A second scenario in which the awareness of a single number is crucial is a surgical operation and particularly a major one lasting many hours. Any change in vital human signs such as blood pressure is a cause for concern, especially for the anaesthetist (as well as the patient), so it is essential that any change in such a vital sign be noticed immediately. The anaesthetist is, of course, presented with comprehensive visual information in the manner illustrated in Figure 3.10, but the effort of paying careful attention to such a presentation for many hours would be considerable. The simple solution is to encode vital signs in sound (Watson *et al.*, 1999; Watson and Sanderson, 2004): whatever the anaesthetist is attending to (perhaps on the telephone to respond to an urgent enquiry from a ward), he or she will immediately be aware of a change in the pitch or repetition rate of a constant sequence of 'beeps'. Another reason for introducing this example is that it emphasizes the fact that data can be encoded in ways other than visual and that the 'visual' in visualization should not be misinterpreted.

3

4

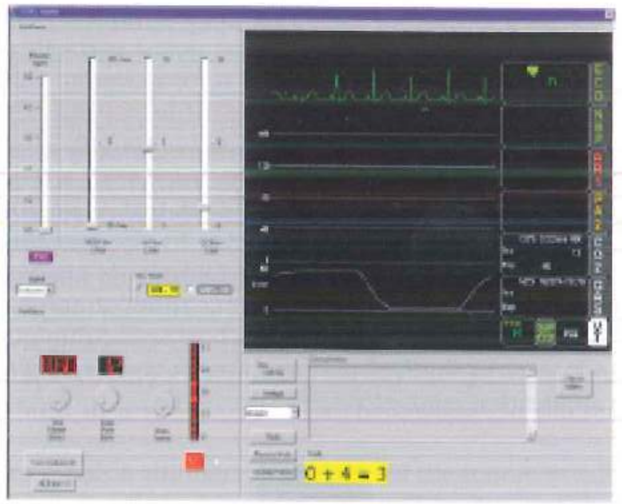


FIGURE 3.10
Representations of the vital signs of a patient during an operation. The difficulty of paying constant attention to such a display throughout a long operation has led to the encoding of vital signs in the pitch of a frequently repeated 'beep'. A change in pitch is immediately apparent wherever the gaze of the anaesthetist is directed
(Image by kind permission of Marcus Watson)

A collection of numbers

Notwithstanding the importance of representing a single number, as in an altimeter, a more common situation is one in which univariate data about a *number* of objects is of interest. Not surprisingly, well-established techniques are available (Cleveland, 1994; Tufte, 1983), though the field of information visualization is such that new ones continue to be invented.

Price data for a number of cars can, for example, be represented as dots on a linear scale (Figure 3.11). But how effective is this representation? A quick overview shows the general distribution of car prices; a more detailed examination will estimate the average price accurately enough to judge 'affordability'; further examination will disclose maximum and minimum prices. A valid question, therefore, for any representation is, 'How will the representation be used?' Many of the *initial* questions in the mind of a car buyer would be answered by a Tukey Box Plot of the same data (Figure 3.12). The box contains the *median* price (the middle line, which divides the number of data items into halves). The two ends of

the box represent the 25 percentile (below which one quarter of the data items are to be found) and the 75 percentile. The 5 and 95 percentiles are indicated by the horizontal bars and the outliers are retained as points. Here, much of the data is being *aggregated*, to reflect the fact that precise detail is often not needed; we are now representing *derived values*.

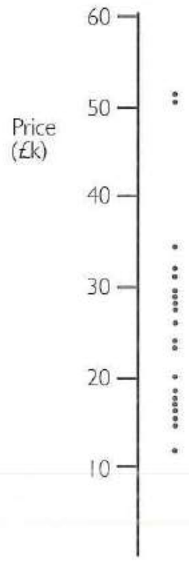


FIGURE 3.11
Each dot represents the price of a car

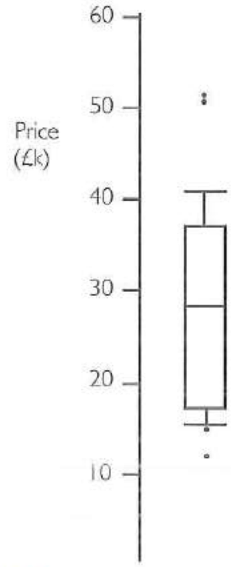


FIGURE 3.12
A Tukey Box Plot of the data represented in Figure 3.11

Another familiar representation of univariate data is the *histogram* (Figure 3.13). Much has been written about the need for appropriate choice of bin sizes and will not be repeated here. The essential point to note is that we are again representing *aggregate* properties – or derived values – of the data in a manner that can support both 'at a glance' awareness and the need for more precise understanding. The histogram is only one of many useful representations of numerical data that are more concerned with derived values. For example, if we 'push over' the columns of the histogram and join them together we obtain the *bargram* (Figure 3.14), already familiar from Chapter 2. While the relative count in a bin is now reflected in the bin's width, certain characteristics of the data are lost or not obvious at first sight – for example, the existence of outliers and the emptiness of a particular range. This omission may be immaterial in certain applications.

Values need not be numerical – they can be *categorical* or *ordinal*. The makes of cars can include Ford, Nissan, Toyota, Ferrari, etc., so that a bargram appearing in an online car sales display like the EZChooser (Wittenburg *et al.*, 2001) might contain this univariate categorical data as shown in Figure 3.15. Ordinal data can be found in the sales volume of a shop for the various ordered days of the week (Figure 3.16).

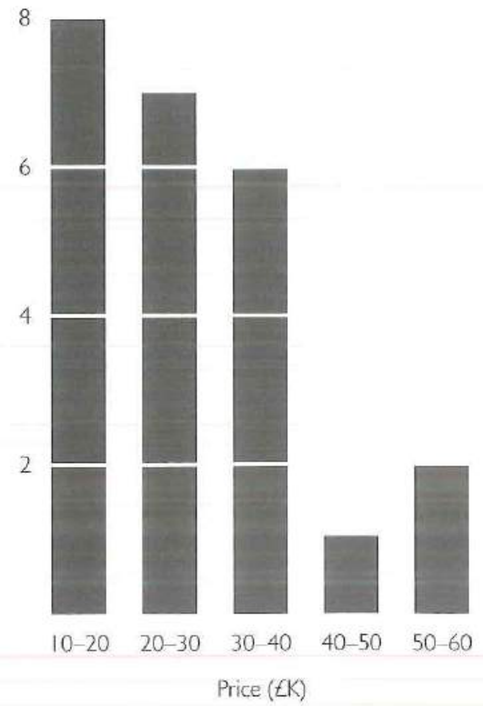
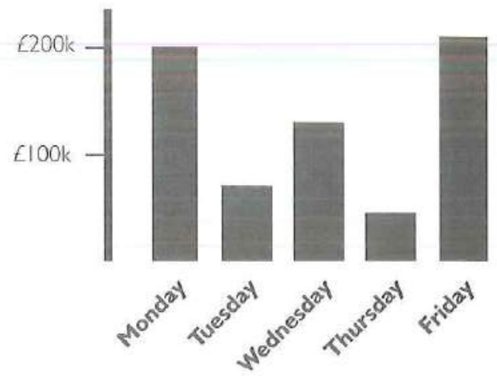


FIGURE 3.13 A histogram of univariate data

FIGURE 3.14 A bargram representation of univariate data, obtained by 'tipping over' the columns (bars) of a histogram and joining them end to end, ignoring any null bins

FIGURE 3.15
A bargram representation of univariate categorical data

FIGURE 3.16
A histogram of univariate ordinal data.



3.1.2 Bivariate data

A conventional approach to the representation of bivariate data is the *scatterplot*. For a collection of houses, all characterized by a price and the number of bedrooms, each house is represented by a point in two-dimensional space with

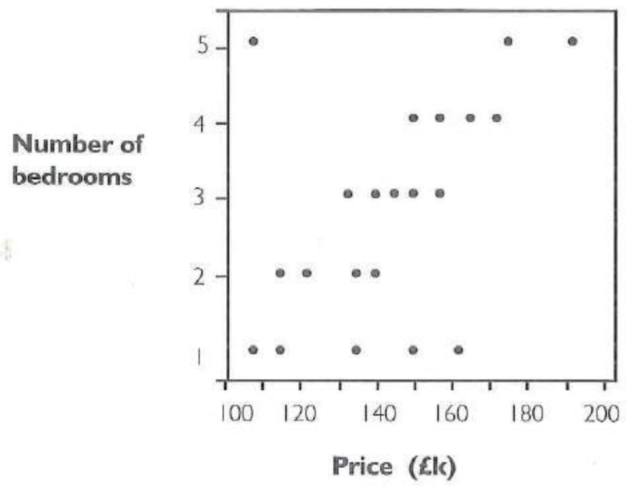


FIGURE 3.17
A scatterplot of bivariate data. Each point indicates the price and number of bedrooms associated with a house

axes associated with these two attributes (Figure 3.17). This representation affords an awareness of a general trend (more money, more bedrooms), of local trade-offs (less money, more bedrooms) and of outliers that may be interesting and might not have been anticipated (a five-bed house for £110,000). Any conventional query system requiring a precise formulation of housing needs would not encourage the specification of such an unanticipated result and is one of the many reasons why information visualization can be so valuable.

A special case of the scatterplot is a time series, in which one axis represents time and the other some function of time. The importance of time-varying data, for example in medical and climate studies, is such that many representation techniques as well as visualization tools have been developed to allow understanding to be derived from a time-series plot. The performance of one time-series query tool is illustrated in Figure 3.18 in the context of a data set containing 52 weekly stock prices for 1,430 stocks (Hochheiser and Shneiderman, 2004). The graph overview of Figure 3.18(a) shows the entire data set, providing some idea of density and distributions. In Figure 3.18(b) a single timebox limits the display to those items with prices between \$70 and \$250 during days 1 to 4. Subsequent queries add additional constraints, selecting items that have prices between \$70 and \$95 during days 7 to 12 (Figure 3.18(c)) and for prices between \$90 and \$115 for days 15 to 18 (Figure 3.18(d)).

An alternative representation of a time series (Figure 3.19), perhaps more suited for gaining an initial impression of data, is illustrated by the level of ozone concentration above Los Angeles, each square associated with one day and coloured to indicate the ozone level. The figure, which represents ten years of data, shows that ozone levels are higher in the summer months and that concentrations during those months have decreased with time.

Perhaps unexpectedly, insight into bivariate data can benefit from an apparent separation of the two attributes. For example, each of two attributes can be assigned to a separate histogram, as shown in Figure 3.20(a). A single house is represented once on each histogram, as illustrated by the highlighted house in Figure 3.20(b). If the histograms are static there is no opportunity to see the relation between the two attributes and there is little to commend the use of two

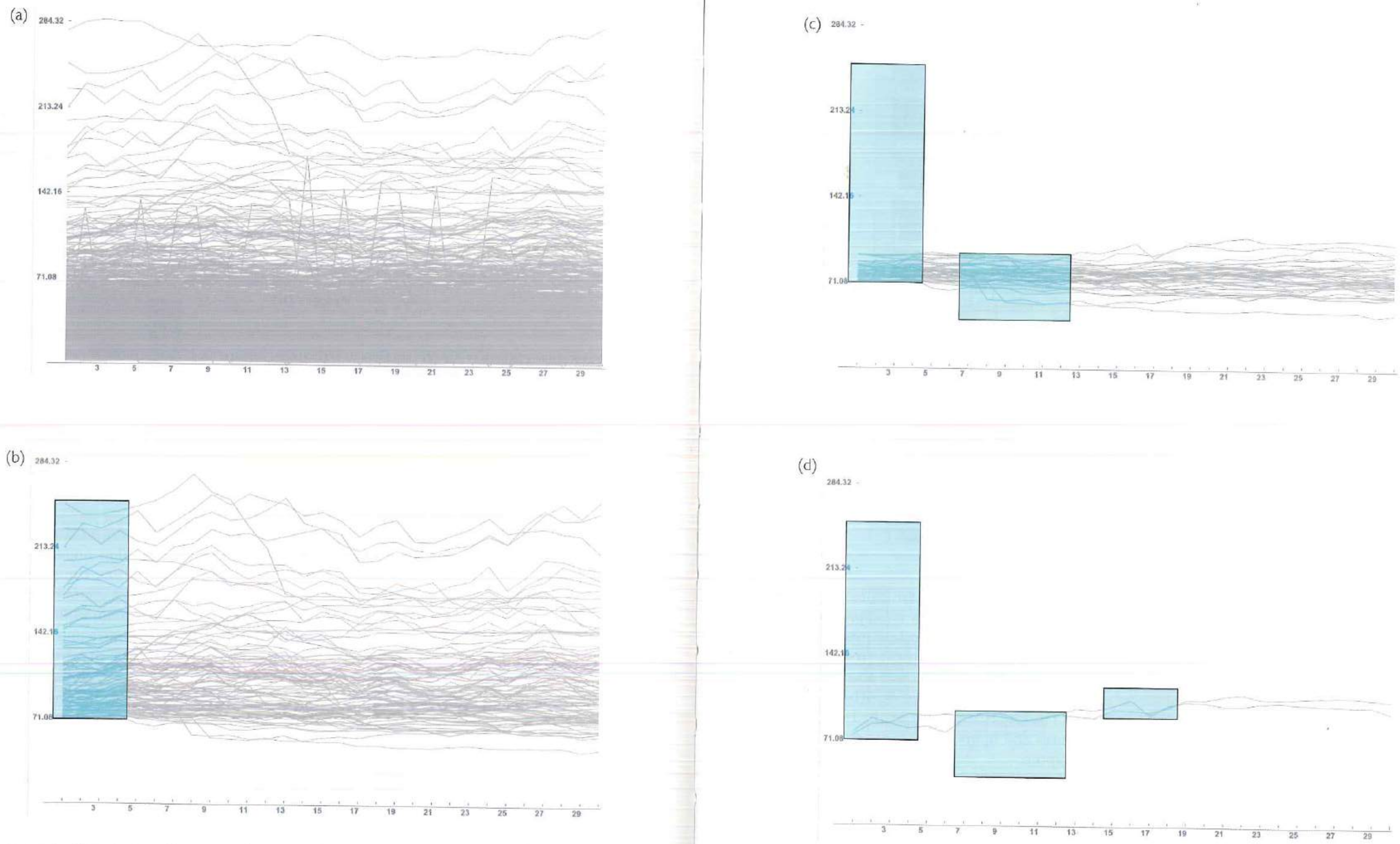


FIGURE 3.18 Four views of a time-series query tool. (a) An overview of the entire data set; (b) a single timebox limits the display to items with prices between $70 and $250 during days 1 to 4; (c) an additional constraint selects items with prices between $70 and $95 during days 7 to 12; (d) yet another constraint concerns prices between $90 and $115 for days 15 to 18 *(Courtesy of Harry Hochheiser)*

separate histograms. However, if interaction allows the placement of limits on one of the attribute scales (Figure 3.20(c)), then the houses thereby identified can usefully be encoded by colour not only on one attribute histogram but on the other as well (Figure 3.20(c)), thereby providing insight into the relation between the two attributes. This is another example of the powerful technique of brushing, whose value is difficult to overstate. Especially if the histogram display is modified to 'range down' selected objects, as in Figure 3.20(d), another

benefit accruing from brushing is illustrated if it is possible to move the entire selected *range* of one attribute from side to side and see how that affects the selected houses on the second attribute. If lower and upper limits on both attribute scales can be adjusted separately, a very flexible exploratory visualization tool results (Tweedie *et al.*, 1994; Spence and Tweedie, 1998; Albinsson *et al.*, 2003).

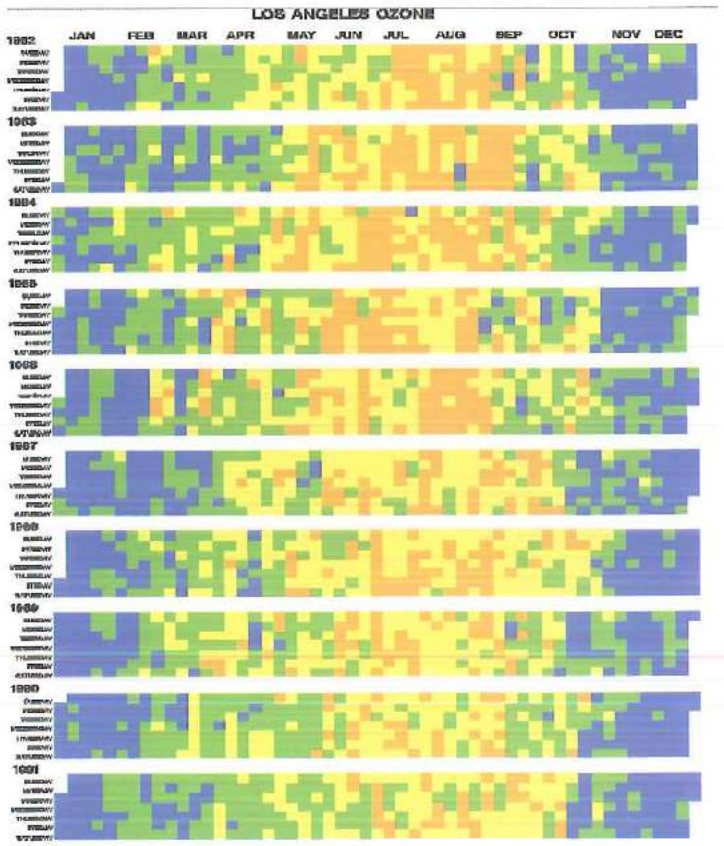


FIGURE 3.19 Representation of the level of ozone concentration above Los Angeles over a period of ten years

It may well be the case that, of two attributes, one is either far more important than the other or must be examined first. In this case it may be appropriate to employ 'logical' or *'semantic' zoom*. If the price of cars is of prime interest, a representation such as that shown in Figure 3.21(a) might be the first to be examined. It is then possible to arrange for a semantic zoom that will show, for a subset of the cars, the make of car in addition to the price (Figure 3.21(b)). This technique, which dates back to 1980 at least (Herot, 1980; Herot *et al.*, 1981), is quite general: it can encompass many attributes and many discrete levels of progressive zoom, as we shall see in the next chapter.

The frequent need for a *qualitative* understanding of data is illustrated in the representation of Figure 3.22. The domain is that of electronic circuit design and the situation is one in which the designer has proposed a design and needs a first appreciation of a particular property of that circuit, for example the magnitude of the voltage at each point in the circuit. These values are encoded by the size of a red square in Figure 3.22. Why? First because the designer will already have a mental model of the expected voltage magnitudes and the display will either confirm that model or, in the case of Figure 3.22, make it clear that a mistake has been made in choosing a value for a particular component (because the two squares at top right are not of equal area). In this example the designer has made a useful discovery, one that might not have been made if the voltage values had

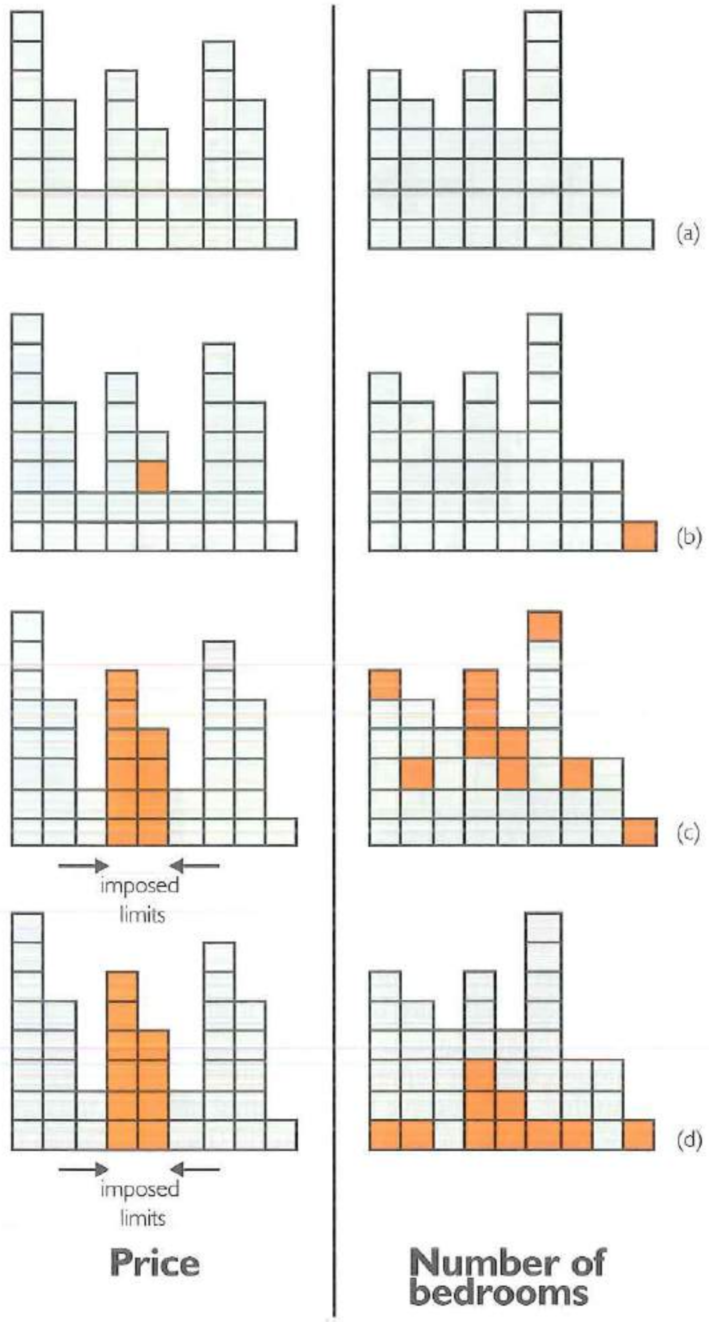


FIGURE 3.20 Linked histograms. (a) The price and number of bedrooms associated with a collection of houses are represented by separate histograms; (b) a single house is represented once on each histogram; (c) upper and lower limits placed on price define a subset of houses which are coded red on *both* histograms; (d) interpretation is enhanced by 'ranging down' the colour-coded houses, especially if exploration involves the dynamic alteration of limits

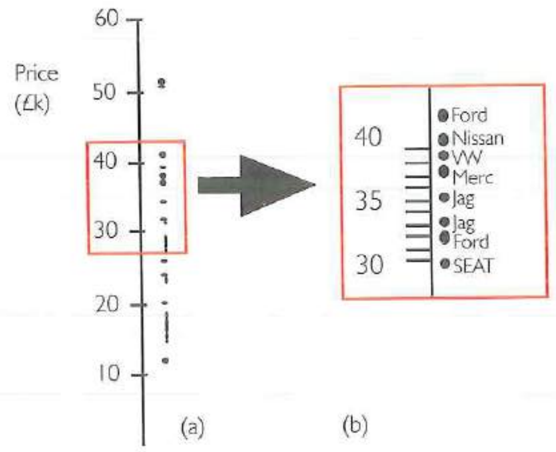


FIGURE 3.21 Semantic zoom reveals data about a second attribute

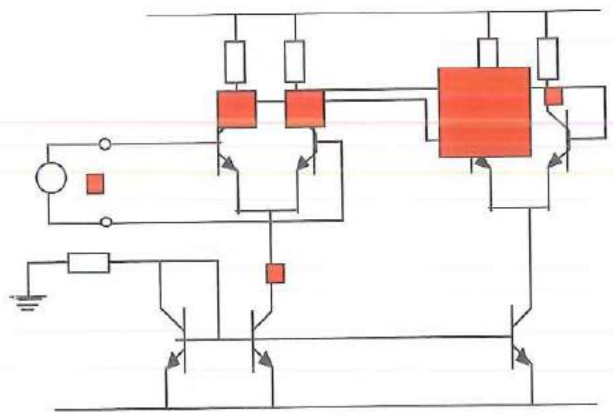


FIGURE 3.22 The area of each red square encodes the value of the voltage occurring at the point in the circuit at which the square is located

been displayed in numerical form. The difference between the two red squares 'pops out'. When numerical values are of interest, a simple mouse click or mouse-over can always replace the red squares with the corresponding numerical values.

A similar technique exists for representing value in such a way as to afford a qualitative understanding. Many people have a mental model of the representation of Australia and New Zealand on a map of the world (Figure 3.23), so that the fact that New Zealanders possess ten times as many bicycles as Australians (if that were true!) could be encoded by correspondingly magnifying the representation of New Zealand on a map, as shown in Figure 3.24. That such magnification encoding can lead to useful insight is demonstrated by *The State of the World Atlas* (Smith, 1999) of which one page, concerned with population density, is shown in Figure 3.25. The familiar large land mass normally located at top left is now replaced by a thin strip representing Canada's low population density, as is the case with Australia. Other pages in this atlas are equally illuminating.

3.1.3 Trivariate data

Since we live in a three-dimensional world there is a temptation to represent data about objects characterized by three attributes by points in three-

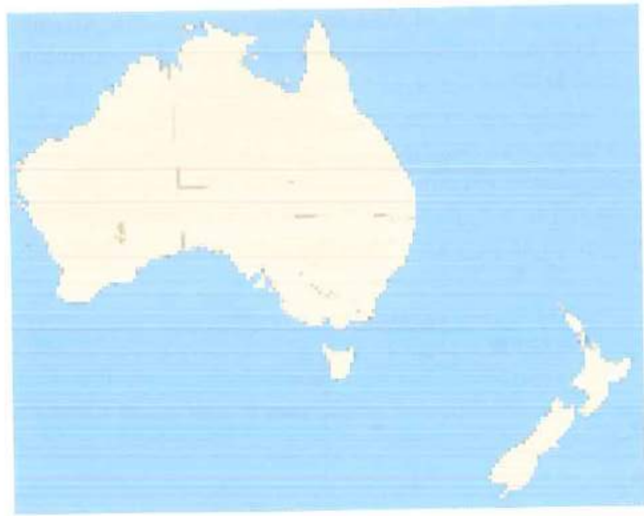


FIGURE 3.23 A representation of Australia and New Zealand on a conventional map

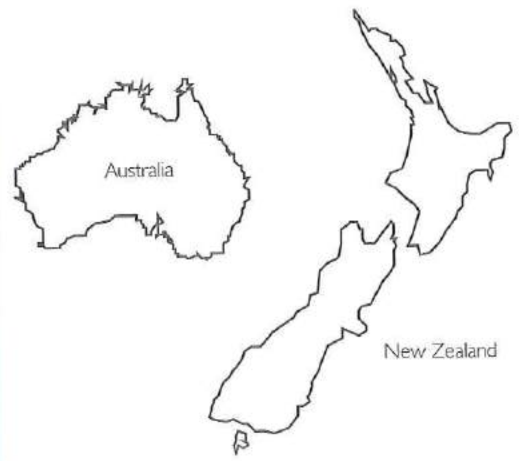


FIGURE 3.24 A representation of Australia and New Zealand indicating that some attribute of New Zealand is ten times its value for Australia

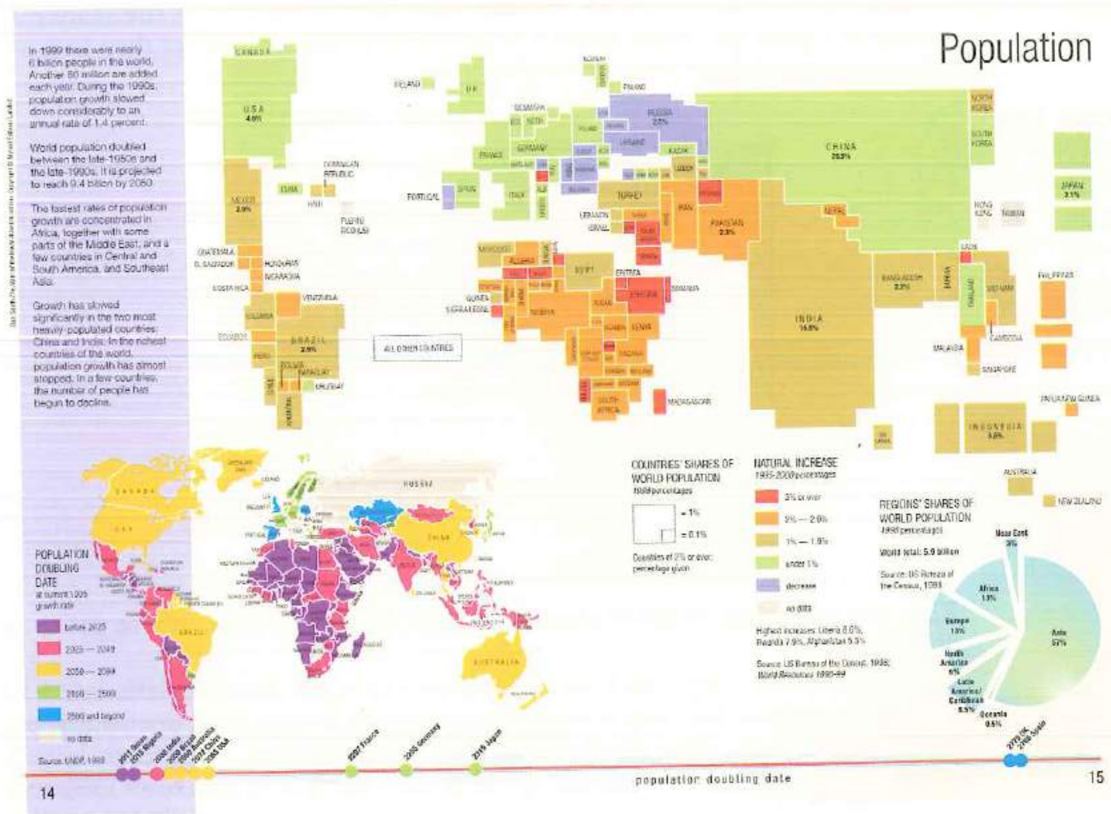


FIGURE 3.25 In *The State of the World Atlas*, magnification encoding is used to give a first impression of population densities. Note the reduced 'size' of Canada and Australia when compared with a conventional map
Source: *Smith (1999)*

dimensional space and to display a 2D view of that space (Figure 3.26). A problem is immediately apparent, however. It is impossible to know, for example, whether house A costs more than house C.

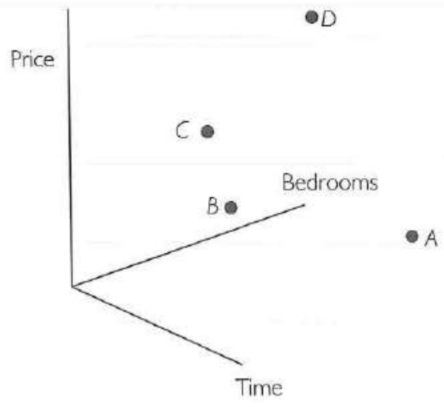


FIGURE 3.26 It is impossible, with a '3D' representation of data, to compare the attributes of Price, Number of bedrooms and Journey time to work values

Scatterplot matrix

One solution is again to bring interaction to the problem and allow a user to reorient the 3D representation so that, for example, the view is the scatterplot shown in Figure 3.27 from which it is clear that house A costs less than house C.

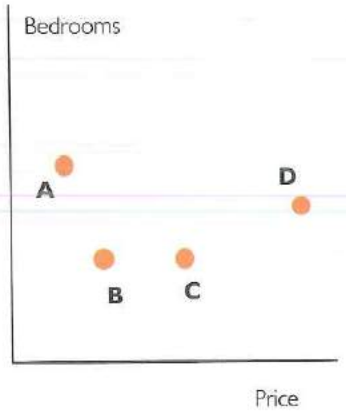


FIGURE 3.27 A projection of the data shown in Figure 3.26, allowing comparison of *price* and *bedrooms* values

This was the origin of the comment that 'for 3D to be useful, you've got to be able to move it'.

For trivariate data – and, as we shall soon see, for hypervariate data – an alternative representation is a systematic structure formed from the three possible 2D views of the data, as shown in the scatterplot matrix in Figure 3.28: in essence, the user has reoriented the 3D plot in three different ways. Although two houses can now be compared easily with respect to three attributes, another problem is present. There are now three times as many points as houses, potentially increasing the cognitive load on the user, and often there will be no room

for labels. Here again interaction can offer a solution by allowing the user to identify, for example, low-cost houses in the Price–Time plane, whereupon the same houses are highlighted in some way in the other two planes (Figure 3.29). Again, we are *brushing* houses from one plane to the other two. But, mindful of the phenomenon of change blindness, we must ensure that the user *notices* the highlighting. Thus, rather than simply changing the colour of a small dot from black to red in all three scatterplots, we might choose to highlight by magnification – even if only temporarily – as well as by colour, as shown in Figure 3.29.

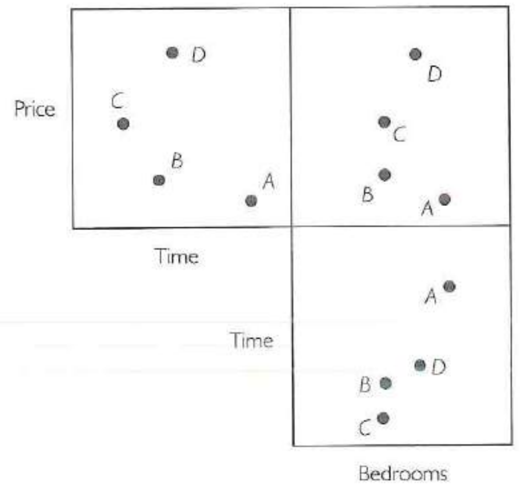


FIGURE 3.28 The scatterplot matrix associated with the data of Figure 3.26

FIGURE 3.29 The highlighting of houses in one plane is brushed into the remaining planes

The above example does not imply that three scatterplots comprising a scatterplot matrix are needed for trivariate data, a fact that can be illustrated by the representation of product fault data. Almost every mass-produced product designed for domestic use will encounter the problem of failure of some kind due to a variety of causes. The fact that a car radio, for example, has failed is easy to record but, and especially in the absence of the faulty product, it is a very expensive matter to find the cause of the fault. Information visualization can help, as shown in the representation of Figure 3.30. Two attributes represent independent variables: one is the month in which the product was manufactured and the other is the number of months it had been in service before the fault occurred. The radius of each circle indicates the number of products manufactured in a particular month and which had been in service for a specific number of months before failing. This representation alone is of considerable help in identifying the types of fault that may be occurring. For example, if many faults occur for one particular month of production irrespective of how long after production the fault occurs (as suggested by Figure 3.30), they are characterized as 'epidemical failure' and can guide the production manager in detecting the underlying cause. If, however, many faults are located on the rising diagonal (again as is the case in Figure 3.30) then a seasonal influence is indicated, since points along such a diagonal refer to products which all fail at roughly the same time.

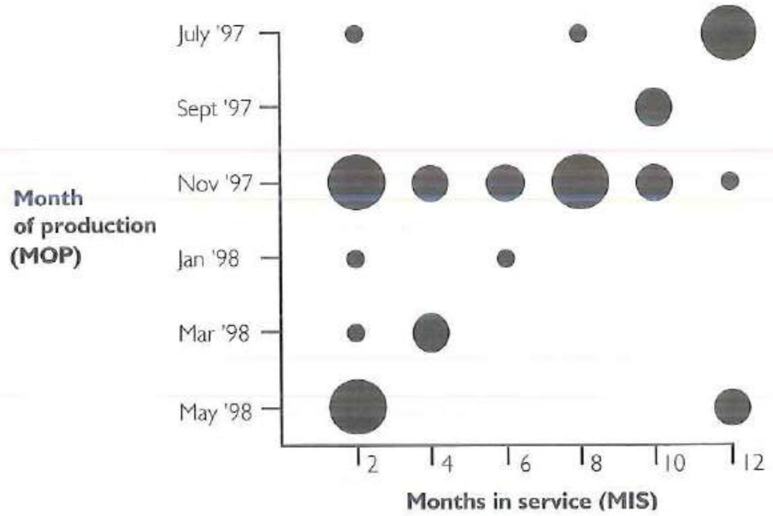


FIGURE 3.30 A representation of reported product failure, based on month of production (MOP) of the failed product and total months in service (MIS) before the fault occurred. The radius of each circle indicates the number of faults reported for a given MOP and MIS

Another example of the representation of trivariate data, and which is similar to Figure 3.22, is concerned with a property of an electronic circuit. Figure 3.31 shows a diagram of an electronic circuit in which circles are superimposed on components to provide some idea as to their importance for – that is, their effect upon – a critical property of the circuit. But if the circuit is a hi-fi amplifier, the designer needs to know this information at a number of frequencies between the bass and treble extremes. It is a simple matter to animate the representation, with circle sizes changing accordingly as a pointer moves continuously up and down the frequency scale.

1

A special category of trivariate data is to be found in maps showing the location (latitude and longitude) and value of some object or attribute. Figure 3.32 shows the population of major cities in England, Wales and Scotland. By its use of circle size to encode each city's population, we can quickly gain an impression of how population is distributed. We might say that the information in this figure 'pops out'. We do not have to examine numbers to discover where the largest city is and what the relative sizes of the populations are. Thus, if we can arrange that information of interest to a user 'pops out' *without* the need for

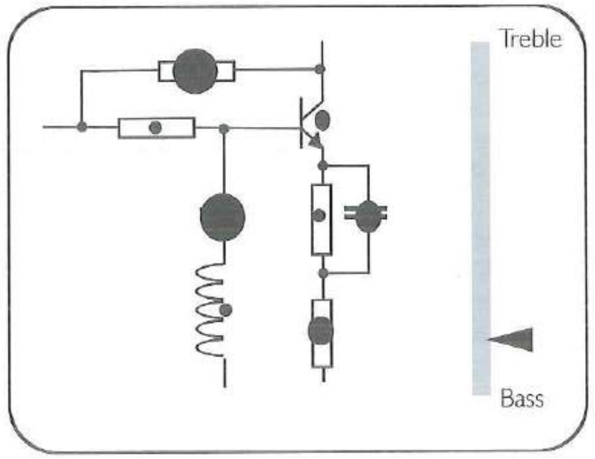


FIGURE 3.31 Circles indicate the extent of the effect of a component on some property of the circuit and change in size as the frequency cycles up and down the range from bass to treble

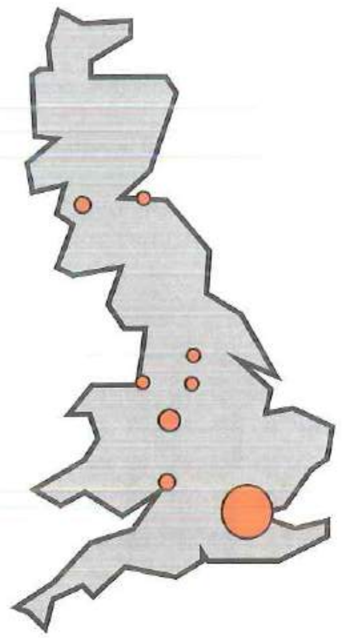


FIGURE 3.32 A representation of the population of major cities in England, Wales and Scotland. Circle area is proportional to population

cognitive effort, then the user is at a huge advantage compared with a situation wherein careful attention has to be given to acquire that insight. How information can be encoded to result in essentially immediate comprehension is closely associated with the human visual system and is discussed immediately below.

Pre-attentive processing: 'things that pop out'

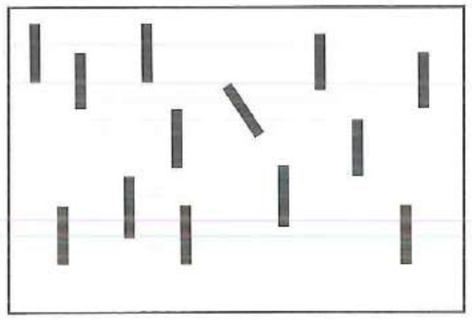
The desirability of information 'popping out', without the need for careful scrutiny, has been mentioned frequently. As Colin Ware (2004) remarks, 'We can do certain things to symbols to make it much more likely that they will be visually identified even after a very brief exposure. Certain shapes or colours "pop out" from their surroundings. The theoretical mechanism underlying pop-out is called **pre-attentive processing** because, logically, it must occur prior to conscious attention.' The 'very brief exposure' mentioned by Ware is typically in the region from 30 to 300 milliseconds. The importance of pre-attentive processing is reflected in Ware's comment (Ware, 2004, page 149) that:

> *An understanding of what is processed pre-attentively is probably the most important contribution that visual science can make to data visualization.*

In view of the immense advantage of deriving insight without the need for conscious attention, an obvious question for the interaction designer is, 'How do we *make* information pop out?' An in-depth discussion of pre-attentive processing in the context of the human visual system is available elsewhere (Ware, 2004): here we examine some examples which are sufficient to provide an initial understanding for an interaction designer.

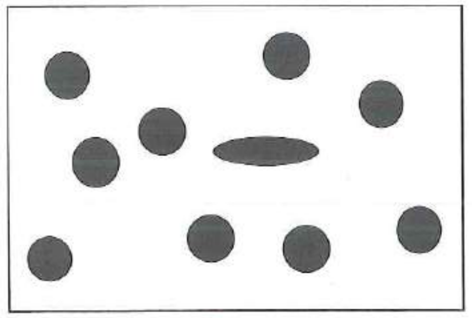
A glance at Figure 3.32 immediately establishes where, in the United Kingdom, the greatest density of city population is to be found (and, for users familiar with a map of the UK, the fact that it *is* a map of the UK). Similarly, in Figure 3.33, the 'odd one out' can quickly be identified, as it can in Figures 3.34 and 3.35. Pre-attentive processing also allows us to quickly identify where, in Figure 3.36, the blue square is located. A cautionary example, however, is provided by the task of locating, in Figure 3.37, the red square in a collection of items which can be either red or blue *and* a square or a circle. Here, pre-attentive processing does not occur; careful scrutiny is required to locate the red square. This is because something which is both square (as opposed to circular) and red (as opposed to blue) is identified by 'conjunction encoding'.

FIGURE 3.33 The 'odd one out' can quickly be identified, by pre-attentive processing



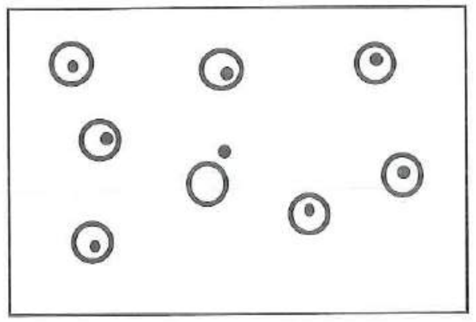
Orientation

FIGURE 3.34 Different shapes can often pop out



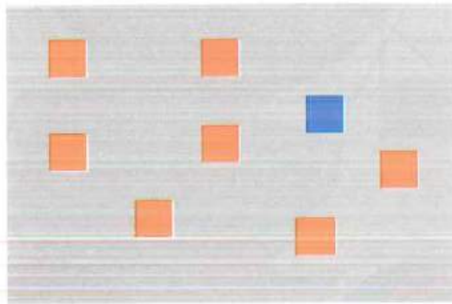
Shape

FIGURE 3.35 A single lack of enclosure can quickly be identified pre-attentively



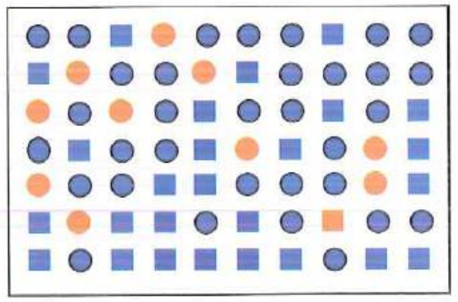
Enclosure

FIGURE 3.36 A different colour can be pre-attentively identified



Colour

FIGURE 3.37 With conjunction encoding the red square is not pre-attentively identified



The last example does not imply that what might be called 'multiple pop-outs' are not possible. Evidence is provided by a novel representation (Irani and Eskicioglu, 2003) of mobile telephone network cell performance. Each cell is represented as shown in Figure 3.38, with colours and dimensions encoding critical performance measures. The outer boundary of a cell representation is chosen to allow their 'packing', as shown in Figure 3.39. As soon as a specialist user decides upon a specific interest – new call blockage rate, for example – the relevant information pops out without interference from other attribute encodings. An interesting and advantageous aspect of the chosen cell representation is that the potential for and general location of additional cells 'pop out' as a result of the arrows, as seen in the upper region of Figure 3.39.

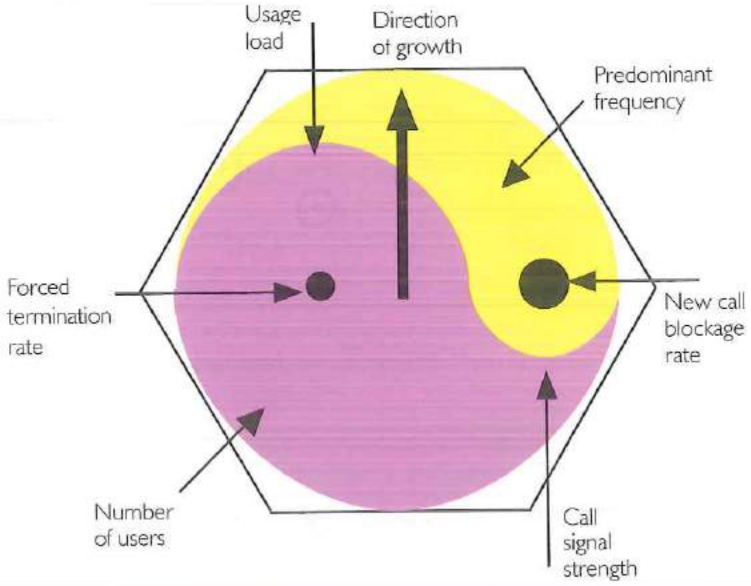


FIGURE 3.38
Representation of attributes associated with a mobile telephone network cell

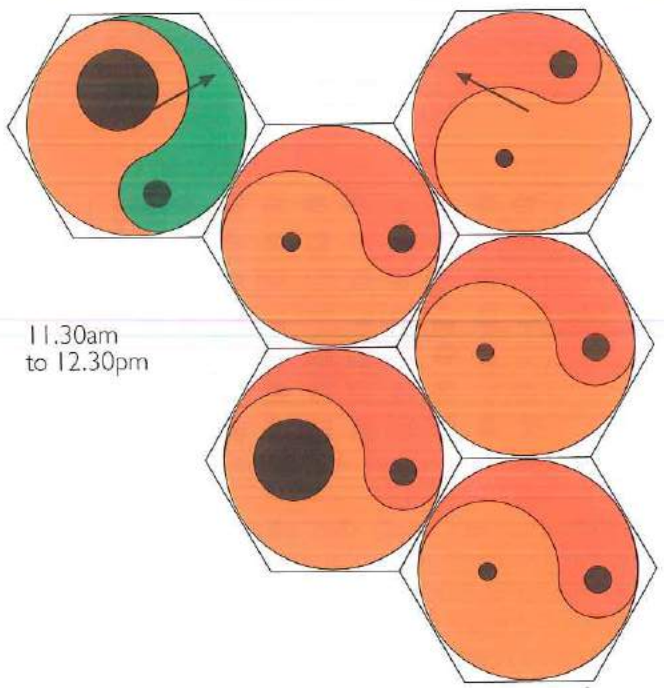


FIGURE 3.39
Representation of attributes associated with a network of mobile telephone cells, averaged over one hour

The examples provided above involve *static* representations of data. Fortunately, the advantages to be gained from pre-attentive processing can be extended through the use of animation, whether automatically or manually controlled. For example, Figure 3.40 shows the map display of Figure 3.32 but with

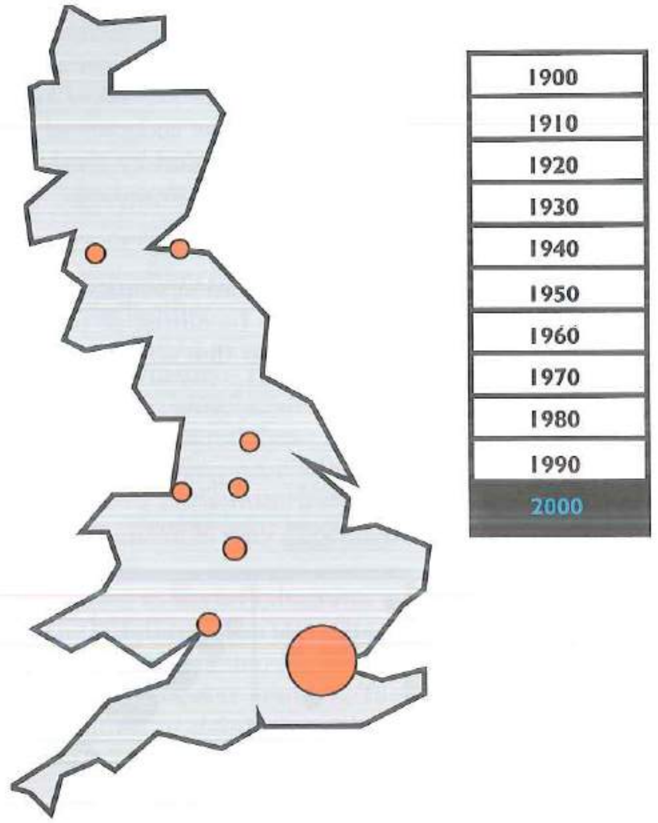


FIGURE 3.40
Circles change in size as the decades are animated, so that sudden changes in population 'pop out'

the addition of a scale calibrated in the decades since 1900 AD. If successive steps on the scale are now highlighted in sequence at a rate of about 5–10 per second, and circles are correspondingly sized, then a mental model of the changes in city populations could quickly be formed. Especially, the emergence of new centres of population would immediately be apparent and, just as with a static representation, 'pop out'. The idea that animation can facilitate 'pop-out' is examined further in Chapter 5.

Although it seems reasonable to employ circle size to encode city populations in Figure 3.40, it is natural to ask 'Why?' Why not use shading? Or bars? Or colour? Or shape? We have seen examples (e.g. Minard) where circle or line size or thickness are suited to the encoding of quantity, and where colour or shape would be inappropriate; colour and shape are perhaps more suited to situations where objects must be perceived as similar or different. Obviously guidance is needed as to the most appropriate encoding for a given task, though it must be realized that the most appropriate encoding depends upon many factors. We briefly discuss this matter below.

Choice of encoding

Although, as Mackinlay (1986) stated, 'there does not exist an empirically verified theory of human perceptual capabilities that can be used to prove theorems about the effectiveness of graphical languages', guidance regarding a choice of encoding is available from many sources, of which three are considered here. The first example is provided by Bertin (1967, 1983), a French pioneer of information visualization. The second comes from the statisticians Cleveland and McGill (1984) following an experiment in which they assessed the accuracy of judging variously encoded quantitative variables. The third is provided by Mackinlay (1986) who considered ordinal and categorical as well as quantitative data.

Bertin's guidance

Bertin identified four *tasks* common to information visualization and identified the encoding mechanisms he considered to be suited to those tasks. He presented his conclusions in a diagram similar to that shown in Figure 3.41, using

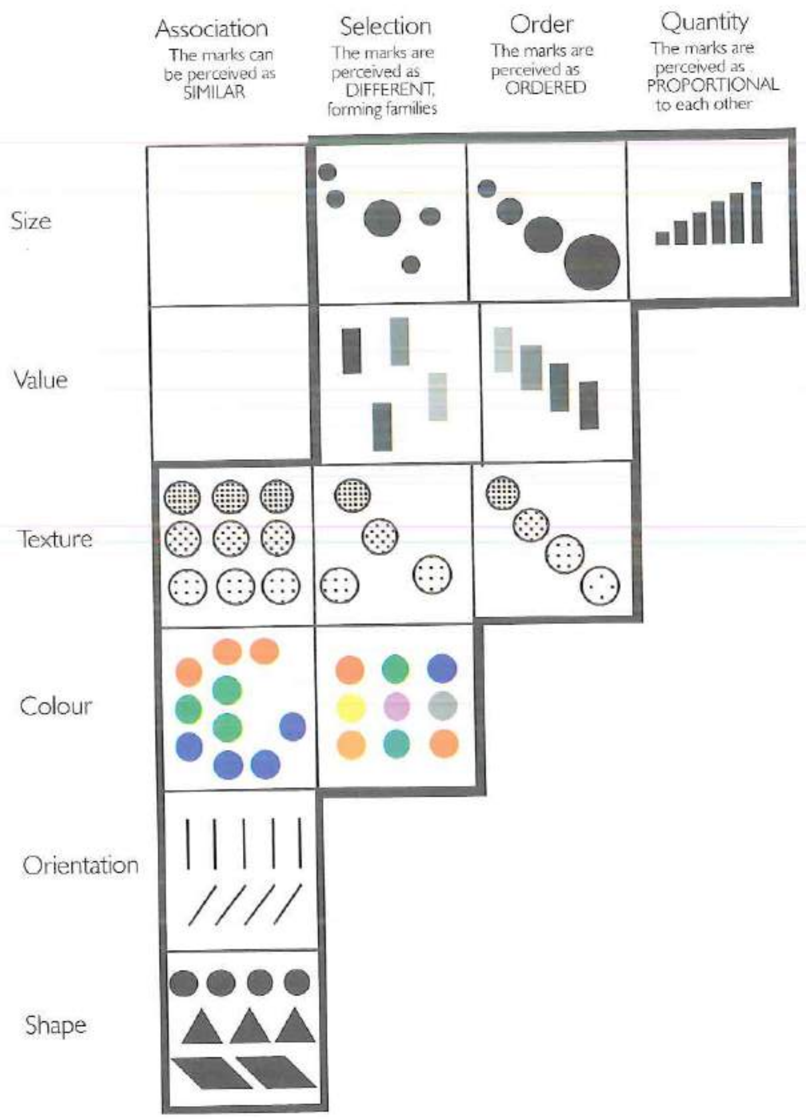


FIGURE 3.41 Interpretation of Bertin's guidance regarding the suitability of various encoding methods to support common tasks

the term 'mark' to denote the result of encoding (e.g. a line or a coloured circle). There are four columns associated with some fundamental tasks. They are:

ASSOCIATION: the question here is how well the marks can be perceived as similar.

SELECTION: here Bertin was concerned with whether the marks can be perceived as different, 'forming families'.

ORDER: Can the marks be perceived as ordered?

QUANTITY: The question here is whether the marks can be perceived as proportional to each other.

The encoding mechanisms – *retinal variables* as they are often called – that he considered were size, value,[2] texture, colour, orientation and shape. As seen from Figure 3.41, Bertin ordered these roughly according to the number of tasks each encoding mechanism can usefully support. Detailed examination of Figure 3.41 shows the guidance to be intuitively reasonable. However, the most appropriate encoding mechanism is very dependent upon context and the guidance is certainly open to debate.

An example related to Bertin's diagram, and illustrating the influence of task, is provided by the TileBars interface (Hearst, 1995). The task supported by TileBars is that of identifying, among a collection of documents, a subset that may be relevant to a particular enquiry. TileBars accepts, from a user, a set of topics and a collection of documents that may or may not be relevant. For example, an investigator interested in research into the prevention of osteoporosis would make the entry shown in Figure 3.42. In response, the TileBars system would return, for each document, a representation of the form shown in the upper portion of Figure 3.43. On the left is a colour-coded reminder of the topics. The TileBar itself contains the same number of rows as there are topics and columns corresponding to segments of the document: these can be paragraphs, pages or chapters. By its density of shading, each rectangle shows, for the corresponding topic and segment, the relative frequency of occurrence of that topic word. Thus, the example shows that in the first segment of the document 'Recent advances in the world of drugs' there is mention of *prevention* but little of *research* or *osteoporosis*. The fifth segment of the document, however, contains substantial mention of all three topics, from which it may be judged

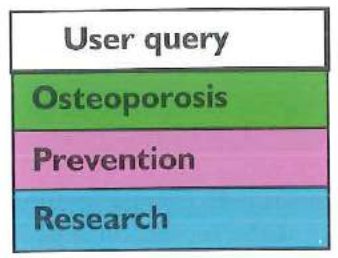


FIGURE 3.42 The specification of three topics of interest

[2] As defined by Bertin (1967; 1983, page 73), 'value' is not restricted to a grey scale. As Bertin remarks, 'One can pass from black to white by grays, by blues or by reds . . .'

that the document may be worth examining, perhaps beginning with that fifth segment. Interaction with the fifth column then leads to the representation in the lower part of Figure 3.43 showing the relevant paragraph with the topic words highlighted. TileBars provides yet another example of the benefit to be obtained from derived values.

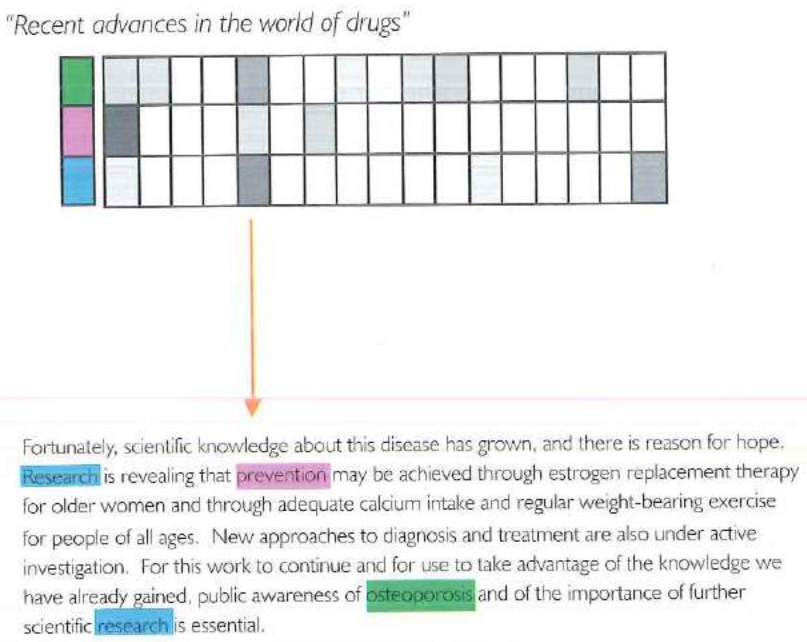


FIGURE 3.43 (top) The TileBar representation of the relevance of paragraphs to the topic words; (bottom) a selected paragraph with topic words highlighted

Accuracy of judgement of encoded quantitative data

The statisticians Cleveland and McGill (1984) addressed a different problem, concerned only with the accuracy with which subjects could assess *quantitative* data encoded in a variety of ways, some identical to the encoding mechanisms considered by Bertin. In one sense they investigated the last column of Bertin's matrix, concerned with quantity and which contained only one entry. Their result is shown in Figure 3.44, where the encoding mechanisms are ordered according to the accuracy with which they were found to support judgement of quantity.

Quantitative, ordinal and categorical data

Mackinlay (1986) addressed the encoding of non-quantitative as well as quantitative data and presented the rankings shown in Figure 3.45. The different positions of a given encoding mechanism on the three rankings are due, as Mackinlay points out, to 'the fact that additional perceptual tasks are involved'.

Representations we have already encountered have been located within Mackinlay's rankings in Figure 3.45, partly as illustrative examples but also because they do not appear at the top of the rankings; they illustrate the fact that many factors influence the choice of encoding. Minard, for example, might have had difficulty, in the context of other desirable features of his map, in encoding the population of the army by position as opposed to length (i.e. the width of a line) and Beck might have faced a similar problem in encoding differ-

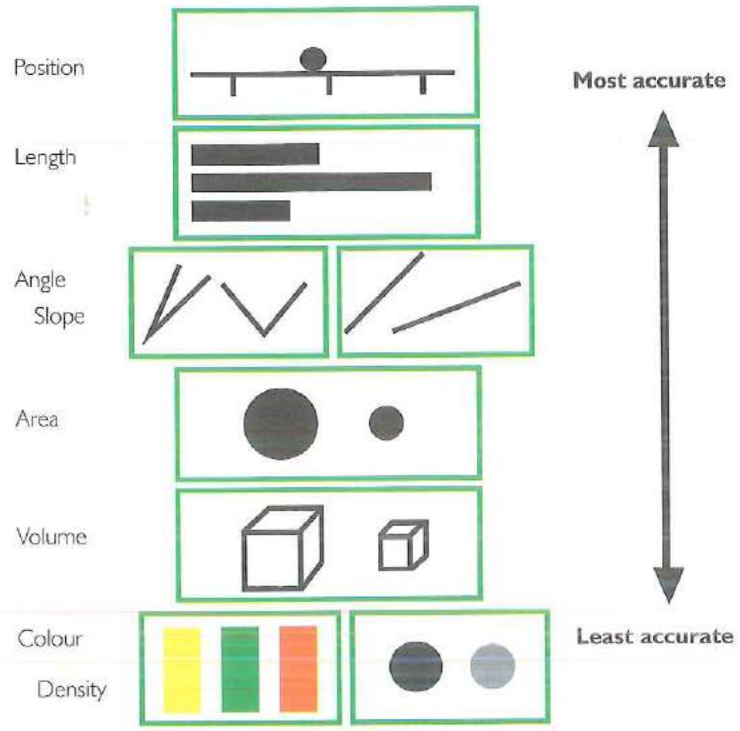


FIGURE 3.44 The relative difficulty of assessing quantitative value as a function of encoding mechanism, as established by Cleveland and McGill

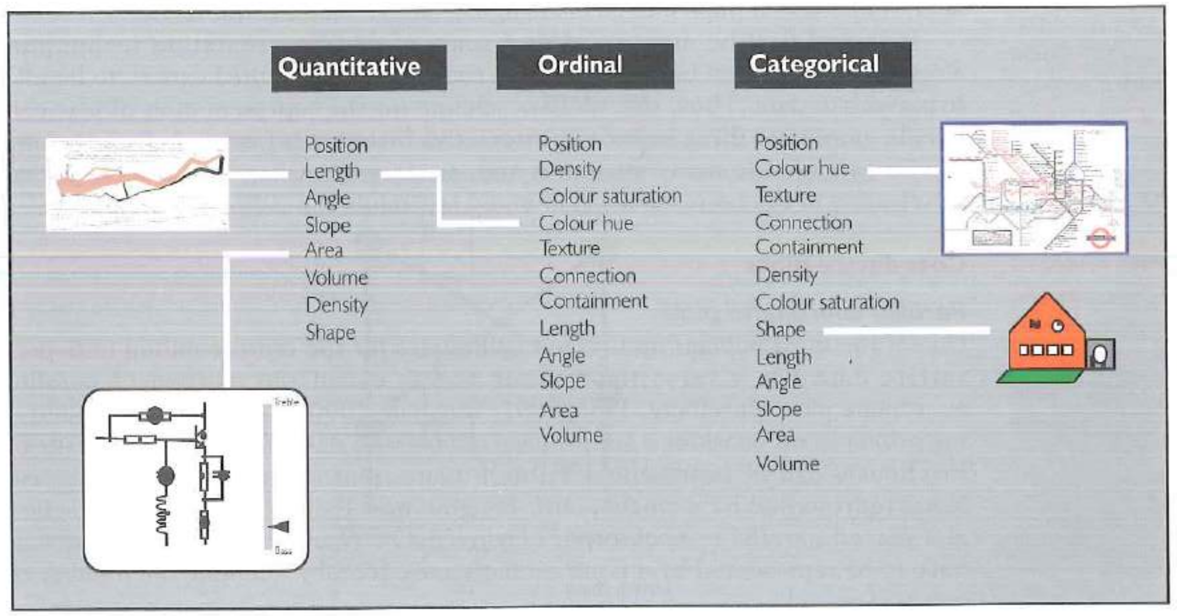


FIGURE 3.45 Taken from Mackinlay's guidance for the encoding of quantitative, ordinal and categorical data

ent Tube lines in anything but colour. It is also difficult to see how Minard could have employed 'position' to indicate the ordering 'advance, retreat'. One must additionally bear in mind that the map was designed to ensure that information 'pops out', with an emphasis on qualitative understanding (e.g. 'Wow! Only a fraction of the soldiers who departed actually returned'), which, if interest persists, can be converted to quantitative understanding by close examination of numbers. Similarly, the 'circle size' representation (Figure 3.31) of the effect of components in an electronic circuit was chosen primarily because a designer is, at least in the initial stages, principally interested in whether some effect is large or small. A different consideration is associated with the multidimensional icon employed to represent the attributes of a house and which we discuss later in this chapter. Here we encounter *iconic encoding*, and this introduces yet another factor to be taken into consideration in the choice of encoding for a multi-attribute representation.

3.1.4 Hypervariate data

The challenge of representing hypervariate (also termed multivariate) data is substantial and continues to stimulate invention. It is an important challenge because so many real problems that are potentially amenable to at least partial solution via information visualization are of high dimensionality. The design of even the simplest silicon chip, for example, can involve over 100 components whose value has to be chosen by the designer in such a way that over 200 performance limits are satisfied. A decision regarding an investment portfolio is equally complex, as is the decision whether or not to continue with the development of a new drug. Even a basis for representing only eight scholastic achievements of a pupil can be challenging, as Exercise 3.3 will reveal.

It should first be mentioned that some of the representation techniques already discussed can be scaled, though sometimes to a limited extent, to handle hypervariate data. Thus, the TileBars scheme for the representation of text can handle more than three keywords, interactive histograms (as we shall see below) can be extended to many attributes and, as shown in Chapter 2, bargrams can be effective in the representation of many attributes of a collection of objects.

Coordinate plots

Parallel coordinate plots

One of the most popular and valued techniques for the representation of hypervariate data has a very simple basis and is called the method of parallel coordinate plots (Inselberg, 1985, 1997; Wegman, 1990). To explain the underlying principle we consider a simple case of bivariate data for which the details of two houses can be represented within a scatterplot (Figure 3.46), each house being represented by a single point. Imagine now the two axes to be detached and placed parallel to each other (Figure 3.47). Necessarily, each house will have to be represented by a point on both axes, thereby doubling the number of points required.

A further disadvantage would appear to be the need to show the relation between the two points characterizing a given house, here achieved (Figure 3.48) for each house by a straight line together with an identifying label. What is to be gained? For the case of two attributes, nothing. However, if we are dealing with objects characterized by more than two or three attributes the so-called parallel coordinate plot offers many advantages. Figure 3.49 shows the parallel

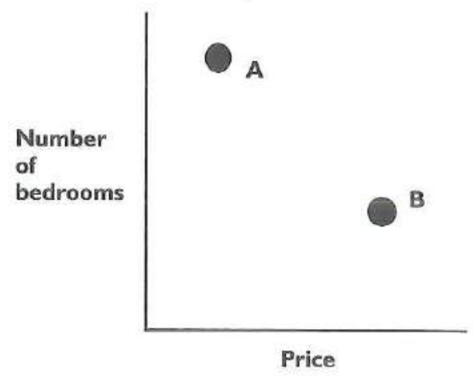


FIGURE 3.46 A simple scatterplot representing the price and number of bedrooms associated with two houses

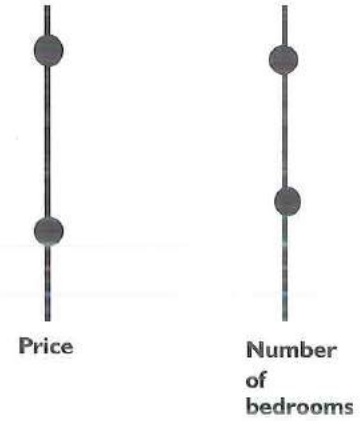


FIGURE 3.47 An alternative representation to the scatterplot in which the two attribute scales are presented in parallel, thereby requiring two points to represent each house

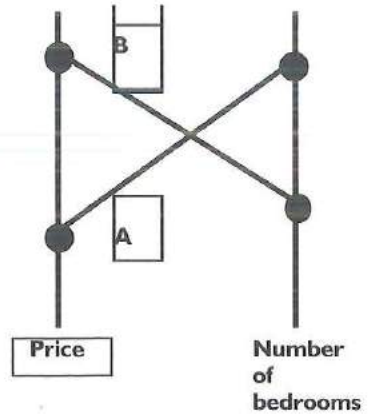
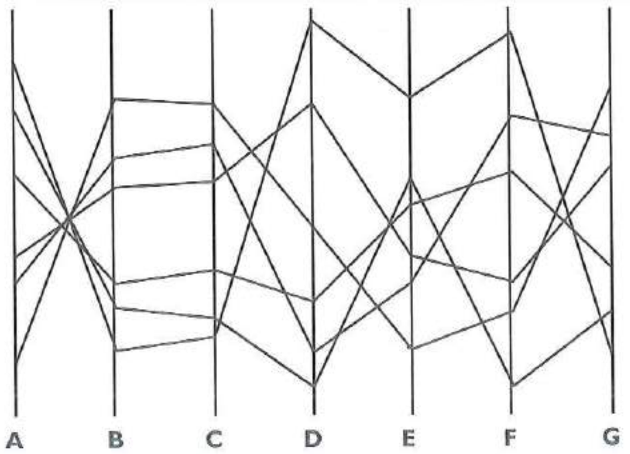
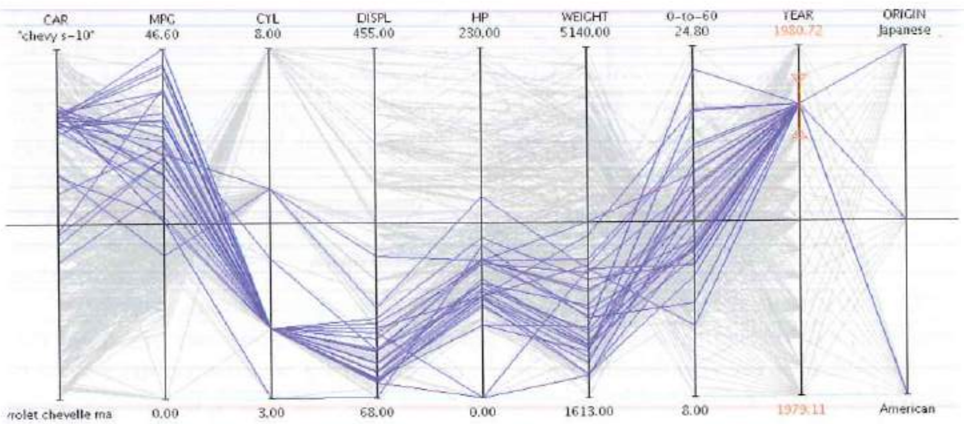


FIGURE 3.48 To avoid ambiguity the pair of points representing a house are joined and labelled

coordinate plot for six objects, each characterized by seven attributes A to G. Each object is represented by a point on each axis and hence by the piecewise linear line ('polyline') joining them. It is immediately (i.e. pre-attentively) apparent that there is a 'trade-off' between attributes A and B, as well as a strong correlation between B and C. Nevertheless, even though a parallel coordinate plot facility is available in many commercial information visualization packages, limitations can be identified. For example, for the data shown in Figure 3.49, it is not apparent that there is also a 'trade-off' between B and E and a strong correlation between C and G; in other words, the ordering of the attributes can significantly affect the ease with which relationships can be identified.



As with many other visualization tools, the potential offered by interaction is considerable (Siirtola, 2000). For example, a range of one attribute can be identified, thereby highlighting all the object lines which pass through that range (Figure 3.50); furthermore, that range can be dynamically explored manually, allowing a user to gain quick insight into relationships between different attributes. Other facilities are normally available, including averages, standard deviations and Tukey box plots. Selected ranges of two different attributes can, for example, be 'ANDed' or 'ORed' highlighting, respectively, only those object lines that pass through *both*



selected ranges and those which pass through either or both.[3] The parallel coordinate plot technique of representation has found a wide range of application. Inselberg (1997), for example, has shown how such plots can be used to enhance the manufacturing yield of a process of making silicon chips.

It is useful to be able to characterize the sort of insight that can readily be gleaned from a parallel coordinate plot. Whereas with the scatterplot each object was easily identified and discriminable in the company of other objects, that is not the case with parallel coordinate plots because each object is now represented by a polyline which usually intersects with many other such curves. What *is* particularly visible from a parallel coordinate plot are the characteristics of the separate attributes and, in some cases, the nature of the relation between them. Thus, remarks such as 'the majority of cars appear to have four cylinders' and 'it seems that low MPG inevitably means high price' can be based on rapidly acquired 'pop-out' insight. Thus, we may observe that the parallel coordinate plot technique can support **attribute visibility**. By visibility we mean the ability to gain insight pre-attentively or without involving a great detail of cognitive effort. We return later to the concept of visibility, and the associated concept of correlation, when more encoding examples have been accumulated.

A major attraction of parallel coordinate plots, which they share with some but not all other techniques, is that their complexity (here, the number of axes) is directly proportional to the number of attributes. They also have the advantage, which is sometimes not present with other techniques (Feiner and Beshers, 1990), that all attributes receive uniform treatment. The literature (e.g. Bendix *et al.*, 2005; Siirtola, 2006)) provides many examples of the continual development of the parallel coordinate technique.

Star plots

A star plot (Coekin, 1969) has many features in common with parallel coordinate plots in that an attribute value is represented by a point on a coordinate axis and, for a given object, those points are joined by straight lines. The difference is that attribute axes now radiate from a common origin. Thus, a star plot of my school report (Figure 3.51) shows, relative to a class average indicated by the extremities of the grey region, good performance in mathematics and chemistry but very poor performance in sport and literature. To make a comparison with the talents of my friend Tony, a separate star plot (Figure 3.52) can be employed. The shape associated with a star plot can provide a reasonably rapid appreciation of the student's achievement and permit comparison with another student. It can be argued that a star plot offers '**object visibility**'. Thus, unlike parallel coordinate plots which are especially suited to the identification of relations between attributes, star plots are perhaps better for comparing specific objects. Star plots have been used to compare objects as different as police forces and mortgage options. Other encoding techniques, such as colour and thickness, can provide additional flexibility.

---

[3] The observant reader may have noticed cars apparently characterized by zero HP or zero miles per gallon, neither constituting a very desirable attribute of a car. In fact, these are misleading representations associated with missing data and illustrate a challenge to the designer of a visualization technique.

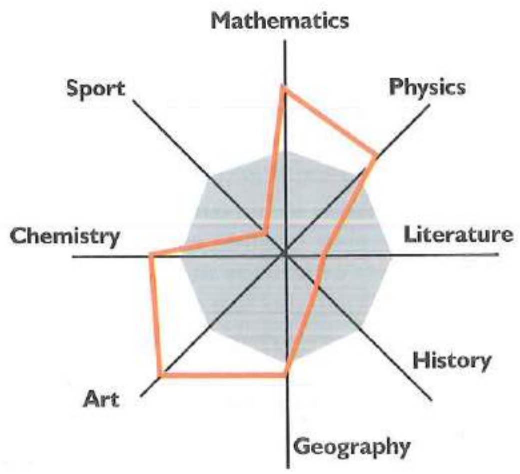


FIGURE 3.51 In a star plot attribute scales radiate from a common origin. Because shape can often effectively represent the combined attribute values of a single object, the points on each attribute scale can usefully be joined. Other useful information such as average values or thresholds can be encoded on the star plot

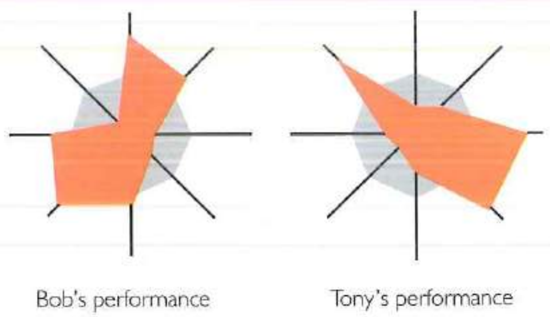


FIGURE 3.52 Star plots can be used to compare the attributes of two different objects, here the exam performance of two people in the subjects identified in Figure 3.51

Scatterplot matrix

The scatterplot matrix discussed in the context of trivariate data is equally applicable to higher dimensions, but with one major disadvantage arising from the fact that, as the number of attributes increases, the number of different *pairs* of attributes increases rapidly. With two attributes we needed one scatterplot, with three we needed three and with four attributes there are six unique pairs. Thus, for 100 houses each associated with four attributes, the scatterplot matrix would contain 600 points. This is in contrast to the linear increase in complexity associated with parallel coordinate plots and, as we shall see, the Attribute Explorer to be discussed immediately below. While there is no theoretical limit to the number of attributes a scatterplot can handle, this unwelcome dependence of complexity on dimension is always present.

The use of a single scatterplot together with other encoding techniques to represent hypervariate data was demonstrated very effectively by Ahlberg (Ahlberg *et al.*, 1992; Ahlberg and Shneiderman, 1994) with an interactive representation designed to allow a user to select a film to watch on video (Figure 3.53). On the main (scatterplot) display each coloured square identifies a film. Colour encodes type (horror, musical, etc.), horizontal position indicates the year of production and vertical position indicates duration. On the right, sliders can be used to specify other attributes of a film such as director or actor. Scroll bars can be used to confine attention to a particular span of years and film length, whereupon more detail can be displayed (Figure 3.54). The Film Finder, as the interface was called, provides a good illustration of the potential of combining different representation techniques.

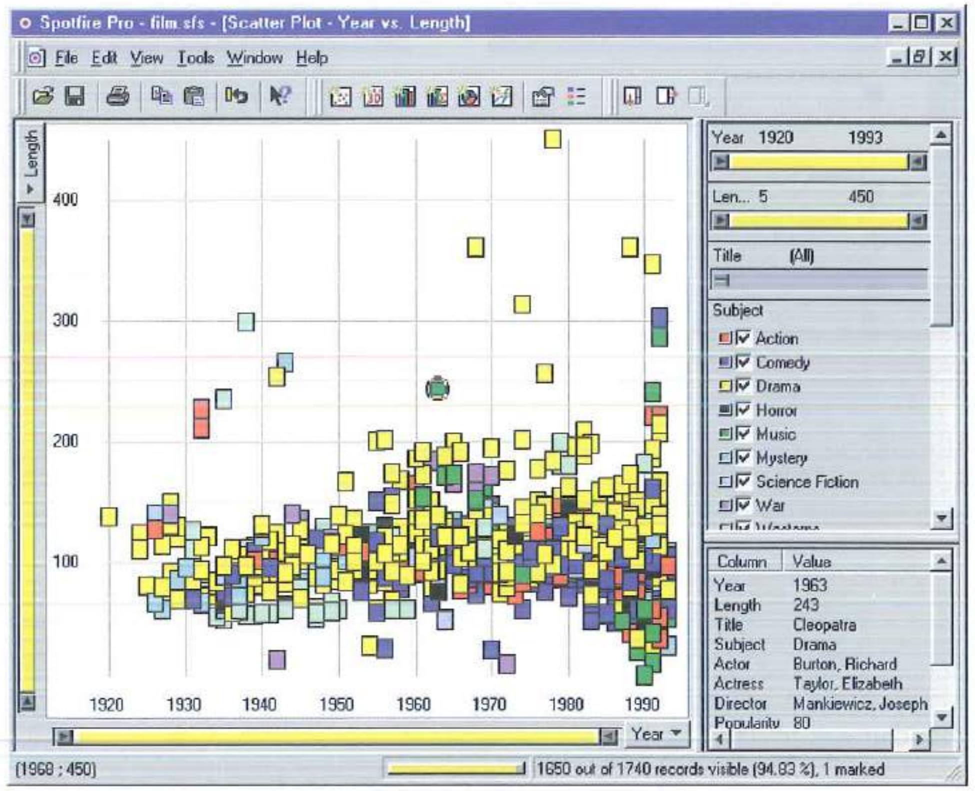


FIGURE 3.53 A scatterplot enhanced by additional and selective encoding, allowing the selection of a film on the basis of type, duration, year of production and other attributes

Linked histograms

The technique of linked histograms presented in the discussion of bivariate data (see Figure 3.20) can be extended to hypervariate data and, in the Attribute Explorer, can be considerably enhanced in value by additional encoding (Tweedie *et al.*, 1994; Spence and Tweedie, 1998). We shall first illustrate the technique in the context of buying a house, a task which, like that of buying a car, can usefully be generalized:

> *Given a collection of objects, each described by the values associated with a set of attributes, find the most acceptable such object or, perhaps, a small number of candidate objects worthy of more detailed consideration.*

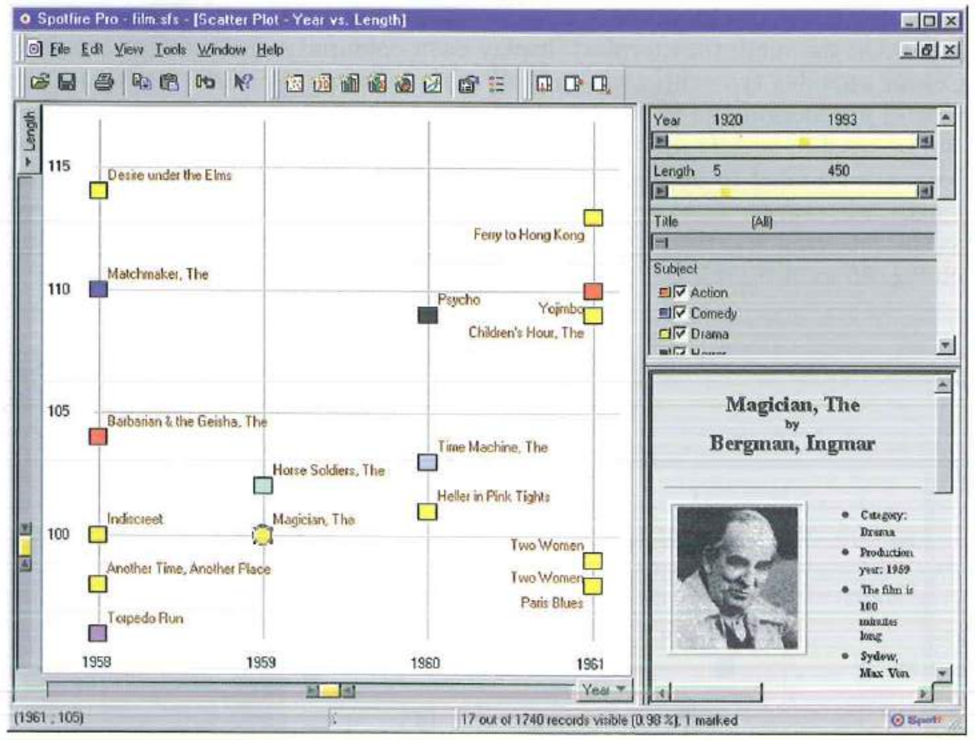


FIGURE 3.54 The automatic display of additional detail following the selection of narrower limits on years of production and film length

However, in the course of illustrating the use of the Attribute Explorer, we shall see that an equally important task for which it (and many other techniques) is suited is:

> *the acquisition of insight into multivariate data.*

We begin by examining a histogram (Figure 3.55) of one attribute of a collection of houses, that of *Price*. Here, each house contributes one small rectangle to the histogram. Upper and lower limits to *Price* can easily be positioned (Figure 3.56) to identify a subset of houses that may initially and experimentally be regarded as affordable. The result is that the houses so identified are encoded green. For reasons soon to be apparent, houses outside the limits continue to be displayed. Many house attributes will normally be of interest and we shall consider just three to establish the concepts underlying the Attribute Explorer: *Price*, *Number of bedrooms* and *Garden Size*. Figure 3.57 shows the corresponding attribute histograms, but with the same limits on *Price* as in Figure 3.56. Those houses satisfying the limits on *Price* are now coded green not only on the *Price* histogram but also on the other two histograms, another example of brushing. Immediately, the consequences of the *Price* limits for the availability of houses with given garden sizes and numbers of bedrooms are apparent and can, in fact, be dynamically explored by moving the range bar between the *Price* limits to and fro: some idea of any correlation between *Price* and *Number of bedrooms* could easily be acquired. If limits are now additionally placed on the remaining attributes (Figure 3.58), green encoding applies only to those houses which sat-

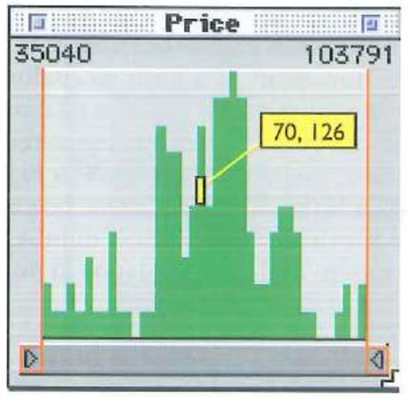


FIGURE 3.55 A histogram representing the prices of a collection of houses. The contribution of one house is shown in yellow

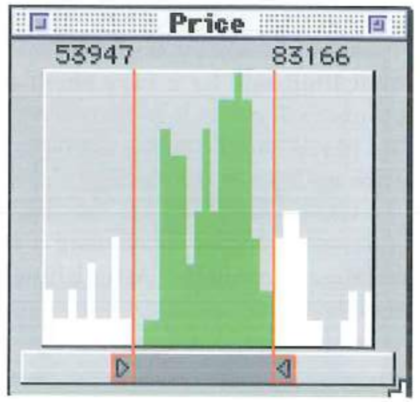


FIGURE 3.56 Limits on *Price* identify a subset of houses, coded green

FIGURE 3.57 Houses defined by the limits on *Price* in Figure 3.56 are coded green in other attribute histograms

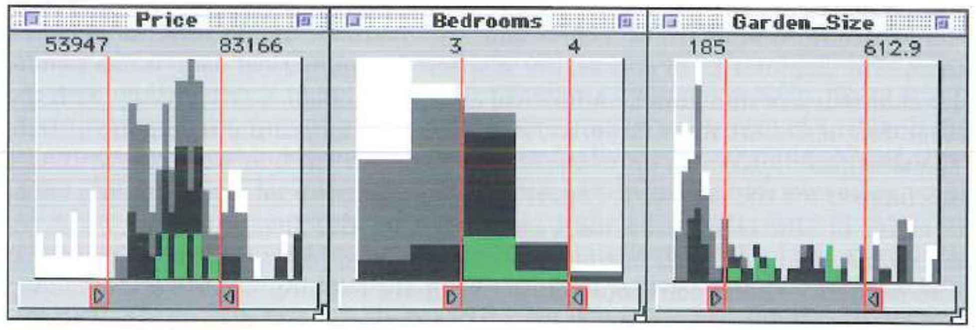


FIGURE 3.58 Green coding applies only to houses which satisfy all attribute limits. Houses which fail one limit are coded black, so if a black house is positioned outside a limit it will turn green if the limit is extended to include it

isfy *all* the attribute limits. Again, dynamic exploration achieved by adjusting either range positions or individual limits can help the user to gain insight into the house data and gradually come to a decision about which house or houses are worthy of more detailed consideration. It is because houses can be explored on the basis of their attributes that the technique described is called the Attribute Explorer.

Of major importance in the Attribute Explorer is the colour coding involved. Black houses are those which fail only one limit; therefore, if a black house is located outside a limit, that must be the limit it fails. This can be very useful information. If a user has set an upper limit on the price they are prepared to pay, they would almost certainly wish to know about any house which failed only that limit and by a very small amount. Movement of a limit to enclose a black house will cause it to turn green, confirming that it now satisfies all limits.

The black encoding is extremely valuable in those circumstances wherein there are no houses that satisfy the stated limits, a situation depicted in Figure 3.59. In this example we see that by spending a little more, one house becomes available, whereas by accepting a house with three bedrooms a number of houses become available. Astonishingly, it is rare to find such guidance in online house sales.

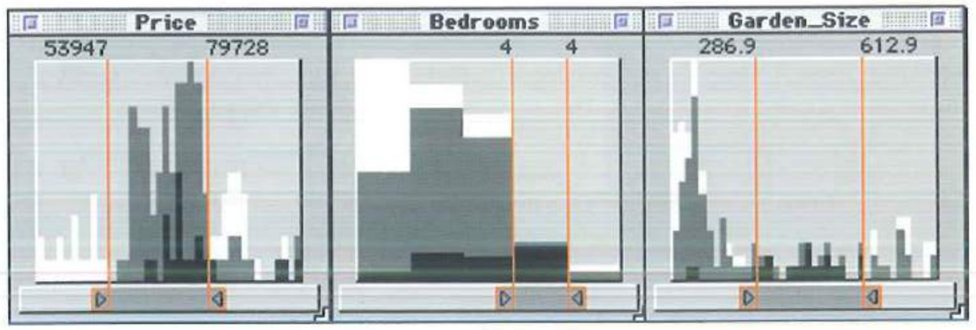


FIGURE 3.59 Even if no houses satisfy all attribute limits, black houses, which fail only one limit, provide guidance as to the effect of relaxing limits

Potential applications of the Attribute Explorer technique will no doubt occur to the reader, especially after experiencing, for example, the frustration of trying to arrange flights on specific dates at convenient times or trying to decide which mobile telephone or washing machine to buy. Combination with other techniques could prove useful, such as providing an Attribute Explorer histogram along each axis of a parallel coordinate plot to enable better insight into the distribution of attribute values and, by brushing, the correlation between them. The Explorer is, of course, not restricted to numerical data: it can handle a combination of numerical, categorical and ordinal data. Concurrently with the invention of the Attribute Explorer, Eick (1994) proposed the technique of Data Visualization Sliders. He drew attention to the fact that, as well as providing a mechanism for the interactive specification of limits to an attribute, 'the effectiveness of sliders may be increased by using the space inside . . . as an interactive colour scale, a bar plot for discrete data, [or] a density plot for continuous data'. The effectiveness of the Attribute Explorer has been studied (Li and North, 2003) by comparing it with a Dynamic queries interface (see Chapter 5, Figure 5.23) in the context of a number of tasks. It was observed that the 'brushing histograms', as the Explorer was called, tended to offer advantages with the more complex tasks. It was commented that brushing histograms could be extended in various ways, for example to be zoomable and to possess granularity controls.

While originally invented to support the selection of one object from many on the basis of its attributes, the Attribute Explorer also offers considerable potential to support exploration with the goal of acquiring **insight** into data. An example is provided in its use as an investigative tool by the Swedish Defence Research Agency and illustrated briefly in Figure 3.60. (A more extensive

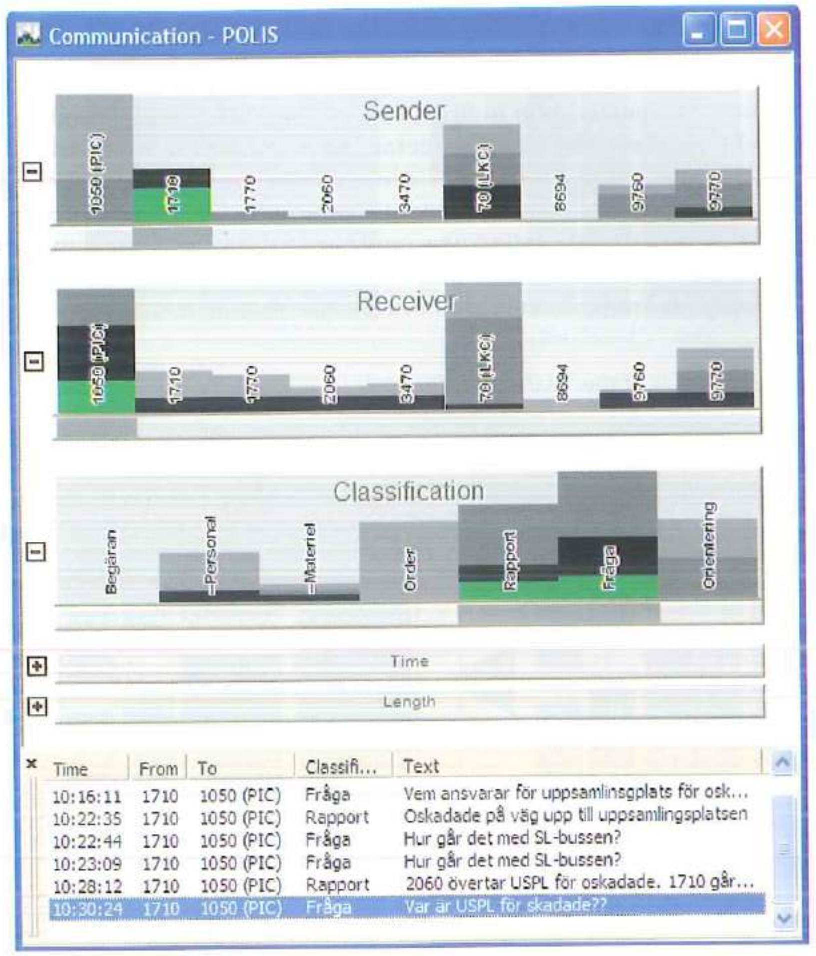


FIGURE 3.60 An Attribute Explorer representation of three dimensions of communication data captured during an emergency services exercise, supporting interactive exploration by an analyst

account is the subject of one of the case studies in Chapter 6.) The figure relates to communication data accumulated during a practice scenario of a train derailment on the Stockholm subway (Morin and Albinsson, 2005) involving three data dimensions. A single sender (1710) has been identified and a single receiver (1050), the on-site commander, as well as two categories of classification (reports and questions). The green 'hits' indicate that a large part of the total communication initiated by 1710 is indeed reports and questions directed to 1050. These 'hits' are also shown in the list below the Attribute Explorer.

Mosaic plots

The idea behind mosaic plots (Friendly, 1994) can be demonstrated by an example related to the Titanic disaster of April 1912, when 1,731 of the 2,201 passengers and crew were lost (Dawson, 1995). Table 3.1 shows the raw numerical data, involving four attributes: gender, survival, class and adult/child. Whereas it is difficult without careful examination to obtain much insight from the table, a so-called mosaic plot (Friendly, 1992, 1994, 2000) can profitably be generated as follows. We start with a rectangle whose area is proportional to the number of passengers and crew (Figure 3.61(a)). We then break down that area according to class of travel (Figure 3.61(b)), again representing number by area. Next we break down the resulting rectangles according to gender (Figure 3.61(c)), from which we can already begin to gain immediate insight into, for example, the male/female ratio in first-, second- and third-class accommodation. Finally, we break down the existing rectangles according to two attributes: survival (green) or otherwise (black), and adult/child (Figure 3.61(d)). The resulting mosaic plot is capable of providing rapid insight into the nature of the Titanic disaster and, of course, generating new questions – Who were the females travelling third class? Why did the vast majority of first-class females survive? Why, proportionately, did more female children survive than male children?

TABLE 3.1 Details of the Titanic disaster

Survived	Age	Gender	Class			
			1st	2nd	3rd	Crew
No	Adult	Male	118	154	387	670
Yes			57	14	75	192
No	Child		0	0	35	0
Yes			5	11	13	0
No	Adult	Female	4	13	89	3
Yes			140	80	76	20
No	Child		0	0	17	0
Yes			1	13	14	0

The order in which the original rectangle is broken down is, of course, open to choice, and the reader may wish to experiment with different orders to gain some feeling for the potential offered (see Exercises). Many other examples, and a fuller discussion, are provided by Friendly (2000) whose website is well worth visiting in this respect.

Icons

Object visibility is a property of representation techniques in which a single object is so portrayed that a number of its attributes can easily be assimilated, qualitatively if not quantitatively. Two examples are discussed here. One was originally invented to characterize the properties of geological samples, whereas the other was evaluated in the context of house selection.

Professor Chernoff, a statistician at Stanford University, observed that human beings are very sensitive to a wide range of human facial characteristics and suggested that facial features – such as eye size and the length of a nose – are not only numerous (he identified 18) but could, in a cartoon face, take on a sufficiently large number of 'values' to offer a useful encoding mechanism (Figure 3.62). He applied this idea (Chernoff, 1973) to the study of geological

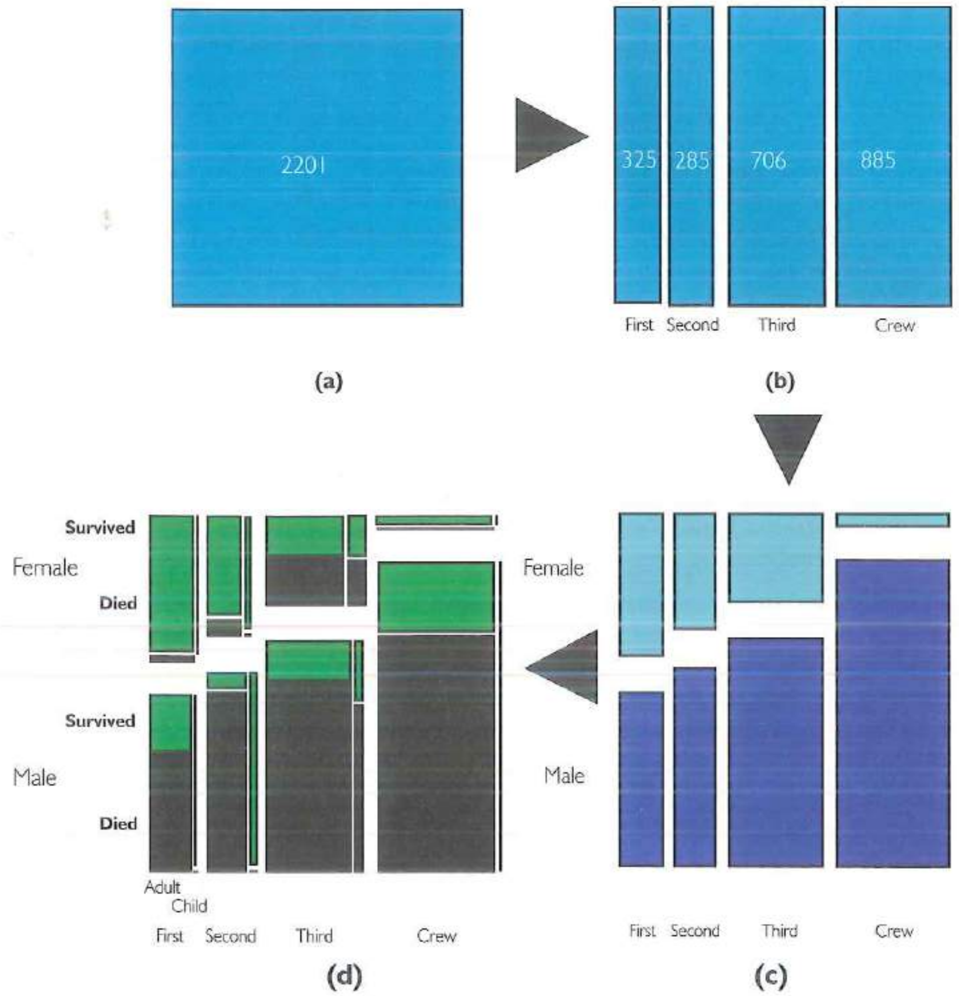


FIGURE 3.61 Steps in the creation of a mosaic plot representing the Titanic disaster

FIGURE 3.62 Chernoff Faces allow attribute values to be encoded in the features of cartoon faces

samples, each characterized by 18 attributes such as salt content and water content, and found that the display of so-called Chernoff Faces facilitated the identification of interesting groups of samples. If this use of computer-generated cartoon faces should appear frivolous it should be remarked that accountants, not usually associated with frivolity, have nevertheless explored the use of Chernoff Faces to display accountancy data (Stock and Watson, 1984). Since one would expect some facial features to be more informative than others (Morris *et al.*, 1999), it is useful to note that a study by De Soete (1986) established the relative value of various facial features. Ware (2004, page 253) comments that we may have special neural hardware to deal with faces.

In the context of house purchase the multidimensional icons shown in Figure 3.63 attempted to portray attributes more directly, for example representing a houseboat by a shape suggestive of a boat and a garage by a rectangular box containing a simple representation of a car. A total of eight house attributes were represented in this way. A question that immediately arises, of course, is whether such multidimensional icons offer any advantage over an equivalent textual description, shown in Figure 3.64. To investigate this issue, controlled experiments were undertaken (Spence and Parr, 1991). Subjects were given a number of tasks to complete, of which the following is representative:

> *You can spend up to £200,000 on accommodation. Locate the best you can with regard to the number of bedrooms and the size of garden, but it must have central heating.*

FIGURE 3.63 Multidimensional icons representing eight attributes of a dwelling

FIGURE 3.64 Textual descriptions of the dwellings represented by the multidimensional icons shown in Figure 3.63

house	flat	houseboat
£400,000	£300,000	£200,000
garage	no garage	no garage
central heating	central heating	no central heating
four bedrooms	two bedrooms	three bedrooms
good repair	poor repair	good repair
large garden	small garden	no garden
Victoria 15 mins	Victoria 20 mins	Victoria 15 mins

The subjects were presented with two conditions. Details of 56 dwellings were presented in either iconic (see Figure 3.63) or textual (see Figure 3.64) form, the layout in both cases being seven rows of eight dwellings (Figure 3.65). The 56 dwellings were so designed that there was a unique solution to all the tasks. Overall it was found that the time taken to identify the appropriate house using icons was about half what it was when using textual description. Observation of the subjects while they carried out the tasks strongly suggested that, if the display had been an interactive one rather than printed on paper, some means of tagging individual dwellings – either to 'discard' them or 'include' them – would have simplified the task. With the paper display, Post-it stickers were employed for tagging.

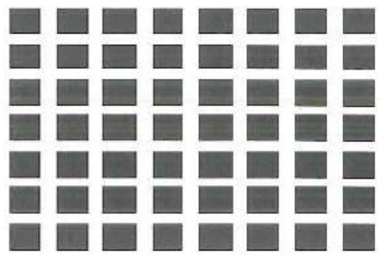


FIGURE 3.65 Layout employed in the experiment carried out to compare the relative merits of iconic and textual descriptions of dwellings

We have discussed two examples of metaphorical icons: in one there is a direct relation between icon and object (the house icon) while in the other there is no direct relation between facial features and the geological attributes they represent. Siirtola (2005) has observed that experiments support the idea that a data-related 'glyph' (icon) is more favoured by users and, with respect to the task of information acquisition, leads to about 13 per cent more accuracy. Direct metaphorical icons find wide application (see, for example, Miller and Stasko, 2001).

7

Object and attribute visibility

We return briefly to the concept of visibility, discussed earlier in the context of different encoding techniques (Teoh and Ma, 2005).[4]

We can say that the property of object visibility is such that each object is represented as a single and coherent visual entity such as a point. Object visibility is desirable when a user is interested in knowing an object's attribute values in many different dimensions and how different objects relate to one another. Examples include Chernoff Faces, multidimensional icons and star plots (Figure 3.66). We do not include static parallel coordinate plots in this set because an object is then represented by a collection of points complemented by a polyline joining those points and it is extremely difficult either to discern a single object's attribute values or to see how objects relate to each other. The same comment applies to the Attribute Explorer.

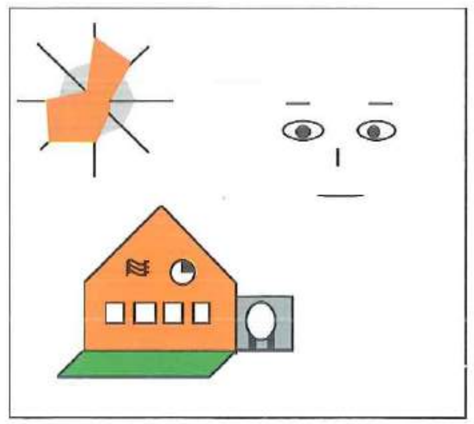


FIGURE 3.66 Representations of multi-attribute objects supportive of object visibility

[4] We use the term 'visibility' in place of 'coherence' used by Teoh and Ma.

A similar form of description applies to attribute visibility, also known as dimension visibility. In a representation with attribute visibility the distribution of objects' attribute values in each dimension is clear. Attribute visibility is desirable when the user wants a clear picture of how objects are distributed in an attribute dimension – for example when there are clusters present. Attribute visibility is associated with the Attribute Explorer and parallel coordinate plots (Figure 3.67), but not with Chernoff Faces, multidimensional icons and star Plots. Teoh and Ma (2005) also introduced the concepts of *object correlation* and *attribute (dimension) correlation*. An example of the latter is provided by a parallel coordinate plot in which points on adjacent axes are connected by lines which clearly indicate a trade-off or correlation. Object correlation may be facilitated by representations such as star plots.

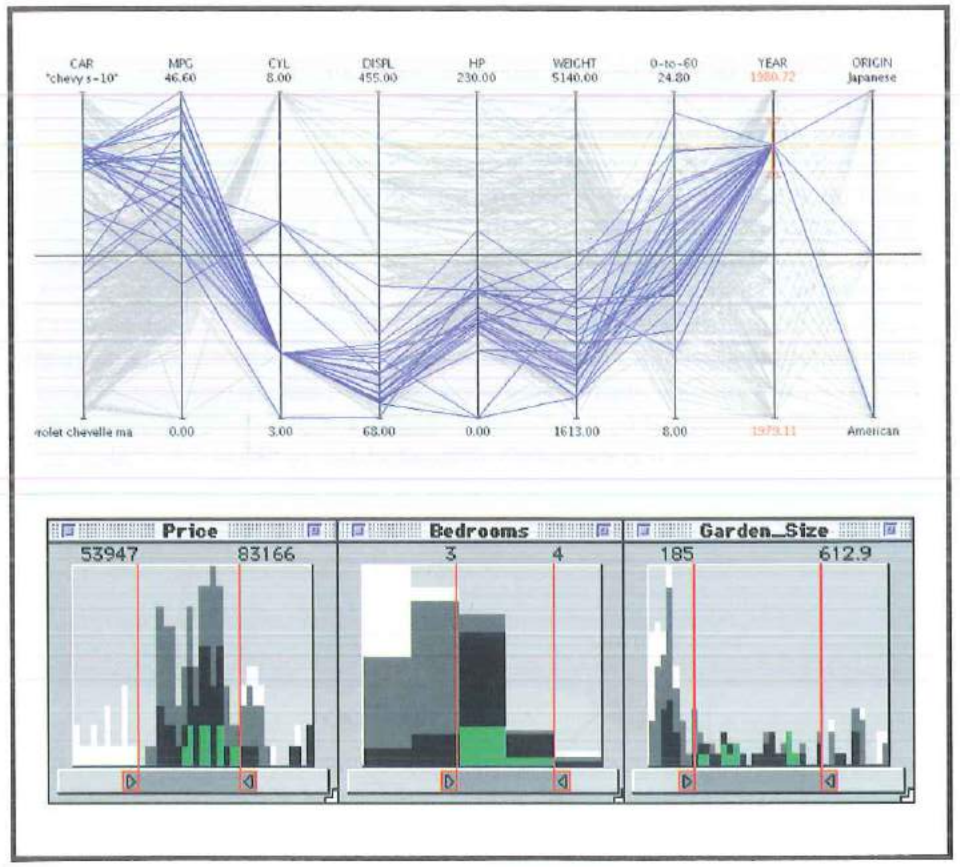


FIGURE 3.67
Representations of multi-attribute objects supportive of attribute visibility

Although the emergence of visibility and correlation as useful characterizations of a representation is quite recent and remains to be validated, it already appears subjectively to be useful and, in addition, seems to have some relation to the 'popping out' property discussed earlier.

3.2 The encoding of relation

In our earlier discussion of data it was established that two types of data exist: *values* and *relations*. We have examined many ways of encoding values; we now ask, 'What do we mean by a relation?' and 'How do we represent it?'.

The dictionary definition says:

relation (n): a logical or natural association between two or more things; relevance of one to another; connection.

A relation between two or more things can be represented in many different ways. A simple straight line or lines can be used to show that John Smith married Mary Robinson (Figure 3.68) or that John borrows money from the Stingy Bank to purchase a 1930 Bentley (Figure 3.69). The relation could be mathematical, of the form $y = f(x)$, and, in turn, could be represented by a node and directed link diagram (Figure 3.70). As shown earlier in Chapter 2, the relation between a car and its price can be represented by a common colour (Figure 3.71). Figure 3.72 summarizes aspects of the warfare in Anglo-Saxon England between 550 AD and 700 AD far more effectively (for me) than an equivalent area of text – the warlike nature of the West Saxons immediately stands out (Arnold, 1997), as does the fact that the Britons continually took quite a beating. Colour and line thickness provide additional scope for encoding aspects of a relation.

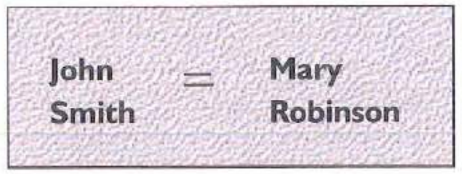


FIGURE 3.68
A simple symbol indicates the relationship of marriage

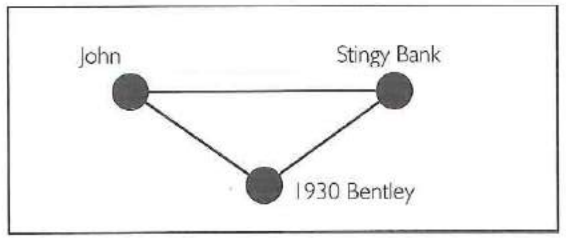


FIGURE 3.69
Lines indicate relationship

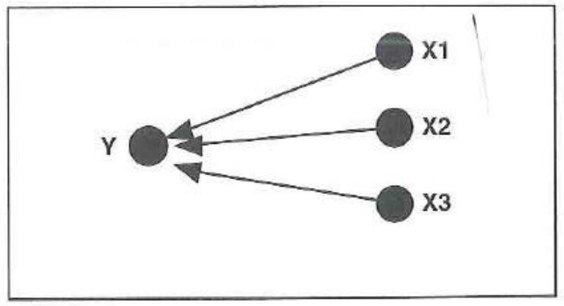


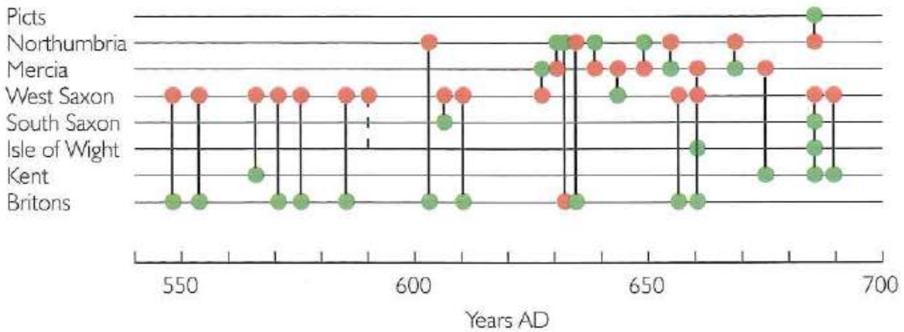
FIGURE 3.70
Arrows indicate unique unilateral functional relations

3.71

a

3.72
ence
e in
o-
gland

d
ed
he



With a relation there is the implication that we are dealing with *discrete* rather than continuous attributes. Whereas a car can have any price, an hotel either has or has not got a swimming pool. To buy my Bentley I apply to Bank A rather than Bank B. The suspect in a fraud case is either in Los Angeles conferring with a colleague or he is not. We are speaking about binary properties. It is therefore not surprising that different techniques may be appropriate to the representation of relations rather than values, since the latter can usually assume more than two conditions. However, the prime consideration in the choice of such a representation is identical to that which applied to value representation: an understanding of the task that is being undertaken, the insight that is sought and what questions might therefore be asked of the relation.

Some representations of relations are very simple, but nevertheless extremely powerful. We begin by examining some of these simple representations and the insights they can provide. As the definition above emphasizes, a relation exists between two or more *things*, so in discussing the representation of relation we must inevitably be concerned with representing the 'things' that are related. 'Thing representation' and 'relation representation' must be considered together.

3.2.1 Lines

Perhaps the simplest way of representing a relation between two entities is to draw a straight line between representations of those two entities. Even a short record of telephone calls (Figure 3.73(a)) is easier to comprehend if represented by a node-link diagram (Figure 3.73(b)), especially if disconnected subsets are present (Figure 3.73(c)). A node-link representation was in fact valuable in the mortgage fraud example mentioned in Chapter 1. If a large number of house purchases take place, involving relations between various lenders of money, solicitors, surveyors, etc., then a representation such as that shown in Figure

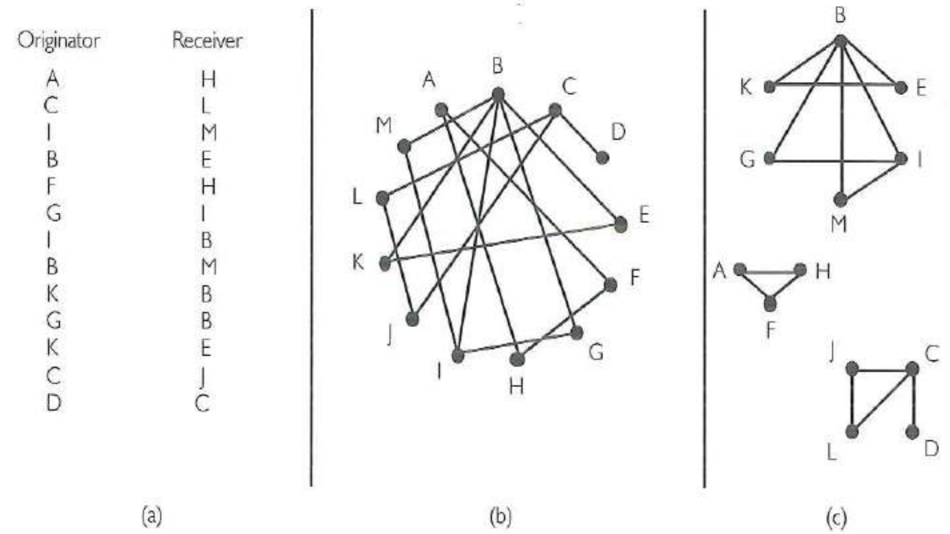


FIGURE 3.73
Insight into even a short list of telephone calls, (a), is enhanced by their node-link representation, (b), especially if disconnected subsets can be identified, (c)

3.74(a) might emerge, where each person or institution involved is represented by a small segment of an annulus, as shown in Figure 3.74(b). By itself the pattern within the inner circle might provide little insight, but if a threshold is now imposed that excludes normal house purchases, a pattern begins to emerge (Figure 3.75) that can provide an investigator with evidence that leads to the arrest of a person perpetrating mortgage fraud (Davidson, 1993; Westphal and Blaxton, 1998).

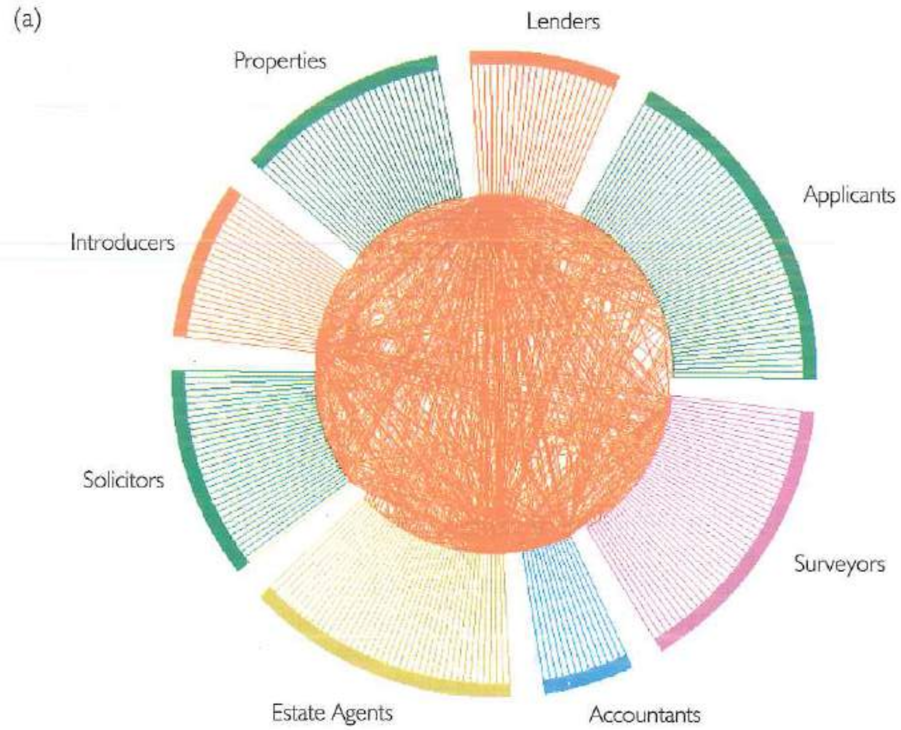
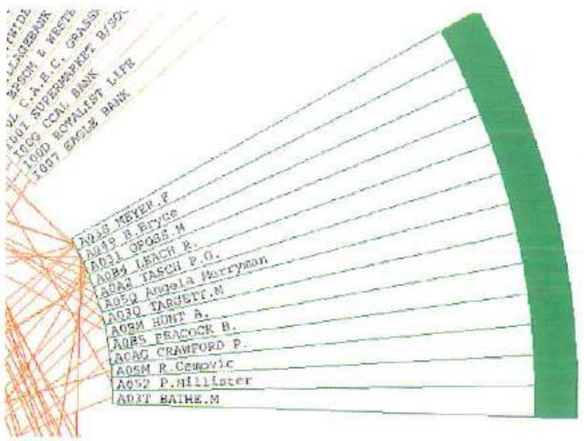
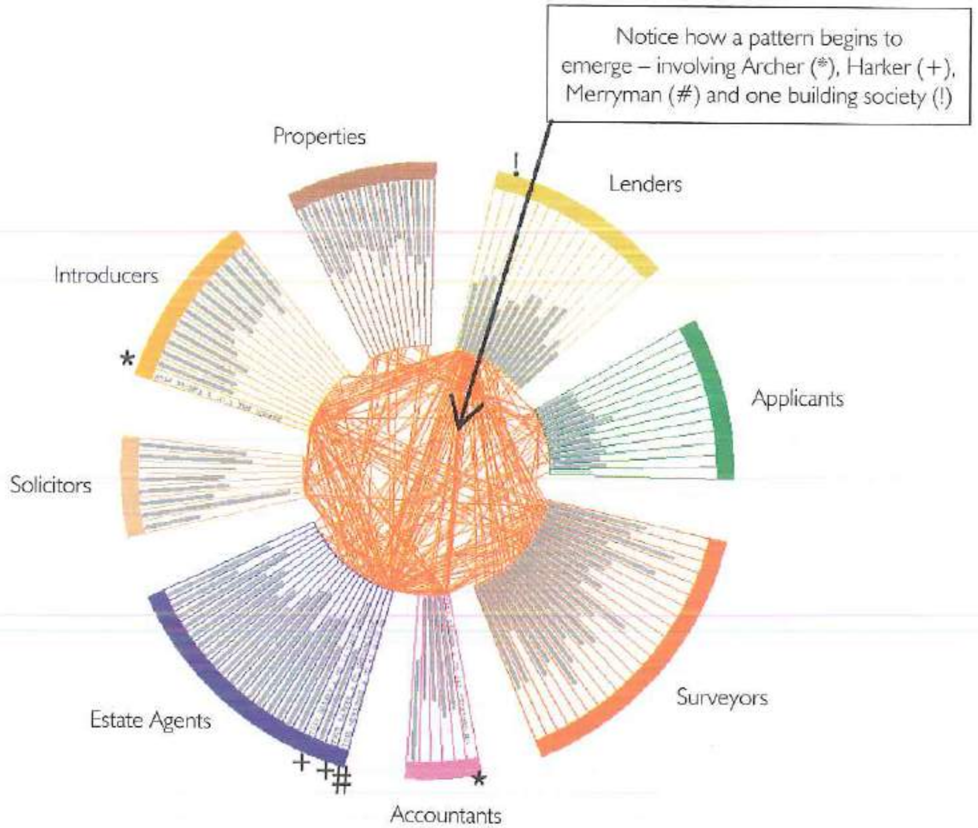


FIGURE 3.74
A representation of mortgage activity (a). Lenders, properties (houses), buyers, etc. are represented by small radial segments of an annulus as shown in (b) overleaf, and their relationships denoted by straight lines



E 3.74
tation
ants by
lial
s within

tation
age
nown in
eaf



E 3.75
data
d in
.74 a
d has
posed
ess the
f

ur. As a
nusual
ur is
by the
formed
nes

Over the last decade the value of representing connections between people, things and institutions has been fully recognized, and exploited by means of powerful visualization tools which support tasks falling under the umbrella heading of intelligence analysis. A good example is provided by Analyst's Notebook (i2, 2006) which is used widely in such areas as law enforcement and forensic investigations. Visualization software of this kind allows the interrogation of data in complex scenarios, the analysis of volumes of seemingly unrelated data and,

not least, the communication of actionable intelligence in a visual format. For some investigations the principal interest may lie in the connection between people, social networks, locations and property, for which an 'association' style chart such as that shown in Figure 3.76 would be appropriate. If – alternatively or additionally – a comprehensive chronology of events over time is of interest in view of the potential for highlighting temporal coincidence which may not be evident in an association style chart, then a 'timeline' style chart such as that in Figure 3.77 may be useful.

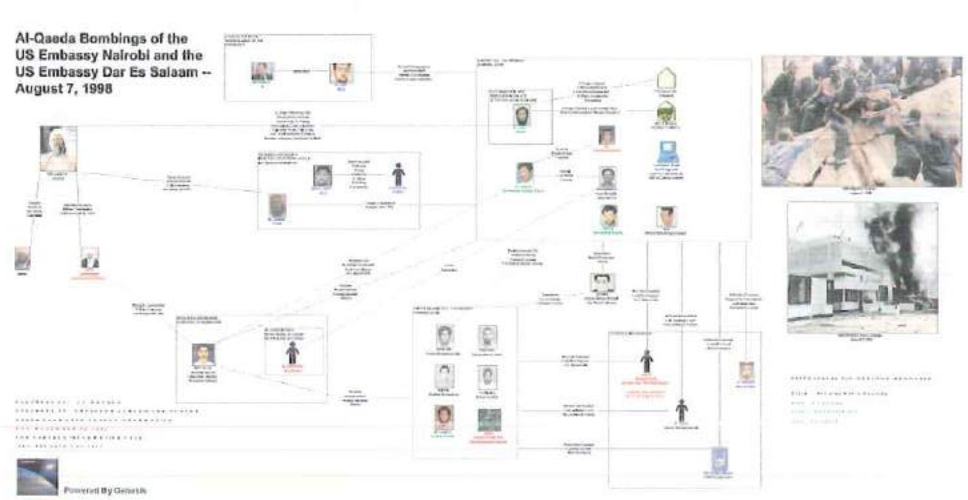


FIGURE 3.76
An 'association' style chart depicting the African bombings
(Courtesy i2 Ltd.)

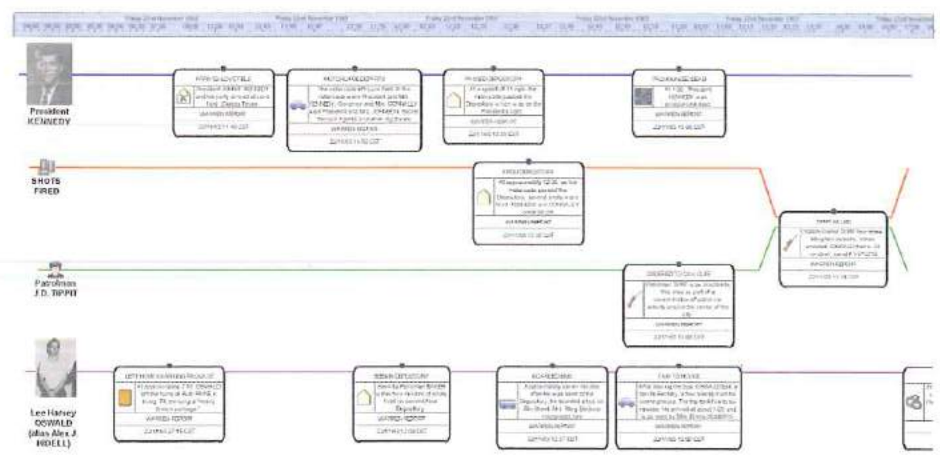


FIGURE 3.77
Part of a 'timeline' style chart depicting the Kennedy assassination
(Courtesy i2 Ltd.)

Perhaps the most familiar use of lines to represent relations, reflecting the term 'connection' in the definition above, is the map of the London Underground (Figure 3.78), a form of representation judged to be so effective that it is now employed by virtually every transportation authority in the world (Ovenden, 2003). It undoubtedly benefits from the shapes into which the lines connecting stations are arranged (indeed, Harry Beck was supposedly influenced by electrical circuit diagrams), as well as from the use of colour to denote differ-

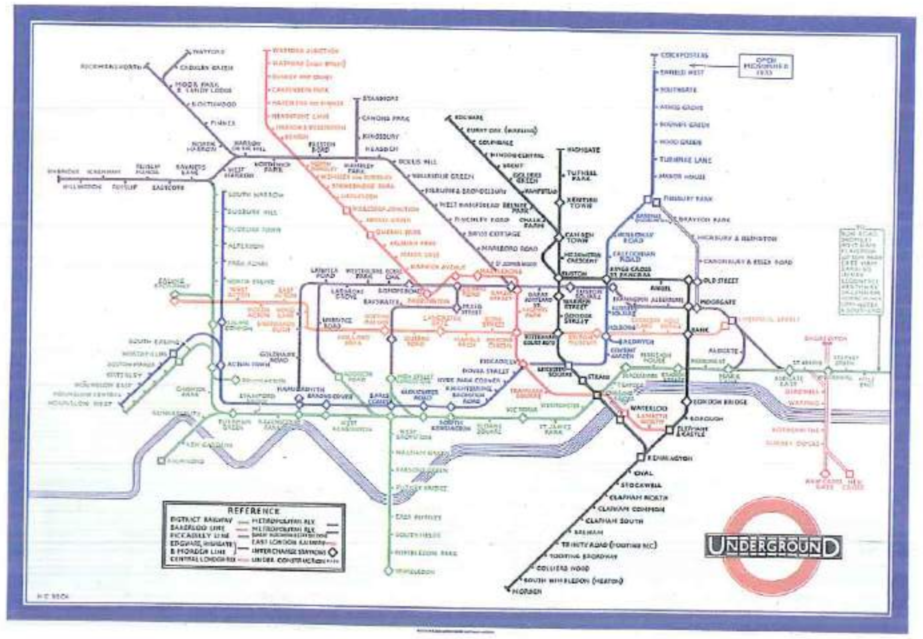


FIGURE 3.78 Harry Beck's original London Underground map (© Transport for London)

FIGURE 3.79 The Underground map in use prior to the introduction of Harry Beck's version (© Transport for London)

ent underground lines and symbols to denote both ordinary and interchange stations (Garland, 1994). Often, coloured regions having implications for fares payable can be provided as background. Another influence on the nature of the map is its general – though not necessarily accurate – geographical veracity: it would not matter if the separation between stations was not proportional to their geographical separation but it would significantly reduce the value of the map if, for example, Wembley (of football fame) in North London were placed towards the bottom and Wimbledon (known for tennis) in South London were placed towards the top. Although I know of no relevant usability study, the current map would appear at first sight to be easier to comprehend than the one in use immediately before Harry Beck's design was introduced (Figure 3.79).

The term *relation* often carries the connotation of interactions between human beings. We are familiar, for example, with the conventional family tree, but that focuses principally on formal relationships cemented by marriage, including the birth and death of children. The use of simple lines can, however, be broadened to usefully represent what Freeman (2000, 2005) has termed 'the consequences of the social nature of social animals'. Social network analysis is concerned with the structural patterning of the ties that link social actors. Two kinds of patterns are of particular concern: (1) those that reveal cohesive social groups, and (2) those that reveal the social positions, or roles, of individual actors.

The representation of social relations by a network is not new: over 70 years ago Moreno (1934) presented a network (Figure 3.80) showing the social choices made by fourth graders in a school, from which it is clear that, overwhelmingly, boys chose boys and girls chose girls (a more recent comparative study would be interesting!). More recently Marbella Canales, a colleague of Freeman (2005), collected data on the recreational social connections among the employees in the cosmetics department of an upmarket department store. Her data revealed the more or less linear pattern of connections shown in Figure 3.81(a) from which it is clear that the interaction was not random but rather patterned

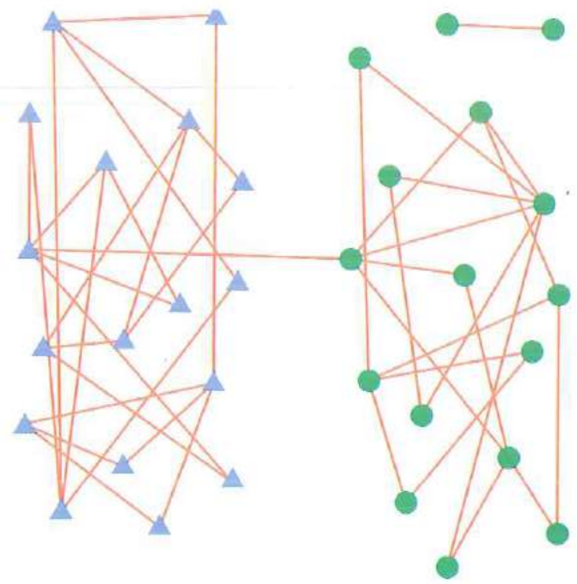


FIGURE 3.80 The social choices of fourth-grade students (after Moreno, 1934)

(Freeman, 2005). She then sought to discover the underlying basis of that pattern by colouring the nodes in the network according to various characteristics of the individuals involved. She tried sex, ethnicity, marital status and other traits that might lead to individuals choosing one another as recreational partners. Most did not work: Figure 3.81(b), for example, shows the individuals coloured in terms of their marital status: married individuals are yellow, singles are red. Because both red and yellow are spread throughout the image it is clear that these individuals did not choose others according to their marital status. When the investigator explored the age of the individuals, however, a very different picture emerged (Figure 3.81(c)). Blue points represent people who are 30 or younger, those between 30 and 40 are yellow and those 40 and older are red. Their connectivity shows that the individuals chose recreational partners on the basis of similarity in age.

FIGURE 3.81(a)
Social choices among department store employees
Source: *L.C. Freeman*

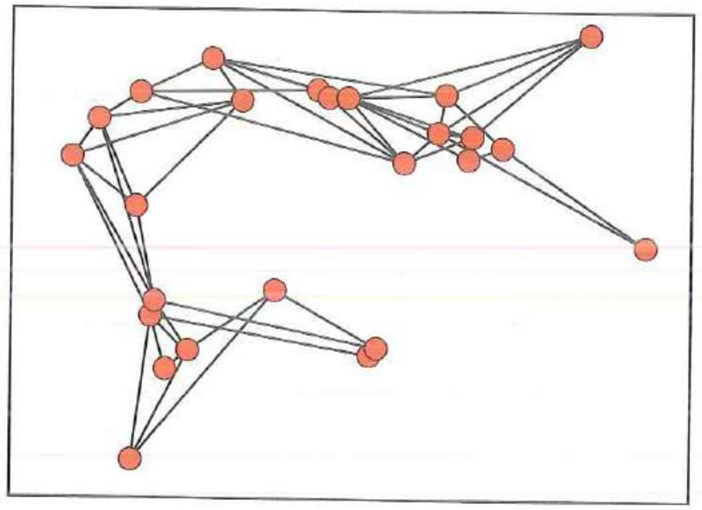
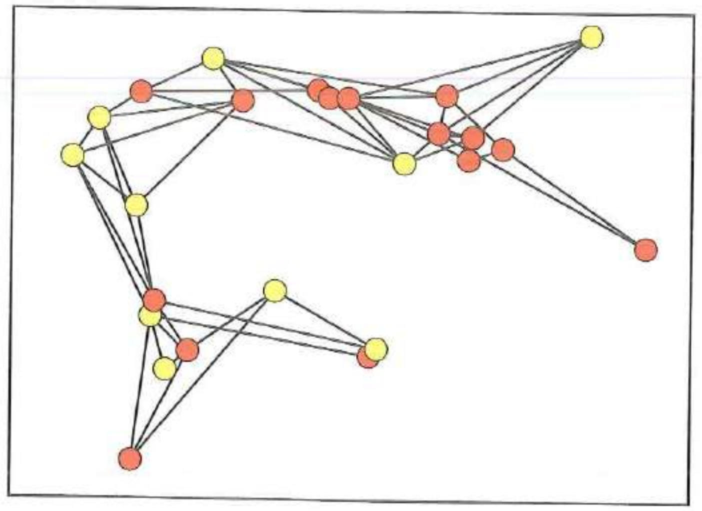


FIGURE 3.81(b)
Social choices among department store employees, with marital status encoded
Source: *L.C. Freeman*



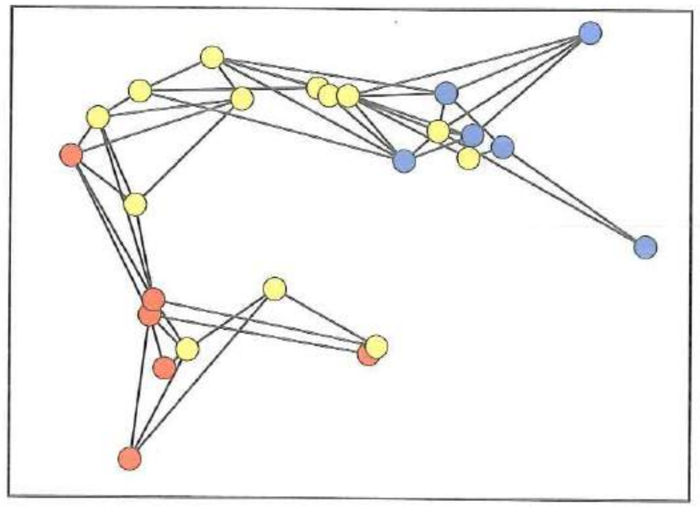


FIGURE 3.81(c)
Social choices among department store employees, with age range encoded (blue <30, 30<yellow <40, red >40)
Source: *L.C. Freeman*

3.2.2 *Maps and diagrams*

Venn diagrams

The definition of relation refers to 'logical association'. A very simple set of logical relations is contained in Table 3.2 which indicates, for example, that hotel B has a swimming pool *and* a golf course but *not* a restaurant. A search for a desirable hotel within a much larger table would be tedious and time consuming and could be much easier through the use of a familiar Venn diagram. The Venn dia-

TABLE 3.2 Facilities offered by eight hotels

Hotels	Swimming pool	Golf course	Restaurant
A	●	●	●
B	●	●	
C	●		●
D	●		
E		●	●
F		●	
G			●

gram pertinent to the hotel information in Table 3.2 is shown in Figure 3.82: for a much larger collection of 24 hotels the corresponding Venn diagram (Figure 3.83) *might* quickly allow a traveller to identify the hotels which satisfy his or her needs. For example, it can be seen that if all facilities are required (swimming pool + golf + restaurant) then four hotels are candidates. Interaction can add considerable value to a Venn diagram.

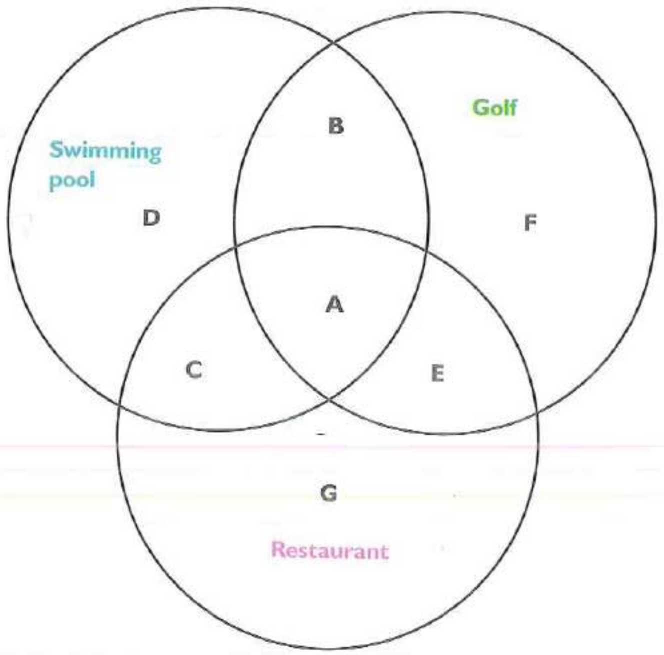


FIGURE 3.82 A Venn diagram representation of the hotels listed in Table 3.2

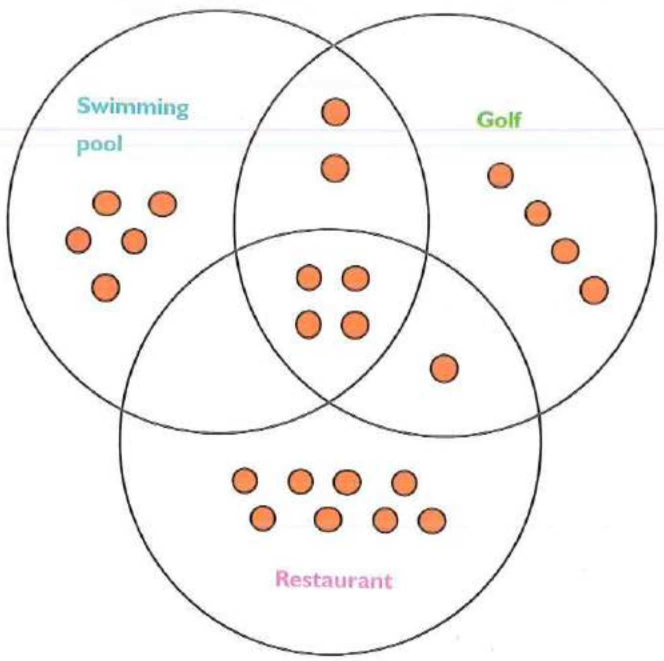


FIGURE 3.83 A Venn diagram representation of the attributes of 24 hotels

InfoCrystal

An improvement on the Venn diagram for the representation of logical relations was proposed by Spoerri (1993). Figure 3.84 shows a progression from a three-attribute Venn diagram to a three-attribute InfoCrystal. All possible logical relations between the three attributes are represented by regions within the crystal. The same colours are used to represent groups of houses as are used in the Attribute Explorer example of Figure 3.58. Thus, the region indicated by an asterisk in Figure 3.84 corresponds to, and can be used as a means of selecting (i.e. as a visual query), those houses which satisfy requirements on price and garden size, but not the number of bedrooms. Spoerri, however, complemented the simple representation of Figure 3.84 by adding interior icons (Figure 3.85) which indicate, by their shape, the number of criteria satisfied (circle = 1, rectangle = 2, triangle = 3) and, by an inscribed number, the number of items in that class. Thus, for the 24 hotels represented in Figure 3.83, four offer all facili-

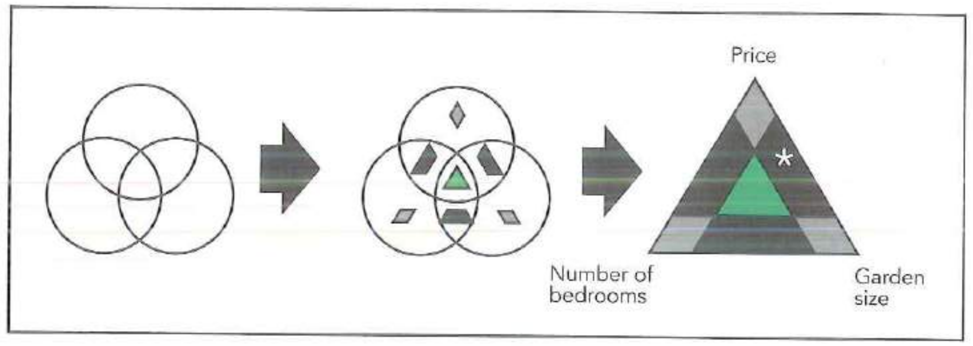


FIGURE 3.84 The development leading from a Venn diagram to an InfoCrystal. The InfoCrystal illustrated allows visual queries to be made concerning price, garden size and number of bedrooms (see the Attribute Explorer of Figure 3.58) The asterisk represents houses satisfying criteria on price and garden size but not number of bedrooms

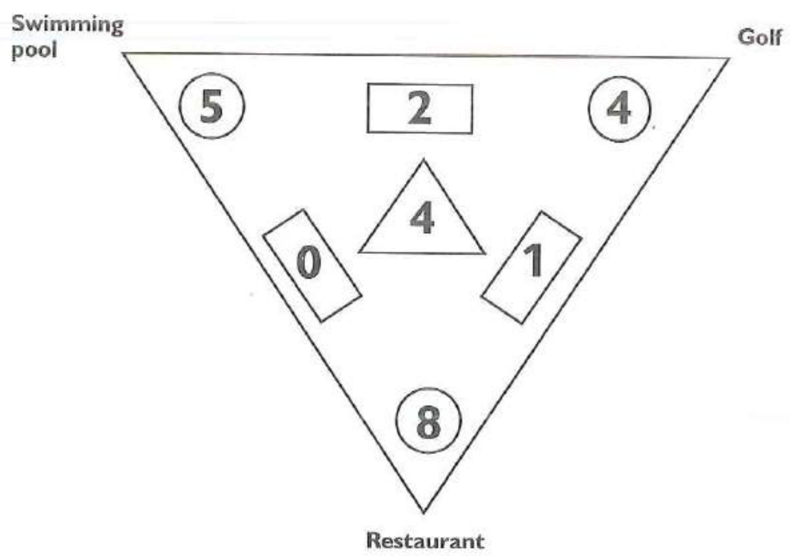


FIGURE 3.85 An InfoCrystal representation of the hotel data of Figure 3.83

ties and the number which offer both golf and a restaurant but no swimming pool is one. Spoerri's InfoCrystal can be extended to more than three attributes, but above four the effort needed to comprehend the representation and to readily formulate a query is probably such that only the specialist user would be able to do so.

Cluster maps

Figure 3.86 shows a cluster map (Fluit *et al.,* 2003) relevant to the hotel data represented by the Venn diagram of Figure 3.83 and the InfoCrystal of Figure 3.85. The labeled nodes represent the three attributes and carry a label identifying that attribute (e.g. 'restaurant') followed by a number indicating how many hotels possess that attribute. Note that the three numbers do not add up to 24, the number of hotels, because a hotel with a restaurant and a gym, for example, will contribute to the count on both those nodes. Associated with each node are one or more circles containing a number of yellow circles, each representing a

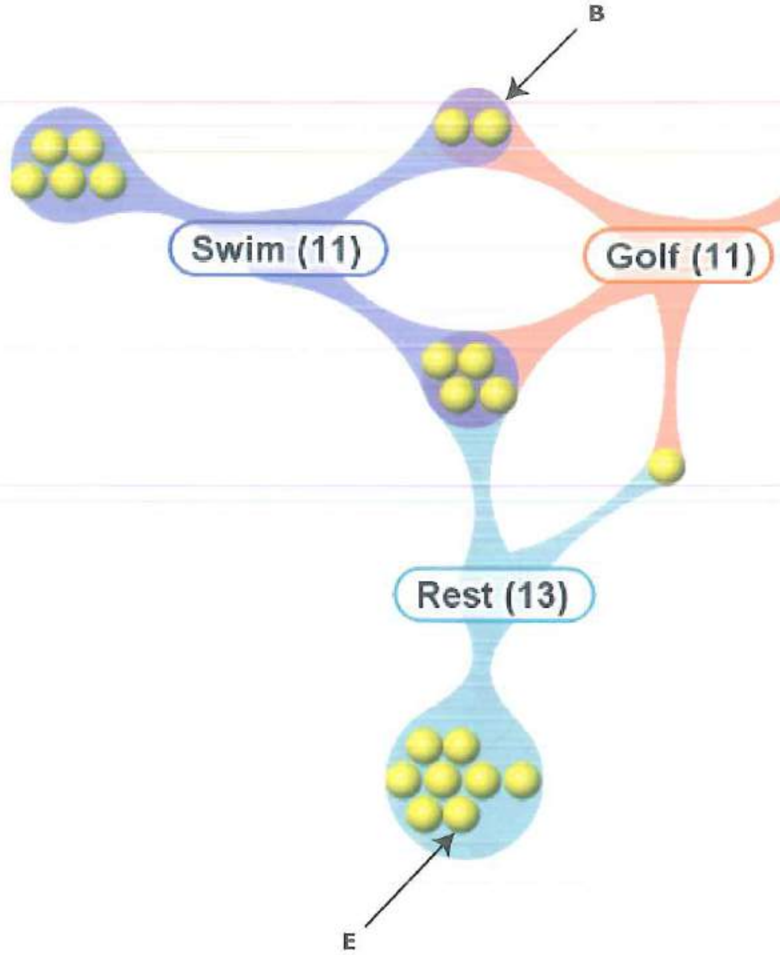


FIGURE 3.86 A cluster map representation of the 24 hotels represented in the Venn diagram of Figure 3.83 and the InfoCrystal of Figure 3.85 *(Courtesy Christiaan Fluit, Aduna)*

hotel. Each hotel is placed in a circle connected to all the attributes that the hotel possesses. Thus, hotel E has only a restaurant, so is placed in the circle associated only with the restaurant node, but hotel B appears in a circle attached to both the swimming pool node and the golf node because it possesses those two, and only those two, facilities. It is therefore reasonably obvious from the cluster map of Figure 3.86 that there are no hotels which possess both a swimming pool and a restaurant but no golf course. Mouse-over action can disclose more details of each hotel.

The test of any technique which represents logical data is how well it scales to many classes. Some impression of the cluster map's capability in this respect can be gained from an example representing 24 hotels each characterized by *four* attributes (Figure 3.87). The absence of any circular area connected solely to the gym node means, for example, that any hotel that has a gym will always possess at least one of the other three facilities. A cluster map can support many different kinds of interaction including the highlighting of all circles associated with a class node and the animated modification of a map following the inclusion of a new class (Fluit *et al.,* 2003).

3.2.3 *Tree representations*

In all the examples of relation representation we have discussed so far there has been no restriction upon what is connected to what other than it be meaningful.

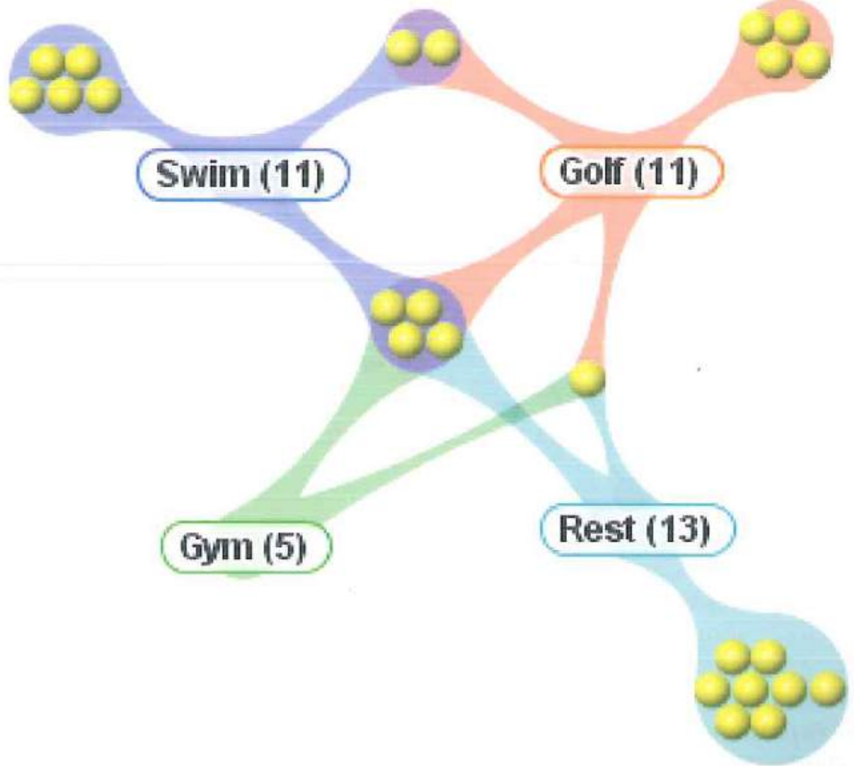
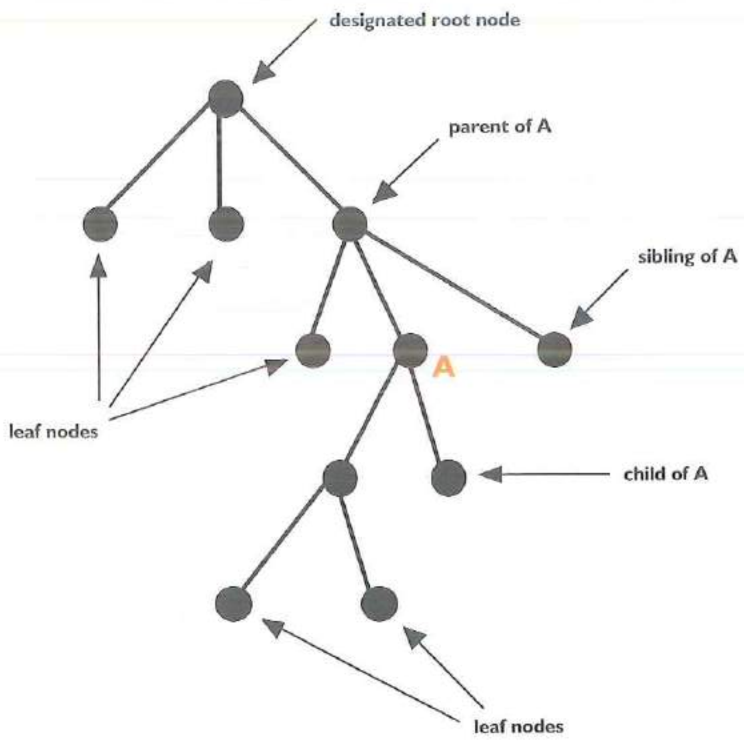


FIGURE 3.87 A cluster map representation of 24 hotels, each described by four attributes *(Courtesy Christiaan Fluit, Aduna)*

The term 'network' is often applied to a node-and-link type of relation to which no topological restrictions apply. Examples include the simple network of Figure 3.73 representing telephone calls and the far more complex network of lines representing mortgage arrangements in Figure 3.74. There is, however, a class of networks having significant practical application in which the topological restrictions that apply allow a new family of representations to be derived. The restriction is that the network be a tree.

The mathematical definition of a tree is a network of nodes and links so connected that no loops are present. In other words you cannot start at a node and trace a path which arrives back at the same node without a direct retrace of steps. The reason for the considerable interest in trees is that many situations are described by hierarchical relations. To introduce relevant terminology we examine the simple tree of Figure 3.88. One node has been designated the root node: we use the term 'designated' because you can take hold of *any* node of a tree and call it the root node. In practice, however, a tree often refers to a hierarchy in which the designated root node is the president of a company or perhaps a department store comprising a number of product lines. If there is indeed a hierarchy, then all nodes except the root node are associated with a superordinate node and, if associated with subordinate nodes, these are often referred to as 'children'. If a node has no subordinate node it is called a 'leaf node'.

FIGURE 3.88
A tree



If some data describes a tree, the immediate question is how such data can usefully be represented and explored. What is wrong, however, with the representation shown in Figure 3.88? If it is satisfactory we need look no further!

Cone tree

The problem with the tree representation of Figure 3.88 is that in most practical cases the tree has many levels: even with an average fan-out (i.e. the number of subordinate nodes associated with a node) as small as three, the space needed to display a tree can become huge, mainly in the horizontal direction. A solution is to imagine the original tree (Figure 3.89(a)) to be rearranged into a 3D structure such that all nodes subordinate to a given node are arranged in a horizontal circle which, together with the superordinate node, forms a cone, as shown in Figure 3.89(b). The resulting 2D view of that structure, called a cone tree (Robertson *et al.*, 1991), is now more compact than the original representation of Figure 3.89(a) and, notwithstanding some occlusion, is easier to handle. The user of a cone tree may, for example, wish to see the reporting path of an employee within the organization represented by the cone tree, in which case entry of the employee's name by some means will bring about any cone rotations needed to position that employee and his reporting path to the foreground. A horizontal orientation of the cone tree, called the cam tree, may be more convenient for the display of node names (Figure 3.90). No record is available of any evaluation of the usability of the cone tree concept.

9

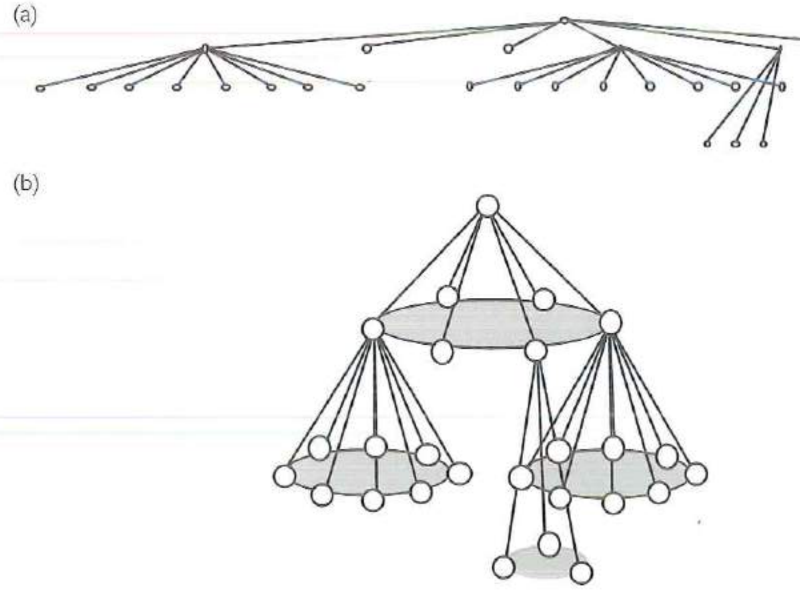
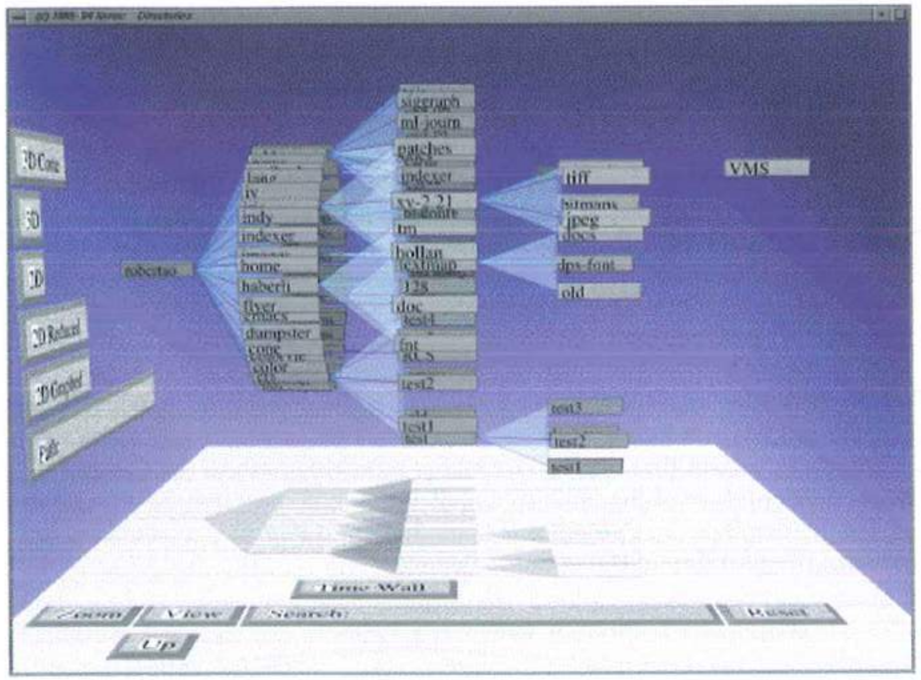


FIGURE 3.89
(a) A tree, (b) the corresponding cone tree

Tree maps

An alternative representation of a tree is the tree map (Johnson and Shneiderman, 1991). Its derivation from the original tree representation is straightforward and is illustrated in Figure 3.91. Starting with the designated root node one draws a rectangle passing through that node. Usually, to make efficient use of display area, the rectangle will be made as large as conveniently possible. Within that rectangle are smaller rectangles, one for each of the nodes that are immediately subordinate to the root node. This construction is repeated until all nodes are accounted for. There is no constraint (except the resolution of

RE 3.90
rientation
cone
more
enient for
xtual
ing of
s

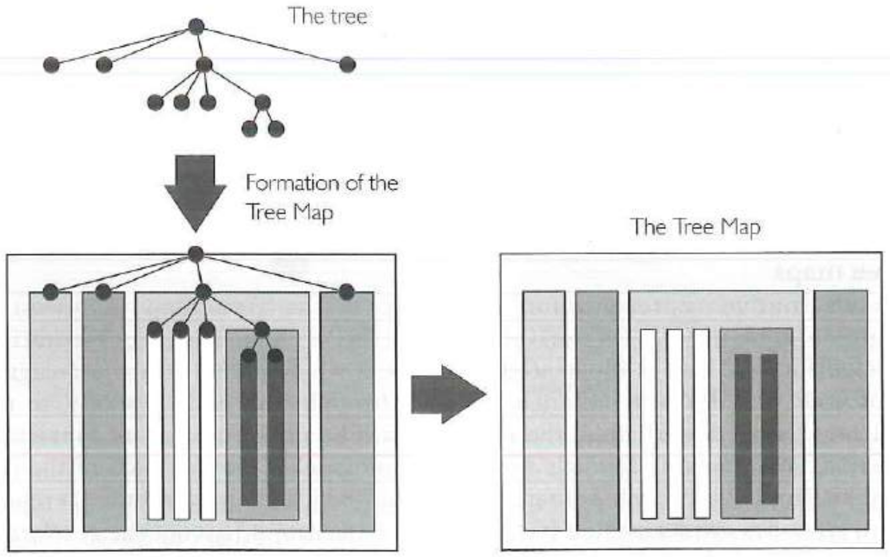


the display) on the depth of the tree and no requirement that all leaf nodes are at the same level or that the fan-out of every node is the same. Once the tree map is derived, colour coding, for example, can be employed to characterize different parts of the tree according to the use to which it will be put.

The main disadvantage of the tree map construction illustrated in Figure 3.91, but one which is easily overcome, is that for typical trees the result is a large number of very thin rectangles in which it is difficult, for example, to display text. The simple solution, illustrated in Figure 3.92, is a 'slice-and-dice'

RE 3.91

truction of
e Map



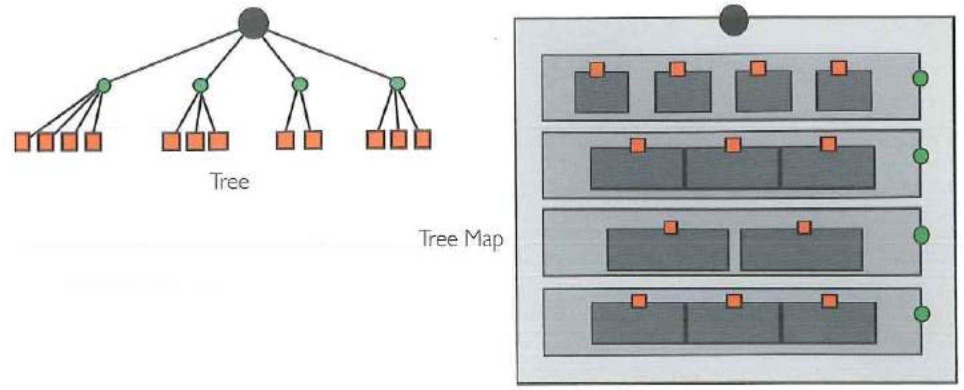


FIGURE 3.92
The 'slice-and-dice' construction of a Tree Map to obtain leaf nodes represented by rectangles more suited to the inclusion of text and images

approach in which the generated rectangles are drawn, alternately, vertically and horizontally from the nodes at successive levels. A typical result is shown in Figure 3.93 which is an author's collection of reports. An advantage of a tree map is that, as in Figure 3.93, it supports an awareness of leaf nodes: a disadvantage is that the hierarchy is not easy to discern.

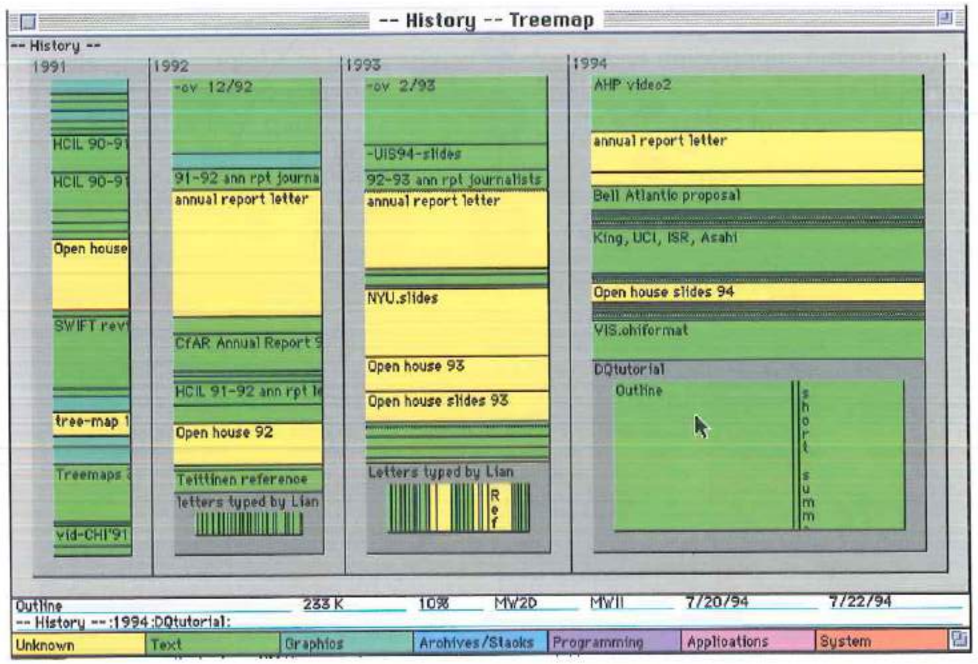


FIGURE 3.93
Tree map display of an author's collection of reports
(Courtesy of Ben Shneiderman)

The tree map technique has found a variety of applications. One is the representation to be found on the website Smartmoney.com (Figure 3.94). Major industrial sectors (e.g. energy, healthcare) are each represented by an area which, in turn, contains other areas representing relevant companies. As illustrated, there are many opportunities for encoding by colour and area, as well as for interaction by mouse-over and the selection of further detail by mouse click.

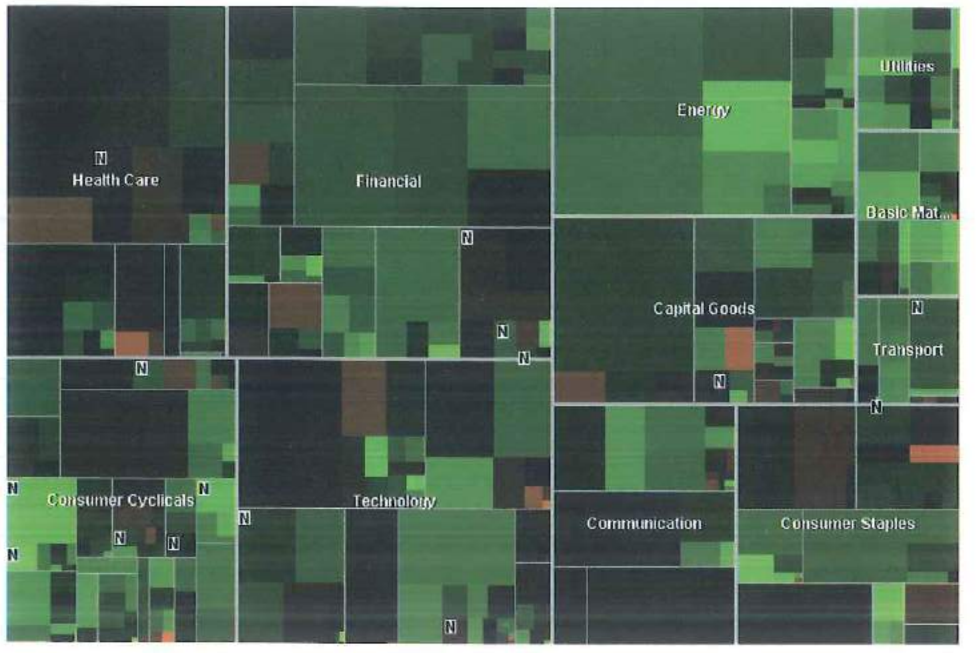


FIGURE 3.94 The Smartmoney.com website showing the status of companies within a number of sectors

Distortion and suppression techniques to be discussed in Chapter 4 may also be appropriate. Figure 3.95 provides an example showing how the selection of attribute limits can control the view provided by a tree map, and illustrates once

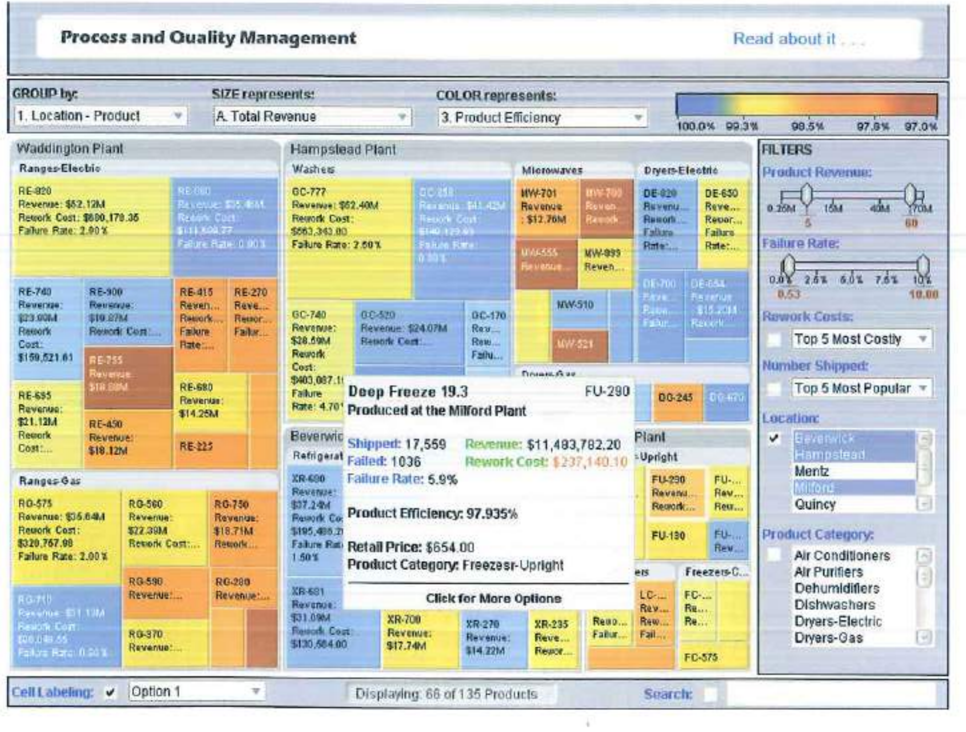


FIGURE 3.95 A tree map presentation of economic information, allowing filtering according to attributes *(Permission from HIVE)*

again the potential offered by combinations of techniques (see, for example, Csallner *et al.*, 2003).

Hyperbolic browser

An ingenious technique whereby an entire tree can be kept within the confines of a circular area on a conventional display screen was invented in 1994. Without going into the sophisticated mathematical detail involved (Lamping *et al.*, 1995; Lamping and Rao, 1994, 1996), the method is based on a hyperbolic geometric transformation which leads to all nodes of the tree being located within a specified area: the resulting appearance is the form shown in Figure 3.96. The designated root node is initially in the centre of the display, its immediate subordinate nodes are distributed around it at a particular distance, but as the number of levels separating a node from the root increases, the separation between a node and its parent decreases and the size of the node also decreases, in such a way that all nodes fall within the circular display area. A practical limit to the display of all nodes is imposed by the resolution of the display: drawing of the tree stops below one-pixel resolution.

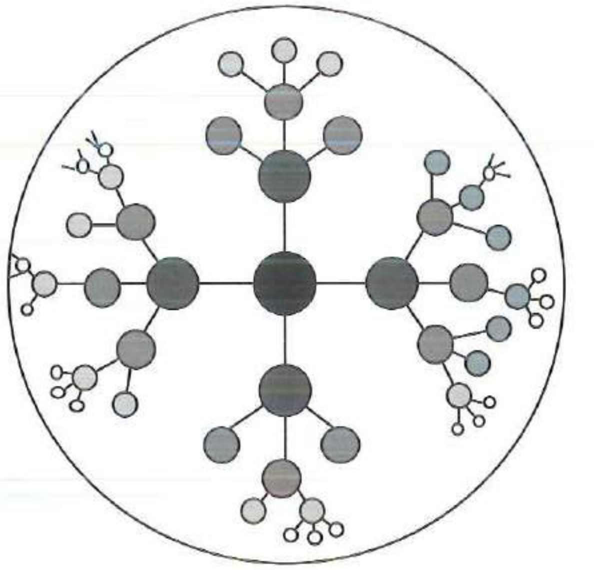
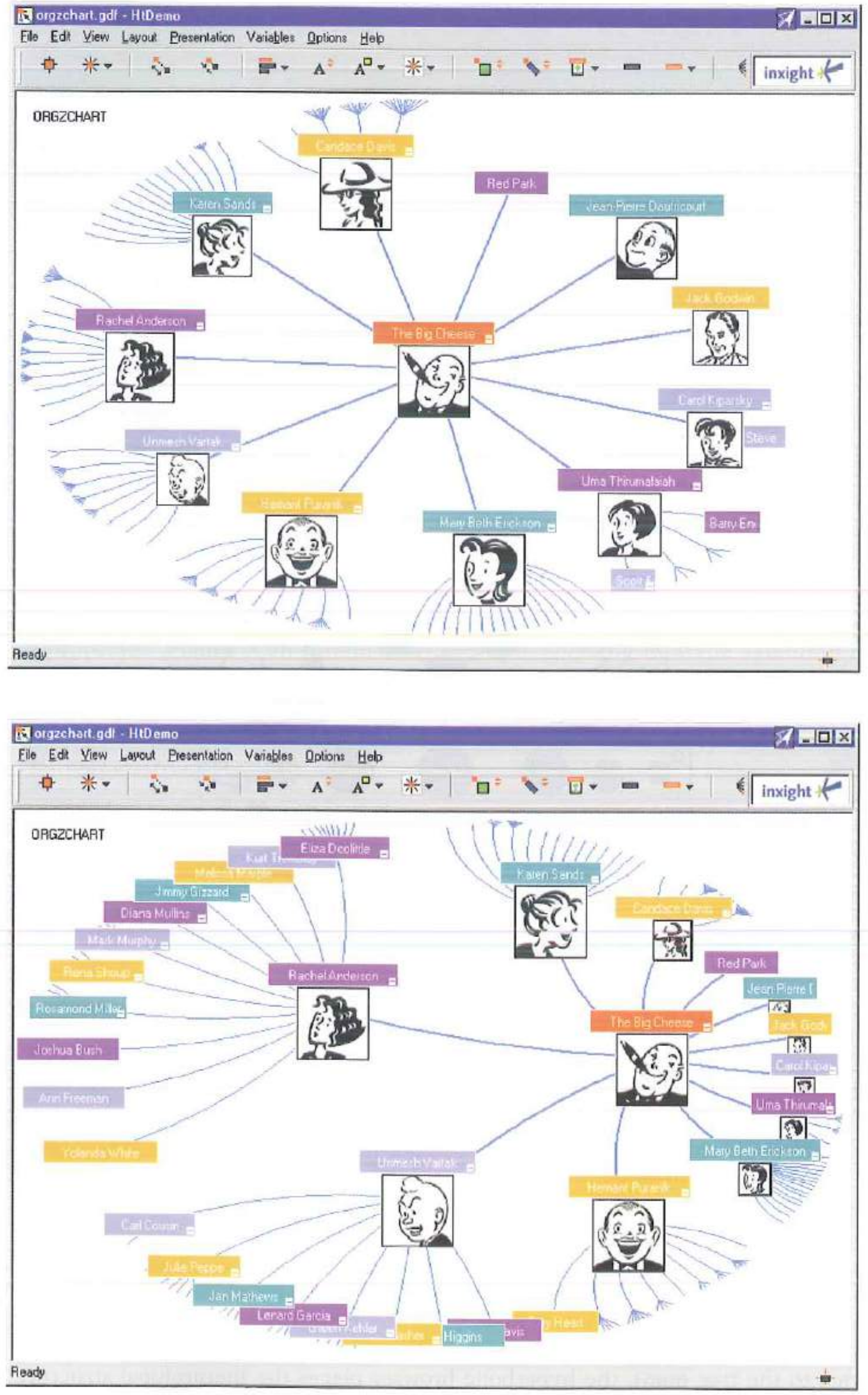


FIGURE 3.96 A sketch illustration of the hyperbolic browser representation of a tree. The further away a node is from the root node, the closer it is to its superordinate node, and the area it occupies decreases

It is the interactive nature of the hyperbolic browser that is its principal advantage. Any node of interest can be moved smoothly towards and into the central position, whereupon its subordinate nodes follow in roughly the same direction. Thus, by *smooth* movements of the tree within the available display area, relevant regions of interest can easily be explored. Figure 3.97(a) shows a hyperbolic display for the employees in an imaginary department: interactive movement of Rachel Anderson towards the centre pushes her superior away from the centre and reveals the employees who report to Rachel (Figure 3.97(b)). In turn, the movement of Eliza Doolittle towards the centre would reveal details of the people who report to her. As with the cone tree (but in contrast to the tree map), the hyperbolic browser places the hierarchical structure

FIGURE 3.97
(a) The reporting structure of the employees of a company.
(b) One employee of interest, Rachel Anderson, has been moved towards the centre, revealing her subordinates



(a)

(b)

in evidence. But again in contrast to the tree map, the leaf nodes are not primarily in evidence unless dragged towards the centre. The application of the hyperbolic browser to the Library of Congress is illustrated in Figure 3.98.

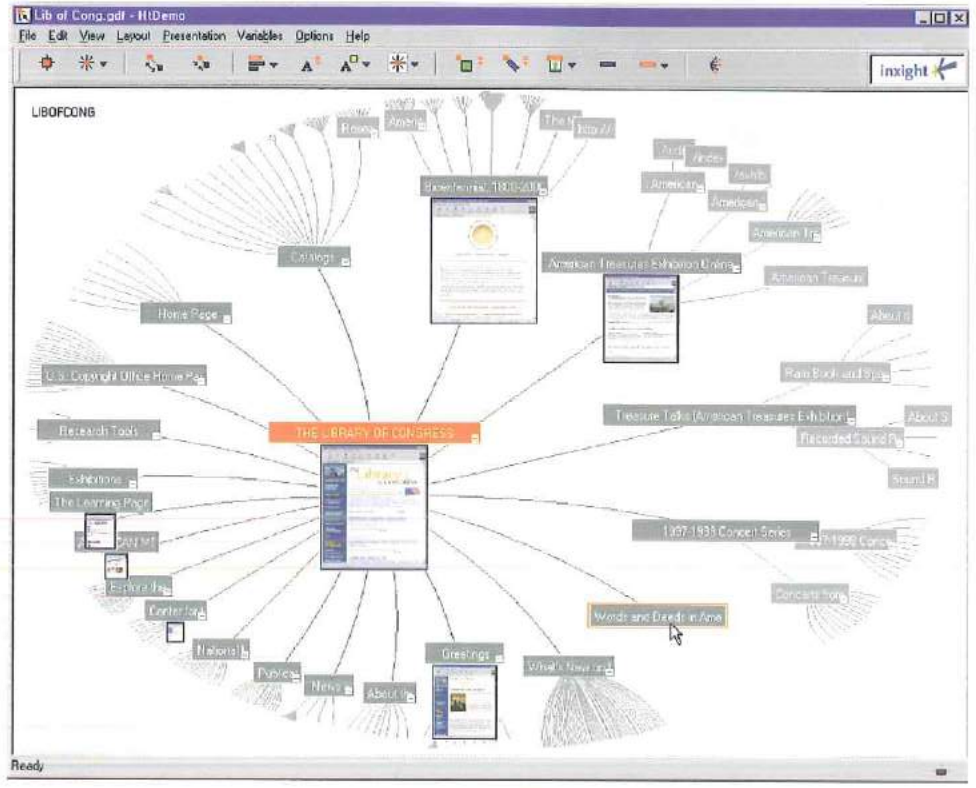


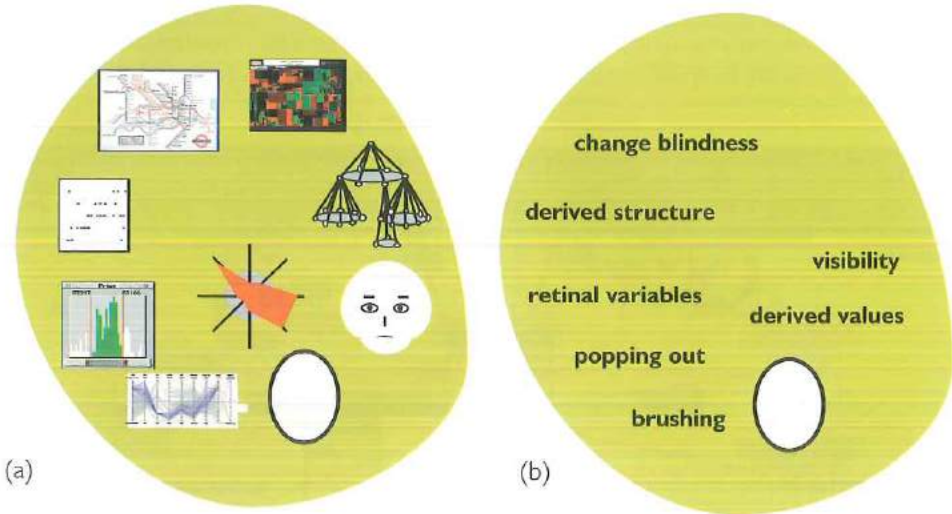
FIGURE 3.98
Representation of the Library of Congress by the hyperbolic browser

3.3 Support for design

The design of even the simplest visualization tool involves creativity. And, like the conventional artist who works in oils, the interaction designer has a number of palettes available. On one (Figure 3.99(a)) are some techniques drawn from an extensive collection that can be used for the representation of values and relations, each having its own advantages and disadvantages. On another (Figure 3.99(b)) is a collection of concepts of potential value when selecting a representation technique appropriate for a given application.

But two other palettes are needed before effective interaction design can occur. One, to be discussed in the following chapter, has to do with *presentation*, concerning the manner in which representations can be presented within the constraints of limited display area and time. The other, the subject of Chapter 5, concerns *interaction* – again, concepts and techniques are available for selection by the interaction designer to complement the benefits to information visualization offered by representation and presentation.

E 3.99
s'
e for the
tion
er:
niques,
cepts



(a) (b)

Exercises

The value of many of these exercises can be enhanced by discussing the results in open class.

Exercise 3.1 (Review)

Make a list of all the methods you can recall for encoding data visually and write brief comments about the circumstances in which they are particularly valuable. Comment upon whether, in their static form (i.e. without interaction), they support object and/or attribute visibility and correlation.

Exercise 3.2

Five students have taken exams in eight subjects and for each subject a mark out of ten has been assigned. Make a list of the questions that might be asked of this data by (a) a student, (b) a parent, (c) a subject teacher and (d) the headteacher. Aim for a total of at least ten questions. Write the questions on Post-its and stick them on the wall for reference during Exercise 3.3.

Exercise 3.3

The performance of the students mentioned in Exercise 3.2 is as follows:

5 STUDENTS (A, B, C, D AND E) SIT 8 EXAMS.
THE MARKS OUT OF 10 ARE AS FOLLOWS:

Art	10	1	5	3	2
Science	1	10	5	4	8
History	8	5	7	1	1
Sport	2	9	5	10	4
Physics	–	1	2	3	1
English	2	8	6	8	5
Chemistry	4	1	1	1	4
Mathematics	10	1	5	4	2
	A	**B**	**C**	**D**	**E**

Without using a computer, sketch one static representation of this data: no interaction with the representation is to be considered. Then see whether it answers any of the questions identified in Exercise 3.2.

For the questions identified in Exercise 3.2, and whose answer cannot easily be found, see whether you can make a useful modification to the representation, still without employing interaction.

Exercise 3.4

Compose a mosaic plot representation of the Titanic data (Table 3.1) but using a different sequence of steps (for example: survival -> gender -> class -> adult/child). List the observations that can readily be made from these representations. Are they different from those triggered by the representations derived in Figure 3.61?

Exercise 3.5

The London Underground map contains no distance or journey-time encoding. With sketches, suggest how this data can be represented. Would it be useful? What other data could usefully be encoded?

Exercise 3.6

At first sight, Florence Nightingale's 'rose plot' appears to have much in common with a star plot. Distinguish between the two, commenting upon the significance that can, or cannot, be associated with the enclosed area of a star plot.

Exercise 3.7

By means of sketches, suggest alternative ways in which the 'hits' returned by Google could be represented.

Exercise 3.8

For your school or university or department (real or imaginary), design a representation of scholastic achievements (e.g. marks in 12 subjects in five year groups) that will show not only the general level obtained but also (1) the way in which the achievements are changing (i.e. first derivatives), (2) the proportion of students obtaining better than a pass mark and (3) the number of students taking a particular subject. Design the representation so that it can be printed on a card that slides easily into the pocket (e.g. one-third of A4).

Exercise 3.9

Study and understand one of the following representation techniques that have not been discussed in this chapter, and prepare a ten-minute presentation in which its features are described and critically evaluated:

(1) Keim, D.A., Hao, M.C., Dayal, U. and Hsu, M. (2002) Pixel bar charts: a visualization technique for very large multi-attribute data sets, Information Visualization, 1, 1, pp. 20–34.

(2) Havre, S., Hetzler, E., Perrine, K., Jurrus, E. and Miller, N. (2001) Interactive visualization of multiple query results, IEEE, Proceedings Information Visualization, pp. 105–112.

(3) Yang, J., Ward, M.O. and Rundensteiner, E.A. (2002) InterRing: an interactive tool for visually navigating and manipulating hierarchical structures, IEEE, Proceedings Information Visualization, pp. 77–84.

(4) Havre, S., Hetzler, B. and Nowell, L. (2000) Theme river: visualizing theme changes over time, IEEE, Proceedings Information Visualization 2000, pp. 145–154.

(5) Geons, as described in Colin Ware's book *Information Visualization: Perception for Design* (2004).

Exercise 3.10

Suggest possible static (not interactive) representations for human relationships (including marriage, births, deaths), test them on real examples and identify the advantages and disadvantages of each.

Exercise 3.11

Bus, metro and train routes are typically represented by lines between nodes. However, with some large cities (London, for example) there are so many routes that it is not easy to plan a journey, especially if it involves intermediate changes. Explore the potential of adding, to the node-link route representation, some overall directional indicators to give a 'first glance' suggestion as to which route might be appropriate.

Exercise 3.12

Select one of the folders on your laptop which contains at least two levels and draw a tree map representation of its contents.