9. SCIENTIFIC VISUALIZATION

Chaired by Susan Chipman, *Office of Naval Research*

Hinton diagrams: Viewing connection strengths in neural networks

FREDERICK J. BREMNER, STEPHEN J. GOTTS, and DINA L. DENHAM
Trinity University, San Antonio, Texas

A graphical analysis similar to that used by Hinton and Shallice (1991) was applied to the hair cells of a simulated auditory transducer. The graphical analysis made it apparent that there were hair cells that responded to a narrow range of frequencies, as would be predicted by the tonotopic organization of the real physiology. In short, this study demonstrates the efficacy of using graphic techniques to examine the nature of the autopoietic organization of the hidden layer of back-propagation networks.

Modern neurophysiology has made great strides since the common use of single-neuron recording by way of microelectrode techniques (Kandel, 1982). By expanding these techniques to look at simple neuronal networks, an accurate picture of both the physiological and the behavioral functioning of the nervous system has begun to emerge (Mpitso, Creech, Cohan, & Mendelson, 1988). As a companion to the strictly physiological approach, some investigators have used computer simulation to model functional neural networks (e.g., Anderson & Hinton, 1981; Willows, Dorsett, & Hoyle, 1973). When using computer simulation to study neuronal circuits, it is important to know what each cell in the simulated network is doing. Unfortunately, many investigators who publish simulations of neural circuits do not look at the functioning of individual nodes (neurons). In fact, there is a consensus among connectionists that, since it is the sum total or mosaic of the entire network that dynamically solves the task, the functioning of any particular neuron is immaterial (Adey, 1972; Mpitso, 1989).

Nevertheless, Hinton (Hinton, McClelland, & Rumelhart, 1986; Hinton & Sejnowski, 1986; Hinton & Shallice, 1991) has published a technique that graphically represents the potency or contribution of each neuron in a neural network (using Hinton diagrams). It is of particular importance that the Hinton technique is a graphic one, because a table of numbers of a dynamically organizing network with hundreds or even thousands of connections is meaningless both visually and statistically. The purpose

of the present study is to demonstrate the usefulness of applying Hinton diagrams to a set of neural transducer elements in a functioning neural network.

The neural system picked for this demonstration is that of the transducer cells of the human auditory system (Wever, 1949). This system was chosen for several reasons. First, it is a well-documented system with published algorithms (Greenwood, 1961; Zwislocki, 1965) that describe its functioning. Second, there are not very many hair cells to simulate, because 3,200 inner hair cells transduce all the sounds that humans can discriminate (Bremner & Denham, 1993; Netter, 1986). Furthermore, Bremner and Denham showed that pitch discrimination could be successfully modeled by computer-simulated transducer cells (hair cells). This tonotopically organized computer simulation supported a place theory for auditory discrimination of both high and low frequencies. A computer simulation presented by Denham (1993) extended this work by showing that a back-propagation model could dynamically organize itself to discriminate a wide variety of pure tones. Therefore, the present study applied Hinton diagrams to the back-propagation network reported by Denham to see if there was any evidence for tonotopic self-organization.

METHOD

Apparatus

A CompuAdd 386SX computer with 4 MB of RAM, a math coprocessor, and an 80-MB hard disk was used as the vehicle for all software. The study described here relied on three software packages: Brainmaker Professional (1990), Wordperfect (1989), and CompuAdd Windows (1990). The program of principal interest to this study is Brainmaker, a version of the back-propagation model (Bremner & Yost, 1992).

Correspondence should be addressed to S. J. Gotts, Department of Psychology, Trinity University, 715 Stadium Drive, San Antonio, TX 78212.

Input Signals

The inputs to this three-layered network were displacements of the basilar membrane. The correct displacement of the basilar membrane for 34 pure tones was mapped onto a 28×70 matrix. The 70 columns of the matrix represented the length of the basilar membrane so that every column equalled .05 cm of basilar-membrane length. The 28 rows represented the amount of displacement the basilar membrane would experience for each pure tone used. The correct place and amount of displacement for each input tone were computed by applying a modification of published equations (Zwislocki, 1965). A more detailed account of the appropriate mathematical manipulation was given by Denham (1993). The 28×70 matrix constituted the input layer of the back-propagation network. A diagram of the 28×70 matrix showing the mapping of a very low tone is presented in Figure 1.

Hair Cells

There were 70 simulated hair cells representing the 3,200 inner hair cells of the human auditory transducer (Bremner & Denham, 1993). Thus, the ratio of simulated to real hair cells was roughly 1:46. The network was designed so that every input tone stimulated every hair cell, just as a traveling wave in the cochlea would stimulate every hair cell. The number of hair cells used was determined by practical rather than theoretical considerations, namely, that Hinton diagrams more than 70 columns wide are difficult to present in standard computer format. Figure 2 gives a schematic diagram for the organization of the entire neural network. The hair cells were represented by the hidden layer of the back-propagation network.

Output Patterns

The output of the network was an array of length 5, each box of the array representing a category of pitch discrimination. The categories were: very low (20–600 Hz), low (601–1800 Hz), medium (1800–6500 Hz), high (6501–9000 Hz), and very high (9001–20000 Hz). The task for the back-propagation network was similar to asking a human subject to classify pure tones on a five-point Likert scale from very low to very high. The five categories constituted the output layer of the three-layer back-propagation network.

Procedure

The network was allowed to self-organize until it correctly mapped the comprehensive training set of 34 pure tones. The network con-

verged when it successfully classified the 34 stimuli with an error less than .05 for each stimulus. The network was then tested by the presentation of 10 new pure tones. This was necessary to see if the network could correctly categorize these new tones without any further training. Following the completion of this test, the network connections were repeatedly randomized using a different random number seed until 20 unique networks were produced. The training and testing procedures were performed on all of the new network generations. In order to establish that any tonotopic self-organization has occurred, it is necessary to make sure that the same tonotopic order results across distinctly different weight randomizations. This approach is analogous to starting with 20 neonate human infants, each with its own genetically programmed hair cells; throughout life, the hair cells self-organize to discriminate the real world (Aslin, 1987).

Data Analysis

The test patterns for the 20 randomizations of the network described above were analyzed vis-à-vis Hinton diagrams. The diagrams showed each hair cell as a square; the darker the shading, the stronger the relationship to the input tone. For each network randomization, each one of the 10 rows of graphic representation corresponds to one of the pure tones presented to the network during the testing phase of the modeling. Each row showed the activation levels for each of the 70 hair cells. This 10×70 matrix was considered as one Hinton diagram. (The Brainmaker program (1990) and the van Camp, Plate, & Hinton program (1993) do not use identical techniques to generate the diagrams, but the end result is the same.) Because there were 20 different randomizations of the network, there were 20 Hinton diagrams in all.

RESULTS

Every network accurately discriminated all 10 test tones, regardless of the randomization used. Figure 3 is one of the 20 Hinton diagrams of the 70 hair cells responding to the 10 test tones. Notice that three hair cells (12, 17, and 69) respond strongly to low frequencies (601–1800 Hz) only. There are two other very narrow-bandwidth hair cells, one of which responds to very high frequencies (Cell 33), while the other responds to very low frequencies (Cell 62). These five hair cells for a particular randomization discriminated input frequencies for one output tone category only. Eighteen out of the 20 unique networks used in this study contained narrow-bandwidth hair cells.

Figure 4 is another Hinton diagram of the 70 hair cells responding to the 10 test tones, in which three hair cells (1, 13, and 25) responded to two adjacent frequency categories. All 20 randomizations of the network in this study contained at least one hair cell that responded to two adjacent categories, most often to the very low (20–600 Hz) and the low categories (601–1800 Hz).

DISCUSSION

The value of using Hinton diagrams to determine the behavior of simulated cells in a neural network is readily apparent in this study. The finding of unique hair cells that respond to a narrow range of frequencies would have been difficult to determine without this graphic technique. The use of Hinton diagrams provided a vehicle to look for narrow-bandwidth hair cells in our simulation. This finding has theoretical implications in that it demonstrates

Figure 1. The 28×70 input array showing the mapping of the vertical displacement of the basilar membrane produced by a 200-Hz tone (Zwislocki, 1965) and its proper discrimination into one of the five output categories, VLOW (20–600 Hz), LOW (601–1800 Hz), MED (1801–6500 Hz), HIGH (6501–9000 Hz), and VHIG (9001–20000 Hz).

Figure 2. A schematic diagram of the organization of the layers that make up the neural network. The input layer is a 28×70 matrix representing the relationship between length of the basilar membrane (horizontal dimension) and the displacement of the membrane for a given pure tone (vertical dimension). The hidden layer is an array of length 70 representing the hair cells. The output layer is an array of length 5 representing the categorization of the pure tone.

how neural networks can mimic physiological systems. The finding that some generations of the network had multiple examples of such hair cells (in some cases as many as 10) was quite interesting. This was true even though the simulation was structured so that every hair cell was potentially triggered by every input tone, as happens in the human cochlea (Zwislocki, 1965).

While this simulation does not demonstrate tonotopic organization in the strictest sense, it does contain many tonotopically constrained hair cells. One wonders if the power of tonotopic organization is so strong that it overrides the tendency of back-propagation networks to self-organize in an arbitrary fashion. The long-term goal of

the project is to determine how many constraints are necessary for a randomized set of hair cells to tonotopically organize itself.

The next step in investigating tonotopic organization is to observe a genetic neural network. These networks have a kind of evolutionary strategy, in that they compare the offspring of two sets of parents and choose the better set to propagate. Next, they compare the F2 generation and pick the best one to propagate, and so on. Applying such a technique to the arrangement of simulated hair cells could lead to a more tonotopic arrangement if the genetically coded structural development of the cortex is a salient factor in the development of tonotopicity.

Figure 3. A Hinton diagram of the 70 simulated hair cells responding to the 10 test tones. Note the five very narrow-bandwidth hair cells (12, 17, 33, 62, and 69) that respond solely to one category of frequencies. Each row shows the activation levels of each of the 70 hair cells to an input tone as labeled. A star indicates no activation; the darker the shading, the stronger the activation.

Figure 4. A Hinton diagram of the 70 simulated hair cells responding to the 10 test tones. Note the three narrow-bandwidth cells (1, 13, and 25) that respond to two adjacent frequency categories. Each row shows the activation levels of each of the 70 hair cells to an input tone as labeled. A star indicates no activation; the darker the shading, the stronger the activation.

REFERENCES

- ADEY, W. R. (1972). Organization of brain tissue: Is the brain a noisy processor? *International Journal of Neuroscience*, **3**, 271-284.
- ANDERSON, J. A., & HINTON, G. E. (1981). Models of information processing in the brain. In G. E. Hinton & J. A. Anderson (Eds.), *Parallel models of associative memory* (pp. 9-44). Hillsdale, NJ: Erlbaum.
- ASLIN, R. N. (1987). Visual and auditory development in infancy. In J. D. Osofsky (Ed.), *Handbook of infant development* (2nd ed.), (pp. 5-97). New York: Wiley.
- Brainmaker Professional, Version 2.0 [Computer program] (1990). Grass Valley, CA: California Scientific Software.
- BREMNER, F. J., & DENHAM, D. L. (1993). Lotus 1-2-3 simulation of frequency encoding on the basilar membrane. *Behavior Research Methods, Instruments, & Computers*, **25**, 208-211.
- BREMNER, F. J., & YOST, M. (1992). Assessing the simulation of chaotic systems. *Behavior Research Methods, Instruments, & Computers*, **24**, 336-339.
- CompuAdd Windows, Version 3.0 [Computer program] (1990). Austin, TX: CompuAdd.
- DENHAM, D. L. (1993, November). *Computer simulation of frequency discrimination in the human auditory system*. Paper presented at the Texas Psychological Convention, Austin, TX.
- GREENWOOD, D. D. (1961). Critical bandwidth summation and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America*, **33**, 1344-1356.
- HINTON, G. E., MCCLELLAND, J. L., & RUMELHART, D. E. (1986). Distributed representations. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 77-109). Cambridge, MA: MIT Press.
- HINTON, G. E., & SEJNOWSKI, T. J. (1986). Learning and relearning in Boltzmann machines. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 282-317). Cambridge, MA: MIT Press.
- HINTON, G. E., & SHALLICE, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, **98**, 74-92.
- KANDEL, E. R. (1982). The origins of modern neuroscience. *Annual Review of Neuroscience*, **5**, 299-330.
- MPITSOS, G. J. (1989). Chaos in brain function and the problem of non-stationarity: A commentary. In E. Basar & T. H. Bullock (Eds.), *Springer series in brain dynamics 2* (pp. 521-535). Berlin and Heidelberg: Springer-Verlag.
- MPITSOS, G. J., CREECH, H. C., COHAN, C. S., & MENDELSON, M. (1988). Variability and chaos: Neurointegrative principles in self-organization of motor patterns. In J. A. S. Kelso, A. J. Mandell, & M. F. Shlesinger (Eds.), *Dynamic patterns in complex systems* (pp. 162-190). New York: World Scientific Publishers.
- NETTER, F. H. (1986). *The CIBA collection of medical illustrations*. New York: CIBA Pharmaceutical.
- VAN CAMP, D., PLATE, T., & HINTON, G. E. (1993). *The Xerion Neural Network Simulator*, Version 3.1 [Computer program]. University of Toronto, Canada.
- WEVER, E. G. (1949). *Theory of hearing*. New York: Wiley.
- WILLOWS, A. O. D., DORSETT, D. A., & HOYLE, G. (1973). The neuronal basis of behavior in Tritonia: I. Functional organization of the central nervous system. *Journal of Neurobiology*, **4**, 207-237.
- Wordperfect, Version 5.1 [Computer program] (1989). Alpine, UT: Wordperfect.
- ZWISLOCKI, J. (1965). Analysis of some auditory characteristics. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 3-97). New York: Wiley.