8 Neural networks in the search for similarity and structure—activity
D.J. LIVINGSTONE and D.W. SALT

8.1 Introduction

This book is concerned with molecular similarity, what it is, how it may be quantified and how it can be used in the design of new drugs. Perhaps one of the most important aspects of similarity, however it is defined, is the question of how it may be perceived. One of the skills that we humans can lay particular claims to is our ability to recognize patterns, in other words to perceive similarity, and thus many of the similarity tools used in drug design are intended to express similarity, leaving the task of perception to the human 'expert'. Unfortunately, our ability to recognize patterns is generally restricted to relatively low dimensional data sets (i.e. 2, 3 or 4D) and thus we need help when similarity is described by hundreds or thousands of variables. This help can take the form of a variety of statistical methods, some of the most useful being pattern recognition techniques (Livingstone, 1991a) which literally set out to identify any underlying patterns in sets of data. An alternative approach in trying to tackle the problem is by the use of comparative molecular field analysis (CoMFA) (see Chapter 12).

One of the sources of pattern recognition methodology is artificial intelligence (AI) research which has aimed to reproduce human intelligence by imitating 'what we do'. The linear learning machine, for example, is an algorithm which learns to distinguish between different classes of samples (Nilsson, 1965) and expert systems are excellent examples of computer programs which are designed to mimic decision making by human experts (Ayscough et al., 1987; Jakus, 1992; Cartwright, 1993). A more recent AI approach to the simulation of intelligence is an imitation of 'the way that we do it', in other words an attempt to construct artificial brains. The rationale underlying this approach is that since we do not know precisely how a brain works then perhaps a model of a brain, physical or otherwise, may exhibit some of the intelligent properties of a brain. How can this be achieved? Figure 8.1 shows a simplified picture of the basic building block of biological brains, the neuron, along with its connections by means of which it forms networks of neurons. The figure also shows an illustration of an artificial neuron which, like its biological counterpart, forms the basic building block of systems designed to simulate such biological networks. These systems are known as
artificial neural networks (ANN) and the whole area has received much 'hype' as the following extract from a recent book\(^1\) shows (Eberhart and Dobbins, 1990):

\(^1\) These authors are at pains to distinguish between real substance and hype and present a very useful set of case studies along with source code listings (in C) for neural network programs.