A quantitative theory of neural computation

Leslie G. Valiant

Abstract We show how a general quantitative theory of neural computation can be used to explain two recent experimental findings in neuroscience. The first of these findings is that in human medial temporal lobe there exist neurons that correspond to identifiable concepts, such as a particular actress. Further, even when such concepts are preselected by the experimenter, such neurons can be found with paradoxical ease, after examining relatively few neurons. We offer a quantitative computational explanation of this phenomenon, where apparently none existed before. Second, for the locust olfactory system estimates of the four parameters of neuron numbers, synapse numbers, synapse strengths, and the numbers of neurons that represent an odor are now available. We show here that these numbers are related as predicted by the general theory. More generally, we identify two useful regimes for neural computation with distinct ranges of these quantitative parameters.

1 Introduction

A central problem of neuroscience is to provide explicit mechanistic explanations for the basic tasks of cognition such as memorization and association. The methodology of computer science suggests that a solution to this problem would consist of three parts: a concrete model of neural computation, explicit specifications of the cognitive tasks that are to be realized, and explicit

L. G. Valiant (✉)
Harvard University, Division of Engineering
and Applied Sciences, Cambridge, MA 02138, USA
e-mail: valiant@deas.harvard.edu

algorithms that can be demonstrably executed on this model to realize the tasks. Also, one expects that the outcome will include a coherent class of scalable algorithms that demonstrably work for large ranges of such resource parameters as the number of neurons.

A theoretical investigation using this methodology can be expected to uncover how the brain actually works if the brain is computationally so constrained that there are few solutions consistent with those constraints. We take this observation as our methodology.

It appears that the brain is indeed highly constrained in how a neuron can have a purposeful effect on an arbitrary other neuron: In a system of n neurons each one typically receives inputs from a much smaller number d of other neurons. For example estimates of $n = 1.6 \times 10^7$ and $d = 7,800$ have been given for the mouse (Braitenberg and Schüz 1998) and of $n = 10^{10}$ and $d = 2.4 \times 10^4 - 8 \times 10^4$ for humans (Abeles 1991). Also, it is believed that even when a neuron does receive an input directly from another, the *mean* strength of such an influence is weak. A recent estimate for rat visual cortex (Song et al. 2005) gave the mean value for the excitatory postsynaptic potential (EPSP) of a synapse to be 0.77 mV. This compares with a typical estimate of the threshold voltage of 20 mV, and implies a rough estimate of $k = 26$ for the mean number of presynaptic neurons that are needed to cause a postsynaptic action potential. An earlier estimate (Abeles 1991) of the general range of k was 5–300. A further constraint is that the brain can perform significant tasks in 100–200 ms, which allows perhaps only 10–20 and certainly fewer than 100 basic steps (Feldman and Ballard 1982). Basic algorithms need to work in very few steps.

The model of computation used in this paper is characterized by the above defined three parameters

(n, d, k), by the requirement that basic algorithms run in a handful of steps, and by the restriction that the computational transitions of a neuron be simple enough that there is little doubt that actual neurons, in all their complexity, are capable of performing them. Graph models based on n and d have been considered widely before (Griffith 1963; Braitenberg 1978; Feldman 1982; Palm 1982) and sometimes in conjunction with the third physical parameter of synaptic strength (Griffith 1963; Abeles 1991). In general, we seek detailed quantitative explanations of how networks with given parameters can support algorithms that can perform sequences of intended actions as required, without unintended interference among these actions. Also, unlike previous authors, we use a hierarchical notion of memory allocation, which would appear to be more useful for building up complex data structures that reflect the complexities of knowledge, rather than being merely sufficient to support data lookup. The specific model and theory we use appears in (Valiant 2005), where it is discussed more fully. For completeness we restate here the most relevant aspects of the theory before we move on to discuss the particular experimental findings (Quian Quiroga et al. 2005; Jortner et al. 2006) that we seek to explain.
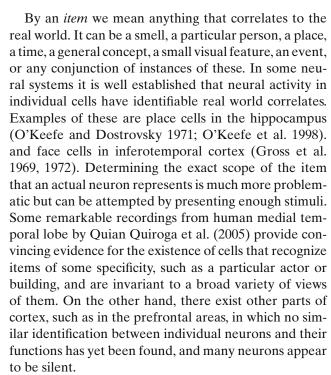
We emphasize that our assumptions are extremely mild. In particular there is no assumption about the nature of the firing mechanisms, which may be governed by various nonlinearities. The assumption that some value of k exists asserts merely that some minimum total presynaptic activity is required to cause an action potential in a neuron.

2 Representation

Before we discuss tasks we need to take a specific position on the representation of information. We define two representations: *positive shared* and *positive disjoint*. They have in common the properties that:

1. Every real world "item" corresponds to the firing of a set of r neurons,
2. The items are memorized in a hierarchical fashion, so that a new item C when first stored will be stored as a conjunction of two items already memorized, say, A and B.
3. The representation may be graded so that the recognition of the item corresponds to a large enough fraction but not necessarily all of the corresponding neurons firing.

The representation is called *disjoint* if each neuron corresponds to at most one item, and *shared* if it can correspond to more than one.

By an *item* we mean anything that correlates to the real world. It can be a smell, a particular person, a place, a time, a general concept, a small visual feature, an event, or any conjunction of instances of these. In some neural systems it is well established that neural activity in individual cells have identifiable real world correlates. Examples of these are place cells in the hippocampus (O'Keefe and Dostrovsky 1971; O'Keefe et al. 1998). and face cells in inferotemporal cortex (Gross et al. 1969, 1972). Determining the exact scope of the item that an actual neuron represents is much more problematic but can be attempted by presenting enough stimuli. Some remarkable recordings from human medial temporal lobe by Quian Quiroga et al. (2005) provide convincing evidence for the existence of cells that recognize items of some specificity, such as a particular actor or building, and are invariant to a broad variety of views of them. On the other hand, there exist other parts of cortex, such as in the prefrontal areas, in which no similar identification between individual neurons and their functions has yet been found, and many neurons appear to be silent.

The issue of neural representations has been widely discussed (Barlow 1972; Page 2000). Our representation can be viewed as particularly specific and simple. In general, the larger r the denser the representation, where the sparse/dense distinction corresponds to the fraction of neurons that are active at a typical instant. If r is large then we have an instance of what some would call a distributed representation. Here we provide a combinatorial explanation for the need for whatever particular value of r is used, and do not assume or require any other assumptions.

3 A basic relationship

In terms of the four parameters n, d, k and r, we can deduce relationship

$$rd/n \lesssim k \tag{1}$$

that holds generally for algorithms of the nature that we shall describe, and which can be tested experimentally. Further, as we shall discuss, it is consistent with recent findings for the locust olfactory system (Jortner et al. 2006). With regard to this relationship we make the following remarks. If r neurons are active and each one is connected by synapses to d other neurons then each neuron receives inputs from an average of rd/n of the active ones. Hence, k cannot be much less than this quantity rd/n, for then all the other neurons would typically become active, precluding meaningful computation. The force of the relationship (1) is that when k is in a narrow

band of values, as large as or a little larger than rd/n, then a rich range of computations of a unified nature can be supported, based though on somewhat complex statistical phenomena (Valiant 2005).

4 Basic tasks and algorithms

Since our methodology seeks to exploit constraints, the tasks we seek to explain are those that are most challenging for networks in which connections may be relatively few and individually weak. We characterize these most challenging tasks as those where arbitrary pairs of distinct previously made memories may have to be related somehow, such as by forming a conjunction of them or making an association. We call these *random access* tasks.

We are interested in basic tasks that have the capability of forming rapidly evaluatable neural circuits that reflect the potentially arbitrary and complex interrelationships among different pieces of knowledge that are required in describing the world. We regard hierarchical memory formation in combination with association to be a viable basis for building such circuits. The widely investigated "associative memories" (Marr 1969; Kanerva 1998; Graham and Willshaw 1997), which have the functionality that information keyed with a fixed key, possibly degraded, can be entered and retrieved, appear to be less effective for this purpose.

Our approach is exemplified by the archetypal random access task of *hierarchical memory formation*, which we define as follows. Given two items A and B already represented, the task is (a) to assign a representing set of neurons for a new item C, and (b) to make further modifications to the circuits as necessary, in the form of state changes at the neurons or weight changes at the synapses, so that the circuit will behave as follows. If at some later time the representatives of both A and B fire then the representatives of C will fire also.

In addition, we also need some non-interference conditions, which ensure that C will not fire if it is not intended to. For example, suppose we determine that more than 95% of a representing set should fire for the associated item to be considered recognized, fewer than 50% if it is not, and that intermediate situations never occur. Then we want less than 50 % of C to fire if less than 50% of A fires even if 100% of B does. *Condition X* represents this condition. *Condition Y* is a condition that ensures that if A and a different item D fire then this will not cause C to fire. (The somewhat more precise description is as follows: For condition X, since the size of C is r only in expectation the technical condition we use is not that fewer 50% of C but that fewer than

$3r/10$ candidate nodes fire. This allows for the fact that C may be smaller than the expected value r. For condition Y the technical condition is that the number of nodes corresponding to $A \wedge B$ and not $A \wedge D$ should be at least $2r/3$. In general, the capacity can be computed by showing that the probability of the specific non-interference conditions that are to be avoided is small enough.)

Our proposal for the mechanism to realize memory formation is the simplest imaginable. *The neurons that are to represent C will be simply those that are richly enough connected to both A and B that activity at A and B can cause each such neuron to overcome its threshold*. We assume a random network of n neurons, each receiving an input from any other independently with probability d/n. In any act of memorization the representatives of A and B are assumed to be fixed, either by preprogramming or by previous acts of memorization. The choice of C is then the result of a process determined by the random connections. Such algorithms may be called *vicinal* to emphasize the fact that their effect and execution is governed to a large extent by the immediate neighborly connections among the neurons.

While testing directly whether the brain executes the above algorithm may be currently impractical, the theory makes some quantitative predictions that are testable. In particular, there is a governing equation that relates n, d, k and r and whose solutions are robust to variations in the equation. Consider the function $\mathcal{B}(m, p, s)$ defined to equal the probability that, in m tosses of a coin that comes up heads, or has success, with probability p, there will be at least s heads or successes. This is simply the upper tail of the Binomial distribution. We will always use it in cases in which s is a little larger than mp, the expected number of successes. The basic property we use is that as s exceeds mp by a larger and larger amount, the tail $\mathcal{B}(m, p, \text{s})$ will get rapidly smaller, and at an exponential rate once this excess is larger than a standard deviation $(mp(1 - p))^{1/2}$.

First we consider the case that A and B are represented by disjoint sets of r neurons. Let us define C to be exactly those neurons that have at least k connections from A and also at least k connections from B. Then the probability that an arbitrary neuron has these required connections to A is $\mathcal{B}(r, d/n, k)$. The same holds for B. It then follows that the probability that an arbitrary neuron has these connection properties to both is $(\mathcal{B}(r, d/n, k))^2$. Now if we want item C to be an "equal citizen" with items A and B in the sense that it has r representatives, at least in expectation, then we want that among the n candidate neurons there be an expected number r with these properties. (Here, as elsewhere, it is convenient to view the random graph as being

Table 1 The entries are solutions for x $(= r/n)$ to equation $(\mathcal{B}(xn, d/n, k))^2 = x$ for various combinations of d and k that hold for all values of $n \geq 10^6$

	$k = 4$	$k = 8$	$k = 16$	$k = 32$	$k = 64$	$k = 128$	$k = 256$	$k = 512$	$k = 1024$
$d = 64$	.0349[6]	.100	.244[9]	.555[9]					
$d = 128$	.0146	.0442	.109[7]	.247[9]	.537[9]				
$d = 256$	.00626	.0198	.0502	.115[9]	.249[9]	.526[9]			
$d = 512$	.00270	.00898	.0234	.0542	.118[9]	.249[9]	.518[9]		
$d = 1024$	.00118	.00410	.0110	.0258	.0568	.120[9]	.250[9]	.512[9]	
$d = 2048$	.000516[6]	.00189	.00517	.0123	.0275	.0586	.122[9]	.250[9]	.509[9]
$d = 4096$	.000227[6]	.000872	.00245	.00594	.0133	.0286	.0598	.123[9]	.250[9]
$d = 8192$	.000101[6]	.000404	.00116	.00286	.00649	.0140	.0294	.0606	.123[9]
$d = 16384$	.0000446[6]	.000188	.000554	.00138	.00316	.00688	.0145	.0299	.0612
$d = 32768$	.0000198[8]	.0000876[6]	.000265	.000670	.00155	.00338	.00715	.0148	.0303
$d = 65536$	.00000886[9]	.0000410[6]	.000127	.000325	.000756	.00166	.00353	.00734	.0151

The solutions given are for $n = 10^9$ and were found numerically. They are accurate to three significant figures for $n = 10^9$ and 10^8, to two for $n = 10^7$ and to one for $n = 10^6$. The superscripts give the greatest power $x \leq 9$ such that for all integer powers $n = 10^y$ for $x < y \leq 9$ the non-interference conditions X and Y hold. For $k = 1$ and 2 condition X is violated for all $n \leq 10^9$ for all values of d shown, and therefore the entries are omitted

generated after the A and B nodes are fixed.) Hence the governing equation is:

$$(\mathcal{B}(r, d/n, k))^2 = r/n. \qquad (2)$$

Qualitatively this equation means that we want the number of successes k to be sufficiently above the expected number of successes rd/n that the resulting probability is r/n. (We think of r/n as being typically small, say smaller than 0.1.) As we have observed, the value of $\mathcal{B}(r, d/n, k)$ drops very rapidly as k increases above the mean rd/n. It follows that the rule of thumb (1) previously stated, that $rd/n \lesssim k$, needs to hold, where the relation $\lesssim$ denotes "is of the same order of magnitude and a little smaller than".

We note that relation (1) holds equally for several important variations of this formulation (Valiant 2005).

In Table 1 we tabulate solutions of $(\mathcal{B}(r, d/n, k))^2 = r/n$. We note that if k and d are fixed then as r and n grow in proportion to each other, in particular, if $r = xn$ for a fixed constant x, then $\mathcal{B}(r, d/n, k) = \mathcal{B}(xn, d/n, k)$ asymptotes to a fixed value dependent only on the two quantities xd and k, in the form of the Poisson distribution. Then equation (2) becomes $(\mathcal{B}(xn, d/n, k))^2 = x$, and its solution for x is essentially invariant to n once n is large enough. Hence we can regard these systems as being governed by just the three parameters d, k and $x = r/n$.

The discussion so far establishes conditions for the wiring diagram that are sufficient to support hierarchical memory formation. To complete the theory one also needs to have a model of computation that enables the neuron and synapse updates to be specified explicitly. Such a model is described in reference (Valiant 2005). There it is also shown that the task of association can be realized with the same parameters as those needed for memory formation, where *association* is defined as

follows: given two items A and B already represented the task is to make modifications in the circuit so that whenever in the future A fires so will B also.

5 Correspondences with experimental findings

The classical model of vision in cortex is as a hierarchy. As one ascends it the complexity of the items represented by a neuron increases, as does their invariance to size, translation, etc. We hypothesize that the higher levels of the vision hierarchy require the capabilities of some form of hierarchical memory formation. Unfortunately, we do not know of any such system for which all of the parameters d, k and r/n, have been measured. However, the results of Quian Quiroga et al. (2005) do imply estimates for r/n in human medial temporal lobe, and one can ask whether these are consistent in our table with reasonable values of the anatomical parameters d and k. Here is a brief summary of their experiments and findings: They recorded from 998 units (346 single- and 652 multi-neurons) across a number of human patients variously in hippocampus, parahippocampal gyrus, amygdala and entorhinal cortex. Each morning they presented each patient with 82–110 (mean 94) visual stimuli (Kreiman, personal communication). For each patient they selected some 7–23 (mean 14) stimuli that provoked significant or the strongest responses, and in the afternoon for each such stimulus they presented 3–8 different views of the item that the experimenter had predetermined to be represented in that stimulus. For example, one chosen item was the actress Halle Berry, and different views included a portrait of her, a portrait of her in her recent role in a mask, and her name written as a string of letters. They

found that 52 of the units did respond to the experimenters' prechosen item, and of these 8 responded to two items. Thus we can conclude that there were 998×94 attempts to identify a neuron-item correspondence, and of these 60 succeeded. This implies an estimate for r/n of $60/(998 \times 94) = 0.00064$. Looking at our Table 1 we note that there are some plausible values of k and d, such as $k = 8$ and $d = 8,192$ that correspond to an r value in that range. The majority of entries there have higher values of r/n, and it is indeed probable that this calculation for r/n gives an underestimate. While the fact that some of the neurons recorded from were multiple neurons would suggest that this calculation overestimates, there are two other implicit assumptions that suggest the opposite. First, it was assumed that all the neurons (or units) in those areas represent some visual item, while it is quite possible that a significant fraction have a different purpose altogether. Second, it was assumed that the patients were familiar with and had representations for all the items presented to them.

The one system we know in which all four parameters n, d, k, and r have been measured is in the olfactory system of the locust as investigated by Jortner et al. (2006). This system consists of a set of about 830 projection neurons (PN) in each antennal lobe, which in turn connect to a set of about 50,000 Kenyon cells (KC). The system is subject to 20–30 Hz cycles. Within each cycle each cell undergoes 0 or 1 spikes. Within one such cycle a natural odor stimulus will cause about 100–150 of the 830 PNs to fire, and a comparatively smaller fraction of the KC's, thought to be in the range 0.01–0.5% (Laurent personal communication). The circuit sparsifies the representation of the odors, apparently to facilitate processing upstream.

We shall use distinguishing notation for the two levels of neurons, lower case (r, n) for the PNs and upper case (R, N, D, K) for the KCs. From what we have already said the parameters of the PNs are $n = 830$ and r in the range 100–150. For the KCs $N = 50,000$ and r, to which our theory is rather insensitive, is in the range 5–250. Further, the estimate (Jortner et al. 2006) for the fraction D/n of PNs from which a KC receives connections is 0.50 ± 0.13. Finally, the estimate (Jortner et al. 2006) of K, derived from the ratio of the voltage threshold of a KC to the mean postsynaptic potential caused by a single PN is 100.

Our interpretation of this circuit in our theory is the following. Its connections form a random bipartite graph where each PN is connected to each KC with probability D/n. An arbitrary odor causes an arbitrary subset of r of the PN's to fire. Given a fixed chosen KC, what is the probability that it will be caused to fire by that odor. The answer to this is simply $\mathcal{B}(r, D/n, K)$. Here each toss

corresponds to one of the r PNs that is activated by the odor, and success for a toss corresponds to that PN being connected to the chosen KC. Further we want the probability that this chosen KC fires to be R/N. *Then the R KCs that will represent the odor will be simply those that are connected by at least K connections to the specific PNs that are activated by the odor.*

Hence, the six measured quantities should satisfy the equation:

$$\mathcal{B}(r, D/n, K) = R/N. \qquad (3)$$

As far as the approximate rule of thumb (1) derived above, we note that taking the mean values of 125 and .5 of the estimates for r and D/n gives an estimate of 62.5 for rD/n, which conforms roughly to the stated rule that rD/n should be of the same order of magnitude and a little smaller than K, which is estimated to have value 100.

For a more precise fit we note that the range estimate for D/n is [0.37, 0.63] and for r is [100, 150]. The theory predicts that there should be solutions to equation (3) for values in this range that give realistic values of R/N, which are in the range [0.0001, 0.005]. It is easy to verify that this is indeed the case. Taking the value $D/n = 0.60$ and $K = 100$ the following are pairs of solutions for $\{r, R/N\}$ of the equation $\mathcal{B}(r, 0.60, 100) = R/N$:

$\{130, 0.000035\}, \{132, 0.000102\}, \{134, 0.000272\},$

$\{136, 0.000671\}, \{138, 0.001529\}, \{140, 0.003244\},$

$\{142, 0.006436\}, \{144, 0.011993\}, \{146, 0.021084\},$

$\{148, 0.035101\}, \{150, 0.055538\}.$

These solutions span the estimated range of 0.0001–0.005 for R/N. Jortner et al. (2006) note that their estimate of $K = 100$ may be an overestimate because of nonlinear effects. Values of r and D/n in the middle of their estimated ranges of these would then be consistent with appropriate lower values of K.

If instead of the above deterministic view we regard each of the r PNs as firing with probability $p < 1$, then equation (3) becomes $\mathcal{B}(r, pD/n, K) = R/N$ and the rule of thumb (1) becomes: $prd/n \lesssim k$. Experimental evidence suggests that PNs are probabilistic in this sense. The 100–150 estimate for r refers to the number of the more reliable ones, and a further large number is activated less reliably.

6 Discussion

In conclusion, we shall now argue that for neural systems generally that compute random access tasks, two distinct regimes are discernible. From Table 1 we see that if k is

large then r is a significant fraction of n. For example for $k = 32$ and $d = 8192$ we get $r/n = 0.00286$. This immediately suggests that if we are to memorize more than $1/.00286 \sim 350$ items then we need shared rather than disjoint representations. A large value of k therefore forces the combination of (1) large k, (2) large r, and (3) shared representations. This combination we shall call *regime-α*. In fact, taking the non-inteference conditions into account one can argue that any value of k larger than 1 can be considered to be large in this context. In this regime we have discussed the task of memory formation. Also, it has been shown that the same parameters support association (Valiant 2005).

In earlier work (Valiant 1994) we considered the combination of (1) strong $k = 1$ synapses, (2) small r, and (3) disjoint representations. We shall identify this as *regime-β*. In this regime, while the paucity of interconnections is still a serious constraint, the system is closer to digital computation than is regime-α in the following senses: Information is localized in that representations are disjoint, and influence is localized in that a single neuron is able to cause an action potential in a neighbor. This regime has been shown to support a broader set of random access tasks (Valiant, 1994), including appropriate formalizations of: supervised memorization of conjunctions and disjunctions, inductive learning of conjunctions, disjunctions and appropriate linear separators, correlational learning, memorization and inductive learning of conjunctive expressions involving multiple objects and relations.

It seems that in systems in which neuron-item correspondences can be discovered by probing a few hundred or a few thousand neurons, r must be a significant fraction of n for otherwise there would be little chance of such a correspondence being found. We suggest that within such systems, which include the hippocampus and inferotemporal cortex, regime-α is in operation. In these areas, which do not represent the highest cognitive areas, the tasks of memory formation and association, for which we have quantitative mechanistic explanations, may be a sufficient basis for performing the required processing.

On the other hand, there are areas such as prefrontal cortex, where item-neuron correspondences have proved more difficult to identify. A natural theory is that in such areas regime-β has some significant role. Any failure to find item-neuron correspondences can be then explained by the small values of r, and the computation of apparently more complex tasks by the possibly greater computational capability of regime-β. Of course, the price to be paid is strong synapses. Synapses there need to have a large dynamic range so that they can change in strength between an average weak level to $k = 1$ strength

through learning. Whether neurons do this is currently unknown. Strong synapses, however, have been observed (Thomson et al. 1993; Markram and Tsodyks 1996; Ali et al. 1998) and one recent study shows that the distribution of synaptic strengths is heavy-tailed at the strong end (Song et al. 2005).

Our analysis treats the human temporal lobe results as exemplars of regime-α. The locust olfactory system appears to be an interface between regime-α and regime-β, since it translates between large r to small r in a single step. There is evidence that different parts of cortex have widely different cortical parameters, as would be expected if they employed such different regimes. It has been found (Elston et al. 1999) that in macaques the number of spines in a polysensory area exceeds that in the primary visual area by a factor of thirteen.

We suggest that this paper is a first step in a research program to validate a general quantitative algorithmic theory of neural computation against experimental evidence. The numerous immediate directions of further work include pursuing the following questions: Can the values of the four parameters n, d, k, and r be measured for systems other than the locust olfactory system? Can it be shown that synapses can change between average strength to $k = 1$ strength by learning, as required in regime-β? Can direct experimental evidence be found for some system that it works in regime-β? Is regime-β provably more powerful than regime-α in some significant computational sense?

References

Abeles M (1991) Corticonics: neural circuits of the cerebral cortex. Cambridge University Press, Cambridge

Ali AB, Deuchars J, Pawelzik H, Thomson AM (1998) CA1 pyramidal to basket and bistratified cell EPSPs: dual intracellular recordings in rat hippocampal slices. J Physiol 507:201

Barlow HB (1972) Single units and sensation: a neuron doctrine for perceptual psychology. Perception 1:371

Braitenberg V (1978) In: Heim R, Palm G (eds) Theoretical approaches to complex systems. Lecture notes in biomathematics, vol. 21. Springer, Berlin Heidelberg New York, pp 171–188

Braitenberg V, Schüz A (1998) Cortex: statistics and geometry of neuronal connectivity. Springer, Berlin Heidelberg New York

Elston GN, Tweedale R, Rosa MGP (1999) Cortical integration in the visual system of the macaque monkey: large-scale morphological differences in the pyramidal neurons in the occipital, parietal and temporal lobes. Proc R Soc Lond B 266:1367

Feldman JA (1982) Dynamic connections in neural networks. Biol Cybern 46:27

Feldman JA, Ballard DH (1982) Connectionist models and their properties. Cog Sci 6:205

Graham B, Willshaw D (1997) Capacity and information efficiency of the associative net. Network: Comput Neural Syst 8:35

Griffith JS (1963) On the stability of brain-like structures. Biophys J 3:299

Gross CG, Bender DB, Rocha-Miranda CE (1969) Visual receptive fields of neurons in inferotemporal cortex of the monkey. Science 166:1303

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J Neurophysiol 35:96

Jortner RA, Farivar SS, Laurent G (2006) Dense connectivity for sparse coding in the locust olfactory system. Manuscript

Kanerva P (1988) Sparse distributed memory. MIT, Cambridge

Markram H, Tsodyks M (1996) Redistribution of synaptic efficacy: A mechanism to generate infinite synaptic input diversity from a homogenous population of neurons without changing absolute synaptic efficacies. J Physiol (Paris) 90:229

Marr D (1969) A theory of cerebellar cortex. J Physiol 202:437

O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely moving rat. J Brain Res 34:171

O'Keefe J, Burgess N, Donnett JG, Jeffery KJ, Maguire EA (1998) TI – Place cells, navigational accuracy, and the human hippocampus. Philos Trans R Soc Lond B 353:1333

Page M (2000) Connectionist modelling in psychology: a localist manifesto. Behav Brain Sci 23:4443

Palm G (1982) Neural assemblies: an alternative approach to artifcial intelligence. Springer, Berlin Heidelberg New York

Quian Quiroga R, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. Nature 435:1102–1107

Song S, Sjostrom PS, Reigl M, Nelson S, Chklovskii DM (2005) Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol 3:3, 507

Thomson AM, Deuchars J, West DC (1993) Large, deep layer pyramid-pyramid single axon EPSPs in slices of rat motor cortex display paired pulse and frequency-dependent depression, mediated presynaptically and self-facilitation, mediated postsynaptically. J Neurophysiol 70:2354

Valiant LG (1994) Circuits of the mind. Oxford University Press, New York

Valiant LG (2005) Memorization and association on a realistic neural model. Neural Comput 17:3, 527