

Binary Encoding of a Class of Rectangular Built-Forms

09

Philip Steadman
University College London, UK

In 1999 I made an abortive attempt to get to the Space Syntax conference in Brasilia, but was turned back at the airport. I had planned to present a paper entitled 'Every built form has a number', which is still unpublished. Since then, my colleague Linda Waddoups and I have continued to work on the same ideas. So I thought I would take this opportunity to offer a resume of that earlier paper, and to report on the progress we have made over the last two years.

In fact the work goes back a little further, to a three-year-old paper called 'Sketch for an archetypal building' (Steadman 1998), in which I suggested a theoretical approach to the classification and enumeration of rectangular built forms. The term 'built form' is used here in Lionel March's sense, to refer to mathematical models for representing buildings 'to any required degree of complexity in theoretical studies' (March 1972). In the present work these are abstractions from the geometrical complexity of real buildings, in which all articulations of facades are ignored, as is the detailed planning of individual rooms. Instead the interior is represented as being divided into *zones* of different kinds, as I will explain.

In computer-aided design the process of defining the geometrical form of a building is generally one of *composing* elementary forms together. In my 1998 paper I proposed a diametrically opposite kind of approach, in which one would start always from the same large and complex 'archetypal form', and generate other forms by *selecting* suitable parts. To draw an analogy with sculpture, this is a method of carving rather than a method of modelling. Figure 1a shows the archetypal form. It has an arbitrary number of storeys. The lower floors are continuous, while the upper floors are punctuated by an array of courtyards (Figure 1b). The diagrams show 3 x 3 courts, but there could be more.

As mentioned, the archetypal building takes no account of how the space inside might be sub-divided in detail. The representation is at a higher level of abstraction, into three types of zone, distinguished by the nature of their *lighting*:

- 1) Space adjacent to the external facades, and around the courtyards, which has the potential to be daylit via windows (the middle tones in Figure 1).
- 2) Space immediately below the bases of the courtyards, and on the topmost level of the courtyard floors, which has the potential to be daylit by roof-lights (the darkest tone).
- 3) All other space in the interior of the archetype, which must of necessity be lit by artificial light (the lightest tone).

Keywords:
Archetypal building,
Boolean description,
daylighting,
dimensionless
configuration, formal
typology, classifica-
tion of built forms

Professor Philip Steadman,
Bartlett School of
Graduate Studies,
University College
London, (Torrington
Place Site), Gower
Street,
London WC1E 6BT
j.p.steadman@ucl.ac.uk

09.1

Notice how the diagrams show strips of artificially-lit space between the sidelit zones surrounding the courtyards. In real buildings these might correspond to internal corridors, or to central strips of service accommodation, as for example rows of internal bathrooms in hotels. The whole of the interior of each of the lower continuous floors must obviously be lit by artificial light, although a zone around the perimeter can be daylight. (There could also be basement floors, which would lack daylight altogether; but they are not shown here.) Note that the archetype is not conceived as being indefinitely extensible in the horizontal direction, like for example the arrays of built forms considered in the work of Martin and March (1966). On the contrary, the archetype is bounded on its four sides by the outward-facing ring of daylight space.

In effect, the archetype is a kind of ‘maximal’ built form in which, within the confines of a rectangular geometrical discipline, as much accommodation as possible is fitted onto a given site area. Should it be acceptable for all this accommodation to be lit artificially, then obviously all floors can fill the site completely, and the result will be a solid rectangular block. If it is required that all the accommodation be daylight, then the courtyard floors provide that configuration

in which daylight floorspace is maximised within the given site area. Differing proportions of daylight to artificially lit space can be accommodated, clearly, by varying the ratio of continuous floors to courtyard floors. I will come back to this point.

The archetypal form is to be imagined as a *dimensionless* configuration, to which dimensional parameters can be assigned in the x , y and z directions. Dimensional values can be assigned in z to correspond to storey heights. Dimensional values in x and y specify the widths in plan of strips of accommodation across the form, whether these be daylight or artificially lit; and they specify the widths of courtyards and the zones that flank them. In practice

there might for example be an effective maximum plan depth for any strip of sidelit space, of around 6 or 7m. The width of a central strip of artificially-lit circulation space might be set at say 2m. And so on. Overall, the archetypal form can be represented as a matrix of cuboids in which, on any one floor level, each court is represented as a single cuboid, and the respective strips of accommodation are made up of rows of cuboids. Should any storey height parameter in z be set to zero, then the entire floor in question will disappear. Should dimensioning parameters in x or y be set to zero then the strip of accommodation in question - possibly including a court or courts - will be suppressed. This is how parts can be selected from the archetype to make different, smaller built forms.

In my 1998 paper I showed how the forms of an eclectic variety of real buildings could be approximated through appropriate transformations of the archetype, in three stages. First, entire floors, and entire strips of accommodation across the form, are eliminated. Second, the remaining parts are joined. Third, values are set for the vertical and horizontal dimensions. Figure 2 demonstrates the process for one bedroom floor of George Post’s Roosevelt Hotel, built in New York in the 1920s. Other examples in the paper included a factory, a theatre and a town hall.

09.2

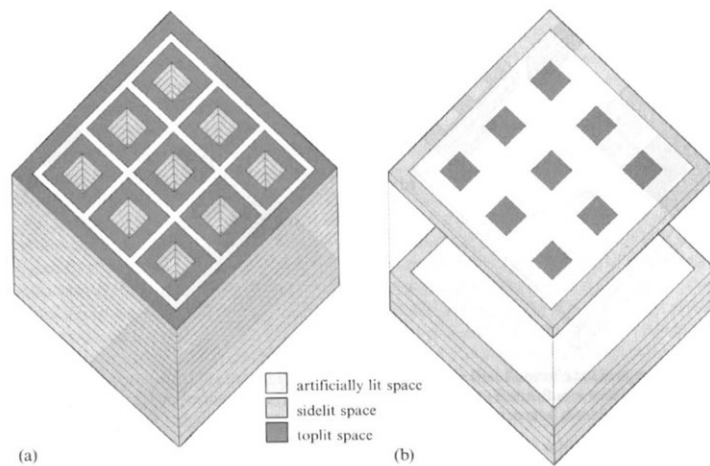


Figure 1: a) The archetypal form. b) The upper courtyard floors of the archetype

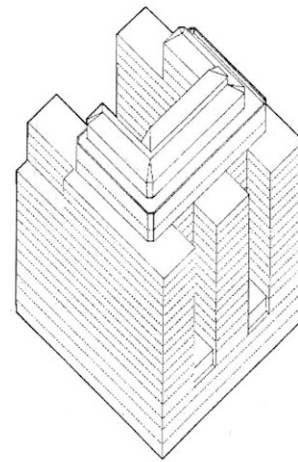
The paper also envisaged the possibility that every underlying dimensionless configuration derivable from the archetype might be described by a binary code. This code would list, in some conventionally-defined order, all strips of accommodation in x and y and all floors in z . If a strip or floor was present, this fact would be signalled by a 1; if absent, by a 0. For an archetypal form with a given number of storeys and a given number of courts in x and y , the resulting codes would be always of the same length. All *possible* forms derivable from the archetype might be enumerated by permuting strings of 0s and 1s of the relevant length. These codes might be set in ascending order to create a comprehensive catalogue.

This technique of binary encoding had its inspiration in some proposals by Lionel March for 'A Boolean description of a class of built forms' (March 1976). March's method required that the envelope of some rectangular building be enclosed in a bounding box. This box was then subdivided with a series of orthogonal planes, coinciding with all major external surfaces of the building, to create a three-dimensional array of cuboids. Any cuboid that corresponded with a part of the built form was coded with a 1. Any cuboid that corresponded to empty space outside the form was encoded with a 0. March proposed a convention for 'unpacking' the cuboids, so that the 0s and 1s could be listed in a single string. The binary encoding served, as with the archetype, to represent the configuration of the built form in question, independent of its metric dimensions. March illustrated the method with the example of Mies van der Rohe's Seagram Building (Figure 3).

There are nevertheless some important differences between March's approach and a method of binary encoding based on the archetypal built form. With March's technique the length of the code is dependent on the complexity of the built form under consideration, and the positions of 0s and 1s in the resulting string are not especially significant. With the coding of the archetype and its transformations, by contrast, the lengths of codes are always the same, and all the 0s and 1s carry definite meanings by virtue of their positions in the string, as we will see. For forms of a given complexity, what is more, the binary codes derived from the archetype are generally shorter than March's equivalents. It is this fact that makes it practical, from a combinatorial point of view, to list them exhaustively. The penalty is a certain inflexibility compared with March's approach, whose cost we will look at in due course.

In 'Every built form has a number' I explored this approach to coding, working by hand. To limit the scale of the task, I confined my attention to an archetypal built form on a single floor level, with a single courtyard. The z component of every code was thus a single 1 and could be ignored. The court was represented as an array of 5×5 cuboids, to correspond to the central court, a ring of inward-looking space sidelit from the court, and a ring of outward-looking space

Figure 2: The Roosevelt Hotel, New York, designed by George Post: general view (left) and a typical bedroom floor (right)



09.3

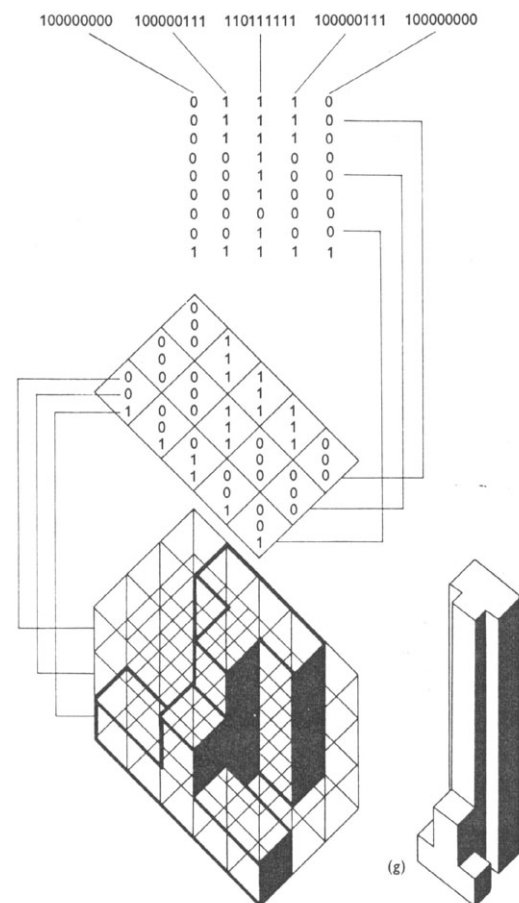


Figure 3: Binary encoding of the form of Mies van der Rohe's Seagram Building, [from March (1976)]

sidelit from the exterior (Figure 4). (There was no artificially-lit space.) This gave a 5-digit x sub-string and a 5-digit y sub-string, creating a binary code of 10 digits in all. The digits are listed by convention with the x sub-string first, reading from left to right, followed by the y sub-string, reading from top to bottom:

x	y
1 2 3 4 5	6 7 8 9 10

Figures 5 and 6 illustrate the derivation of built forms and their corresponding codes,

approximating to a detached house and a terraced house respectively, both with simple rectangular plans. The code 10001 10001 selects the four corner cuboids to create the detached plan, daylit from all four sides (Figure 5). The code 00010 10001 selects two cuboids from opposite sides of the archetype to create a terraced plan that is daylit from the two ends, while the remaining sides - shown in heavy line - are 'blind' (Figure 6). The paper showed how this

Figure 4: Array of 5 x 5 cuboids, represented by a 10-digit binary string, where the first five digits represent the columns and the last five digits represent the rows. The central cuboid marked X represents the court. All other cuboids are daylit

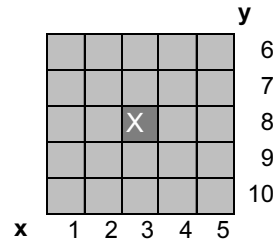


Figure 5: Four corner cuboids selected from the 5 x 5 array with the binary code 10001 10001, to give a detached plan lit on all four sides

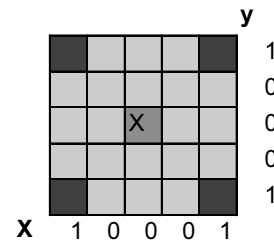
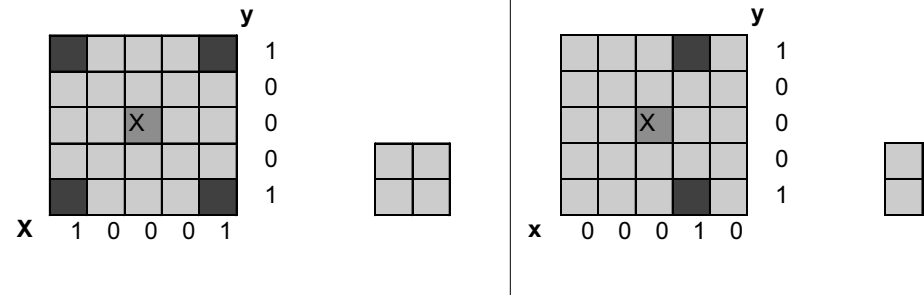


Figure 6: Two cuboids selected from the 5 x 5 array with the binary code 00010 10001, to give a terraced plan lit at two ends only

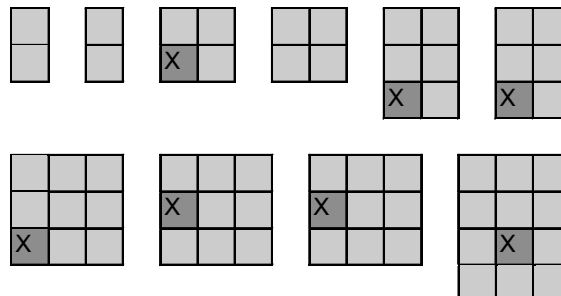


simple encoding could be used to capture the basic plan arrangements of a sample of English public houses. Figure 7 illustrates a selection of frequently-occurring plan configurations for pubs, including L-shapes and U-shapes suitable for corner and mid-terrace sites. See how, in these instances, the 'courtyard' serves as an external zone around which the L or U-shape is bent.

It might be thought, at first sight, that the number of distinct possible configurations derivable from this 5 x 5 archetype would equal the number of different possible 10-digit strings, which is $2^{10} = 1024$. This however is not the case, for several reasons:

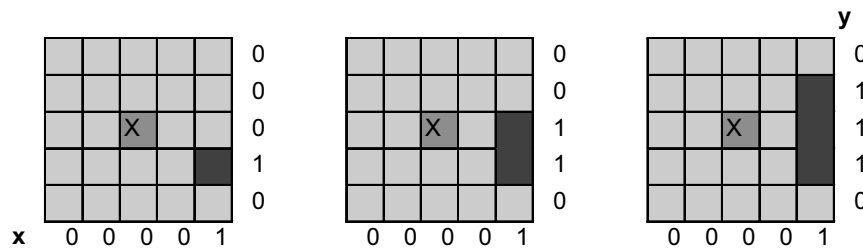
1) If the five digits in either the x sub-string or the y sub-string are all zero, then no cuboid is selected, and there is no corresponding form.

Figure 7: Frequently-occurring plans for English public houses, derived from the 5 x 5 array



2) Many pairs of binary strings (or groups of four strings) correspond to configurations that are isomorphic by reflection and/or rotation. Take the ‘terraced house’ example 00010 10001 of Figure 6. The string 01000 10001 represents an identical configuration, as do the strings 10001 00010 and 10001 01000 (Figure 8). These four configurations differ only by virtue of being rotated relative to the coordinate system. In other cases there are different left and right-handed versions of the same configuration (enantiomorphs). It seems reasonable to select just one isomorph to stand for all the others in every such instance. It is convenient to choose always that isomorph which has the *lowest* binary code. In the example of Figure 8 this would mean selecting 00010 10001.

3) It is possible in certain instances for different binary strings to correspond to configurations that are effectively indistinguishable once they are dimensioned. Consider for example the single cuboid selected by the code 00010 00001. This corresponds to a simple rectangular plan, daylit from one side only. (It could be the plan of one floor of a ‘back-to-back’ house, of the type built in some cities in the north of England in the 19th century.)



Now consider the configurations represented by the codes 00110 00001 and 01110 00001 (Figure 9). These are effectively identical to 00010 00001, it merely requiring the one, two or three cuboids to be given *y* dimensions which sum to the same value, to produce the very same dimensioned plan. In such cases it seems reasonable again to pick the configuration with the lowest binary code, to stand for all the rest.

4) Should any court not be selected, then obviously the cuboids which would otherwise be adjacent to that court cannot be daylit. Any configuration in which these ‘sidelit’ cuboids are selected therefore, but the court is not selected, is inadmissible.

With suitable filtering rules applied to eliminate these various duplicates and forbidden cases, it is possible to find all legitimate codes and list them in ascending order. For the 5 x 5 array the number of distinct codes is 65 (that is, a mere 6% of the 1024 distinct 10-digit binary strings). My colleagues Linda Waddoups and Jeff Johnson at the Open University have developed a computer algorithm which identifies legitimate codes for larger (single-storey) arrays. For the 7 x 7 single-court array, in which a ring of artificially-lit space is introduced between the two rings of sidelit space (Figure 10), the number of distinct configurations is 675. Figure 11 reproduces a sample page from a complete catalogue prepared by Waddoups

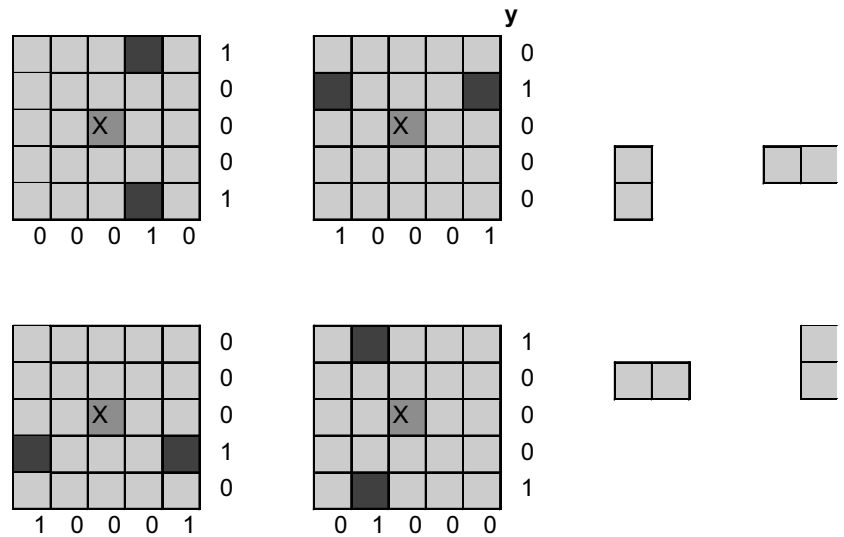


Figure 8: Four configurations that are isomorphic under rotation: 00010 10001 (top left); 00001 01000 (top right); 10001 00010 (lower left); 01000 10001 (lower right)

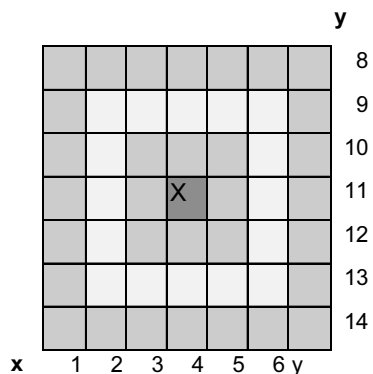


Figure 10: 7 x 7 array of cuboids, encoded with a 14-digit binary string. The central cuboid marked X is the court. Light tone indicates artificially-lit cuboids, darker tone indicates daylight cuboids.

09.6

which shows diagrammatic plans accompanied by their codes. In each case the square marked by a cross is the courtyard; artificially-lit space is in light tone; and daylight space is in darker tone. Sidelit facades are shown by dotted lines, and blind facades by solid lines. (Of course if these were indeed single-storey buildings, then *all* floorspace could in principle be toplit. We might imagine that they are, rather, intermediate floors of multi-storey buildings.)

A four-court archetype can be represented in plan by an 11 x 11 array, allowing again for artificially-lit strips between the sidelit zones (Figure 12). The Waddoups-Johnson algorithm generates 37,137 distinct codes for this 11 x 11 arrangement (that is, less than 1% of the 4,194,304 possible 22-digit strings). The algorithm has yet to be applied to the nine-court archetype (as in the courtyard floors in Figure 1), but Waddoups estimates that the corresponding number of distinct codes will be around 2 million. This is certainly a large number. But the task of searching a catalogue of 2 million 30-digit codes by computer is hardly a daunting one.

What is more, once the codes are arranged in ascending order, it turns out that the corresponding built forms themselves become ordered into some potentially interesting groupings. Overall, it will be appreciated that the ordering must correspond to a progressive increase in the 'size' of forms, understood as the number of 1s in their codes. The sequence starts from forms whose codes contain just two 1s (the minimum) and ends with the complete archetype, whose code, uniquely, consists entirely of 1s. But there is further ordering within this larger sequence. I mentioned earlier that the positions of 0s and 1s in any binary code are always meaningful and convey information about the corresponding form. This can be demonstrated by reference to the 14-digit strings which encode single-court forms derived from a 7 x 7 array of cuboids (refer to Figure 10). It is clear, for example, that digits in the fourth and eleventh places represent the court. If both these digits are 1s, the court is present.

1 ***1*** The court is included

If one or the other (or both) of these digits is 0, the court is absent. Similarly, 1s or 0s in the second, sixth, ninth and thirteenth places represent the presence or absence of artificially-lit strips.

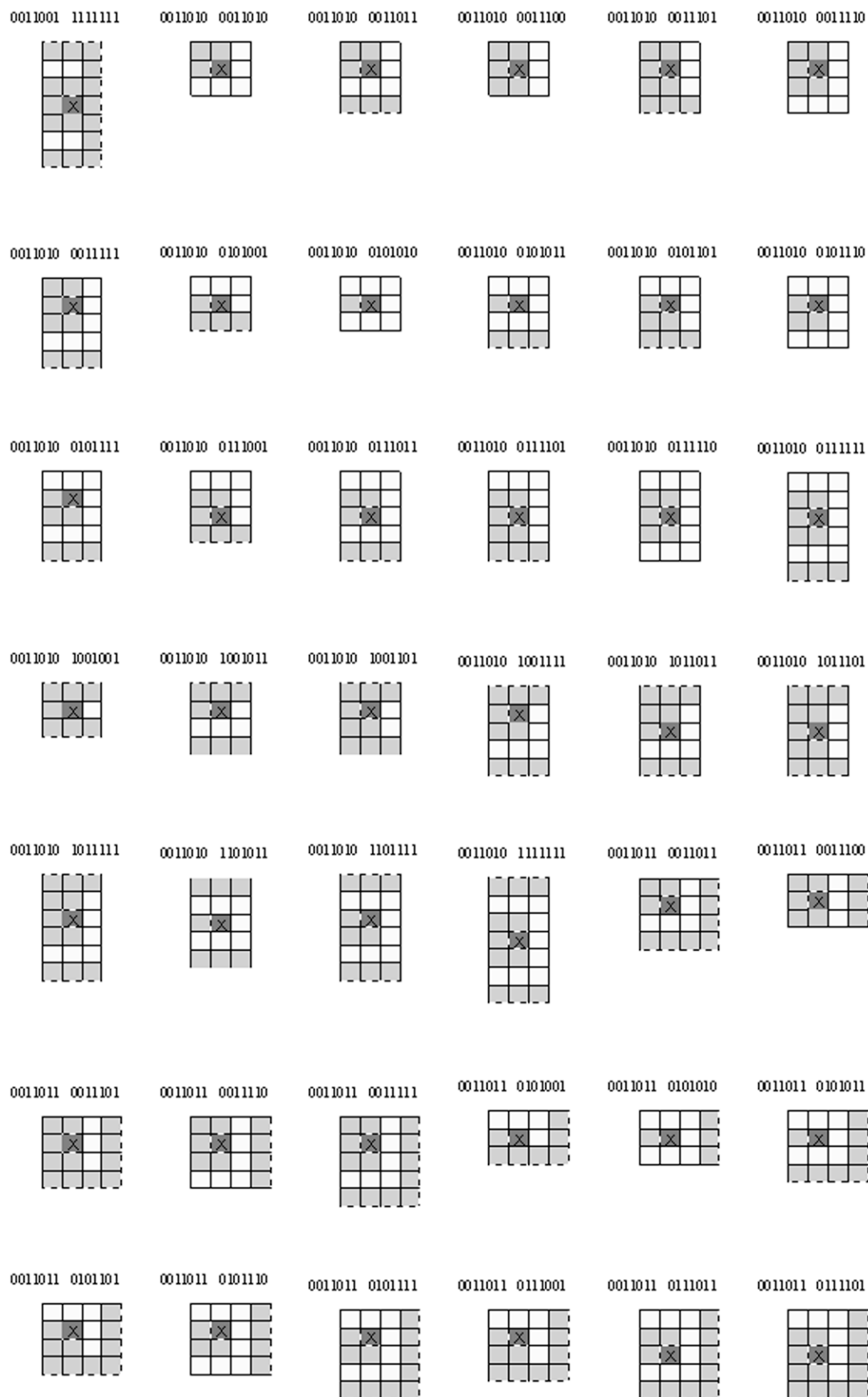
*1***1* *1***1* Artificially-lit strips

1s in any other positions represent sidelit strips. If digits in either the first, seventh, eighth or fourteenth places are 0s, then the respective external facades are 'blind'. And so on.

0*****0 0*****0 'Blind' facades

It follows that codes with certain *patterns* of 0s and 1s must correspond to built forms with specific *shapes*. For example L-shapes are given by codes of the general form 0001*** 0001***, to be understood as meaning that in each sub-string 0001*** there are 1s in one or more of the positions marked by *. This can be seen in Figure 13, which shows how the 1s in the two sub-strings select the court, and the groups of 000s remove one large L-shape at top left, to leave a smaller L at lower right. Any combination of 1s in the two groups of ***s must now select some or all of this second remaining L-shape. Other shapes are given by codes of general forms as follows:

Figure 11 (right): Sample page from a catalogue of all 675 distinct configurations derivable from the 7 x 7 array, with their binary codes. Courts are marked X; the light tone indicates artificially-lit cuboids; and the darker tone indicates daylight cuboids. Daylit facades are shown in broken line and 'blind' facades in solid line



09.7

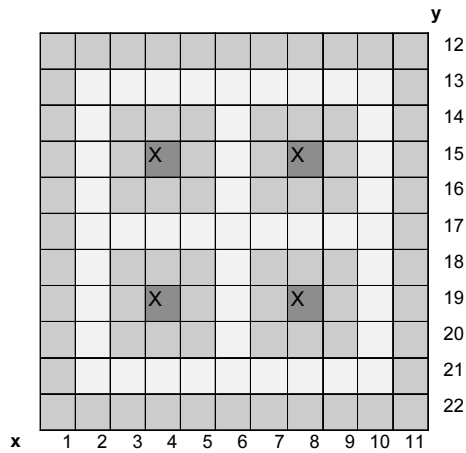


Figure 12: 11 x 11 array of cuboids, encoded with a 22-digit binary string

```
0001000 ***1***
0001*** ***1***
***1*** ***1***
```

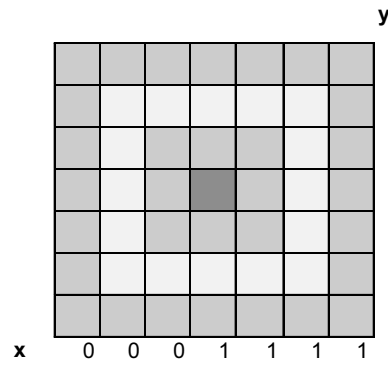
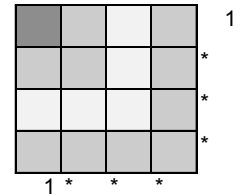


Figure 13: L-shapes created in the 7 x 7 array by binary codes of the form 0001* where *** means that at least one 1 is present in any of the three positions**

'Broken Is'
Us
Os



09.8

The 'Broken Is' are forms with what might be called 'degenerate' courts, bordered by built space either on two opposite sides (so that the built form is broken) or on one side only. The overall shape, including the court, is rectangular in both cases. Such forms approximate the plans of certain kinds of courtyard houses on narrow sites. With a four-court archetype, whose codes are 22 digits long, suitable patterns of 0s and 1s will generate sixteen different possible plan shapes, resembling letters of the alphabet or combinations of letters (Fig. 14).

Examples of any given shape are not scattered randomly throughout the catalogue of all codes, but are found clustered together. Appendix A lists all 675 codes for the one-court 7 x 7 archetype, prepared by Waddoups. The plan shapes are indicated in every case. (SB signifies 'Simple Block' forms and BI marks the 'Broken Is'.) In general the simple rectangular shapes are found at the start of the catalogue, then the Ls and the Us, and finally the Os. It will be clear that, should plans of a given shape be required, it would be possible to direct a search selectively towards the appropriate area of the catalogue.

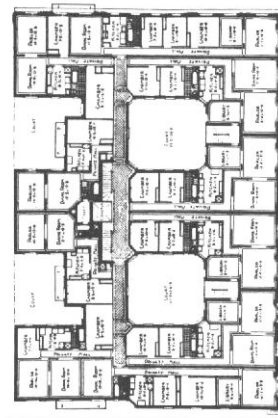
There is another important geometrical property of built forms, again signalled by their codes. This is the property of *bilateral symmetry* in shapes, whether about axes in x or y , or both, or about a diagonal axis (Figure 15). If either the x or the y sub-string in a code is palindromic (reading the same backwards as forwards) then the corresponding built form shows bilateral symmetry about a single axis. If both sub-strings are palindromes, the form has bilateral symmetry about two perpendicular axes. Should both sub-strings be the *same* (but not necessarily palindromes) then the form has bilateral symmetry about a diagonal axis. (It should be emphasised that these are symmetries in the undimensioned configurations. The symmetries could be destroyed by the assignment of unequal dimensions. One should perhaps speak, rather, of *potential symmetries*.)

The purpose of a catalogue of this kind, obviously, is to make exhaustive searches for forms with desired combinations of characteristics. There might be applications in the early strategic stages of design, in building science, and in architectural history. I will describe briefly a few illustrative examples. Waddoups and I are planning to develop fast computer methods for searching catalogues with codes corresponding to nine-court archetypes or even bigger.

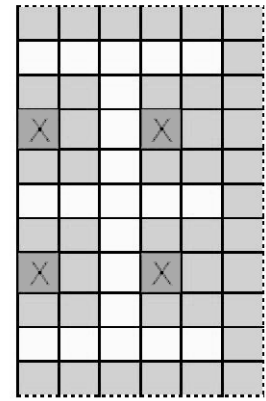
$IX = \begin{array}{|c|} \hline \vdots \\ \hline \end{array}$ for X = all other shape generators
 $LL = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$
 $LU = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$ $UU = \begin{array}{|c|} \hline \square \\ \hline \end{array}$
 $LT = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$ $UT = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$
 $LF = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$ $UF = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$
 $LE = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$ $UE = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$

 $TT = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$ $FF = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$
 $TF = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$ $FE = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$
 $TE = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$ $EE = \begin{array}{|c|} \hline \text{┐} \\ \hline \end{array}$

a)



b)



For the present however we are limited to making searches by hand, or using spreadsheets, and so our examples here make use only of the catalogue of 675 forms derived from the one-court, 7 x 7 array, reproduced in Appendix A

Let us look first at some of the potential of this approach in building science. The fact that the catalogue of built forms is complete and comprehensive - within the terms of the archetypal representation - means that geometrical properties relevant to various aspects of physical performance can be compared across the entire population. Such properties can include for example the ratio of total floorspace to site area ('floor space index'); the ratio of external surface to volume, with obvious relevance to heat loss and energy use; or the total length of horizontal circulation in relation to floor area (giving some measure of 'circulation efficiency'). Such comparisons of metric properties depend, naturally, on assigning dimensions to all parameters of the archetype. In principle, these parameters might take *any* values. In practice, for the results to be 'building-like', the values would have to be restricted within realistic ranges. Thus storey heights in real buildings would never be much lower than 2m; there would be an effective limit on the depth in plan of sidelit zones - depending on the storey height and glazing ratio - of something like 7m; and so on. Nevertheless, within such ranges there could legitimately be a wide variation found in practice. To illustrate the kinds of experiments that are possible, we have therefore chosen a set of *default* values for these dimensional parameters, corresponding to norms found empirically in samples of actual buildings.

In the results given below, storey heights are taken to be 2.5m and sidelit strips are taken to be 7m deep. It is assumed that the central strips of artificially-lit space serve as corridors only, and are made 2m wide throughout. There remains the question of the widths of courtyards in plan. In practice, if these are to provide daylight, they are likely to be made wider, the greater the height of the building. In other studies of built form, and in legislation governing the bulk of buildings, this relationship of height to width in courtyards has often been represented in terms of a 'cut-off angle'. This is the angle that a line, joining the top of one facade of the court to the bottom of the opposing facade, makes with the ground. The cut-off angle can be varied experimentally and the consequences for other metric properties calculated. The graphs in Figure 16 plot values for a series of geometrical properties, calculated in a spreadsheet, for all L-shaped forms derivable from the 7 x 7 archetype. In each case, the 28 built forms are arranged along the *x* axis in order of their binary codes. The forms are all

Figure 14: Possible plan shapes generated from the 11 x 11, four-court array, denoted by letter pairs

09.9

Figure 15: a) Crescent Court, New York 1905 (from Holl 1980) and b) its representation in the 11 x 11 array by the code 0001110110111111111111. The fact that the *y* sub-string is palindromic indicates bilateral symmetry about an axis in *x*. The fact that the *x* sub-string is not palindromic indicates the lack of bilateral symmetry about any axis in *y*.

09.10

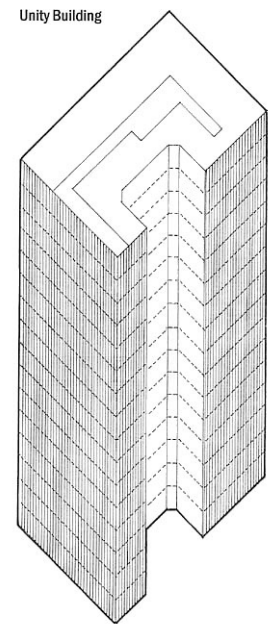
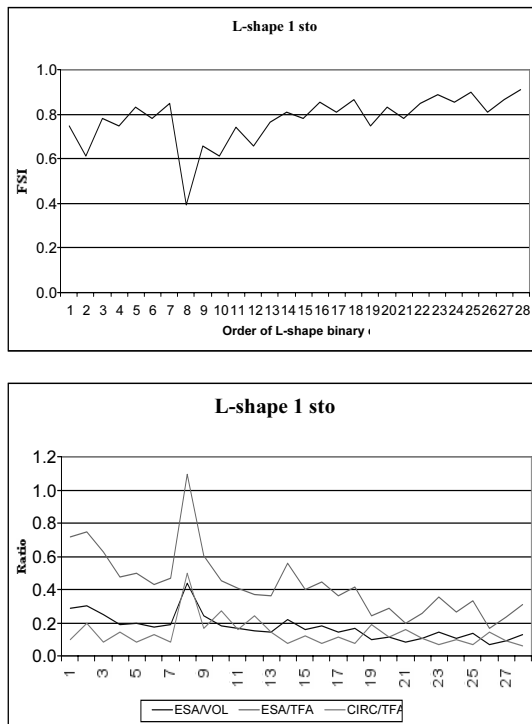


Figure 16: Variation in floorspace index (left); and variation in external surface area/ volume, external surface area/ floor area and circulation length/ floor area (right), for 28 single-storey L-shaped built forms, arranged in ascending order of their binary codes

Figure 17: Axonometric view of Unity Building, Chicago (Clinton J. Warren, 1891-92)

single-storey. The properties in question are floor space index (fsi), surface area/ volume, surface area/ floor area and total circulation length/ floor area. In calculating floor space index, the site area is assumed to be the area of the minimum rectangle within which the building footprint can be contained. The area of surface is taken to include all external walls, plus the roof (assumed to be flat).

The calculation of total circulation length is slightly more complicated. Where sidelit space is immediately adjacent to artificially-lit space, the latter is assumed to provide the circulation (in the form of corridors), and the lengths of the relevant strips are totalled. However in many configurations there is daylight space which is *not* adjacent to artificially-lit strips. Here the circulation is assumed, in effect, to be provided within the sidelit zones themselves, and to run along the interior edges of these zones, away from the window walls. (All vertical circulation is ignored.) This is obviously a rather rough-and-ready method of calculation; and the circulation in real buildings with the same plan shapes might well take different routes. It does nevertheless provide some indication of the extent to which the circulation system is likely to be straggling, or compact.

Figure 16 is intended simply to provide illustrative examples of the systematic variation of such geometrical indicators of performance. There is a general, slow increase in the value of floorspace index for successive L-shape binary codes. This is because those codes contain increasing numbers of 1s, and so the forms become 'fuller' and more complete, culminating in 0001111 0001111 which provides the maximum floorspace possible in an L-shaped plan. The extreme low value for fsi (the eighth code in sequence, 0001010 0001010) is for a somewhat pathological plan consisting *only* of an artificially-lit corridor along two sides of the court. One might be tempted to conclude that, if fsi is to be maximised, then the 'fullest' possible form should always be selected. However this is not necessarily the optimal plan, in fsi terms, on sites with shared boundaries and with restricted dimensions.

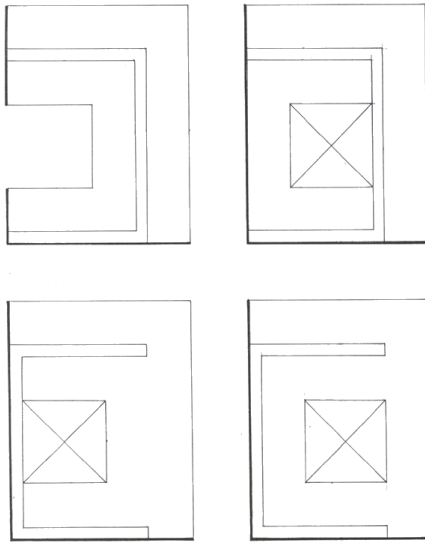
These fsi calculations reproduce, in effect, the original work of March and Trace (1968) on 'the land use performances of selected arrays of built forms'. They extend that research, by covering all plan shapes (within the archetypal representation) where March considered just built forms with square, rectangular and cruciform shaped plans - the last-mentioned joining together to form continuous systems of square courts. (On the other hand, no allowance is made here at present, as in March's work, for land area around the edges of forms with sidelit space on their external facades. This can and should be remedied.)

The graphs for the exposed surface area ratios and for circulation length/ floor area all follow a pattern that is inverted in relation to that for fsi. All these measures decline, generally, as the sequence of codes progresses, with strong peaks at the 8th position for code 0001010 0001010. There is a certain periodicity within the overall trend. In the catalogue, L-shapes are found in seven separate groups containing respectively 7, 6, 5, 4, 3, 2 and 1 codes. This means that, in the graphs in the figure, new sequences of codes start in the 8th, 14th, 19th, 23rd, 26th and 28th places - a fact that is visible, to an extent, in the changes in value of the various measures.

The default dimensional values used here have been based loosely on typical norms for modern daylit office buildings. It would be possible to base metric values, more systematically, on empirical evidence from surveys of the existing building stock or samples of historical buildings. A case in point is the survey of non-domestic buildings of all types made at 3500 addresses in four English towns in 1989 and 1992 (Brown et al 2000). These data were all entered to a geographical information system (Holtier et al 2000), and automatic measurements made of floor areas, storey heights, external wall areas and depths in plan. The results provide ranges of values over which these measurements are observed to vary, for different built form types, in typical English building practice of the 19th and 20th centuries.

In the analyses of Figure 16 the catalogue of binary codes has been used to plot variation in performance across the entire range of allowable forms having the same general shape. My concluding example shows the catalogue being used in a different way. Now specific values for some set of metrical and shape properties are set in advance, and the catalogue is searched to discover whether any form or forms exist which meet those specifications. This kind of search begins then to approximate something like a simple - and highly constrained - design process. In this illustration, Waddoups and I have chosen however to investigate not a contemporary problem, but the historical design process of a particular office skyscraper, the Unity Building, erected in Chicago in 1891-92. The U-shaped plan of the Unity Building, like many in the city, was constrained by the typical size of block in the 'Loop', the central business district of Chicago (Willis 1993). It was set on a corner site with dimensions 24.5 x 36.5m (Figure 17). The practical challenge facing the architect Clinton J. Warren was to pack as much office accommodation onto this site as possible, under a series of constraints.

The maximum height of the building was limited, both by technological constraints and by a legal limit of 130 feet (39.6m) established by the Chicago city council in 1893. Typical storey heights for such buildings at this time were around 4m, to allow for natural ventilation and daylight (supplemented by gas lighting). It was only in the late 1920s and early '30s, with the introduction of air conditioning and fluorescent lighting, that the ceilings in office buildings could be lowered to today's norms of 2.5 or 3m. The need for daylight also constrained the depths of offices in plan. The rule of thumb quoted in the 19th century literature was that an effective limit existed at about 20 to 25 feet (6 to 7.5 metres) from the windows (see for example Hill 1893). Beyond this distance, it was believed that office space would be too



09.12

Figure 18: Four plans selected from the catalogue of 14-digit binary codes, meeting a set of configurational and metric constraints approximating those applying to the Unity Building. The plan at top left, 0001111 1111110, is very close to that of the actual Unity Building. The other plans are variants with the same overall dimensions.

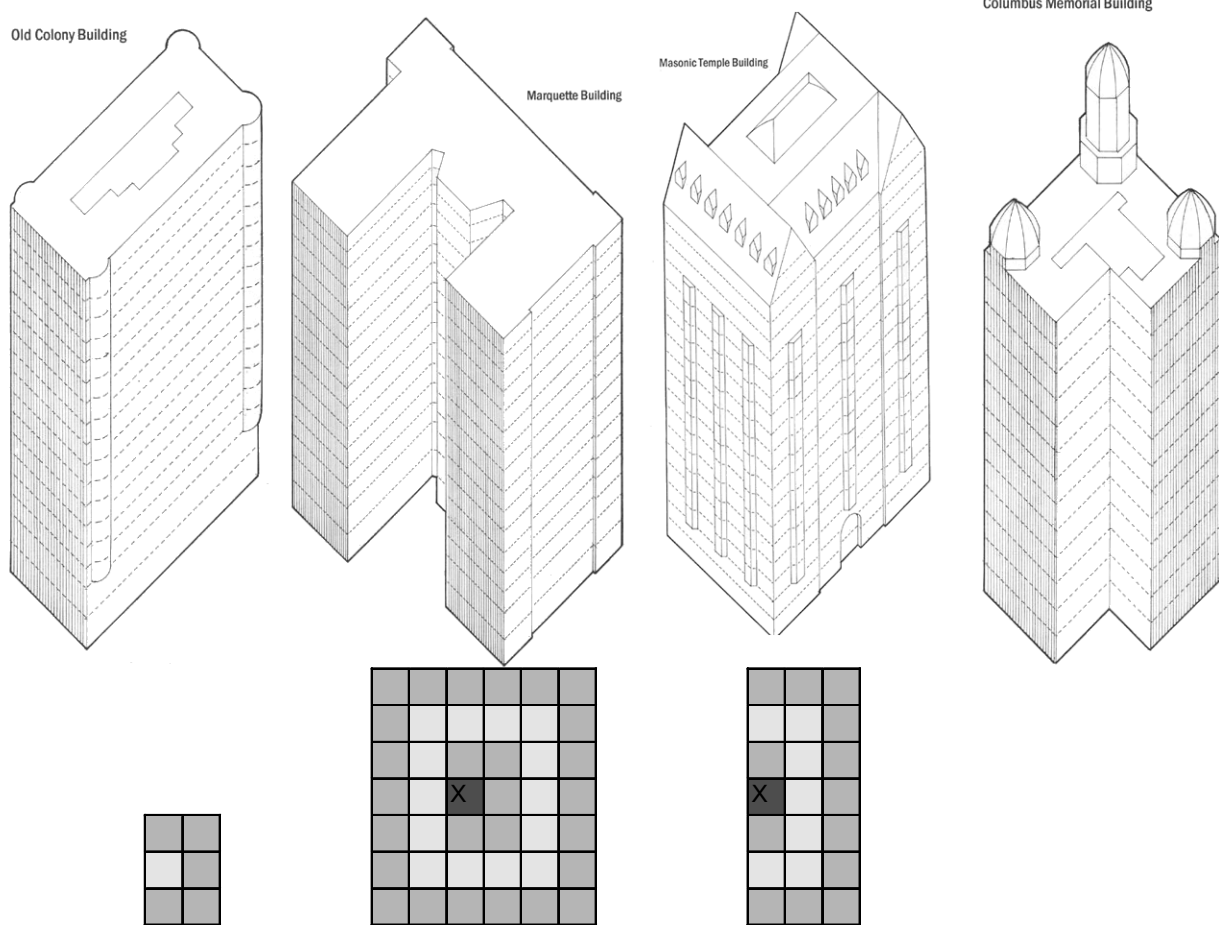
The fact that the building is to fit on a corner site, and that two adjacent facades are to be daylit, the other two ‘blind’, means that only codes of the general form 0*****1 0*****1 need to be inspected. The result is that just four dimensioned configurations are found which meet the requirements, as shown with their

corresponding binary codes in Figure 18. One of these (top left, code 0001111 1111110) approximates closely Clinton J Warren’s actual U-shaped design, suggesting that he was indeed successful in finding an ‘optimal’ arrangement in these terms. (The plan here is mirrored in relation to the real Unity Building, because this is the isomorph with the lowest binary code.) The other three options have the same overall plan dimensions and floor area, but the positions of court and daylit strips are rearranged. The (artificially-lit) corridors do not however reach all office areas in these plans. Only in the plan at top left is the circulation arranged satisfactorily and efficiently.

It will be clear that a similar approach could be applied to other buildings of this period, occupying sites of different shapes and sizes. Figure 19 illustrates plans of a selection of late 19th and early 20th century Chicago skyscrapers, with their equivalent undimensioned configurations in the 7 x 7 array. Many such buildings had shops or banks on the ground floor and sometimes also the first floor, filling the entire site. Either the shops were artificially lit at the centre of the plan, or else a central concourse or banking hall, under a lightwell, was toplit via a glass roof. These lower floors thus resembled the lower deep-plan storeys of the archetypal form. The upper floors were wholly devoted to daylit offices however, as in the Unity Building. Some contemporary Chicago hotels took similar forms. The plans of other larger office blocks and hotels (not illustrated) can be approximated by configurations derived from two-court or four-court arrays. At a more strategic level, it is possible to imagine a study of the *generic* constraints that the size of the standard Chicago block placed on the overall variety of possible plan shapes for these buildings. The results might be compared say with the variety of plans resulting from the constraints of the typical Manhattan block, which is much more elongated.

poorly lit to be lettable - although one does find that in practice some buildings were actually constructed with office ‘strips’ deeper than 7.5 metres, including some with rows of wholly internal rooms allocated to secretaries. The typical cut-off angle for interior light and ventilation courts in Chicago office towers seems to have been about 70° (although this was not controlled by law).

For our sample search, we have therefore set default values for dimensional parameters as follows: storey heights 4m, maximum number of storeys 10, width in plan of daylit strips 7.5m, width of artificially-lit corridors 2.5m, and cut-off angle in courts 70°. Some simple arithmetic is sufficient to show that, given the site dimensions, no built forms with two courts (or more) can be accommodated. It is therefore sufficient to consider just one-court forms. Within the given metric constraints, the catalogue (as in Appendix A) is then searched for forms in which total floor area is maximised.



I have already mentioned how Waddoups and I are planning to extend this work by creating catalogues of built forms with up to nine courts, and by developing automated methods of searching for forms with required combinations of shape, symmetry and metric properties. We plan to display the results not just as codes and statistics, but also as simple 3D models. I may perhaps have given the impression with the Chicago examples that the technique, and method of representation, are intended to be applied predominantly to office buildings in

central urban positions. But I see applications to several other building types, with particular promise in relation to residential buildings including houses, public houses (as in 'Every built form has a number'), apartments and hotels. Nor do I see that all search processes have necessarily to be quite so narrowly focussed, or as driven by the optimisation of a single objective function, as in the example of the Unity Building. A tool of this kind could be used in a much more exploratory way, for comparing many forms visually as well as quantitatively.

This said, it is certainly true that the archetypal representation lends itself, preferentially, to the description of built forms on tight urban sites, where the constraints of daylighting are at their most demanding. These are, as one might say, the 'fullest' and most 'economically-packed' of all forms of building. In general however it is possible to take any built form derived directly from the archetype, and to remove certain parts of it, producing a smaller form (not derivable from the archetype) that - so long as structural stability is not prejudiced

Figure 19:
Axonometric
views of 19th
century Chicago
skyscrapers:
(from left to
right) the Old
Colony Building,
the Marquette
Building, the
Masonic Temple
Building and the
Columbus
Memorial Build-
ing. Representa-
tions of their
upper floor plans
in the 7 x 7
archetype are
also shown. The
respective binary
codes are:
Old Colony,
0000011 1000101;
Masonic Temple,
1101111 1111111;
Marquette,
0001011 1111111;
Columbus
Memorial,
0001111 0001011

- is still quite workable as a building design. In the bedroom floor of the Roosevelt Hotel illustrated in Figure 3 for example, the daylight strips are missing from two sides of one of the four open-ended courts. In other building types, on large sites, parts of the daylight strips around a central zone of artificially-lit space may become 'unwrapped' and stretched out into extended wings.

Such forms are impossible to represent in full detail by means of the archetype as it stands. One of its intrinsic limitations is that it treats entire *rows* of cuboids. When a row is selected, then by implication *all* cuboids in the row are present, and all share the same dimension of width. This is the reason that the binary codes are more compact than codes that follow March's convention. But the price is that it is not possible to signal the fact that individual cuboids are missing, or are dimensioned differently so as to recede or protrude from the remainder of the facade or the roof surface. There may nevertheless be ways of coping with such difficulties. In some instances the 'negative form' missing from a 'full' built form can itself be directly represented within the archetype. Waddoups and I are investigating whether it is possible to develop rules describing a kind of 'subtraction' of forms derived from the archetype, one from another.

There are other weaknesses of the archetypal representation. It is not possible for example to derive forms with L, U or comparable plan shapes, with daylight strips across the ends of the arms. These end facades are always blind. Some more complex plan shapes - although consisting of rectangular configurations of daylight strips, artificially-lit strips and courts - cannot be generated. For example E-shapes are possible, but not S-shapes. An S-shape would have to be produced from a pair of U-shapes, one of them mirrored. Waddoups is developing an algebra of possible such combinations of forms, set side-by-side. Finally, all our examples in this paper have been confined either to single-storey cases, or to prismatic forms with identical plans on every floor level. The archetype itself allows for two shapes of floor plan (in the deep-plan and courtyard levels) to be superimposed. But we have not yet begun to examine the issues raised by built forms where the rows of cuboids selected are different on more than two successive levels - as often occurs in real buildings.

References

- Brown F E, Rickaby P R, Bruhns H R, Steadman J P (2000) 'Surveys of nondomestic buildings in four English towns', *Environment and Planning B: Planning and Design* **27**, pp.11-24
- Hill G (1893) 'Some practical conditions in the design of the modern office building', *Architectural Record* **2**, pp.222-68
- Holtier S, Steadman J P, Smith M G (2000) 'Three-dimensional representation of urban built form in a GIS', *Environment and Planning B: Planning and Design* **27**, pp.51-72
- March L (1972) 'Elementary models of built forms' in L Martin and L March (eds), *Urban Space and Structures*, Cambridge University Press, Cambridge pp.55-96
- March L (1976) 'A Boolean description of a class of built forms' in L March (ed) *The Architecture of Form*, Cambridge University Press, Cambridge pp.41-73
- March L and Trace M (1968) *The Land Use Performances of Selected Arrays of Built Forms*, Land Use and Built Form Studies Working Paper No.2, Cambridge University Department of Architecture
- Martin L and March L (1966) 'Land use and built forms', *Cambridge Research*, April
- Steadman P (1998) 'Sketch for an archetypal building', *Environment and Planning B: Planning and Design*, 25th Anniversary Issue, pp.92-105
- Steadman P and Waddoups L (2000) 'A catalogue of built forms, using a binary representation', *Proceedings*, 5th International Conference on Design and Decision Support Systems in Architecture, Nijkerk, The Netherlands pp.353-373
- Willis C (1993) 'Light, height and site: the skyscraper in Chicago' in J Zukowsky (ed) *Chicago Architecture and Design, 1923-1993: Reconfiguration of a Metropolis*, Prestel, Munich and Art Institute of Chicago, Chicago, pp.119-39

Catalogue of all 675 binary codes for the one-court 7 x 7 archetype, prepared by Waddoups. The plan shapes are indicated in every case. (SB = 'Simple Block' forms, BI = 'Broken I' forms.)

SB-shapes	1	0000001	0000001	57	0001001	0011111	L-shapes	113	0001011	0001011	169	0001100	1001011	
	2	0000001	0000010	58	0001001	0101001		114	0001011	0001100	170	0001100	1001101	
	3	0000001	1000001	59	0001001	0101010		115	0001011	0001101	171	0001100	1001111	
	4	0000010	0000011	60	0001001	0101011		116	0001011	0001110	172	0001100	1011011	
	5	0000010	1000001	61	0001001	0101101		117	0001011	0001111	173	0001100	1011101	
	BI-shapes	6	0000010	1000011	62	0001001	0101110	U-shapes	118	0001011	0011001	174	0001100	1011111
		7	0000011	0000011	63	0001001	0101111		119	0001011	0011010	175	0001100	1011011
		8	0000011	1000011	64	0001001	0111001		120	0001011	0011011	176	0001100	1011111
		9	0001000	0001001	65	0001001	0111011		121	0001011	0011100	177	0001100	1111111
		10	0001000	0001010	66	0001001	0111101		122	0001011	0011101	L-shapes	178	0001101
11		0001000	0001011	67	0001001	0111110	123	0001011	0011110	179	0001101		0001110	
12		0001000	0001100	68	0001001	0111111	124	0001011	0011111	180	0001101		0001111	
13		0001000	0001101	69	0001001	1001001	125	0001011	0101001	U-shapes	181		0001101	0011001
14		0001000	0001110	70	0001001	1001011	126	0001011	0101010		182		0001101	0011010
15		0001000	0001111	71	0001001	1001101	127	0001011	0101011		183	0001101	0011011	
16	0001000	0011001	72	0001001	1001111	128	0001011	0101101	184		0001101	0011100		
17	0001000	0011010	73	0001001	1011011	129	0001011	0101110	185		0001101	0011101		
L-shapes	18	0001000	0011011	74	0001001	1011101	130	0001011	0101111	186	0001101	0011110		
	19	0001000	0011100	75	0001001	1011111	131	0001011	0111001	187	0001101	0011111		
	20	0001000	0011101	76	0001001	1101011	132	0001011	0111011	188	0001101	0101001		
	21	0001000	0011110	77	0001001	1101111	133	0001011	0111101	189	0001101	0101010		
	22	0001000	0011111	78	0001001	1111111	134	0001011	0111110	190	0001101	0101011		
	23	0001000	0101001	79	0001010	0001010	135	0001011	0111111	191	0001101	0101101		
	24	0001000	0101010	80	0001010	0001011	136	0001011	1001001	192	0001101	0011110		
	25	0001000	0101011	81	0001010	0001100	137	0001011	1001011	193	0001101	0101111		
	26	0001000	0101101	82	0001010	0001101	138	0001011	1001101	194	0001101	0111001		
	27	0001000	0101110	83	0001010	0001110	139	0001011	1001111	195	0001101	0111011		
U-shapes	28	0001000	0101111	84	0001010	0001111	140	0001011	1011011	196	0001101	0111101		
	29	0001000	0111001	85	0001010	0011001	141	0001011	1011101	197	0001101	0111110		
	30	0001000	0111011	86	0001010	0011010	142	0001011	1011111	198	0001101	0111111		
	31	0001000	0111101	87	0001010	0011011	143	0001011	1101011	199	0001101	1001001		
	32	0001000	0111110	88	0001010	0011100	144	0001011	1101111	200	0001101	1001011		
	33	0001000	0111111	89	0001010	0011101	145	0001011	1111111	201	0001101	1001101		
	34	0001000	1001001	90	0001010	0011110	L-shapes	146	0001100	0001100	202	0001101	1001111	
	35	0001000	1001011	91	0001010	0011111		147	0001100	0001101	203	0001101	1011011	
	36	0001000	1001101	92	0001010	0101001		148	0001100	0001110	204	0001101	1011011	
	37	0001000	1001111	93	0001010	0101010		149	0001100	0001111	205	0001101	1011111	
38	0001000	1010101	94	0001010	0101011	U-shapes		150	0001100	0011001	206	0001101	1101011	
39	0001000	1011101	95	0001010	0101101		151	0001100	0011010	207	0001101	1011111		
40	0001000	1011111	96	0001010	0101110		152	0001100	0011011	208	0001101	1111111		
41	0001000	1101011	97	0001010	0101111		153	0001100	0011100	L-shapes	209	0001110	0001110	
42	0001000	1101111	98	0001010	0111001		154	0001100	0011101		210	0001110	0001111	
43	0001000	1111111	99	0001010	0111011	155	0001100	0011110	U-shapes		211	0001110	0011001	
44	0001001	0001001	100	0001010	0111101	156	0001100	0011111			212	0001110	0011010	
45	0001001	0001010	101	0001010	0111110	157	0001100	0101001			213	0001110	0011011	
46	0001001	0001011	102	0001010	0111111	158	0001100	0101010		214	0001110	0011100		
47	0001001	0001100	103	0001010	1001001	159	0001100	0101011		215	0001110	0011101		
U-shapes	48	0001001	0001101	104	0001010	1001011	160	0001100	0101101	216	0001110	0011110		
	49	0001001	0001110	105	0001010	1001101	161	0001100	0101110	217	0001110	0011111		
	50	0001001	0001111	106	0001010	1001111	162	0001100	0101111	218	0001110	0101001		
	51	0001001	0010101	107	0001010	1011011	163	0001100	0111001	219	0001110	0101010		
	52	0001001	0011010	108	0001010	1011101	164	0001100	0111011	220	0001110	0101011		
	53	0001001	0011011	109	0001010	1011111	165	0001100	0111101	221	0001110	0101010		
	54	0001001	0011100	110	0001010	1101011	166	0001100	0111110	222	0001110	0101110		
	55	0001001	0011101	111	0001010	1101111	167	0001100	0111111	223	0001110	0101111		
	56	0001001	0011110	112	0001010	1111111	168	0001100	1001001	224	0001110	0111001		

Continued on next page

[illegible]