into one of the function spaces. The distribution survives if we add four linear strips around the plan to represent the outside world (since the relation to the outside is often a critical aspect of domestic space organisation), and reanalyse for integration (3.9vi and vii). The offset salle commune space is still stronger than the central corridor element.

We now overlay the convex elements on the tessellation shape, connecting each to all the tessellation squares that lie immediately under it, and re-analyse the two layers as a single system, so that each convex element is affected by the number of tessellation elements it is directly connected to, and each tessellation element is affected by the links made to other tessellation elements through the pattern of convex elements. Not surprisingly, we find that each layer has affected the distribution of universal distances in the other. Figures 3.9viii and ix show each layer of the two-layer system separately. 3.9viii, the convex layer of the two-level analysis, shows that compared to 3.9v, the large space on the left, the 'best' room, has become relatively more 'integrated' than the work spaces on the right and the office. This is an effect of scale. The fact that the much larger convex area of the ' best' room overlays far more tessellation squares than the small work rooms has the effect of drawing integration towards the 'best' room in direct proportion to its metric scale, and conversely for the small rooms. In effect, the convex layer of the two-level system shows how the pattern of integration of the convex elements is affected by their area, as measured by the number of uniform tessellation elements each overlays. This effect is clarified in figure 3.9ix the tessellation layer of the two-layer system. Comparing this to figure 3.9iii, we see that the overlaying of the larger convex element on the tessellation squares within the 'best' room has the effect of making them more integrated and more uniform. These results show that metric scale, shape, and spatial configuration can all be expressed in the common language of universal distances, or integration, in layered spatial representation considered as unified systems.

We may take this a little further. Another potential 'layer' in the plan is the system of lines of sight linking the convex elements together through the doorways, assuming for this purpose that they are open. We can represent this layer by drawing axial 'strips' corresponding to lines of sight as in figure 3.9x and analyse its pattern of integration, figure 3.9xi. We find that the front 'axis' passing through the salle commune, the salle and the corridor is now the most integrating element but the main entrance front-back line mid-left and the salle commune front-back line mid-right are almost as strong.

We may then superimpose the linear elements on the convex elements and reanalyse these as a single two-level system in which the line elements are all directly connected to the convex elements that lie immediately under them. The effect of this simultaneous analysis of the two layers will be to show how integration is shared between convex and linear elements. We find that the front corridor is still strongest, followed by the front-back line through the salle commune, followed closely by both the front back line through the main entrance and by the convex space of the salle commune itself. These results can be shown by keeping the line and convex system together, as in figure 3.9xii, but also by showing them separately for greater clarity.

Finally we can assemble all three layers into a single system in which both convex and line elements are directly connected to all the tessellation elements that lie immediately under them. We then analyse and print out the three layers separately, first the tessellation layer, figure 3.9xiii, then the convex layer, figure 3.9xiiii, and finally the line layer, figure 3.9xv. The final pattern emerging from the three-layer analysis is that the 'front axis' linking through all the front space is the strongest integrator, followed by the salle commune, the grande salle, the line to the back through the salle commune and the main entrance line and the secondary entrance line.

Compared to the purely convex analysis outlined in Chapter 1, then, a number of new subtleties have been added. For example, it has become clear that the potential line of sight linking rooms through the corridor at the front of the house is a more critical element than appeared in the earlier analysis, and in effect imparts to the house a front-back organisation that had not emerged from the earlier analysis. Also, we can see that the relation between what we might call the 'energy economy' of the house plan, that is, the amount of effort needed to go from one location to another as shown in the metric tessellation, and the higher-level organisation is quite subtle. In effect, convex space integration for the major spaces is displaced from the metric centre of gravity, and the degree of displacement is to some extent compensated by size. Thus the grande salle is more displaced than the salle commune, but compensates for this greater displacement by its greater size. Multi-layered analysis suggests then that we should not see a system of space as one thing. A spatial layout is a shape which contains many configurational potentials, each of which seems to relate to a different aspect of function. These potentials may be treated as independent systems of space by choosing to analyse the layout on the basis of one particular representation rather than another, or they may be treated in selective combinations, or even altogether. It all depends on what we are trying to find out.

Façades as configurations

If the distribution of the various layers of integration in a shape relates to the ways in which we use shapes, then an intriguing possibility might be that it could also be implicated in how we understand shapes. For example, building façades seen as shapes seem capable of being 'understood' as communicators of information in some sense. Could configuration be involved in this type of apparent communication?

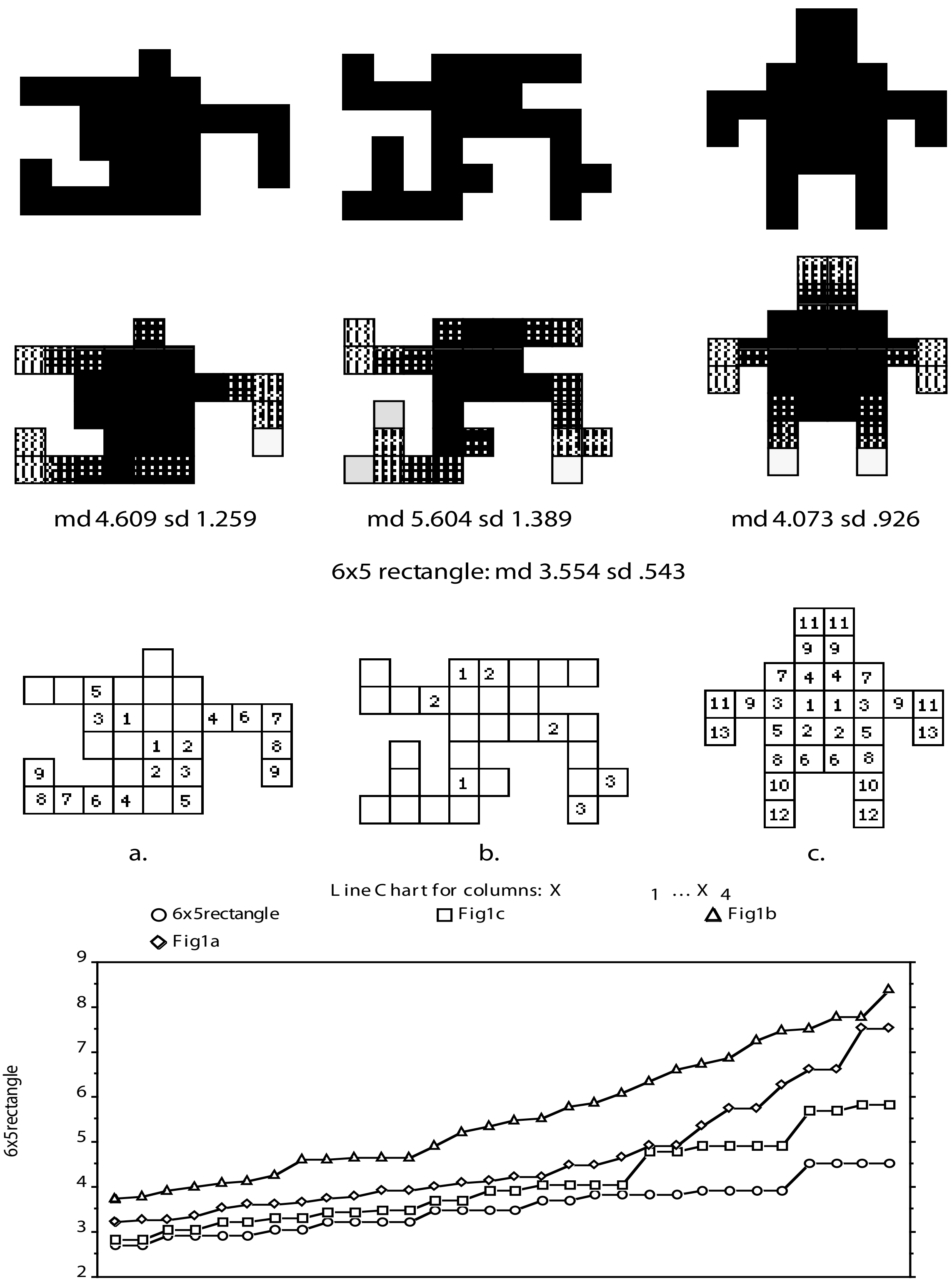
Consider in a very elementary way how we recognise objects. The top row of figure 3.10 shows three figures which are constructed by arranging thirty square elements in different ways. Recognising these figures seems to happen in two stages. In the first stage, we identify a distinct shape, different from others. In the second we assign that shape to a category by giving it a name. In figure 3.10a and b, we see two shapes. We easily recognise the difference between the two shapes, that is, we readily make a pure configurational distinction between the two objects. But we have no category to which we can assign either object. The process of object recognition is therefore ended at the first stage. In figure 3.10c we also see a shape, but this time we conjecture a category: the shape looks like an over-regularised humanoid, so we conjecture it is meant to be either a robot, a caricature human, or perhaps a toy. Of course, the figure does not really bear much resemblance to a human being or humanoid. The evidence on which our category conjecture is based is, to say the least, flimsy. However the nature of the evidence is interesting. It seems to be configurational. Figures 3.10a, b and c are no more than outlines produced by rearranging 30 square cells into different configurations. We have, it seems, a clear ability to distinguish pure shapes or configuration from each other, prior to any intuition of the category of thing to which the configuration might belong.

We can call the first the syntactic stage of object recognition, and the second the semantic stage. The second stage has been extensively dealt with by philosophers and others, but what about the first, 'syntactic' stage, only now being investigated by cognitive psychologists?[21] What does it mean to recognise a configuration? One approach to this is to reverse the question and ask what properties configurations have that might allow them to be recognised. Suppose, for example, we analyse the configurations as distributions of total depth values as in the second row of figure 3.10.

This gives us several kinds of useful information about the configuration. First, there is the distribution of integration in each form, as shown by the dark-to-light pattern. This can be thought of as a structure within the shape. Second, there are the integration characteristics of the form as a whole, as indexed by the mean depth (md) values and their standard deviation (sd) as shown beneath each form. For comparison, the mean depth and standard deviation for a six by five rectangle (that is, a regular form with the same number of elements and approximating a square as closely as possible) is also noted. We see that 3.10c is more integrated than 3.10a, which is more integrated than 3.10b, and that all are less integrated than the six by five rectangle. Standard deviations follow a similar pattern. These depth values seem to correspond to certain intuitions we have about the forms, as do the standard deviations, which shows that 3.10b has greater variation in the mean depths of individual elements than 3.10a, which has more than 3.10c, and all have more than the six by five rectangle.

However, there is another intuition which is not expressed in these measures. It is obvious that 3.10c is more 'symmetric' than either 3.10a or 3.10b, since it has the property of bilateral symmetry, one of the commonest and most easily recognisable types of symmetry found in artefacts or in nature. However, while figures 3.10a and 3.10b both lack formal symmetries, they do not seem to be entirely equivalent from this point of view. In some sense, figure 3.10a seems to be closer to symmetric organisation than 3.10b. There is a possible quantification for this property. To explain

Figure 3.10


md 4.609 sd 1.259
md 5.604 sd 1.389
md 4.073 sd .926
6x5 rectangle: md 3.554 sd .543
a.
b.
c.
Line Chart for columns: X 1 ... X 4
6x5rectangle
Fig1c
Fig1b
Fig1a
6x5rectangle

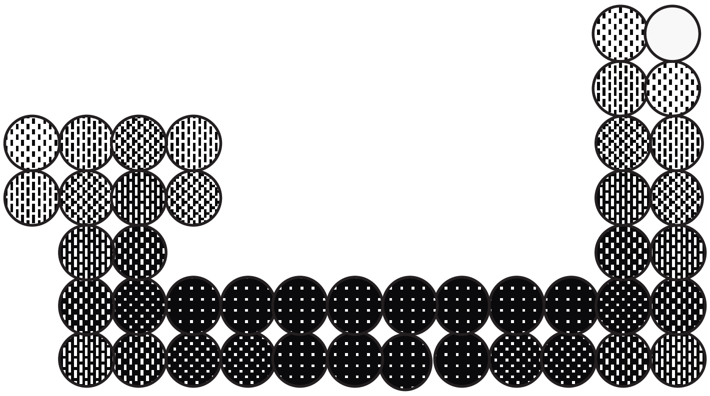
it, we must consider the whole idea of symmetry from a configurational point of view. We have already seen that pure symmetries in shapes could be interpreted as configurational properties, namely j-graph isomorphisms. From an architectural point of view, it is very useful to formulate properties of symmetry in this way, since, unlike the normal 'invariance under motion' definitions of symmetry, it opens the way to weaker definitions of symmetry, and permits an account of intuitively important architectural properties which approach symmetry but cannot be so formally defined. For example, we can specify identity of positional information with respect not to the whole object but to a region within the object, that is, local rather than global j-graph isomorphism, and discuss the relation between local and global j-graph isomorphism. Buildings are full of local symmetries – the form of a window, or of a particular mass within a complex – which sometimes are and sometimes are not reflected in a global symmetry. The relation between local and global symmetry seems a natural way to express this.

Most significantly, we can specify similarity, rather than identity, of positional information, and do so in a precise way. For example, j-graph isomorphism means that j-graphs share not only the same number of elements and the same total depth, but also the same number of elements at each level of the j-graph and the same connections between elements. One way of weakening this property would be to maintain all properties except the requirement that the connections be identical. Another would be to vary the number at each level (from which it follows that connections would be different) but to maintain the total depth the same.[22]

The second of these seems particularly interesting, since it offers a possible formalisation of the property of 'balanced' asymmetry often discussed in the literature in the formal properties of architecture.[23] For example, in figure 3.11 we load a simple linear shape with two sets of four by two cells, one horizontal, the other vertical, but each joined to exactly two cells in the basic form. Although the two end shapes created are different, and in themselves have different distributions of total depth values (or i-values), all the values in the bottom two rows are paired in that each cell has exactly one other cell which is 'symmetrically' located and has the same i-value. This i-value equality seems to give a rather precise meaning to the idea of 'balanced asymmetry'.

We may apply this analysis to the three shapes shown in figure 3.10. The third row shows each shape with cells with equal i-values marked with the same number, from the most to the least integrating. We see that 3.10a has far more equal i-values than 3.10b. Also, in 3.10a the equal values reach well into the integration core of the shape, whereas in 3.10b they are distinctly peripheral. Both of these properties, as well as the degree of integration, can be represented through a simple statistical device: the line chart shown in the final row of 3.10. Here each shape is represented by a series of i-values, plotted from most to least integrated (shown as least to most depth), together with a series representing the six by five rectangle (shown as circles) to provide a baseline for comparison: 3.10a

Figure 3.11



is represented as diamonds, 3.10b as triangles, and 3.10c as squares. Evidently, the overall degree of integration is indexed by the location of the series on the vertical axis. Thus the rectangle is the most integrated, 3.10c next, then 3.10a and finally 3.10b. Also, the shapes diverge as they move from integrated to segregated elements, so that the most integrated elements in each shape are much closer together than the least. The line charts also show the degree of ‘balanced asymmetry’ in the shape by aligning elements with the same i-value next to each other to form a horizontal line. The ratio of the total number of elements to the number of elements that form part of such lines will index the degree of balanced asymmetry in the shape. The simplest index is the number of i-values over the number of elements. Identical i-values will include both those resulting from perfect symmetry as shown by isomorphic j-graphs, and those that only share the same total depth. This summary figure may then be thought of as a broad ‘symmetry index’. Si values for 3.10a, b and c are below the line chart.

Integration analysis of shapes, then, permits us to retrieve some useful descriptions of shape properties in a consistent way, though without any pretence that this is a full account of those properties. One area where this approach is useful, however, is in considering buildings as shapes. The key point here is that buildings are not pure shapes, in the geometric sense of free-standing forms in a uniform context, but oriented shapes, in the sense that they are oriented to and away from the ground on which they stand. If we take this simple fact into account in analysing building façades as shapes then we easily find some very suggestive results. This can be demonstrated by simply standing shapes on a line, which we will call the ‘earthline’. The three figures of figure 3.12 are the square and rectangular forms shown earlier with earthlines added. In the case of the rectangular form, the earthline is added twice, once to create a shape horizontally aligned to the earth and once to create a shape vertically aligned.

The first effect that must be noted is that in the case of the square, adding the earthline has the effect of reducing the original eight symmetries of the square to a simple bilateral symmetry. This can be seen visually if we compare the shading patterns of the square with an earthline to the original square form. The concentric pattern is still quite marked, but now an additional bilaterally symmetric pattern is detectable. This effect results, of course, from the earthline, as it were, drawing integration down towards itself. This confirms intuition. It is clear that we do not regard a square ç as having the symmetries of a free-standing geometrical square. We see it as a form anchored to the earth and having left-right symmetry, but not top-bottom symmetry. Indeed the language in which we describe the form - top and bottom, left and right, shows which relations we see as symmetrical and which asymmetrical.

The 'bilateral effect' of the earthline is far more marked in a square form than in an elongated form, whether we elongate the form horizontally or vertically. In the vertical form, the effect of the earthline is to make integration run from the bottom of the form to segregation at the top. This obliterates any sense of a bilateral symmetric effect in the shading pattern, and substitutes a differentiation from bottom to top. Adding an earthline to a horizontally elongated form, we again find the bilateral effect is barely noticeable in the shading pattern, and instead there is a tendency to form broad layers in the form, but with much weaker differentiation from bottom to top.

In terms of integration and symmetry index the differences between the

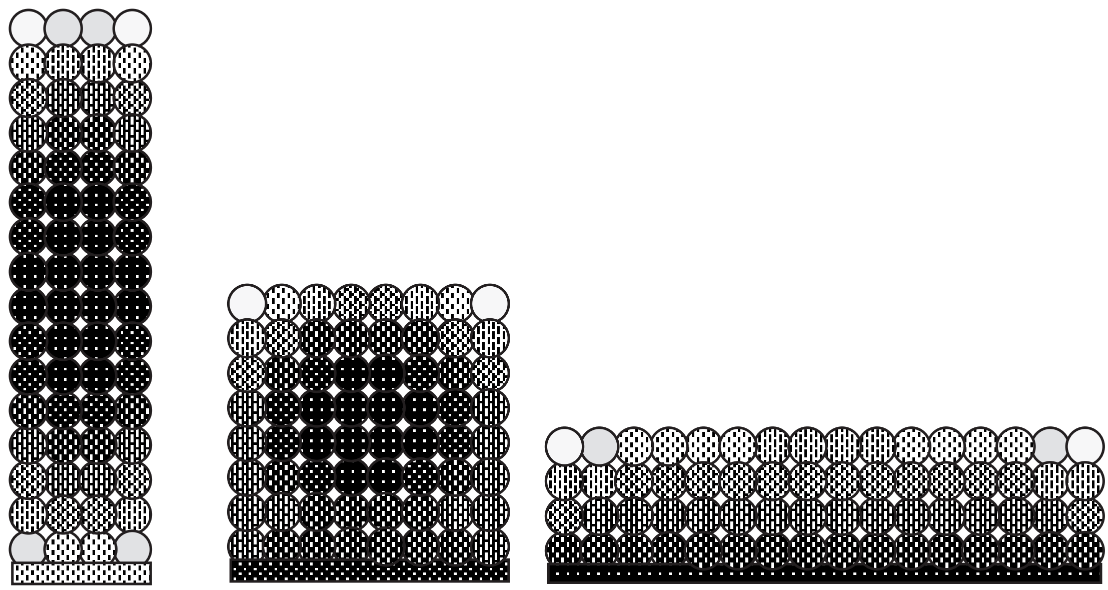


Figure 3.12

vertical and horizontal forms are also striking. The vertical form, because of the greater distance of most elements from the earthline and the fact that far fewer connect directly to it, is almost as segregated as the elongated form without the earth line. In the horizontal form, however, most elements are now closer to the earthline, with many actually touching it, and the effect is that the shape has now become much more integrated than the square form, the opposite of the case without the earthline.

Non-discursive technique

When we consider the symmetry index the effects are no less striking. Whereas in the original shapes, the square form had more 'symmetry' than the elongated form, the addition of the earthline has opposite effects on the vertical and horizontal forms. The vertical form has less symmetry than the square form, because fewer elements are on the same level, while the horizontal form has substantially more, for the contrary reason. Again, there is a common-sense reason for these effects. The addition of an earthline to a vertical form converts a pattern of integration that in the original form went from centre to edge to one that also now goes from the earthline – which, as it were, now anchors the form – upwards through the form, from more integration at the bottom, closest to the earthline, to least at the top, farthest from the earthline. The vertical form in effect now runs vertically from integration to segregation. In the horizontal form, on the other hand, insofar as elements are horizontally related, they will tend to become more similar to each other, by virtue of their closeness to the earthline. This corresponds to the intuition that the more shapes are aligned along a surface, the more equal they become. In contrast, the vertical dimension stresses difference, in that the relations of above and below are asymmetrical. Horizontality, we may say, equalises and integrates, while verticality segregates and differentiates.

The analysis of façades as layers is also suggestive. For example, if we take a simplified representation of a classical façade, we can represent it first as a shape, that is, as a metric tessellation, then, by drawing the dominant elements in the façade, as a pattern of convex elements. By analysing each separately, as in figure 3.13 a and b, we see that the shape, as represented by the tessellation shows a centralised pattern of integration focussed above, and running down into, the central column, giving the distribution a strongly vertical emphasis. In contrast, the convex analysis focusses integration on the frieze, creating a horizontal emphasis. One might conjecture that in looking at a façade we see a shape, and our view of that shape is then modified by the larger-scale organisation of elements imposed on that shape.

These centralised vertical and linear horizontal structures which are revealed by the analysis are, taken separately, among the commonest - perhaps the commonest - formal themes which builders and designers have created in whole classes of building façades across many cultures. The fact that analysis 'discovers' these structures seems, at least, a remarkable confirmation of intuition. The analysis perhaps suggests that one reason why the classical façade has often, from Laugier onwards [24] been argued to constitute a fundamental mode of façade organisation, is exactly because through its shape and convex organisation it both expresses and creates a tension between the two most fundamental modes of façade organisation. If this were the case, then it would suggest that what the human mind 'reads' when it looks at the form of a building is, or at least includes, the pattern of integration at more than one level, and the interrelations between the levels.

Non-discursive technique

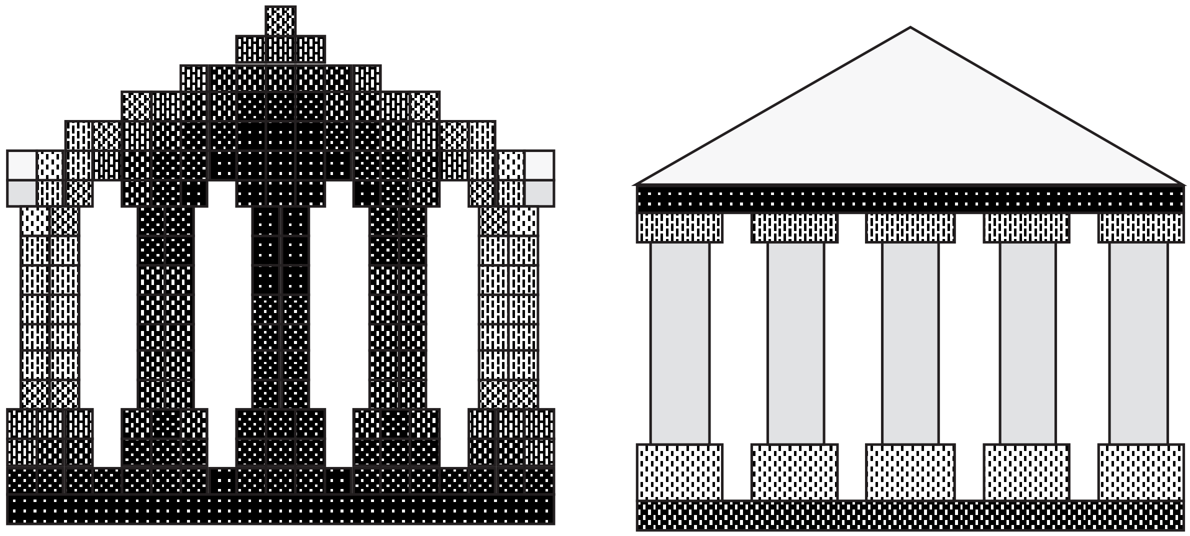


Figure 3.13

Urban space as layers: the problem of intelligibility

Whatever the case with façades, one area where substantial empirical research has established the need to consider layers of configurational potential, and their inter-relations, is urban space. Consider, for example, the two hypothetical urban layouts in figure 3.14a and b. The two layouts are composed of the same 'blocks' or 'islands' of buildings. In the first case, they are arranged in a way which has a certain degree of irregularity, but looks more or less 'urban', in that the pattern of space created by the arrangement of the blocks – and this is all that urban space essentially is – seems to have the right kinds of spaces in the right kinds of relations, and as a result appears 'intelligible' as an 'urban' system. In the second layout, all the 'blocks' are the same but each has been moved slightly with the effect that the system of space seems much less 'urban', and much less easily 'intelligible'. It is clear that any useful analysis of urban space must either capture these intuitions or show why they are illusory. It will turn out that they are not illusory at all, and that they arise from well-defined relations amongst the different spatial potentials that make up the layout.[25]

In one sense, both layouts represent the commonest type of urban space structure. We can call it the 'deformed grid', because while made up of outward facing islands of buildings each surrounded on all sides by continuous space in the manner of a regular grid, the structure of that space is deformed in two ways: it is linearly, or axially deformed, in that lines of sight and access do not continue right through the grid from one side to the other, as they would in a perfectly regular grid, but continually strike the surfaces of the building blocks and change direction as a result; secondly it is convexly deformed in that two-dimensional spaces continuously vary in their dimensions and shape, making a pattern of wider and narrower spaces. The visibility field at any point in the space for someone moving in the grid will be made up of both kinds of element. Wherever the observer is, there will always be a local convex element of some kind, in which every point is visible from every other point, plus the shape made by all lines of sight and access passing through the point. The easiest way to describe the differences between the two layouts intuitively is to say that a moving observer in either layout would experience continuous changes in the visibility field, but that the kinds of visibility field experienced in the first are quite different to those in the second. The apparent differences in intelligibility in the two layouts will turn out to be related to these formal differences in the succession of visibility fields.

We can build up an analysis of the two layouts by investigating these different potentials. First, we will consider the 'overlapping' convex elements that are defined by the surface of this block.[26] Here convex elements are defined by reference to the surface of each block, each of which defines its maximal convex field. These fields will inevitably overlap, and where they do, the area of overlap will itself form a smaller convex element from which both overlapping convex spaces will be fully visible, that is, will be convex, although these spaces are not convex to each other. The same will be true when further overlapping spaces are added. Certain small spaces will indeed be convex to a substantial number of convex spaces because all those spaces overlap in that area. Such areas will as a result have large visibility fields, whereas areas where there is no overlap will tend to have much smaller visibility fields. Overlapping convex elements are virtually impossible to intuit, because the overlapping is so difficult to represent. Computer analysis is therefore required.

Let us look first at the pattern of overlapping convex spaces generated in our two layouts. Figures 3.14c and d, are the result of the analysis of the open-space structure of the two layouts. The computer has first drawn all the overlapping convex elements defined by the faces of each 'block' and then carried out an 'integration' analysis of the pattern, with integration to segregation shown from dark-to-light, as before. In the first 'urban' layout, the darkest spaces of the resulting 'integration core' (the shape made by the darkest areas) cross each other in the informal 'market square', and dark spaces link the market square towards the edge of the 'town'. In the second, there is no longer a strong focus of integration linking a 'square' to the edges of the system and, in effect, the integration core has become diffused. In fact, the most integrating spaces are now found at the edge, and no longer get to the heart of the system. On average, the layout as a whole is much less 'integrated' than the first, that is, it has much greater total depth from all spaces to all others.

In other words, the marginal rearrangement of the urban blocks from the first to the second layout resulting in a spatial structure which is quite different both in the distribution and in the degree of integration. Intuitively, we might suspect that the edge-to-centre integration core structure of the first layout has much to do with the overall sense of urban intelligibility, and its loss in the second layout. Intelligibility is a challenging property in an urban system. Since by definition urban space at ground level cannot be seen and experienced all at once, but requires the observer to move around the system building up a picture of it piece by piece, we might suspect that intelligibility has something to do with the way in which a picture of the whole urban system can be built up from its parts, and more specifically, from moving around from one part to another.

There is in fact a simple and powerful way in which we can represent exactly this property. It is illustrated in the two 'scattergrams' in figures 3.14e and f, corresponding to the two layouts. Each point in the scatter represents one of the overlapping convex spaces in the figure above. The location of the point on the vertical axis is given by the number of other convex spaces that space overlaps with, that is, the 'connectivity' of the space with other spaces, and on the horizontal axis by the 'integration' value of the space, that is, its 'depth' from all others. Now 'connectivity' is clearly a property that can be seen from each space, in that wherever one is in the space one can see how many neighbouring spaces it connects to. Integration, on the other hand, cannot be seen from a space, since it sums up the depth of that space from all others, most of which cannot be seen from that space. The property of 'intelligibility' in a deformed grid means the degree to which what we can see from the spaces that make up the system - that is, how many other spaces are connected to - is a good guide to what we cannot see, that is, the integration of each space into the system as a whole. An intelligible system is one in which well-connected spaces also tend to be well-integrated spaces. An unintelligible system is one where well-connected spaces are not well integrated, so that what we can see of their connections misleads us about the status of that space in the system as a whole.

We can read the degree of intelligibility by looking at the shape of the scatter. If the points (representing the spaces) form a straight line rising at 45 per cent from bottom left to top right, then it would mean that every time a space was a little more connected, then it would also become a little more integrated - that is to say, there would be a perfect 'correlation' between what you can see and what you can't see. The system would then be perfectly intelligible. In figure 3.14e, the points do not form a perfect line, but they do form a tight scatter around the 'regression line', which is evidence of a strong degree of correlation, and therefore good intelligibility. In figure 3.14f we find that the points have become diffused well away from any line, and no longer form a tight fit about the 'regression line'. This means that connectivity is no longer a good guide to integration and therefore as we move around the system we will get very poor information about the layout as a whole from what we see locally. This agrees remarkably well with our intuition of what it would be like to move around this 'labyrinthian' layout.[27]

Now let us explore the two layouts in more detail. In figure 3.14g and h, we have selected a point in the 'square' in the analysis of the first layout, and drawn all the overlapping convex elements that include this point. The scatter then selects these spaces in the scattergram by making them coloured and larger. We can see that the spaces that overlap at this point are among the best connected and most integrated in the layout and that the points also form a reasonable linear scatter in themselves, meaning that for these spaces more visible connectivity means more

Non-discursive technique

Figure 3.14a

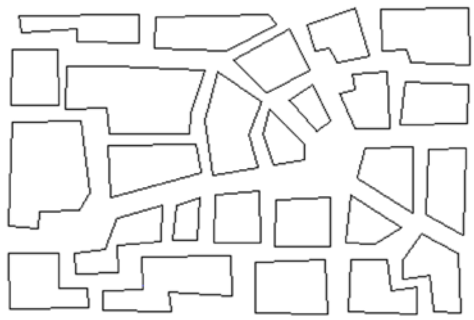


Figure 3.14b

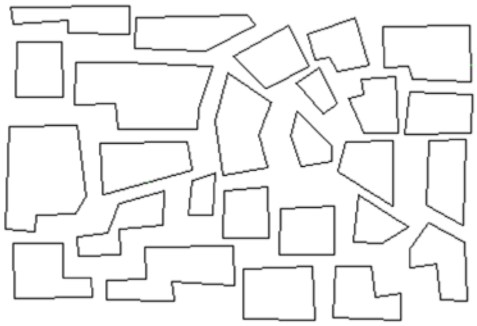


Figure 3.14c

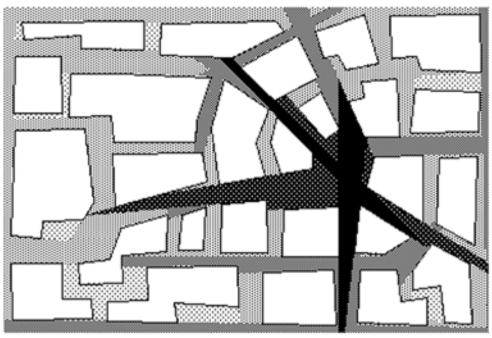


Figure 3.14d

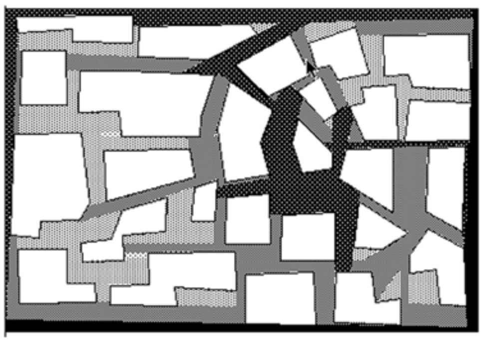


Figure 3.14e

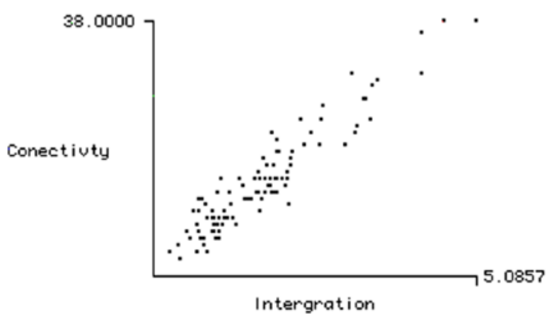


Figure 3.14f

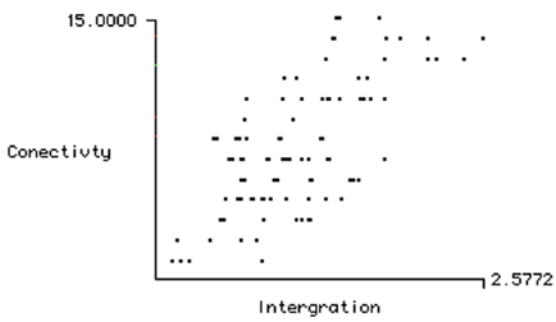


Figure 3.14g

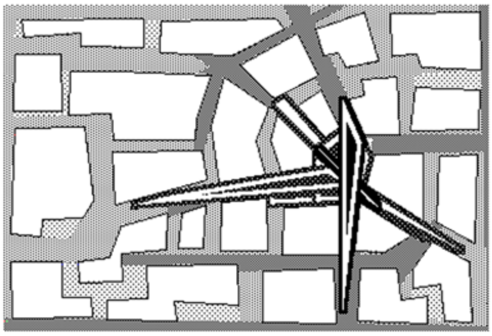


Figure 3.14h

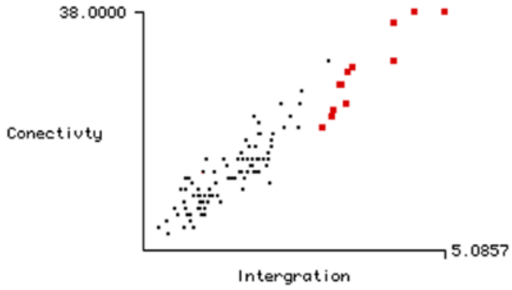


Figure 3.14j

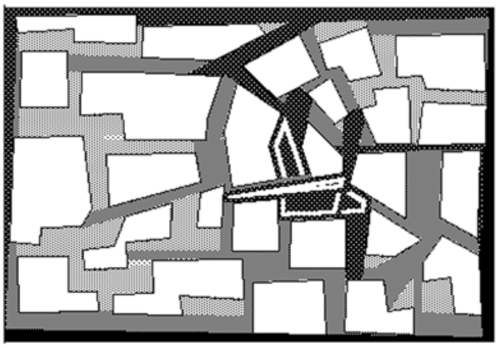


Figure 3.14k

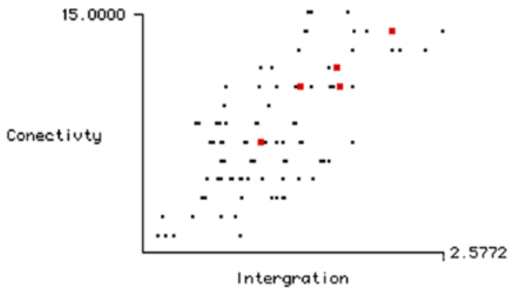


Figure 3.14l

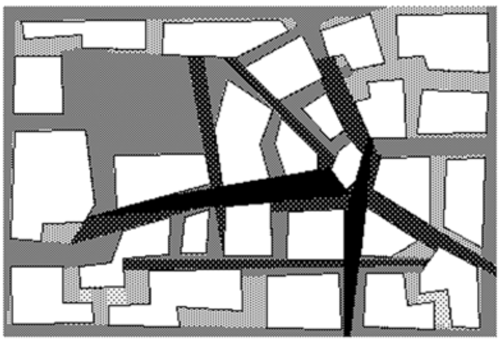


Figure 3.14m

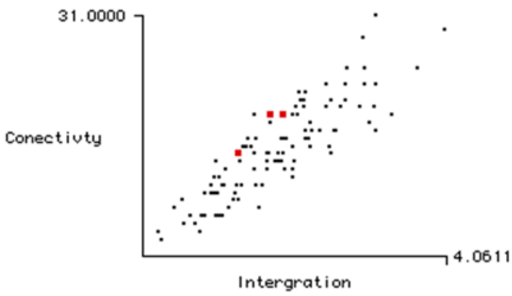


Figure 3.14n

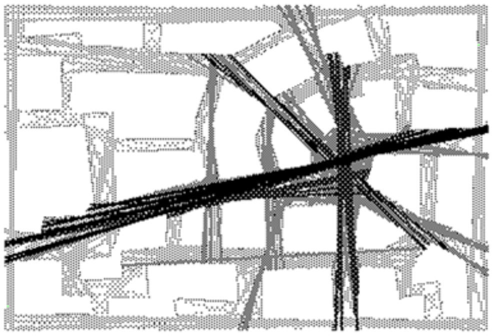


Figure 3.14p

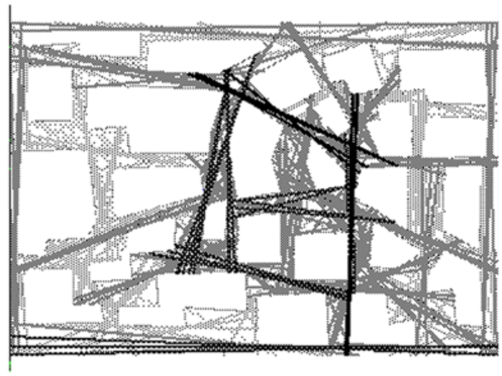


Figure 3.14q

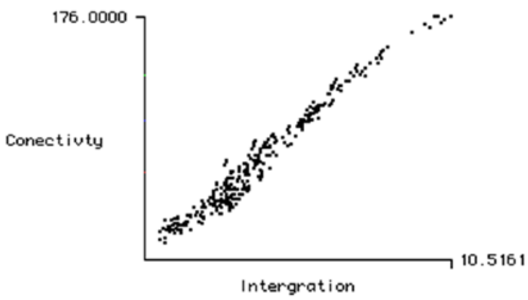
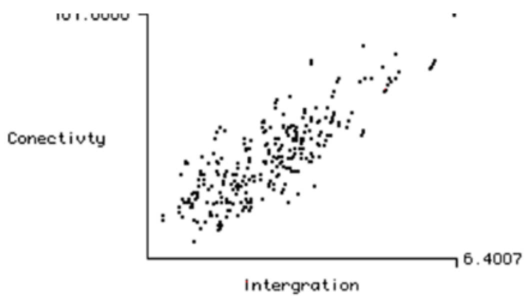


Figure 3.14r



integration. Both the shape made by the set of spaces, reaching out from the square in several directions towards the edge of the system, and the scattergram properties confirm that this point in the 'square' space has a high 'strategic' value in the layout as a whole. If we try to do the same for points in the second layout, as in figure 3.14j and k, we find that the points are buried in the scatter and have no special strategic value. By experimentally clicking on a series of points, and checking both the visual fields and the scattergrams, one can establish that there are no comparable strategic points from which a series of key spatial elements in the layout can be seen.

We may also experiment with the effects of changes to the layout. Suppose, for example, we decide that the current 'market square', although strategically placed, is too small and that it should therefore be moved elsewhere in order to enlarge it. In figure 3.14l and m, the old market square has been built over and a new, larger square has been created towards the top left of the layout. The layout has been analysed and the convex elements overlapping in the new square picked out. In spite of its size, the new square has poor integration, and its overlapping spaces occupy a poor position in the scatter. The most integrated spaces remain those pointing into the old market square. In other words, the spatial configuration as a whole continues to 'point to' the old square. An important conclusion from this, amply confirmed by the examination of real town plans, is that a square is more than a local element. How it is embedded in the configuration as a whole is equally, if not more, important. If we were to seek to exploit this by expanding the old market square by removing adjacent blocks, we would find the square becomes much more dominant, and that the largest space within the square (i.e. as opposed to those entering and leaving which are normally more dominant) is now itself the second most integrated space. In other words, we would begin to shift the emphasis of integration from linear elements to the open space itself. Again, this would distort the essential nature of layout. The size, location, and embedding of major open spaces are all formally confirmed as aspects of what we intuitively read as the urban nature of the layout.

Convex elements are not, of course, the most 'global' spatial elements in a layout, and do not exhaust all relationships of visibility and permeability. These limits are found by looking not at two-dimensional convex elements, but at one-dimensional line elements. In a deformed grid, the elements most spatially extended linearly will be the set of straight lines that are tangent to the vertices of blocks of buildings. Relations between pairs of these vertices in effect define the limits of visibility from points within the system. This can be explored through 'axial' or 'all line' analysis, and in figure 3.14n-r where the computer has found and carried out an integration analysis of all the line elements tangential to block vertices. We find that the intelligibility of the system seen axially is better than seen convexly, because lines are more 'global' spatial elements than convex elements, in that they explore the full limits of visibility and permeability within the layout. Lines therefore make the relation between the local spatial element and the global pattern of space look as good as possible. The differences between the two layouts that we found through the overlapping convex analysis are however more or less reproduced in the all-line analysis. This agreement between the two kinds of analysis is itself a significant property of the layouts.

From the point of view of how layouts work, both types of analysis are important. Movement, for example, can be predicted from a stripped down version of the axial analysis in which only the longest and fewest lines needed to cover the whole system form the line matrix. Similarly, many aspects of 'static' urban behaviours, especially the informal use of open spaces, exploit the two-dimensional 'visibility field' properties of space, with the highest levels of use normally adjacent to the most strategic spaces.

Designing with configurational models

Because these techniques allow us to deal graphically with the numerical properties of spatial layouts, we can also use them creatively in design, bringing in much new knowledge about space and function as we do so. For example, extensive research has shown[28] that patterns of movement in urban areas are strongly predicted by the distribution of integration in a simple line representation of the street grid. By using configurational analysis techniques in simulation mode, we can exploit both this knowledge, and the potential for configurational analysis to give insight into possible urban patterns that will not be at all clear to intuition. This potential has now been exploited in a large number of urban design projects, often involving the modelling of whole cities in order to simulate the effects of new designs.[29]

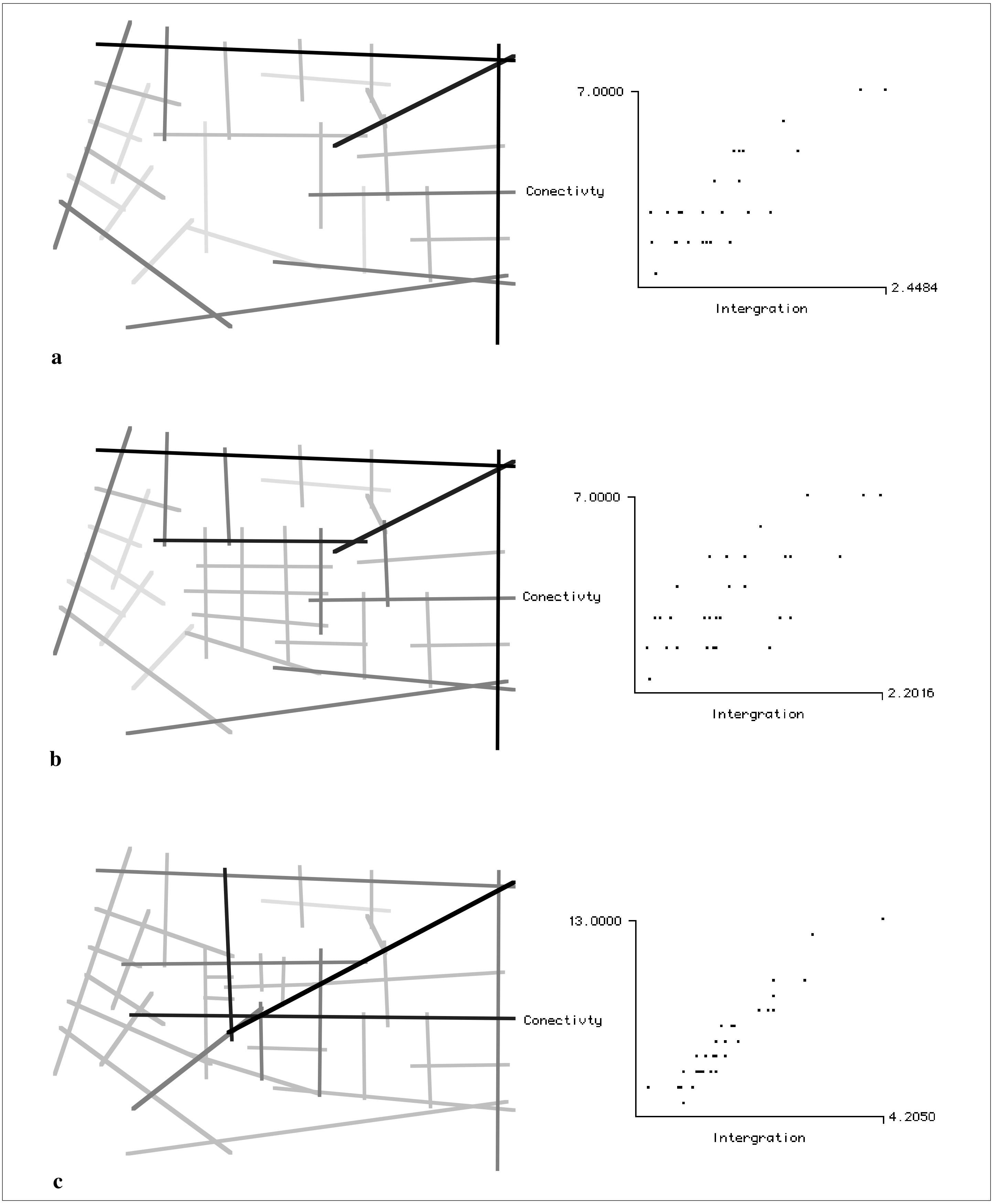
To demonstrate the essentials of the technique, a simplified hypothetical model will suffice. The top left figure of figure 3.15 is an analysed axial map (the longest and fewest lines that cover the street grid) of a small area around a hypothetical redevelopment site, with integration from dark to light as before, with, to its right, the scattergram of its intelligibility, showing a weakly intelligible system. We can experiment by asking, what would happen if, for example, we imposed a regular grid on the site without taking too much account of the surrounding structure, as the second-row figure and scatter. We see that in spite of the geometric regularity, our lack of concern for the global pattern has left us with a rather uniformly segregated space pattern within the site, with too poor a relation to the surrounding areas. As a consequence, we see from the scatter that the area as a whole has become even more unintelligible.

Suppose we then go the other way, and try to design the site by extending strong lines, and linking them to others, as in the third row figure and scatter. The result is an integrating site, and good intelligibility. The spatial structure in the site also has a good range of integrated and segregated space in close proximity to each other. As we will see in later chapters this is an important urban property (see Chapters 4 and 5.) This is a simple example, but it shows the ability of configurational analysis not only to aid the designers' intuition in thinking about patterns, and in particular in trying to understand the pattern consequences of individual design moves, but also its ability to permit the designer to think more effectively about the relation of new and existing patterns, and in general about the relation of parts and wholes in cities.

We may again illustrate this by a simplified simulation. Plate 1 is the axial map of a hypothetical urban system with well-defined sub-areas. Research has shown that the critical thing about urban sub-areas is how their internal structures relate to the larger-scale system in which they are embedded. The best way to bring this out is to analyse the system for its integration at two levels. First we do ordinary integration, which counts how deep or shallow each line in is from every other line. Second we count how deep or shallow each line in is from all lines up to three steps away. The latter we call radius-3 integration, since it looks at each line up to a radius of 3. The former we can call radius-n integration. Radius-3 integration presents a localised picture of integration, and we can therefore think of it also as local integration, while radius-n integration presents a picture of integration at the largest scale, and we can therefore call it global integration.

We will see in due course that local integration in urban systems is the best predictor of smaller-scale movement - that usually means pedestrian movement

Figure 3.15


7.0000
Conectivty
2.4484
Intergration
a
7.0000
Conectivty
2.2016
Intergration
b
13.0000
Conectivty
4.2050
Intergration
c

because pedestrian trips tend to be shorter and read the grid in a relatively localised way - while global integration is the best predictor of larger-scale movement, including some vehicular movement, because people on longer trips will tend to read the grid in a more globalised way. In historical cities, as will be shown, the relationship between these two levels of integration has been a critical determinant of the part-whole structure of cities, because it governs the degree of natural interface there would naturally be between more local, and therefore more internal movement, and more global and therefore more in-out movement and through movement.

Some of the different effects on this relationship that different types of local area design will have can be shown by highlighting the areas in scattergrams of the whole system and examining the scatter of local against global integration. The area shown in the bottom row, for example, is a classically structured area for a European city, with strong lines in all directions from edge to centre, with a less integrated structure of lines related both to this internal core and to the outside. This ensures that those moving in the area will be conscious of both the local and global scales of space as they move around, and there will be a good interface between local and global movement. The scatter formed by the sub-areas is shown to the right. The points of the area form a good linear scatter, showing that local integration is a good predictor of global integration, and cross the regression line for urban area as a whole at a steeper angle, showing that there is a stronger degree of local integration for the degree of global integration. A line on the core of the whole settlement will, in contrast, lie at the top end of the main regression line. This shows how subtly urban areas create a sense of local structure without losing touch with the larger-scale structure of the system. (See Chapter 4 for an examination of real cases).

The area shown immediately above, in the second from bottom row, is typical of the layouts we tend to find in housing estates, with few connections to the edge and little relation between the edge to centre structure and the internal structure of the layout. This type of layout is invariably shown as a series of layers in the red point scatter with virtually no correlation between local and global integration. Such layouts invariably freeze all our natural movement and become structurally segregated lumps in the urban fabric.[30] The areas in the top two rows show other variations on local area structure, one producing effects rather similar to those in the experimental grid in the design experiment of figure 3.16, while the other is a random scatter of lines, showing that in spite of the apparent informality of much good urban design, random lines simply do not work except by chance.

Future urban models: intelligent analogues of cities

In addition to their role in design, configurational models are now being developed as a basis for researching into the multidimensional dynamics of cities. Consider, for example, one of the broadest and least tractable of issues facing the built environment industry: that of the economic, social and environmental 'sustainability' of cities. Even to monitor effectively and compare cities on sustainability criteria,

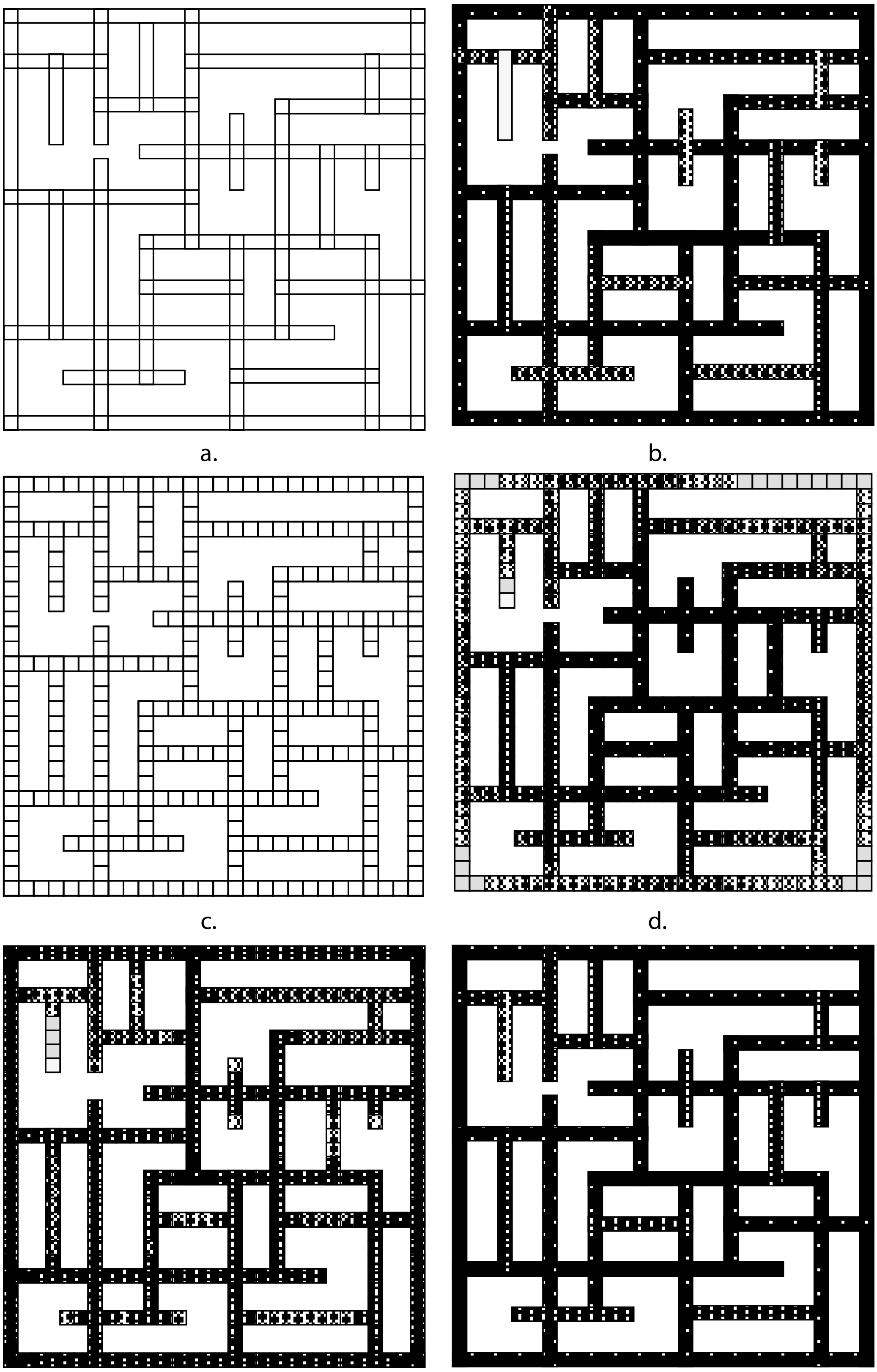
Non-discursive technique

whatever they might turn out to be, we must bring data on the physical and environmental performance of cities together with data on their economic and social performance, and to relate both to some kind of description of the city. For example, energy consumption and pollution production depend, among other factors, on settlement patterns. Should settlements be dense or sparse, nucleated or dispersed, monocentric or polycentric, or a mix of all types? For research to give an answer, measurement data on environmental performance, and data on the implications of different behavioural assumptions (for example about the distribution of work and home) and 'knock-on' effects such as the economic, social and cultural consequences of spatial aggregation and disaggregation policies, must be related to descriptions of the physical and spatial form of cities which reflect the range of variation found in the real world.

To work towards a theoretical model of how this might be done, we may begin with the purely 'configurational' models we have presented, and show how other key spatial attributes such as metric distance, area, density, plot ratios, shape, political boundaries, and so on can be expressed within the configurational model by using the idea of integrating 'layered' representations of space into a single system. For the purposes of illustration we will again use notional, simplified examples. First, we represent a street network as a series of lines or strips, and analyse their pattern of integration, as in figure 3.16a and b. In this analysis, no account has yet been taken of metric distance. However, in some circumstances at least, this seems likely to be an important variable. We can supply this by selecting an arbitrary module - say a ten-metre square – and linking modules into the pattern of the grid and analysing this as a tessellation shape, as in figure 3.16c. On its own, this is not of great interest, since it inevitably reflects the pattern of metric centrality in the grid, as in figure 3.16d, but if we superimpose the line network onto the metric modular system and analyse the two layers as a single system, then the effect is to weight each line with a number of modules directly related to its length. The outcome of this 'length weighted' integration analysis is shown at both levels of the combined analysis: in terms of the modular units in figure 3.16e, and in terms of the 'line superstructure' of strips in figure 3.16f. The strip level is much the same as previously, but the modular elements show an interesting - and very lifelike - localised structure in which greater integration is concentrated at the 'street intersections', with less integrated modules in the centres of links away from the intersections. This immediately enables us to capture a new and functionally significant aspect of space organisation in a representation.

The relationship between metric area and configuration can be dealt with in an analogous way by underlaying convex elements with a two-dimensional modular layer, as in figure 3.17a-f. In a–c we see how a simple system in which four convex spaces of equal size and shape and the connections between them are represented as a layer of modular elements with four convex elements and four strips for the connection superimposed. The two-layer system is then analysed. Whether we look at the result with the convex layer uppermost or the modular layer, the results will

Figure 3.16



a. b. c. d. e. f.

be a symmetrical distribution of integration dominated by the strips. In figure 3.17 d–f we give the convex elements different areas and underlay modular elements accordingly, so that each is now weighted by the number of modular elements it overlays. Analysis separately then together shows that integration is drawn into the convex elements according to their area. Note however that the integration of the two smaller convex areas (on the top) are in the 'wrong' order. This is because the one on the left is closer to the largest-scale convex area (bottom left) and this affects its own integration with respect to the rest of the system. Thus the results show a combination of configurational effects and metric area effects. From this we can see that if we make a large and small square configurationally equivalent in an urban system then the large square will integrate more. Metric area, it turns out is like distance, a property capable of expression as an aspect of configuration.
We may simulate the effect of plot ratios and densities by equally simple means. For example, if we wish to attach a building with a given number of floors to a street network, all we need to do is attach a convex space the size of the ground area of the building to the appropriate position in the street system, then overlay on that a convex element for each floor, making sure that each element above the ground is detached from the street and only connected through the ground layer as it would be in real life. This will not appear visually as a three-dimensional structure, but it will exactly represent the addition of above ground floor space to the urban system.

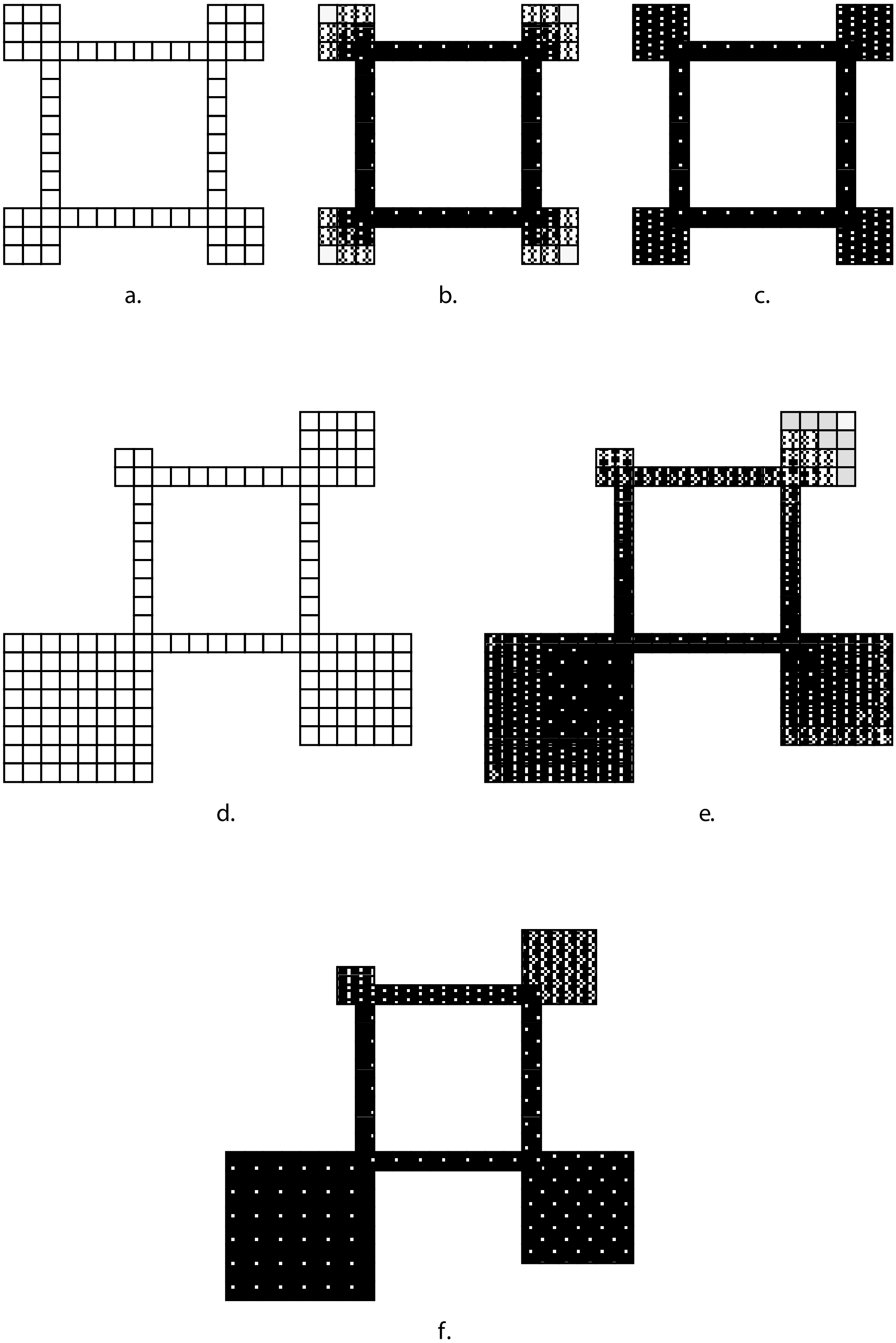
We may now build a model of an urban system in the following way. First, we divide the city up into an arbitrary number of areas and represent them as non-contiguous polygons. These may be as small or as large as we need, according to the level of resolution required by the research question. The polygons may be based on political boundaries, like wards, administrative boundaries like enumeration districts, segments defined by an arbitrarily fine grid, or they may be defined by objective morphological properties of the built environment. These polygons representing areas are the fundamental units of analysis for the technique.

Figure 3.18a shows our imaginary simplified case in which the street network of the city (or part-city) is superimposed on the patchwork of polygons so that each polygon is linked into the urban system by all the streets or part-streets that pass through it or alongside it. This two-level spatial system is analysed 'configurationally' to find the pattern of integration in the whole system. Evidently, the street pattern will tend to dominate the area polygons simply because the streets are connectors. However, the street system can then be 'peeled off' the polygons, as in figure 3.18b, leaving a pattern of polygons with their spatial characteristics in relation to the city area around them, and to the city system as a whole, recorded as a set of numbers.

This basic process of linking areas together by the street network in a single configurational model is the basis of what we call an 'intelligent urban analogue' model. Once this is established, we can then complicate the model in all the ways we have described previously. For example, we can underlay the street network with metric modules so that the analysis of the street system takes distances into account. We can underlay the polygons with metric modules so that the metric area

Non-discursive technique

Figure 3.17



of a polygon is taken into account. We can also, if we wish, superimpose layers on the polygons representing off the ground floor space.

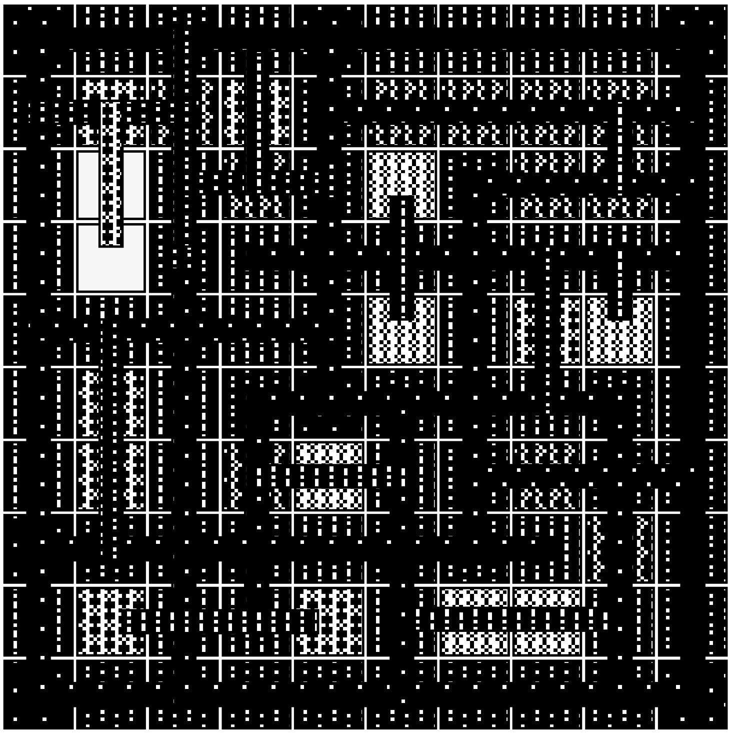
There is also an easy way of further disaggregating any model from the level of resolution originally selected. Each of the original area polygons can be itself subdivided into much smaller polygons and analysed as before. This more localised analysis will give a much richer and denser picture of the detailed characteristics of the area. These may then be fed into a larger-scale model as more detailed environmental descriptors. There is no reason in fact why both levels of the model should not be analysed as a single system. The principal barrier would be computing time. In our experience adding a new level of fine structure to an existing model leaves the larger-scale picture more or less intact provided that the disaggregation is done uniformly and is not confined to particular regions.

At the other end of the scale, we may also derive new measures of the most macro-properties of the city system, such as shape, and shape loaded with different densities in different regions. This can be done by simply linking the area polygons together and analysing the distribution of integration in the system without the superimposed street system. Shape will be indexed by the degree and distribution of integration, and can be shown both by direct graphical representation of the city system, or by statistical representations such as frequency distributions, or simply by numbers. The effects of weighting shapes by loading different regions with higher densities can be explored by simply overlaying the spaces representing the additional densities onto the relevant polygons of the contiguous polygon system, then proceeding as before. By varying the pattern and density of centres we can explore their effects on total distance travelled, other things being equal, in different kinds of three-dimensional urban system. The effects of other nearby settlements can also be investigated by simply adding them as extensions to the model.

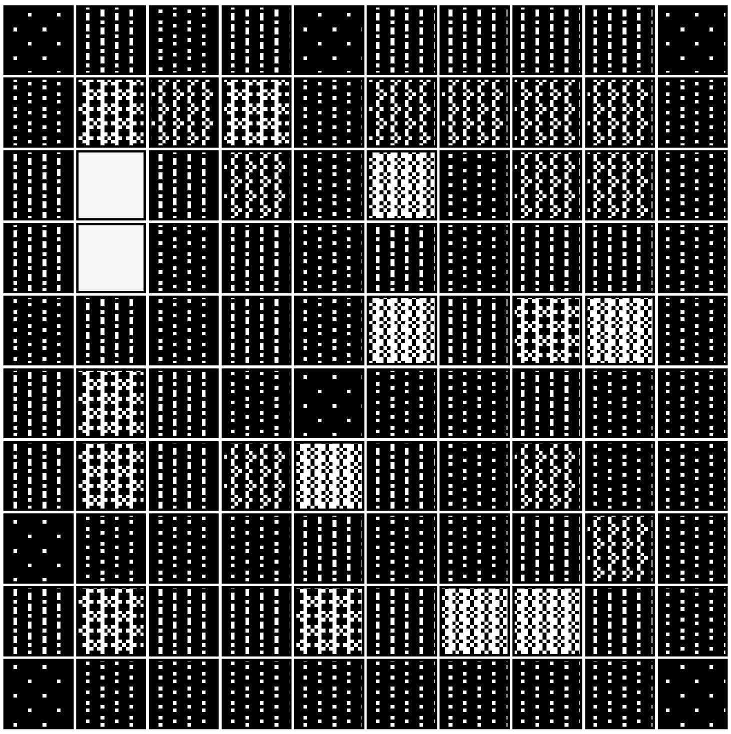
The numerical data resulting from the analysis of the urban system can then be used in a number of ways. First, most obviously, the parametric descriptors for the polygons resulting from analysis, reflecting as they do the position and configuration of each 'finite element' in the city system as a whole, then become the frame for other kinds of data which can be assigned as descriptors to the polygons. This can be done with any functional variable that can be numerically indexed for that area such as population densities, pollution levels, traffic movement, pedestrian movement, unemployment rates, crime rates, council tax banding, and so on. Because spatial and other descriptors are now all in numerical form, simple statistical analyses can begin to reveal patterns. Second, the distribution of any property may be represented graphically in the urban system as a visual distribution of that property in the city system. This means, in practice, that all the visualising and cartographical potentials that have been developed in the past few years through 'geographic information systems' can be interfaced with, and potentially brought within the scope of, an analytic model with proven ability to link morphological and functional properties of built environment systems, hopefully in a more predictive way.

Layered models are the future of configurational modelling of space.

Figure 3.18



a.



b.

Non-discursive technique

These new techniques arise from the results of research over several years in which various types of configurational modelling have been used first to identify non-discursive regularities in the ways in which architectural and urban systems are put together spatially and identify the 'genotypes' of spatial form; second to correlate these non-discursive regularities with aspects of how human beings can be observed to function in space; and third, to begin to build from these regularities a picture of higher generality of how spatial systems in general are put together and function in response to the demands that human beings and their collectivities make of them. In the next chapter we introduce the most fundamental of all correlates with spatial configuration: human movement.

Notes

1 H. Simon H, *The Sciences of the Artificial*, MIT, 1969.

2 F. De Saussure F, *Course in General Linguistics*, McGraw Hill, 1966 translated by C. Bally and A. Sechahaye with A. Riedlinger See pp. 9–15 (originally in French 1915).

3 It has of course become fashionable to follow the later Wittgenstein's *Philosophical Investigations* (Basil Blackwell 1953; Edition used 1968) and deny any systemic properties to such things as languages, and see in them only shifting contingencies. For example: 'Instead of producing something common to all that we call language, I am saying that these phenomena have no one thing in common which makes us use the same word for all, but they are all related to each other in many different ways. And it is because of this relationship, or these relationships, that we call them all "language".' – Wittgenstein, para 65. Or: 'Language is a labyrinth of paths. You approach from one side and know your way about; you approach the same place from another side and no longer know your way about' – Wittgenstein, para 203. The use of the urban analogy is interesting. As we will see in later chapters, this is the one type of artefact where it can be shown quite clearly that Wittgenstein was wrong.

4 The clearest statement is still probably the 'Overture' to Claude Lévi-Strauss's *The Raw and the Cooked*, Jonathon Cape, London 1970, originally in French as *The Cru et le Cuit*, Plon, 1964.

5 Plato, *The Republic*, for example VI, 509–11, pp. 744–7 in Plato, *The Collected Dialogues*, eds. E. Hamilton and H. Cairns H, Princeton University Press, Bollingen Series, 1961. See also ed. F. M. Cornford, *The Republic of Plato*, Oxford University Press, 1941, pp. 216–21.

6 For the clearest formulation, see R. Thom, 'Structuralism and Biology', in ed. C. H. Waddington, *Theoretical Biology 4*, Edinburgh University Press, 1972, pp. 68–82.

7 W. Heisenberg, *Physics and Philosophy* George Allen & Unwin, 1959 p. 57.

8 N. Chomsky, *Syntactic Structures*, Mouton, The Hague, 1957.

9 There are important exceptions to this, for example Lévi-Strauss's attempt, in collaboration with Andre Weil, to model certain marriage systems as Abelian groups. See Lévi-Strauss, *The Elementary Structures of Kinship*, Eyre & Spottiswoode, 1969, pp. 221–9. Originally in French as Les Structures *Elementaire de la Parente*, Mouton, 1949.

10 For example, his ingenious attempt to model the elementary properties of matter through the five regular solids in the *Timaeus*. See Plato, *Timaeus* 33 et seq. p. 1165 in *The Collected Dialogues* (see note 5 above)
11 This process is the subject of Chapter 9.
12 As described, for example, in Chapter 2 of *The Social Logic of Space*.
13 For a lucid summary, see P. Steadman, *Architectural Morphology*, Pion, 1983.
14 L. March, In conversation.
15 I. Stewart and M. Golubitsky, *Fearful Symmetry*, Penguin, 1993, p 229.
16 F. Buckley and F. Harary, *Distance in Graphs*, Addison Wesley,1990, p. 42.
17 B. Hillier and J. Hanson , *The Social Logic of Space*, Cambridge University Press, 1984, p 108. See also note 16 in Chapter 1.
18 Steadman, p. 217.
19 Hillier & Hanson, pp. 109–13.
20 However, see the references in note 16 of Chapter 1.
21 For example, I. Biederman, 'Higher level vision', in eds. D. Osherson et al., *Visual Cognition and Action*, MIT Press, 1990.
22 For a discussion of some of these variations from the point of view of graph theory see Buckley and Harary, *Distance in Graphs*, pp. 179–85.
23 For example, P. Tabor, 'Fearful symmetry', *Architectural Review*, May 1982.
24 Abbe Marc-Antoine Laugier, *Essai sur l'architecture*, Paris 1755.
25 See Hillier & Hanson, *The Logic of Space*, p 90.
26 It should be noted at the outset that these overlapping convex elements are unlike the convex elements described in *The Social Logic of Space*, which were not allowed to overlap. See Hillier & Hanson, pp. 97–8.
27 It is exactly this property that labyrinths exploit. At every point the space you see gives no information – or misleading information – about the structure of the labyrinth as a whole. In general – though not invariably – a good urban form does exactly the opposite.
28 See Chapter 4. Also B. Hillier et al., 'Natural movement: or configuration and attraction in urban pedestrian movement, *Environment & Planning B, Planning & Design*, vol. 20, 1993.
29 As, for example, in the case of the new Shanghai Central Business District on which we collaborated with Sir Richard Rogers and Partners, or the original plan for the Kings' Cross Railways Lands, London with Sir Norman Foster and Partners. See for example B. Hillier, 'Specifically Architectural Theory', Harvard Architectural Review, vol. 9, 1993. Also published as B. Hillier, 'Specifically architectural knowledge', *Nordic Journal of Architectural Research*, 2, 1993.
30 The problems generated by this type of layout are examined in detail in Chapter 5.