•1• Techniques for analysing structure

Perception is active

Our visual environment is made up of objects rather than features. When we look around a room we see different objects, for instance, some books on a desk. It is hard for us to 'see' the pile of books as an area of different hues and shades, although that is what is represented by the pattern of light that is arriving at our eyes. The process of perception is one of structuring the sensory information that we receive from objects in the environment so that we can interact with them. We need to be able to see a set of differently coloured planes and surfaces as belonging to a single object, a book, that is distinct from the other planes and surfaces that represent the desk and the other books. If we pick up a book, we expect all of the parts of the sensory world that 'belong' to it to move together in a predictable way, and for all of the parts that belong to the 'desk' to stay where they are. If we try to pick up a stack of books, we know that the individual books might not remain as a stack, and that the stack cannot be treated in the same way as the individual books that it is made of.

These details about the structure of objects and their inter-relationships are not explicitly contained in the visual information. It must be interpreted, by combining the visual information with knowledge about the world, which we have learnt through our lifelong experience of interacting with it. This is why we can say that perception is active process, blending knowledge and sensation. The structure of the perceived world affects our interactions with it, and our interactions with the world affect our subsequent interpretations of its structure.

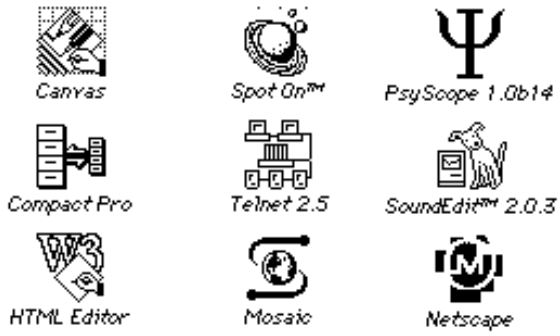

Figure 1.1: an array of different icons

Computer displays are just like the rest of the world in this respect. Figures 1.1 and 1.2 show two groups of icons - one of these is a group of different icons, the other a collection of very similar icons. In Figure 1.1, if the user knows what a particular icon in the first array looks like, the dissimilarity will make it easier to locate - it may even seem to 'pop-out' from the array. But if they do not know what an icon looks like, and have to read all of the names, they may find it difficult to 'separate' the text labels from the 'background' of icons.

In Figure 1.2 the icons are all very similar, and so even if the user knows what the icon of the document they are searching for looks like, they may find it harder to locate than in the previous figure. However, the very similarity of the individual icons makes it much easier for the user to 'group' them as a single, ordered array, and for them to form a 'background' against which the text labels stand out. In this figure, the information provided by the appearance of each icon is less important than the fact that they cluster together to say "we're all documents": this becomes an attribute of the group rather than just an attribute of each separate icon.

Designing a computer display is all about choosing the form of objects and arranging them within a two-dimensional area of the screen. As Figures 1.1 and 1.2 show, a correct choice of form and arrangement can affect the way that objects are perceived and dealt with by the user of the computer. Further than this, you need to think about the structure of tasks, and about the relationships between objects on the screen and any sounds that the system makes.

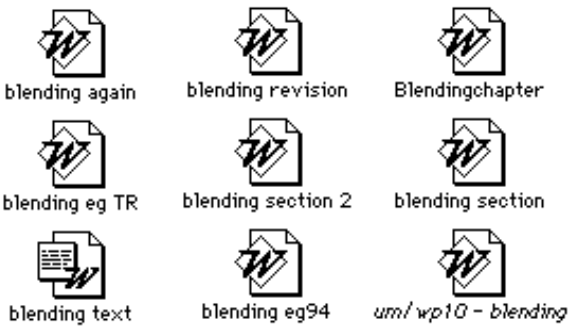

Figure 1.2: an array of structurally similar icons

In this Guide we will explain some ideas that can help you to think about the way that people perceive interfaces, and teach you techniques which you can use to analyse designs. We'll start by looking at visual scenes, and then introduce some psychological ideas to help you understand the way that people perceive, think and act. This will let us extend the techniques to multimodal interfaces.

The structure of visual scenes

Although computer displays are produced on two dimensional, flat screens, we use the same perceptual processes to perceive them as we do to perceive the real, three dimensional world. When we look at a visual scene, whether it is two or three dimensional, the features, colours, and textures in the sensory information that we receive from our eyes group together to form objects. The scene as a whole is a structure of objects. The objects have certain qualities - they stand out from their background and are discrete entities, which can often be named. If we look closely at an object, though, we can see that it also has a structure, and may be composed of other objects. We can perceive the world at several different scales, from a global level, down through many levels of detail. You could stand at the door of the room in Figure 1.3 and see 'an office - a room with objects in it'. You could then focus your attention towards the far wall, which is a plane surface with items of furniture superimposed on it. Within this level you could see the window, a chair and desk. Within the region of the desk you could see a pad of paper. The pad has a pencil resting on it, and is written upon. You could look at the text on the page by moving down into the structure of the pad, and moving down again you could see the individual words that make up the text (if you were near enough).

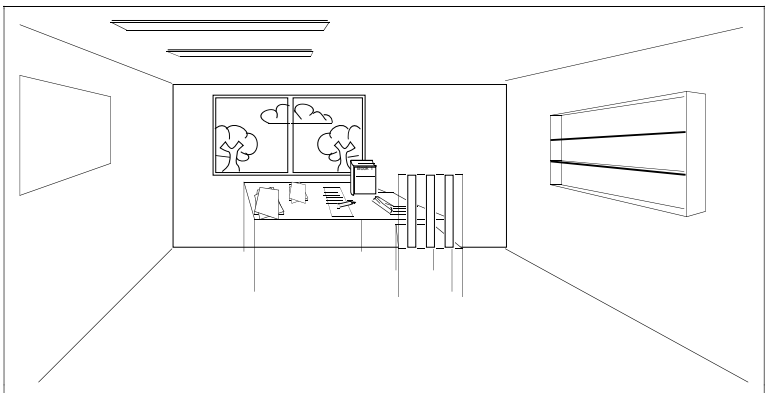

Figure 1.3: a visual scene – an office

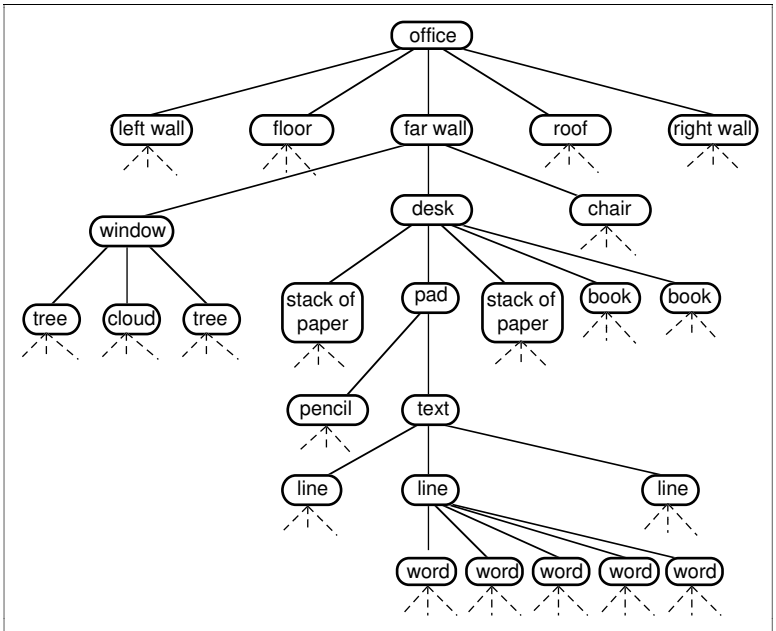

Figure 1.4: a structure diagram of the office scene

This hierarchy can be represented as a structure diagram, as in Figure 1.4, where the different horizontal groups in the figure represent different levels of visual structure. At each of these levels, sensory patterns of light are interpreted as forming a group of individual objects. Each object itself 'contains' visual details that can be further interpreted as another group of objects. The dotted lines indicate that some objects have further structural details that we have not included.

What we actually perceive from moment to moment is limited by the level at which we are analysing the scene. While attending to the pad of paper we can be aware of the relationships it has to the other objects within its own 'group' – the stacks of paper, and the books – and we can be aware of their shared relationship to the desk. We can also be aware that the pad itself has some structural details and, if we wanted to, could attend to some object of this structure; perhaps looking at a line of text.

The hierarchical structure of the visual scene, as represented in Figure 1.4, constrains the direction of visual search. After having attended to the far wall, the words of text cannot be reached by looking at the structure of the window. We have to successively focus in to the desk, the pad and then the text before we can attend to the words. Likewise, after attending to a line of text, attending to a book requires a movement back up the structure, to an object that is at the same level of visual detail as the book (here, the pad of paper).

The **structure** of the scene constrains the way people can search through it.

These two ideas – the structure of visual scenes, and the transitions of attention between objects – are the tools that we will use to analyse the composition of displays. In general, a well composed display will be constructed so that the user can attend to the appropriate object easily. These tools help us to assess the ease with which a user can move their focus of attention around between objects. In the next section we will describe how they are derived.

Psychological subjects and transitions

To describe the way that we change the focus of attention, it is useful to think of the object that is being attended to as the 'psychological subject'. In the office example of Figure 1.3, there are several different objects on the desk. We can focus our attention on any of these objects, and we can shift our attention between them. Any one of them can be the psychological subject at different times. Other objects at the same level of decomposition in the visual scene form its context, and can be used to discriminate it from other similar or identical objects. Because these other objects provide information about the subject, they are collectively called its 'predicate'.

The object that is being attended to is the **psychological subject.** Other objects in the same group form its **predicate**.

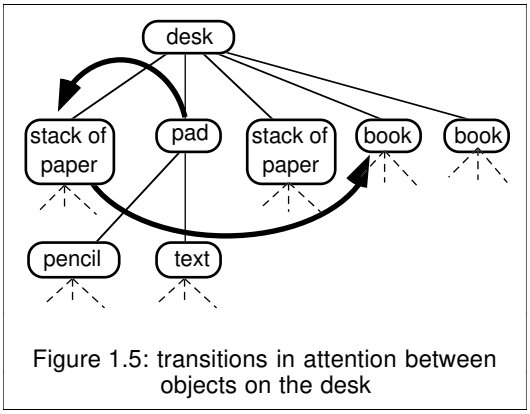

Figure 1.5: transitions in attention between objects on the desk

Figure 1.5 shows part of the office – the group of objects that are on the desk - as attention switches from the pad to a stack of paper, and then to a book, as indicated by arrows. Adding a lot of arrows to the structure diagram would make it rather complicated, especially if attention repeatedly moved back to the same object, and so we need to use a representation that can include time as a dimension.

Figure 1.6 is an example of a 'transition path diagram' that describes the transitions in attention made in Figure 1.5. Each row represents a different moment in time, and a new focus of attention. One object is shown on a black background: this is the psychological subject at that moment. In the first row it is the pad, and the other objects form its predicate, and are listed in a group to its right. As successive transitions are made from object to object, each in turn moves left to become the subject, as shown by the second and third rows. The lines between the rows show the visual transitions that are made as attention shifts between the objects.

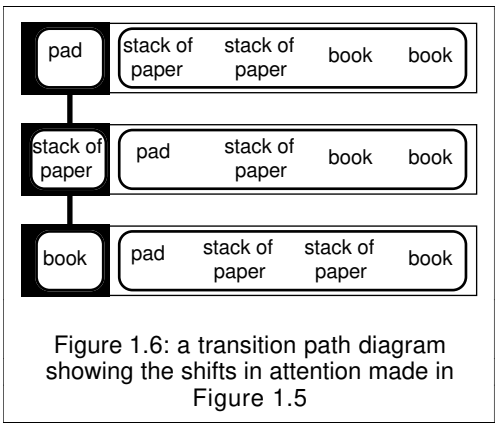

Figure 1.6: a transition path diagram showing the shifts in attention made in Figure 1.5

Figure 1.6 might not seem to offer many advantages over Figure 1.5, but that is because the transitions were quite simple. As well as shifting attention between objects within a group, it is also possible to 'zoom in' and 'zoom out', attending to an object's group - the larger object it belongs to - or to a part of its structure - an object that it visually contains, surrounds, or is made up of. We need to be able to represent these possible transitions as well.

Structure diagrams show the hierarchical relationships between objects.
Transition Path diagrams show changes in the psychological subject and predicate in time.

Figure 1.7 shows how these 'up' and 'down' transitions can also be represented in a transition path diagram. As well as showing the predicate of the psychological subject, each row includes (on the left) the group that the objects belong to, and (on the right) the constituent structure of the psychological subject. This diagram now contains all of the objects that could become the subject following a transition in attention. In the first row the pad of paper is again the subject, but the transition that is made next is 'up' the structure, to the desk. The 'U' shape linking the first row to the desk indicates that the transition is 'up' the structure from many objects (the pad and its predicate) to a single object (the desk).

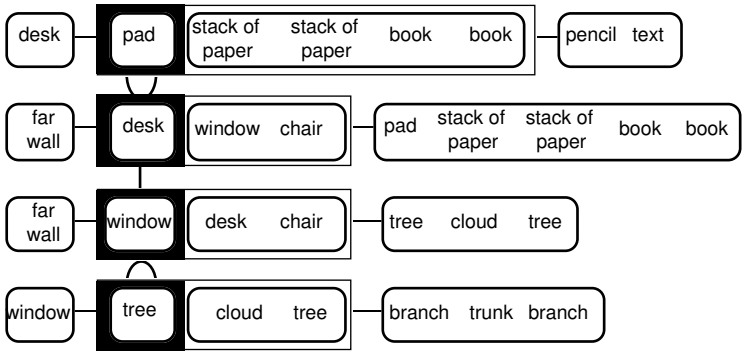

Figure 1.7: transitions as attention 'zooms out' and 'zooms in'

In the second row the desk is shown as the subject. Now the predicate consists of the other objects that are at the same level of decomposition as the desk - the window and the chair – and the 'far wall' is shown as the group that they belong to. The pad, the stacks of paper and the books, that were the active level of the previous representation, are now shown as the desk's constituent structure. They have moved to the right, as has the desk. Each time a transition is made 'up' a structure, the old group moves right to become the new subject, and the old subject-predicate level moves right to become the new constituent structure. In the third row a 'within level' transition is made to the window, so that it becomes the subject: the group remains the same, but the constituent structure changes, to show which transitions 'down' the structure are now possible from the window. The previous subject has become part of the predicate. A point to notice here is that the objects within the predicate are 'unordered' - they are all equally able to become the subject. The second and third rows are linked by a plain line, to show that the transition is just from one object to another, within the same group. Finally, in the bottom row a transition is made to a tree – one of the objects that the window 'contains' in its structure. The group of objects that was on the right of the third row has moved left to become the 'active level' of the representation in the fourth row. The subject is the tree that is being attended to, the predicate consists of the other objects in the window's structure, and the tree's constituent structure must be included on the right of the row. The window has also moved left, to become the group. Each time a transition is made 'down' a structure, the old subject moves left to become the new group, and its constituent structure moves left to become the new subject-predicate level. The 'inverted-U' linking the third and fourth rows now indicates that the transition has been from a single object, the window, to the many objects in its structure.

Transition path diagrams help to make it clear how simple or how complicated it will be for users to move their attention from object to object within a display. On each row, all of the objects that could be attended to following a transition are indicated. A transition 'up' the structure makes the group and subject-predicate move right in the row. A transition 'down' makes the subject and its constituent structure move left in the row. In Figure 1.7, it took three transitions to look up from the pad, and to look at a tree. It might take a user of a computer several transitions to move their focus of attention from the document they are reading on-screen to locate an icon in a menubar, depending on the structure and grouping of all of the objects. In analysing a display, it is helpful to construct a structure diagram first, and then to use it to draw transition path diagrams for particular tasks that a user will want to carry out. The next section shows how this can be done for a typical computer display.

Using diagrams to analyse a display

The 'office' example was a real-world, three dimensional structure, but the structural and transition path diagrams can be used to analyse two dimensional computer displays. The only differences between the two 'control panels' in Figure 1.8 are the boxes that have been drawn around the groups of words and buttons. This might be an aspect of the design that is left to a designer's aesthetic judgement, or it might be constrained by the interface software 'toolbox'.

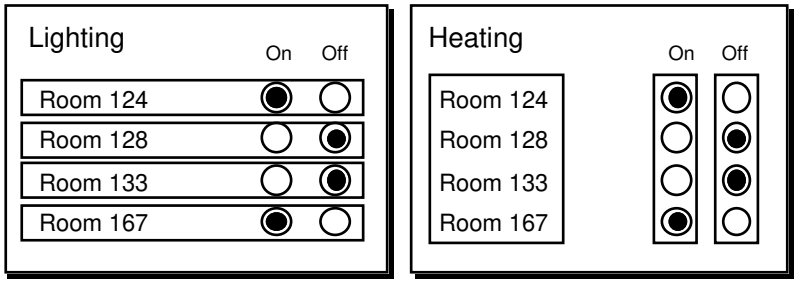

Figure 1.8: two 'control panels' that differ in the way the objects have been grouped by the designer

The 'lighting panel' has boxes that relate objects together functionally, so each room label is linked to its own on and off buttons. In the 'heating panel', the objects have been linked by type, so that all of the room labels, on buttons and off buttons each form different groups. This is a fairly small difference, and if anything, the heating panel looks more aesthetically appealing. The structure diagrams for the two panels (Figure 1.9) show the difference that these boxes make to the grouping of the objects. The lighting panel is made up of four 'groups', one for each room, each containing a room label and an empty and filled button. The heating panel consists of three groups, one of four similar room labels, and two groups of mixed circles.

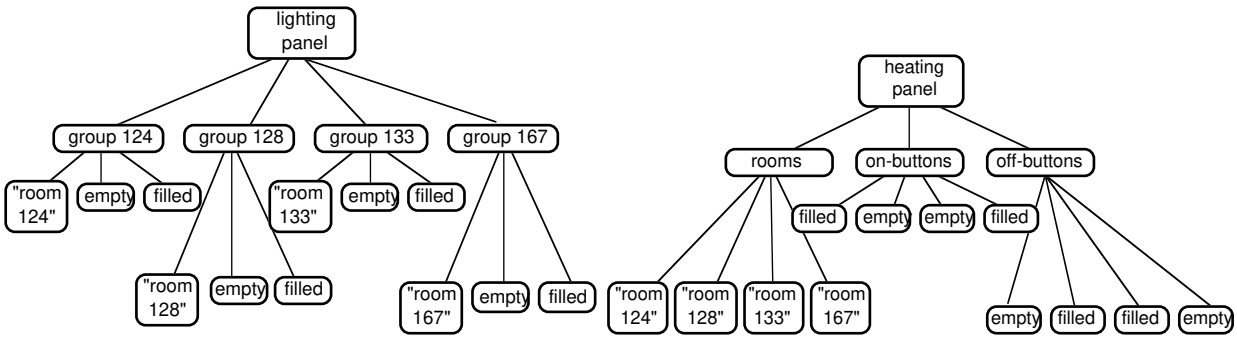

Figure 1.9: structure diagrams of the two control panels

If we draw transition path diagrams for a user who has to turn the lights and heat on in Room 133, we can see the difference that these groupings have made to the panels' 'ease of use'. For the lighting panel, the task is quite simple, as the diagram in Figure 1.10 shows. The button that turns the lights on is part of the predicate of the label for 'room 133' and so only one transition is necessary. The transition is made within a single group, and so the object to the left of the subject does not change. The empty circle has no constituent structure, and so when it is the subject nothing is shown to the right of the row, to indicate that no further transitions could be made 'down' the structure.

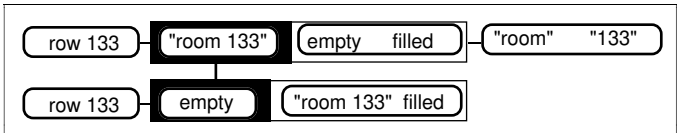

Figure 1.10: a transition path diagram for the lighting panel

The situation is quite different for the 'heating panel'. Now the room label and the button are in different groups, and the user has to momentarily move their attention up the structure to the 'rooms' group, across to the on-buttons, and then down again to the third button (Figure 1.11). Here three transitions are needed instead of one, and so for this particular task, we can say that the 'lighting panel' will be easier to use than the 'heating panel'.

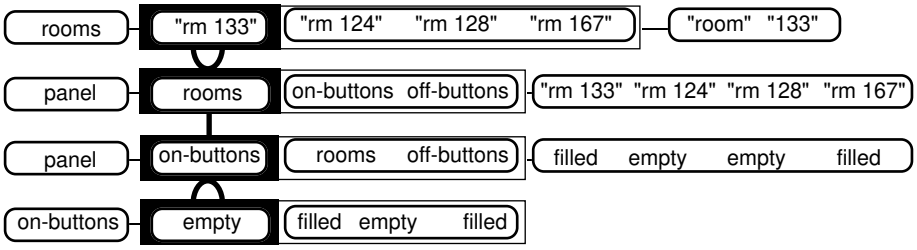

Figure 1.11: a transition path diagram for the heating panel

The idea of 'task' is very important, of course. If instead of a task that required the use of the room label and a button, the user had to operate on each of the on-buttons in sequence, regardless of the room labels, the 'heating panel' might be found to have an advantage. Suppose the user just had to make sure that all the heaters were on. Once the user had located the group of on-buttons, and attended to one of the buttons, the other buttons would all be part of the predicate. This task would require fewer transitions than the equivalent task of turning on all of the lights.

This example shows that it is vital to make the grouping of screen objects correspond to the task that the user is going to perform, because this determines the way that they will have to move their attention between the objects. In choosing between different possible forms for objects and different ways of arranging them, the designer is attempting not just to make an aesthetically pleasing interface, but one which helps the user perform a particular task.