ABSTRACT

An artificial neural network was trained to classify musical chords into four categories—major, dominant seventh, minor, or diminished seventh—independent of musical key. After training, the internal structure of the network was analyzed in order to determine the representations that the network was using to classify chords. It was found that the first layer of connection weights in the network converted the local representations of input notes into distributed representations that could be described in musical terms as circles of major thirds and on circles of major seconds. Hidden units then were able to use this representation to organize stimuli geometrically into a simple space that was easily partitioned by output units to classify the stimuli. This illustrates one potential contribution of artificial neural networks to cognitive informatics: the discovery of novel forms of representation in systems that can accomplish intelligent tasks.

Keywords: artificial neural networks; chord classification; representation

INTRODUCTION

Cognitive informatics is a field of research that is primarily concerned with the information processing of intelligent agents (Wang, 2003). One way to characterize cognitive informatics is in terms of an evolving notion of information (Wang, 2007). When it originated six decades ago, conventional accounts of information were concerned about using probability theory and statistics to measure the amount of information carried by an external signal. This in turn developed into the notion of modern informatics which studied information as “properties or attributes of the natural world that can be generally abstracted, quantitatively represented, and mentally processed” (Wang, 2007, p. iii). The current incarnation of cognitive informatics recognized that both information theory and modern informatics defined information in terms of factors that were external to brains. Cognitive informatics has replaced this with an emphasis on exploring information as an internal property.

This emphasis on the internal processing of information raises fundamental questions about how such information can be represented. One approach to answering such questions—and for proposing new representational accounts—
would be to train a brain-like system to perform an intelligent task, and then to analyze its internal structure to determine the types of representations that the system had developed to perform this intelligent behavior. The logic behind this approach is that when artificial neural networks are trained to solve problems, there are few constraints placed upon the kinds of internal representations that they can develop. As a result, it is possible for a network to discover new forms of representation that were surprising to the researcher (Dawson & Boechler, 2007; Dawson & Zimmerman, 2003).

Cognitive informatics has been applied to a wide variety of domains, ranging from organization of work in groups of individuals (Wang, 2007) to determining the capacity of human memory (Wang, Liu & Wang, 2003) to modeling neural function (Wang, Wang, Patel & Patel, 2006). The research below provides an example of this approach in a new domain, musical cognition. There is a growing interest in the cognitive science of musical cognition, ranging from neural accounts of musical processing (Jourdain, 1997; Peretz & Zatorre, 2003) through empirical accounts of the perceptual regularities of music (Deutsch, 1999; Krumhansl, 1990) to computational accounts of the formal properties of music (Assayag, Feichtinger, & Rodrigues, 2002; Lerdahl & Jackendoff, 1983). Because music is characterized by both many formal and many informal properties, there has been an explosion of interest in using artificial neural networks to study it (Griffith & Todd, 1999; Todd & Loy, 1991). The simulation below illustrates one intriguing possibility for such research: the discovery of previously unknown representations of formal musical structures. As such, it illustrates that artificial neural networks can be used as a medium to explore a “synthetic approach” to psychology and make important representational contributions to cognitive informatics and cognitive science (Dawson, 1998, 2004).

**Chord Classification By Neural Networks**

In a pioneering study, artificial neural networks (ANNs) were trained to classify musical stimuli as being major chords, minor chords, or diminished chords (Laden & Keefe, 1989). Laden and Keefe’s networks used 12 input units, where each unit represented a particular note or “piano key” in an octave range—a so-called pitch class representation (see Figure 1). They created a training set consisting of 36 different chords: the major triad for each of the 12 different major key signatures, as well as the minor triad and a diminished seventh triad for each of the 12 different minor key signatures. They examined a number of different networks by manipulating the number of hidden units (three, six, eight, or nine), and by manipulating the pattern of network connectivity.

In general, Laden and Keefe (1989) found that the performance of their simple networks was disappointing. Their most successful simple network used three hidden units, and had direct connections between input and output units, but was still able to correctly classify only 72 percent of the presented chords. Other small networks had accuracy rates as low as 25 percent. Laden and Keefe improved network accuracy to a maximum level of 94 percent by using a more complex network that had 25 hidden units and which used output units to represent distances between musical notes (i.e., musical intervals) rather than chord types.

The current study is an extension of Laden and Keefe’s (1989) research. It examines a simple network that uses a pitch class representation and three hidden units. Laden and Keefe used a traditional sigmoid-shaped activation function (the logistic equation) in the processing units of their networks. We instead used a Gaussian activation function because previous research has demonstrated that networks that employ this activation function are adept at solving complex problems, and also lend themselves to detailed interpretation (Berkeley, Dawson, Medler, Schopflocher, & Hornsby, 1995; Dawson, 2004, 2005). Our working hypothesis was that this change in network architecture would