3.71

3.72

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

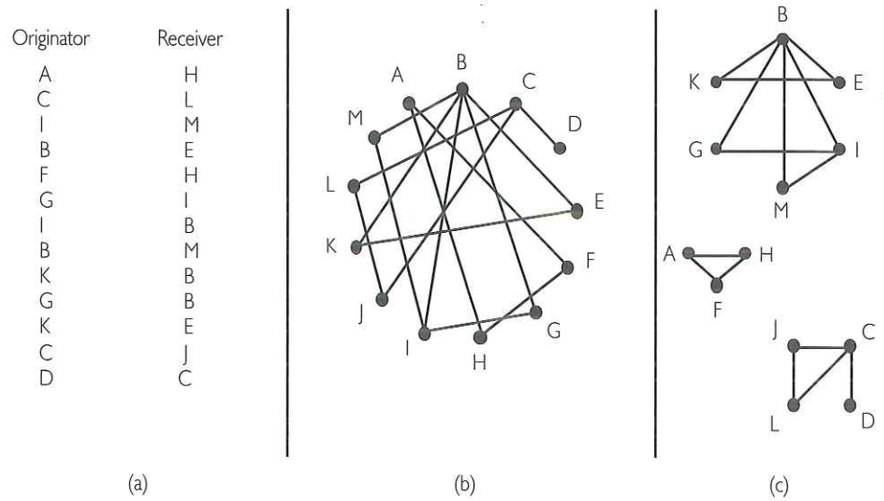


FIGURE 3.73
Insight into even a short list of telephone calls, (a), is enhanced by their node-link representation, (b), especially if disconnected subsets can be identified, (c)

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).

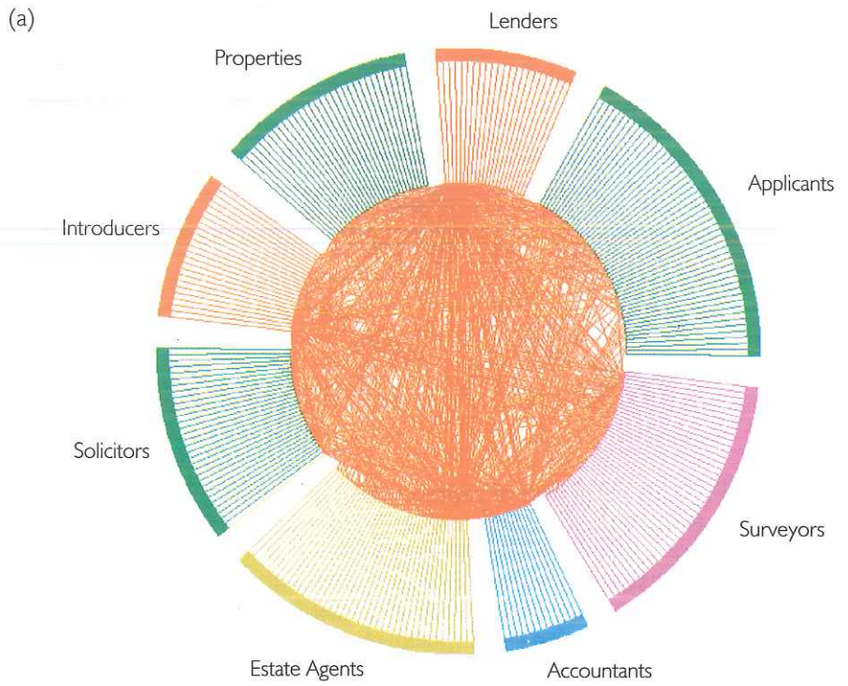


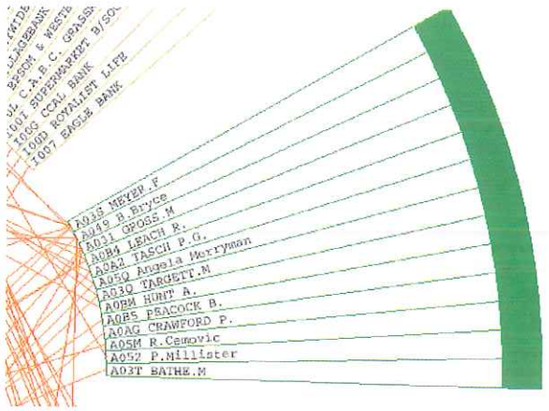
FIGURE 3.74
A representation of mortgage activity (a). Lenders, properties (houses), buyers, etc. are represented by small radial segments of an annulus as shown in (b) overleaf, and their relationships denoted by straight lines

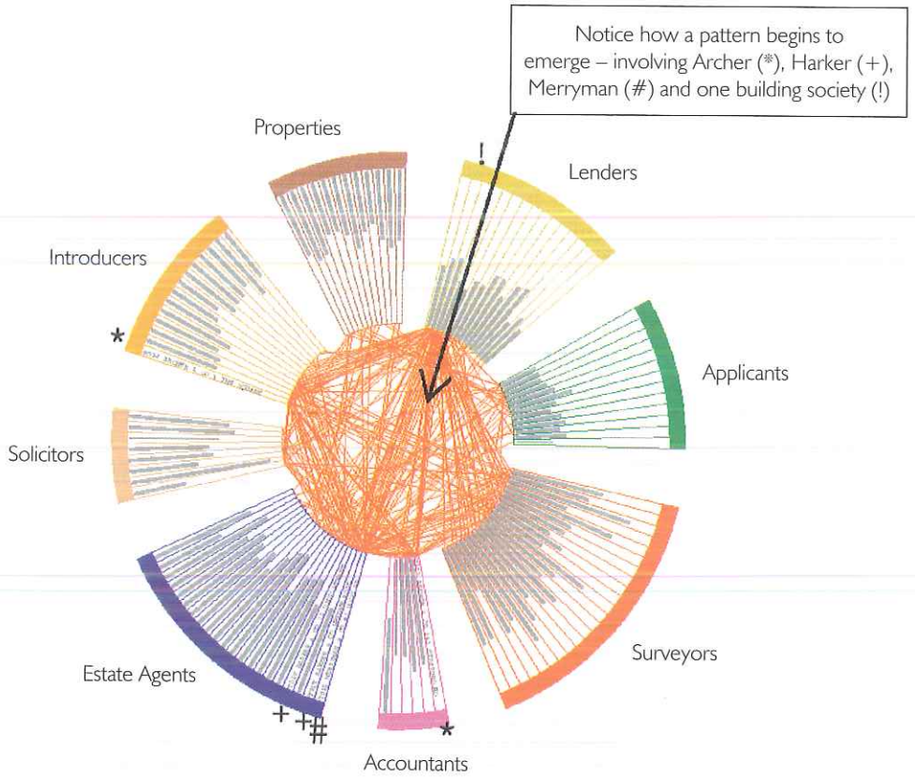
E 3.74
ntation
ants by
lial
s within

ntation
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.

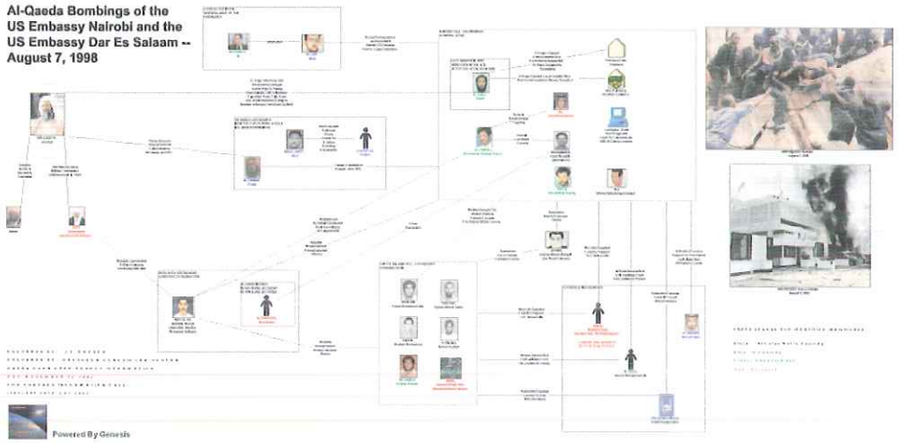


FIGURE 3.76
An 'association' style chart depicting the African bombings
(Courtesy i2 Ltd.)

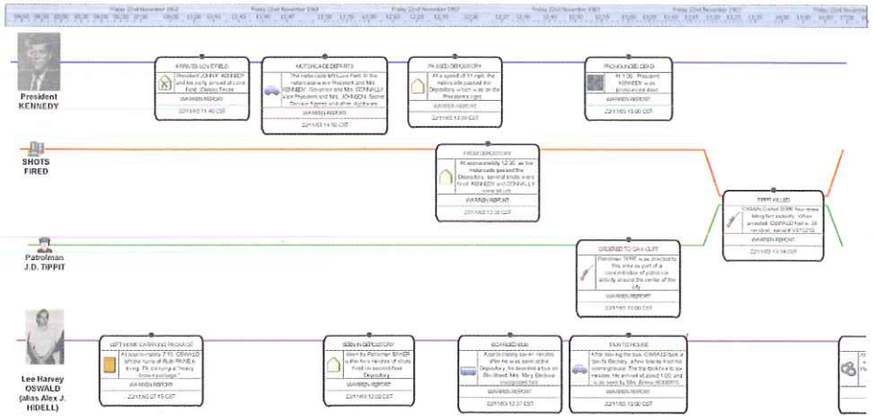


FIGURE 3.77
Part of a 'timeline' style chart depicting the Kennedy assassination
(Courtesy i2 Ltd.)

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-

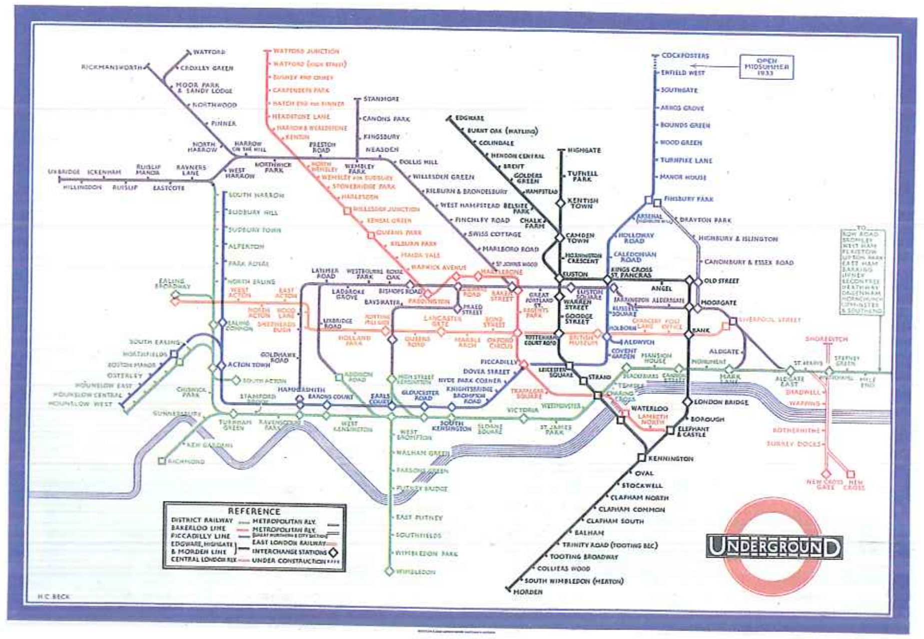


FIGURE 3.78
Harry Beck's original London Underground map
(© Transport for London)

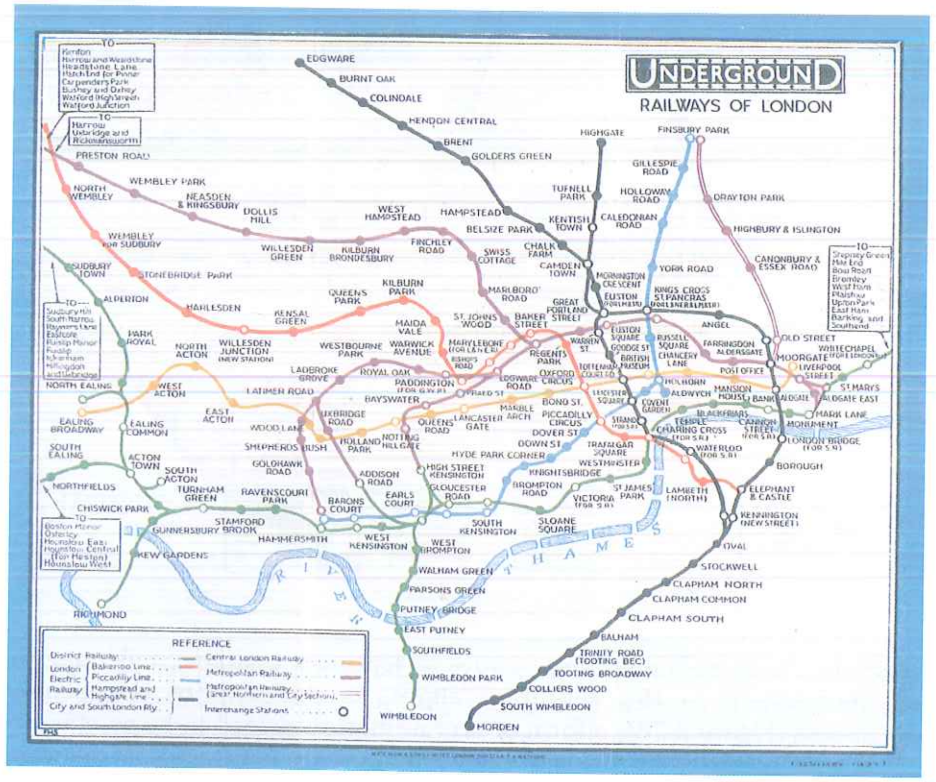


FIGURE 3.79
The Underground map in use prior to the introduction of Harry Beck's version
(© Transport for London)

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

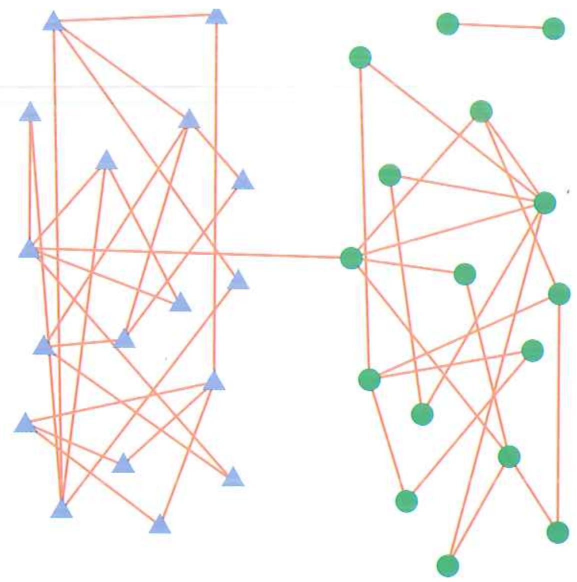


FIGURE 3.80
The social choices of fourth-grade students (after Moreno, 1934)

(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.

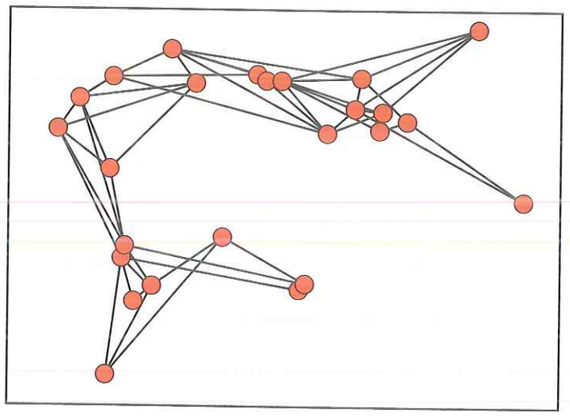


FIGURE 3.81(a) Social choices among department store employees
Source: *L.C. Freeman*

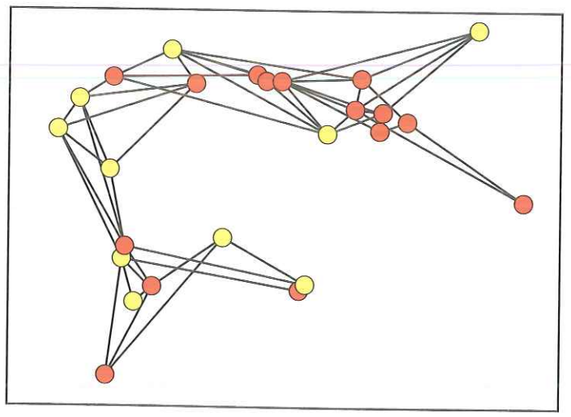


FIGURE 3.81(b) Social choices among department store employees, with marital status encoded
Source: *L.C. Freeman*

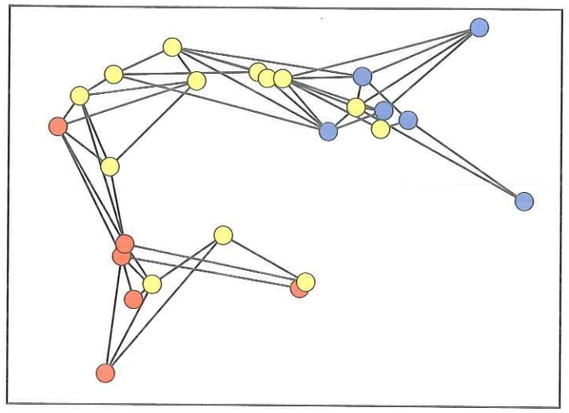


FIGURE 3.81(c) Social choices among department store employees, with age range encoded (blue <30, 30<yellow <40, red >40)
Source: *L.C. Freeman*

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			●