

# SINBAD: A neocortical mechanism for discovering environmental variables and regularities hidden in sensory input

Oleg V. Favorov<sup>1,\*</sup>, Dan Ryder<sup>2</sup>

<sup>1</sup>School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816-2362, USA

<sup>2</sup>Department of Philosophy, University of North Carolina at Chapel Hill, NC 27599-3125, USA

Received: 17 December 2003 / Accepted: 23 December 2003 / Published online: 12 March 2004

**Abstract.** We propose that a top priority of the cerebral cortex must be the discovery and explicit representation of the environmental variables that contribute as major factors to environmental regularities. Any neural representation in which such variables are represented only implicitly (thus requiring extra computing to use them) will make the regularities more complex and therefore more difficult, if not impossible, to learn. The task of discovering such important environmental variables is not an easy one since their existence is only indirectly suggested by the sensory input patterns the cortex receives – these variables are “hidden”. We present a candidate computational strategy for (i) discovering regularity-simplifying environmental variables, (ii) learning the regularities, and (iii) using regularities in perceptual and decision-making tasks. The SINBAD computational model discovers useful environmental variables through a search for different, but nevertheless highly correlated, functions of any kind over nonoverlapping subsets of the known variables, this being indicative of some important environmental variable that is responsible for the correlation. We suggest that such a search is performed in the neocortex by the dendritic trees of individual pyramidal cells. According to the SINBAD model, the basic function of each pyramidal cell is to (i) discover and represent one of the regularity-simplifying environmental variables and (ii) learn to infer the state of its variable from the states of other variables, represented by other pyramidal cells. A network of such cells – each cell just attending to representation of its variable – can function as a sophisticated and useful inferential model of the outside world.

## 1 Introduction

We live in a world that is, to a large degree, orderly. All aspects of our lives – the way our hands interact with objects or our balance and mobility are affected by the surface we stand on, what significant events happen around us and what impacts we can have on them, our overall living conditions, etc. – are to a large degree predictable by us; this is what enables us to have purposeful and successful behaviors. Mammals, in general, are very adept at discovering this order through learning. They are born with only rudimentary innate skills and understanding of the world, but in the course of interacting with their surroundings they gradually discover regularities that are hidden in the spatiotemporal patterns of activities of their sensory receptors, allowing them to acquire great expertise in dealing with the challenges and opportunities posed by their environments.

Discovering hidden order is possibly the most challenging task performed by the brain. Its demands are certain to be among the most fundamental determinants that shaped the brain’s functional design. In this paper we consider obstacles to learning environmental regularities, propose how these obstacles can be overcome, and formulate a general computational approach to discovering deeply hidden regularities. We next suggest how this approach might be implemented in the neocortex. Finally, we conclude with how the neocortex might internally represent the environmental order it discovers and how such an internal model of the outside world might guide interactions of an animal with its environment.

## 2 The core idea: learning the identities of important environmental variables

### 2.1 What are the obstacles to learning regularities?

Interdependence among environmental variables – arising from interactions among the constituents of the

Correspondence to: O.V. Favorov  
(e-mail: favorov@bme.unc.edu,  
Tel.: +1-919-9661291, Fax: +1-919-9662963)

\* Present address:

Department of Biomedical Engineering,  
CB #7575, University of North Carolina,  
Chapel Hill, NC 27599-7575, USA

environment – makes it possible to infer (i.e., compute with some degree of accuracy) the state of one variable from the known states of some other variables. In other words, the variables have *inferential relations* with other variables. The question that we will consider first is how a hypothetical observer can discover as many of these inferential relations as possible.

We will illustrate our considerations with the example of a brain prototype whose task is to learn the orderly nature of kitchen sinks through observation of different sinks in different states. To simplify matters, we will suppose that all sinks in our example receive water from two pipes and are controlled by the following five variables:  $HC$  indicates which of the two pipes carries hot water and which cold water (0 – the left pipe is hot and the right pipe is cold, 1 – vice versa),  $DIR_L$  and  $DIR_R$  are the directions in which the left and the right knobs should be turned to open the pipes (0 – clockwise, 1 – counterclockwise), and  $KP_L$  and  $KP_R$  are the radial positions of the two knobs. Sinks in different kitchens can have different  $DIR_L$ ,  $DIR_R$ , and  $HC$ ;  $KP_L$  and  $KP_R$  can vary across time in the same sink.

These five variables determine the flows of water through the two pipes:

$$\begin{aligned} F_L &= (1 - DIR_L - KP_L + 2 \cdot DIR_L \cdot KP_L)/2 \quad \text{and} \\ F_R &= (1 - DIR_R - KP_R + 2 \cdot DIR_R \cdot KP_R)/2 \quad . \end{aligned} \quad (1)$$

In turn, the flows of water through the two pipes determine the total water outflow from the faucet and its temperature:

$$\begin{aligned} F_T &= F_L + F_R \quad \text{and} \\ t^\circ &= ((1 - HC) \cdot F_L + HC \cdot F_R)/(F_L + F_R) \quad . \end{aligned} \quad (2)$$

$$t^\circ = \frac{(1 - HC)(1 - DIR_L - KP_L + 2 \cdot DIR_L \cdot KP_L) + HC \cdot (1 - DIR_R - KP_R + 2 \cdot DIR_R \cdot KP_R)}{2 - DIR_L - KP_L + 2 \cdot DIR_L \cdot KP_L - DIR_R - KP_R + 2 \cdot DIR_R \cdot KP_R} \quad . \quad (3)$$

To learn the orderly nature of sinks is to learn the inferential relations among these variables. This task may be defined explicitly as learning how to determine the state of each variable from the states of the other variables.

Brains have only limited direct sensory access to things in the world. To reflect this situation, suppose our brain prototype cannot see inside the pipes and therefore does not know about the flows of water through them, i.e., variables  $F_L$  and  $F_R$  are not given to it by its sensors. Luckily, because  $F_L$  and  $F_R$  are implicit in the states of other variables, the relations among the remaining variables can still be defined without  $F_L$  and  $F_R$ . *But these relations will be more complex.* For example, compare equation (3), expressing  $t^\circ$  as a function of the known variables  $HC$ ,  $DIR_L$ ,  $KP_L$ ,  $DIR_R$ ,  $KP_R$  with the much simpler equation (2). The more complex the inferential relations, the more difficult it will be to learn them; at some degree of complexity it will become impossible.

To draw a lesson from this example, *a major source of difficulty in learning regularities (inferential relations) is*

*not knowing about the existence of some of the factors contributing to them.* This is a fundamental problem faced by real brains, considering that their sensory receptors reflect only indirectly the environmental variables that are relevant to behavior. If an environmental variable that plays an important role in some regularity is not among the variables known to an observer but is reflected implicitly in the behaviors of some of the known variables, then the observer can still, in principle, learn the regularity. However, the regularity will now become more complex, involving all these extra variables with their implicit information about the missed key variable. This involvement of more variables and, likely, extra nonlinearities contributed by them will make the regularity more difficult to learn. Given that learning capabilities are not unlimited (no matter how intelligent the observer), a regularity can thus easily be placed beyond an observer's grasp.

This lesson illustrates an argument of Clark and Thornton (1997) about the necessity of representing inferentially important environmental variables explicitly: *to learn regularities, it is crucial first to learn separately the identities of as many environmental factors contributing to those regularities as possible.* Clark and Thornton call this “trading representation against computation”. In the case of kitchen sinks, our prototypical brain should learn about the flow of water through each pipe, i.e., it should learn about the existence of variables  $F_L$  and  $F_R$  from observing the known variables, and it should learn to infer the states of  $F_L$  and  $F_R$  from the states of the other variables, and vice versa. Knowledge of  $F_L$  and  $F_R$  will simplify our brain prototype's task of learning the regularities governing the sink, i.e., how each variable relates to the other vari-

ables. For example, instead of (3), it will need to learn the *simpler* equation (2).

But how can our brain prototype discover the existence of these inferentially useful variables  $F_L$  and  $F_R$ ? It will have to derive them from the known variables, i.e., it will have to learn to compute their states from the states of the known variables.

## 2.2 How can important environmental variables be discovered?

Hidden environmental variables that are involved in regularities but are reflected only implicitly in the known variables can be learned from the known variables by looking for what Barlow (1992) calls “suspicious coincidences” among sensory events. That is, we should look for “something” in one subset of the known variables that correlates with “something” in another such subset. These “somethings” are, mathematically, *functions* of some sort over the two subsets of the known

variables. Thus we should look among various functions over various subsets of the known variables, searching for such pairs of functions over nonoverlapping subsets of variables that will show correlated behaviors. There must be a *reason* why these different functions are correlated, something in the environment that is responsible for their statistical interdependence. This means that by identifying a pair of correlated functions we, in effect, detect some influential variable operating in the environment that is responsible for the correlation (Barlow 1992; Becker and Hinton 1992; Phillips and Singer, 1997; Ryder and Favorov 2001; Ryder 2004). Once we find such correlated functions, we can use them as alternative estimates of the same, likely to be regularity-simplifying, environmental variable.

Such influential variables are very likely to have other effects in the observed environment besides the ones that led to their recognition (Ryder 2004). And once we have learned to recognize an influential environmental variable by some of its effects, it will become easier to notice its other effects. First, it will become easier to learn those inferential relations that depend on this variable. Second, it will also become easier to learn to recognize this variable by its other – subtler – manifestations. Finally, once we have learned to recognize a number of new variables, we can look for suspicious coincidences among *them* (i.e., among functions over these variables), thus discovering higher-order influential variables, etc.

If we apply this method of discovering inferentially useful variables to kitchen sinks, we will observe that the following two functions,  $f_a(DIR_L, KP_L) = (1 - DIR_L - KP_L + 2 \cdot DIR_L \cdot KP_L)/2$  and  $f_b(F_T, HC, t^\circ) = (F_T \cdot (t^\circ - HC))/(1 - 2 \cdot HC)$ , correlate perfectly, i.e.,  $f_a(DIR_L, KP_L) = f_b(F_T, HC, t^\circ)$ . The fact that these two functions correlate perfectly indicates that they compute a variable  $X = f_a(DIR_L, KP_L) = f_b(F_T, HC, t^\circ)$  that plays a significant role in kitchen sinks. Variable  $X$ , in fact, expresses the flow of water in the left pipe,  $F_L$ . As we saw above in comparing (2) and (3), the discovery of this variable will simplify the inferential relations among the sink variables. For examples of applications of this approach, such as discovery of surfaces in stereograms, recognition of moving objects, and speaker-independent vowel recognition, see Becker and Hinton (1992), Becker (1996, 1999), and Stone (1996).

### 2.3 How does the neocortex discover important environmental variables?

Our current knowledge of neocortical neurons and their networks is clearly too limited to answer this fundamental question with any degree of certainty. Still, it is important to consider the possibilities based on what is known because only such considerations will identify what critical knowledge is missing and how to acquire it (by making predictions and by suggesting how these predictions can be tested experimentally). In this spirit, we offer here our exploratory proposal on how the neocortex might discover influential environmental variables, using it to highlight the issues involved in this

task and to identify some major, but previously neglected, questions concerning basic cortical properties.

The general computational approach to finding correlated functions over different inputs has been developed by Becker (1995, 1996, 1999) and Becker and Hinton (1992). The basic idea is to set up two or more computational modules that look at separate but related parts of the sensory input and to make these “search” modules *teach each other* to produce outputs that match their own as closely as possible. In this way the modules will learn to perform different operations (i.e., functions) on their different inputs so as to yield correlated outputs. The interlinked modules will thus learn to “speak with a single voice”, which will carry explicit information about some influential environmental variable, the one that explains the correlation. We call this general computational approach “the mining of correlated functions”.

Becker’s basic concept forms one of the cornerstones of Phillips and Singer’s (1997) view of the cerebral cortex. Phillips and Singer propose that contextual (i.e., lateral, as opposed to afferent) inputs to cortical cells guide them to tune to those stimulus features in their receptive fields that are predictably related to the context in which they occur. In this way, by discovering predictable relations between different inputs (within vs. outside their receptive fields), cortical cells become better indicators of objectively important variables in the external world. Phillips and Singer offered a possible mechanism for how such tuning might be accomplished. Unfortunately, that mechanism is limited in its practical utility due to its inability to search for correlations among *nonlinear* functions, although these are the functions most likely to relate influential environmental variables to each other in the real world. Still, Phillips and Singer’s focus on individual cortical cells as discoverers of correlated functions is very appealing: placing the mechanism in the dendritic trees of individual cells (rather than requiring multicellular modules) greatly expands the number of inferentially useful variables that the brain will be able to discover in the environment. As Barlow (1995) has argued, such an emphasis upon the power of individual cortical cells does not entail any unacceptable commitment to “grandmother cells”.

Looking in the dendritic apparatus of neocortical pyramidal cells for a candidate mechanism, we find the main, or *principal*, dendrites originating from the soma to be the most promising candidates for the role of Becker’s search modules (Ryder and Favorov 2001). Neocortical pyramidal cells have 5–8 principal dendrites: 4–7 basal dendrites and the apical dendrite with its side branches (Feldman 1984). Each principal dendrite sprouts an elaborate, treelike pattern of branches. Just as is required of the search modules, such a dendritic tree is capable of complex forms of nonlinear integration of its synaptic inputs (Mel 1994; Häusser et al. 2000; Koch and Segev 2000; Segev and London 2000). Even more, we observe that (i) the shape of a dendritic tree is a major determinant of its integrative properties, (ii) principal dendrites in different and in the same cells are all very individualistic in the shapes of their trees, and (iii) these trees are eminently malleable in response to

sensory experiences (Quartz and Sejnowski 1997; Woolley 1999; Matus 2000; McAllister 2000). Together these observations suggest that, just as is expected of the search modules, different principal dendrites implement different input-to-output transfer functions and, furthermore, that these functions are shaped by sensory experiences. (For a direct experimental demonstration of stimulation-induced changes in nonlinear integrating properties of pyramidal dendrites, see Wang et al. 2003).

Another feature of principal dendrites that supports their candidacy for the role of the search modules concerns the requirement that the modules *teach* each other to produce correlated outputs. In dendrites, synaptic learning is controlled both by the presynaptic activity and by the output activity of the postsynaptic cell; the latter is signaled to each synapse by spikes that are backpropagated from the soma up through each dendrite (Markram et al. 1997; Stuart et al. 1997). This means that, through its contribution to the cell's output, the output of each principal dendrite can influence the strengths of synaptic connections on the cell's other dendrites. Due to the Hebbian properties of cortical synaptic plasticity (e.g., Singer 1995; Paulsen and Sejnowski 2000), the effect of this influence will be for each principal dendrite to drive the other principal dendrites of the same cell to learn to behave the way it does. Thus it appears that each principal dendrite will, as required of the search modules, teach the other dendrites of the same cell to behave the way it does while also learning from them. If that were the case, then through such mutual teaching and learning all the principal dendrites in a cell would be learning to produce *correlated* outputs.

Now, turning to potential challenges to our choice of principal dendrites to play the role of search modules, there are two major areas of uncertainty. The first one concerns the types of input-to-output transfer functions that a principal dendrite can be made to learn by other dendrites. There is little doubt that principal dendrites can be taught linear functions, but those would be of limited utility in discovering influential environmental variables. Instead, to function effectively as the search modules, principal dendrites in a cell must be capable of teaching each other at least simple, if not moderately complex, nonlinear transfer functions. This is not an implausible expectation given that neocortical dendritic learning is turning out to be a very complex, multifactorial phenomenon, as evidenced by the active and constantly evolving experimental and theoretical research in this area (e.g., Markram et al. 1998; Johnston et al. 1999; Häusser et al. 2000; Segev and London 2000; Paulsen and Sejnowski 2000; Williams and Stuart 2000; Poirazi and Mel 2001; Stuart and Häusser 2001; Vetter et al. 2001; Wang et al. 2003). Unfortunately, it is still a poorly understood phenomenon and, in particular, little is known about whether and how a dendrite can modify its nonlinear integrative properties. Consequently, our expectation that principal dendrites are capable of teaching one another nonlinear transfer functions should be regarded at this point as a necessary, but largely unsubstantiated, prediction in need of continued experimental testing.

Another requirement for pyramidal cells to be effective in mining correlated functions is that diverse sensory information reaches different principal dendrites of the same cell. To function as the search modules, different principal dendrites in a cell must be exposed to different sets of environmental variables that carry implicit information about the same hidden environmental variable(s). In support of this requirement, there are a number of reasons to suppose that neocortical pyramidal cells are exposed to diverse but related information. One source of this diversity is cortical topographic organization. Its significance lies in the fact that most regularities in natural environments (or at least most of the regularities that we are able to appreciate perceptually) are local in one way or another. For example, lower-order regularities in peripheral input patterns involve environmental conditions in close spatial proximity to each other; consequently, exposing different dendrites of a pyramidal cell in primary sensory cortex to raw information from distant spatial locations would be useless. In agreement with this observation, afferent connections to cortical areas do not all contact each and every cell there but have clear topographic organization (e.g., body maps in somatosensory and motor cortices, retinotopic maps in visual cortex, etc.). These maps are created in middle cortical layers by a host of genetic and epigenetic mechanisms (e.g., von der Malsburg and Singer 1988). In the context of this paper's argument, we anticipate that the mechanisms that control perinatal development and adult maintenance of cortical topographic maps are designed to supply each cortical neighborhood with limited but functionally related information in order to improve cells' chances of finding important environmental variables (Barlow 1986).

To suggest one such mechanism, as reviewed elsewhere (Favorov and Kelly 1996), when receptive fields of cortical cells are considered in toto, in all their dimensions, neighboring neurons typically have little in common – a stimulus that is effective in driving one cell will frequently be much less effective in driving its neighbor. Correspondingly, even adjacent neurons carry in common less than 20% of stimulus-related information (Gawne et al. 1996; Richmond et al. 1997; Vinje and Gallant 2000; Reich et al. 2001). Thus it appears that local neuronal populations bring together a variety of different but related sensory information concerning a local region of the stimulus space, with neighboring neurons carrying only minimally redundant information about what takes place in that stimulus space. Because this information comes from a local region of the stimulus space, it is likely to be rich in regularities reflecting the orderly features of the outside world. Thus it appears that local cortical neighborhoods create local informational environments enriched in regularity but low in redundancy (for a possible mechanism responsible for differences among neighboring minicolumns, see Favorov and Kelly 1994a,b). Such local cortical environments are exactly the right informational environments for pyramidal cells, to be mined by them in their search for correlated functions. That is, each basal dendrite of a pyramidal cell will extend through a functionally distinct cortical territory and be

exposed to a unique combination of afferent information via synaptic connections with radially oriented axons of neurons there (DeFelipe and Farinas 1992; Malach 1994). Thus exposed to different but related sensory information, basal dendrites in a cell will have an opportunity to find correlated functions over different environmental variables, tuning the cell to the underlying environmental factor.

Are the inputs to different principal dendrites of neocortical pyramidal cells sufficiently diverse to be consistent with their proposed role as the search modules? The answer to this question depends greatly on the form of nonlinear learning that these dendrites perform, which – as we just discussed – is itself an unanswered question. Because of this uncertainty, all that we can do now is draw attention to the functional significance of the subject of input differences among pyramidal cell dendrites and make the prediction that these differences will turn out to be sufficient to enable pyramidal cells to discover hidden environmental variables.

Thus to conclude, while other possibilities ought to be explored as well, we are inclined toward the possibility that the task of discovering hidden environmental variables is carried out by the principal dendrites of individual neocortical pyramidal cells. In this scenario, different principal dendrites in a cell take different sensory inputs but learn to convert their different inputs to mutually correlated outputs (even though that will often require learning nonlinear operations). By thus discovering correlated functions over different subsets of environmental variables, dendrites tune the cell to the environmental variable responsible for this correlation. We have named this conceptual model of the pyramidal cell “SINBAD” (Ryder and Favorov 2001). This name is an acronym that stands for *Set of INteracting BAcKpropagating Dendrites*, where “backpropagating” refers to the teaching signal (e.g., a spike) that is sent back along the dendrites from the soma as a part of dendritic learning.

Pyramidal cells with access to the original, sensory-receptor-expressed variables can be expected to find only the least hidden among regularity-simplifying environmental variables. However, by feeding the outputs of those cells to another layer of pyramidal cells (e.g., primary visual cortical area V1 projecting to V2), the latter cells will be able to derive the next generation of variables from the variables derived by the first layer, revealing a deeper order (Mel 1997). The next layer of pyramidal cells will discover variables implicit in those found by the preceding layer, etc. This way, a series of layers (i.e., cortical areas) will be able to discover very deeply hidden orderly features of the environment.

### 3 An inferential model: the cortical representation of the world's order

#### 3.1 Learning the environmental regularities

To move from the question of discovering important environmental variables to the one of how these

variables can be used in learning and representing environmental regularities, let us sketch this task first in general terms. Thus the basic problem that we face is that we are given a set of original (sensor-given) variables that are only distantly related to each other. The solution is to proceed by first deriving new variables from the original ones, as described in Sect. 2.2. These new variables, as prominent environmental factors, will have inferential significance for other variables and thus will provide inferential *links* between the variables used to determine their states and the variables whose states can be inferred from them. Thus, by placing the newly derived variables in between the original ones, we will break down the complex inferential relations among the original variables into simpler inferential relations, from the original variables to the derived ones, and from the derived variables to other original ones. The more inferentially significant variables we add to our repertoire (deriving them from the original and the already derived ones), the more distant inferential relations will be broken down into ones that are simpler and easier to learn.

Thus, as a general approach, we should aim to derive as many variables as we can and learn as many ways to infer each variable from the other ones as we can. By thus expressing each variable in many different ways in terms of other variables, which in turn are expressed in terms of yet other variables, etc., we will construct a rich web of inferential relations. In the ideal case, this web (which is grounded in the orderly structure of the observed environment) will link each variable either directly or via intermediaries to every other variable.

In the cortex, such learning of inferential relations can be accomplished by making use of lateral connections among cells within each cortical area and feedback connections from higher-level cortical areas. Via these connections pyramidal cells will have access to the variables derived by other cells in the same and in higher cortical areas. This means that the principal dendrites will have additional information to use in learning how to predict the outputs of the other dendrites on the same cell. This way, each pyramidal cell will learn to determine the state of the environmental variable it represents from many alternative and complementary lines of evidence, some direct, others circumstantial. The cortex, as a whole, will capture the web of inferential relations among the original and derived environmental variables in its pattern of ascending, lateral, and feedback connections.

In this web all the discovered inferential relations will be tied together into a single functional entity – *an inferential model* of the observed environment (Ryder 2004). The basic mode of operation of such an inferential model is what might be called “filling in missing parts of a picture”. The “picture” is the model's representation of the observed environment, cast in terms of the observer-chosen variables (original and derived). Faced with any particular situation, the observer will obtain information about some of the environmental variables from its sensors, thus establishing the picture's factual foundation. This will cause representations of the

other variables to be “filled in” by way of their known inferential relations to the observed variables, adding new details to the emerging picture of the situation or clarifying some fuzzy ones. These new details might make it possible to infer the states of some other variables, which in turn might make it possible to infer yet more variables, and so on, thus gradually elaborating the picture of the situation with more and more details inferred flexibly and opportunistically from whatever circumstantial information is at hand.

If the observer itself is active in the environment, then its own needs, desires, intentions, and motor actions become important factors affecting the environment. Consequently, these factors will be identified as influential environmental variables by the observer and will be incorporated into the web of inferential relations. In this case, the observer becomes a part of its picture of the observed environment, making its actions inferable from what is present and what is desired. That is, the web of inferential relations can be used not only to infer what has happened, what is happening, and what will happen, but also what should be done, i.e., how to reach the desired goals. Possession of an internal inferential model will enable the cortex first to obtain from sensory organs some information about a situation and, next to use its knowledge of inferential relations among environmental variables, to add more and more detail about that situation’s past, present, and (potentially action-involving) future. Representations of inferred actions can then be translated into actual physical actions in the environment.

### 3.2 Kitchen sink example

To illustrate the capabilities of a network of SINBAD neurons to discover prominent factors in its sensory environment and to develop an internal inferential model of that environment, we set up a network of 15 SINBAD neurons and exposed it to kitchen sinks. For simplicity, SINBAD cells in this exercise were given only three dendrites. Our approach to modeling SINBAD cells was shaped by the following consideration (Ryder and Favorov 2001). Since we predict that principal dendrites in a pyramidal cell should be able to teach each other nonlinear input-to-output transfer functions, but at this time we do not know its biophysical mechanism, we cannot rely on a standard compartmental modeling approach (Segev et al. 1989) to render the dendrites. Instead, we represent each principal dendrite by an error-backpropagating network (Rumelhart et al. 1986), with the somal output as the training signal. The reason for this choice is that according to our proposal, each principal dendrite should be capable of learning, under supervision by the other dendrites, functions over its inputs that might be nonlinear. In this regard dendrites resemble artificial backpropagation nets. Thus in our model we substitute dendrites with what we believe are their *functional approximations*. We have already used this modeling approach to demonstrate the SINBAD method of discovering influential environmental variables (Ryder and Favorov 2001).

Thus, representing each principal dendrite by an error backpropagation net, we model a pyramidal cell as a trio of backpropagating networks whose outputs are added together to produce the cell’s output (Fig. 1). The cell’s output is also used as the training signal for each dendrite. In this particular demonstration, each dendrite is modeled as a standard backpropagation network (Rumelhart et al. 1986) with one layer of 50 hidden units and one output unit. Two of the dendrites, representing *basal* dendrites of pyramidal cells, are given *afferent inputs* carrying information about the states of environmental variables. The third dendrite, representing the *apical* dendrite of pyramidal cells, is given both afferent inputs and *lateral inputs* from other SINBAD cells. This greater variety of inputs on the apical dendrite is intended to reflect the prominent convergence of afferent, lateral, and feedback connections on the apical dendrites of the neocortical pyramidal cells, especially on their terminal tufts (McGuire et al. 1991; Cauller 1995).

The activity of a hidden unit  $h$  in dendrite  $d$  is computed as a sigmoid function of the activities of its input sources:

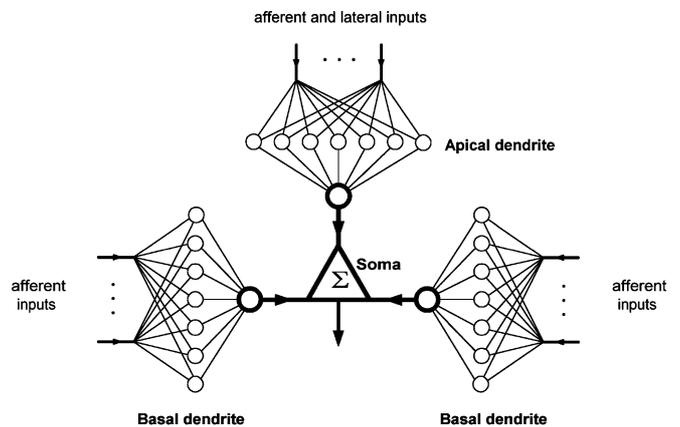
$$H_{d,h} = \tanh \left( \sum_i w_{d,i,h} \cdot A_{d,i} \right), \quad (4)$$

where  $A_{d,i}$  is the activity of input source  $d, i$  and  $w_{d,i,h}$  is the weight of its connection onto the hidden unit  $h$  of dendrite  $d$ . The activity of the output unit, i.e., the output of dendrite  $d$ , is

$$D_d = \sum_{h=1}^{50} w_{d,h} \cdot H_{d,h}, \quad (5)$$

where  $w_{d,h}$  is the weight of the connection from the hidden unit  $d, h$  to the output unit. The outputs of the three dendrites are summated to produce the cell’s output:

$$A = D_1 + D_2 + D_3. \quad (6)$$



**Fig. 1.** The SINBAD model of a neocortical pyramidal cell with three dendrites connected to the soma (shown as a *triangle*). Each dendrite is modeled as an error backpropagation network with one output unit and a single layer of hidden units

The cell's output  $A$  is the principal contributor to the training signal  $T$ , used to adjust the weights of connections on the three dendrites. Additional factors contributing to the training signal are: (i) the average output activity of the cell,  $\bar{A}$ , driving the cell to have  $\bar{A} = 0$ ; (ii) deviation of the current output activity from the average,  $A - \bar{A}$ , designed to expand the dynamic range of output values; and (iii) lateral inhibition from other SINBAD cells,  $I$ . Thus,

$$T = A - \alpha \cdot \bar{A} + \beta \cdot (A - \bar{A}) - I , \quad (7)$$

where  $\alpha$  and  $\beta$  are scaling coefficients. Coefficient  $\beta$  is determined by the variability of the output activity: the smaller the variability, the greater the value of  $\beta$ . It is computed as

$$\beta = \left[ \beta_{\max} - \gamma \cdot \overline{|A - \bar{A}|} \right]^+ , \quad (8)$$

where  $\beta_{\max}$  and  $\gamma$  are controlling parameters and  $[ ]^+$  indicates that if the quantity is negative, the value is to be taken as zero.

The somal inhibition  $I$  is computed as

$$I = \iota \cdot \sum_j w_j^- \cdot (A_j - \bar{A}_j) , \quad (9)$$

where  $\iota$  is a scaling constant and  $w_j^-$ ,  $A_j$ , and  $\bar{A}_j$  are the somal inhibitory connection weight, activity, and average activity, respectively, of SINBAD cell  $j$ . The task of somal inhibition is to drive SINBAD cells to tune to different features of the environment, thus maximizing the number of environmental variables expressed by the network as a whole. To accomplish this task, we follow Foldiak (1990) and make these connections anti-Hebbian (see below). For simplicity and to emphasize its dendrite-training function, somal inhibition in this exercise does not contribute to the cell's output but is used only to adjust the training signal  $T$ .

The connections of the hidden units were adjusted according to the error backpropagation algorithm of

Rumelhart et al. (1986). Specifically, the error signals  $\delta_d$  were first computed for the three dendrites as

$$\delta_d = T - 3 \cdot D_d . \quad (10)$$

For the hidden units,  $\delta$  was backpropagated as

$$\delta_{d,h} = \delta_d \cdot w_{d,h} \cdot (1 - H_{d,h}^2) . \quad (11)$$

Connection weights were adjusted by

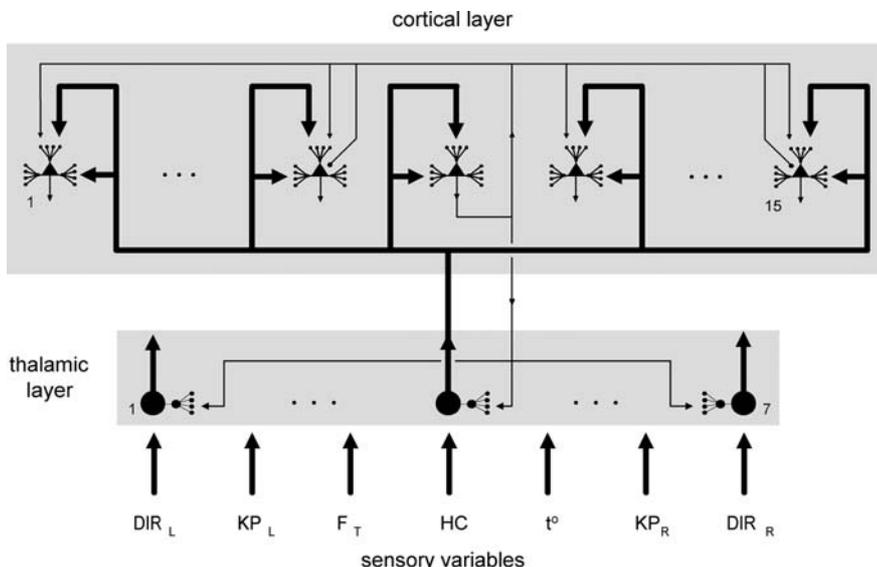
$$\begin{aligned} \Delta w_{d,i,h} &= \mu_i \cdot A_{d,i} \cdot \delta_{d,h} \quad \text{and} \\ \Delta w_{d,h} &= \mu_h \cdot H_{d,h} \cdot \delta_d , \end{aligned} \quad (12)$$

where  $\mu_i$  and  $\mu_h$  are learning rate constants for the input and hidden unit connections. Somal inhibitory connections were adjusted by

$$\Delta w_j^- = \mu_s \cdot \left[ -w_j^- + (A_j - \bar{A}_j) \cdot (A - \bar{A}) \right] , \quad (13)$$

where  $\mu_s$  is a learning rate constant and  $w_j^-$ ,  $A_j$  and  $\bar{A}_j$  are the somal inhibitory connection weight, activity, and average activity, respectively, of presynaptic SINBAD cell  $j$ . This synaptic learning rule is anti-Hebbian in its effect because it makes connection  $w_j^-$  track the covariance in activities of the two connected cells.

Turning to the design of the entire network, 15 such SINBAD cells were organized into a single "cortical" layer (Fig. 2). The network was designed to observe, sequentially, kitchen sinks of various possible configurations and states (as described in Sect. 2.1). Upon an exposure to a sink, the sensory information about that sink is transmitted to SINBAD cells via relay cells of the "thalamic" layer, with each thalamic cell reporting the state of one of the sink variables. The thalamic layer consists of seven cells, representing the seven observable sink variables:  $DIR_L$  and  $DIR_R$  (directions to open the left and right pipes),  $KP_L$  and  $KP_R$  (left and right knob positions),  $HC$  (which pipe is hot), and  $F_T$  and  $t^\circ$  (water outflow and temperature). The two other important sink variables –  $F_L$  and  $F_R$  (flow of water in the two pipes) –



**Fig. 2.** The connectational diagram of the SINBAD network. Thalamic and SINBAD cells are shown as *solid circles* and *triangles*, respectively, with their dendrites drawn as miniature hidden layer-to-output unit nets. For clarity, output connections are shown only for one thalamic and one SINBAD cell

are *invisible*, i.e., no thalamic cells represent them. The central goal of this exercise is to demonstrate that SINBAD cells will discover these two variables, as was explained in Sect. 2, from observations of patterns of activity in the thalamic cells.

For simplicity, in this demonstration exercise we do not take advantage of the cortical topographic mapping mechanisms, which were described in Sect. 2.3, to arrange the thalamic connections among the dendrites of SINBAD cells. Instead, we distribute thalamic connections among SINBAD dendrites randomly, and thus not as efficiently as might otherwise be possible. Specifically, each thalamic cell connects to every SINBAD cell on the apical dendrite and on one of that cell's two basal dendrites, chosen randomly for each thalamocortical pair. Thus, basal dendrites of different SINBAD cells will have afferent connections from different combinations of thalamic cells. Furthermore, the two basal dendrites on the same SINBAD cell will have connections from different thalamic cells and therefore will receive explicit afferent information about *different* sink variables. Consequently, in their search for correlated output functions (as explained in Sect. 2.3), the two basal dendrites will have to discover and learn to respond to whatever *implicit* information they have in common, thus making this information *explicit* in the cell's output. In effect, the basal dendrites will tune the cell to one of the important variables of kitchen sinks.

SINBAD cells are interconnected via lateral connections, with each cell connecting to the apical dendrite of every other cell. That means that each apical dendrite receives all the afferent information available from the thalamic layer about the "visible" sink variables and all the lateral information available from other SINBAD cells about sink features discovered by them. In each SINBAD cell, then, the apical dendrite will have the most diverse and comprehensive information about the sink, which it will use to compute the state of the sink variable chosen for that cell by its basal dendrites. Thus, in this demonstration the web of inferential relations among sink variables (as described in Sect. 3.1) is captured in the SINBAD layer specifically by the net of its apical connections.

SINBAD cells are also interconnected via inhibitory somal connections, with each cell connecting to five other, randomly chosen, SINBAD cells. The purpose of these connections is to make SINBAD cells tune to different environmental variables. Finally, every SINBAD cell has a connection to every thalamic cell. This feedback system of corticothalamic connections implements Mumford's (1991) idea of the thalamus being used by the cortex as a "blackboard," on which the cortex draws its interpretation of the attended subject. In our use of this idea, the web of inferential relations learned by the cortical network acts as an inferential model of the outside world (Sect. 3.1), and this internal model projects its picture of the outside world back onto the thalamus, so that it can be returned again to the cortex for another pass of inferential adjustment and elaboration, and so on. This will enable the cortical inferential model to fill in holes, when they happen, in

the raw picture of the world that the thalamus receives from its sensory channels. We will use this "filling in" function of corticothalamic feedback as a demonstration of the inferential capabilities of the SINBAD network (details in Sect. 3.3). (It should be remembered, however, that filling in occurs not only upon the thalamic blackboard but also within the cortical network itself.)

To implement Mumford's idea, each thalamic cell in our demonstration is given a dendrite in the form of a backpropagation network, identical to that used to represent dendrites of SINBAD cells (4, 5). Hidden units of this dendrite receive connections from all the SINBAD cells. Further, afferent inputs to thalamic cells block their cortical feedback inputs. That is, when a thalamic cell receives direct information from the outside world, it simply relays that information to the cortical layer without any contribution from the feedback dendrite (i.e., the cell's output is equal to the value of the environmental variable it represents, scaled to vary between  $-1$  and  $+1$ ). But when the cell does not receive information from the outside world, then it uses its cortical feedback inputs (i.e., the cell's output is equal to the output of its feedback dendrite). The weights of input connections to hidden units of the thalamic dendrites and of hidden unit connections to the output unit are adjusted by the error backpropagation algorithm (10–12), using as the training signal  $T$  the output of the thalamic cell.

### 3.3 Simulation results

After initially setting all the adjustable connections to randomly chosen strengths  $w$ 's (in  $\pm 1.25$  and  $\pm 0.05$  ranges for input and hidden connections, respectively), the network was exposed to sinks of varying configurations of pipes and valves, knob positions, and resulting water outputs. Each sink exposure lasted eight time steps, during which the states of the SINBAD cells and the thalamic dendrites were computed iteratively from their previous states, thus giving lateral interactions among SINBAD cells time to express themselves. Eight time steps were sufficient for the network to settle into a stable state. After the eighth time step, all the cortical and thalamic connections were adjusted according to the learning algorithms described in Sect. 3.2.

To let the network exercise its inferential skills, on 50% of the sink exposures one of the sink variables – chosen at random – was not shown to the network, i.e., one of the thalamic cells was not given the state of the sink variable it represents, instead being made to use cortical feedback to its dendrite to determine its output. Such partial sink exposures were interspersed randomly among full exposures. Many simulation runs of the network were performed using different randomly chosen distributions of thalamic connections on the basal dendrites of SINBAD cells, different initial connection weights, and different random series of sink exposures. The aim of the simulation runs was to answer the following central questions: (i) Will the basal dendrites of a SINBAD cell learn to correlate their outputs? (ii) Will

any of the SINBAD cells discover the hidden, but causally central, sink variables  $F_L$  and  $F_R$ ? (iii) Will the network develop an effective inferential model of kitchen sinks?

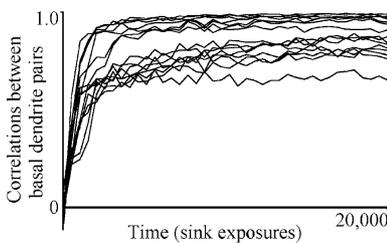
Figure 3 shows the progress of SINBAD cells' learning during a representative simulation run. It plots, for each of the 15 SINBAD cells, the correlation coefficient between outputs of the cell's two basal dendrites as a function of time from the start of the learning period. This plot shows that dendrites in each cell gradually discovered something in common, since their correlations grew with time. Eventually a majority of dendrites learned to match each other fully. The remaining dendrites learned to predict each other's activity only partially, due to limitations on the inputs available to them (for example, the least successful cell by chance was assigned  $F_T$ ,  $t^\circ$ ,  $KP_R$  inputs to go to one basal dendrite and  $HC$ ,  $DIR_L$ ,  $DIR_R$ ,  $KP_L$  to go to the other basal dendrite, thus not giving them enough mutual information to produce fully correlated outputs).

The next question is: What was it that the basal dendrites discovered they had in common? It turns out that nine SINBAD cells tuned to the original sink variables, already represented by thalamic cells, such as  $t^\circ$ ,  $DIR_L$ ,  $KP_L$ , etc. (see Fig. 4 for an example). Two other SINBAD cells discovered the two hidden variables of central significance to sink functioning. That is, as Fig. 4 shows, one cell tuned to the water flow in the left pipe ( $F_L$ ) and the other cell tuned to the water flow in the right pipe ( $F_R$ ). This discovery of hidden, but function-

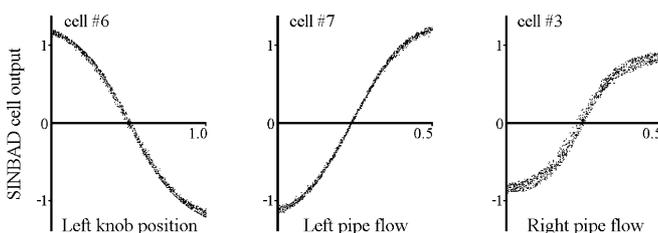
ally and therefore inferentially significant, environmental variables is the central accomplishment of this modeling demonstration. Finally, the remaining four cells failed to tune to a single sink variable and instead tuned to some not easily interpretable nonlinear functions of multiple sink variables. These are the cells whose basal dendrites could not learn to produce fully correlated outputs due to limitations of their inputs (Fig. 3). Apparently, tuning to nonlinear functions of multiple sink variables allowed their dendrites to get the best correlations out of their poorly matched inputs. Which of the 15 SINBAD cells tune to which of the sink variables and which of the cells tune to multivariable functions varies randomly from one simulation run to another as a direct consequence of the random assignment of the thalamic connections on the basal dendrites of the SINBAD cells.

The final question is: Did the SINBAD network develop an effective inferential model of kitchen sinks? Such an internal model will be difficult to demonstrate directly because it will be distributed in the pattern of lateral connections among the SINBAD cells. However, we should be able to observe it through its effects by providing the thalamic layer with incomplete sink information. In such a situation we should expect the internal sink model to fill in the withheld information by inferring it from the provided items. And since the SINBAD layer projects its representation of the outside world back onto the thalamic layer, a thalamic cell that was denied the external information about the state of the sink variable it represents should nevertheless represent that state more or less accurately, informed about it by the SINBAD layer.

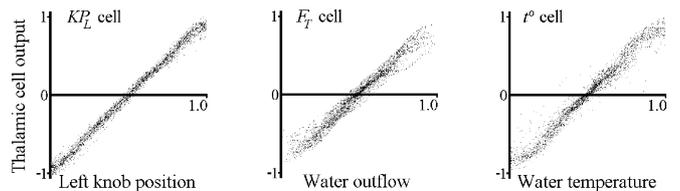
Representative examples of such inferential recovery of withheld information are shown in Fig. 5. Each plot there shows responses of one thalamic cell to sink exposures in which that cell was not given direct input from the outside world. In those exposures, then, the cell's output was determined by its cortical feedback. The cells shown in Fig. 5 are the thalamic cells that received from the outside and relayed to the cortical layer the states of sink variables  $KP_L$ ,  $F_T$ , and  $t^\circ$  (the left knob position, water outflow, and its temperature). Each cell's responses are plotted as a function of the actual state of the withheld sink variable during that sink exposure. The plots show that the cells' inferred representations of



**Fig. 3.** Learning performance of the basal dendrites of SINBAD cells, exposed to a sequence of 20,000 sinks. For each of the 15 SINBAD cells, the correlation coefficient between the outputs of the cell's two basal dendrites is plotted as a function of the number of exposures to sinks. Network parameters used in this simulation:  $\alpha = 0.02$ ,  $\beta_{\max} = 0.6$ ,  $\gamma = 0.5$ ,  $\mu_i = 4$ ,  $\mu_h = 0.002$ ,  $\mu_s = 0.01$ ,  $\iota = 0.5$



**Fig. 4.** Tuning of three sample SINBAD cells to sink variables, acquired in the course of sink exposures shown in Fig. 3. Cell #6 (left plot) tuned to one of the original variables,  $KP_L$ , the left knob position. Cells #7 (middle) and #3 (right) tuned to hidden variables  $F_L$  and  $F_R$ , the water flows in the left and the right pipes, respectively

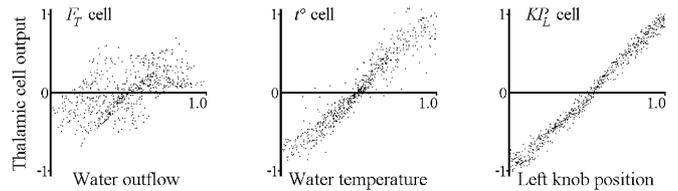


**Fig. 5.** Inferential performance of the SINBAD network, acquired in the course of sink exposures shown in Fig. 3. Each panel shows the responses of one of the thalamic cells during sink exposures in which that cell was not given external input, being driven instead exclusively by its feedback inputs from SINBAD cells. These responses are plotted against actual values of the withdrawn sink variable to show the accuracy of the inferred estimates

the withheld sink variables closely match their actual states (with a correlation coefficient above 0.96). The other four thalamic cells, not shown in Fig. 5, matched just as closely their inferred representations with the actual states of their variables. Thus, whenever the network was not informed about one of the sink variables on a particular exposure to a sink, the SINBAD cells were able to infer accurately the state of that variable from the known states of the other variables and made the affected thalamic cell represent that inferred estimate, thus completing the thalamic representation of the observed sink.

The fully successful network simulation run described in Figs. 3–5 is representative of half of our simulation runs. In the other half, the internal model was less accurate in its estimation of one, or sometimes several, of the sink variables. As an example, in one of the worst simulation runs, five SINBAD cells tuned to the same sink variable  $KP_L$ , one cell tuned to  $t^\circ$ , and one cell tuned to  $DIR_L$ . Three other cells discovered the same hidden variable  $F_L$ , and one cell discovered hidden variable  $F_R$ . The remaining four SINBAD cells tuned to not easily interpretable nonlinear functions of multiple sink variables. An occasional result like this is to be expected, given the random assignment of input connections. The choice of what sink variable a SINBAD cell can tune to is greatly constrained by how the thalamic inputs are split between its two basal dendrites. In this simulation run, it happened that by random coincidence many SINBAD cells were given similar distributions of thalamic inputs on their basal dendrites. This left them no other choice but to tune to the same sink variable, swamping the effect of anti-Hebbian inhibition among them.

Because in this simulation run the SINBAD layer came to excessively overrepresent variables  $KP_L$  and  $F_L$  while losing information about  $DIR_R$  and  $KP_R$ , the network could develop only a limited internal model of kitchen sinks. When this internal model was tested on its ability to infer withheld information, it had a mixed performance (Fig. 6). Specifically, the network was not able to estimate at all the withheld states of variables  $DIR_R$  and  $KP_R$  and gave a poor estimate of  $F_T$ . But, on the other hand, it had fair estimates of  $t^\circ$ ,  $HC$ , and  $DIR_L$  and an excellent estimate of  $KP_L$  (Fig. 6). Such a relatively poor performance was rare, however, among simulation runs; more typically among the less than perfect runs the network suffered degradation of its ability to estimate just one or two sink variables. This indicates that the ability of the SINBAD layer to develop an internal model of the environment is robust and exhibits graceful, rather than catastrophic, degradation under adverse conditions. The adverse conditions in our simulations were a consequence of our random assignment of thalamic connections to pairs of basal dendrites of SINBAD cells, the method that we chose in this study for its simplicity, despite its inefficiency. Whether there are more efficient alternatives to such random connection assignments is an important subject that remains to be explored. (However, for the cerebral cortex it is certainly possible that topographically constrained random



**Fig. 6.** Inferential performance of the SINBAD network in one of the worst simulation runs. Each panel shows the responses of one of the thalamic cells during sink exposures in which that cell was not given external input, being driven instead exclusively by its feedback inputs from SINBAD cells. These responses are plotted against actual values of the withdrawn sink variable to show the accuracy of the inferred estimates

assignment might be unproblematic in natural, informationally rich environments with large numbers of hidden variables to be discovered.)

To conclude, the results of these simulations demonstrate that a network of SINBAD neurons can indeed discover inferentially significant factors in the outside world. It can also learn the nonlinear orderly relations that exist there and can use knowledge of these relations to infer information not provided directly by sensory inputs.

#### 4 Conclusion

The SINBAD hypothesis offers a general solution – in the form of a computational approach (the “mining” of correlated functions) as well as its tentative cellular implementation – for one of the most profound and most challenging problems in cortical function: How does the cortex discover and make use of regularities in its sensory environment? The hypothesis ties together several major ideas in the recent literature: “trading representation against computation” of Clark and Thornton (1997), “suspicious coincidences” of Barlow (1992), mutual information maximization, IMAX, of Becker and Hinton (1992), and “contextual guidance” of learning in pyramidal cells of Phillips and Singer (1997). It addresses the most fundamental impediment to discovering regularities in spatiotemporal patterns of sensory receptor activities, namely, that behaviorally significant regularities are, typically, relations that are orders of complexity removed from raw sensory inputs. The hypothesis identifies the basic computational approach to dealing with this difficulty through the incremental learning of regularity-simplifying environmental variables that are deeply hidden in sensory patterns. At the implementation level, the hypothesis offers a cellular mechanism for implementing this approach in the dendrites of pyramidal cells and suggests how a network of SINBAD neurons can learn to function as a model of the outside world.

The SINBAD view of neocortical pyramidal cells is a hypothesis that is primarily motivated by (i) the fundamental need for the brain to discover and represent explicitly influential environmental variables, (ii) the availability of a computational strategy to discover such variables (the method of mining correlated functions),

and (iii) the suitable candidacy of single pyramidal cells for implementing this mechanism. The core of the SINBAD idea is the proposed presence in pyramidal cells of two or more dendritic sectors that (i) receive different inputs and (ii) teach each other to produce correlated outputs by performing nonlinear integration of their inputs. This last requirement – that the dendritic sectors be capable of learning nonlinear input-to-output transfer functions – is imperative if pyramidal cells are to discover novel orderly features of the environment. Unfortunately, while dendrites are known to be nonlinear integrators (Mel 1994; Häusser et al. 2000; Koch and Segev 2000; Segev and London 2000), and that plasticity affects their nonlinear integrative properties (Wang et al. 2003), the question of whether they can be taught nonlinear functions by spikes backpropagating into them from the soma has not yet been addressed experimentally. For the SINBAD hypothesis this is, at present, the most important experimental question.

The other biological details of our SINBAD model are more suggestive than definitive. They should be viewed as our first attempt at interpreting cortical organization in the light of the computational mechanism, mining correlated functions, that forms the core of the SINBAD proposal. These details depend on the identity of the postulated mutually teaching dendritic sectors of pyramidal cells. In this paper we identify these sectors with the principal dendrites (i.e., each of the dendritic trees that grow from the soma). However, while principal dendrites are our preferred candidates for this role, other possibilities exist as well, among them: (i) proximal vs. distal dendrites, (ii) basal dendrites together as a single sector vs. the apical dendrite, or (iii) the terminal tuft of the apical dendrite vs. the rest of the dendrites. The final resolution of this question will have to be determined experimentally.

Overall, we hope that the questions we raise in this paper concerning the distribution of synaptic inputs on the dendritic trees of neocortical pyramidal cells, dendritic integration, the function of synaptic learning, the aim of cell tuning, and the cortical learning of environmental regularities will lead to a new, regularities-centered conceptual framework for the analysis of cortical structure and function. We also hope this will lead to new ideas, understanding, and experimental exploration of dendritic physiology and dendritic information processing as well as cortical operation in general.

*Acknowledgements.* This work was supported, in part, by US Army Research Office grant P43077-LS.

## References

- Barlow HB (1986) Why have multiple cortical areas? *Vision Res* 26: 81–90
- Barlow HB (1992) The biological role of neocortex. In: Aertsen A, Braitenberg V (eds) *Information processing in the cortex*. Springer, Berlin Heidelberg New York, pp 53–80
- Barlow HB (1995) The neuron doctrine in perception. In: Gazzaniga MS et al (eds) *The cognitive neurosciences*. MIT Press, Cambridge, MA
- Becker S (1995) JPMAX: learning to recognize moving objects as a model-fitting problem. In: *Advances in neural information processing systems 7*. Morgan Kaufmann, San Mateo, CA, pp 933–940
- Becker S (1996) Mutual information maximization: models of cortical self-organization. *Netw Comput Neural Sys* 7: 7–31
- Becker S (1999) Implicit learning in 3D object recognition: the importance of temporal context. *Neural Comput* 11: 347–374
- Becker S, Hinton GE (1992) A self-organizing neural network that discovers surfaces in random-dot stereograms. *Nature* 355: 161–163
- Cauller L (1995) Layer I of primary sensory neocortex: where top-down converges upon bottom-up. *Behav Brain Res* 71: 163–170
- Clark A, Thornton C (1997) Trading places: computation, representation, and the limits of uninformed learning. *Behav Brain Sci* 20: 57–90
- DeFelipe J, Farinas I (1992) The pyramidal neuron of the cerebral cortex: morphological and chemical characteristics of the synaptic inputs. *Prog Neurobiol* 39: 563–607
- Favorov OV, Kelly DG (1994a) Minicolumnar organization within somatosensory cortical segregates: I. Development of afferent connections. *Cereb Cortex* 4: 408–427
- Favorov OV, Kelly DG (1994b) Minicolumnar organization within somatosensory cortical segregates: II. Emergent functional properties. *Cereb Cortex* 4: 428–442
- Favorov OV, Kelly DG (1996) Local receptive field diversity within cortical neuronal populations. In: Franzen O, Johansson R, Terenius L (eds) *Somesthesia and the neurobiology of the somatosensory cortex*. Birkhauser, Basel, pp 395–408
- Feldman ML (1984) Morphology of the neocortical pyramidal neuron. In: Peters A, Jones EG (eds) *Cerebral cortex*, vol 1. Plenum, New York, pp 123–200
- Foldiak P (1990) Forming sparse representations by local anti-Hebbian learning. *Biol Cybern* 64: 165–170
- Gawne TJ, Kjaer TW, Hertz JA, Richmond BJ (1996) Adjacent visual cortical complex cells share about 20% of their stimulus-related information. *Cereb Cortex* 6: 482–489
- Häusser M, Spruston N, Stuart GJ (2000) Diversity and dynamics of dendritic signaling. *Science* 290: 739–744
- Johnston D, Hoffman DA, Colbert CM, Magee JC (1999) Regulation of back-propagating action potentials in hippocampal neurons. *Curr Opin Neurobiol* 9: 288–292
- Koch C, Segev I (2000) The role of single neurons in information processing. *Nat Neurosci (suppl)* 3: 1171–1177
- Malach R (1994) Cortical columns as devices for maximizing neuronal diversity. *TINS* 17: 101–104
- Markram H, Lubke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275: 213–215
- Markram H, Pikus D, Gupta A, Tsodyks M (1998) Potential for multiple mechanisms, phenomena and algorithms for synaptic plasticity at single synapses. *Neuropharmacology* 37: 489–500
- Matus A (2000) Actin-based plasticity in dendritic spines. *Science* 290: 754–758
- McAllister AK (2000) Cellular and molecular mechanisms of dendrite growth. *Cereb Cortex* 10: 963–973
- McGuire B, Gilbert CD, Wiesel TN, Rivlin PK (1991) Targets of horizontal connections in macaque primary visual cortex. *J Comput Neurol* 305: 370–392
- Mel BW (1994) Information processing in dendritic trees. *Neural Comput* 6: 1031–1085
- Mumford D (1991) On the computational architecture of the neocortex: I. The role of the thalamo-cortical loop. *Biol Cybern* 65: 135–145
- Paulsen O, Sejnowski TJ (2000) Natural patterns of activity and long-term synaptic plasticity. *Curr Opin Neurobiol* 10: 172–179
- Phillips WA, Singer W (1997) In search of common foundations for cortical computation. *Behav Brain Sci* 20: 657–722
- Poirazi P, Mel BW (2001) Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* 29: 779–796

- Quartz SR, Sejnowski TJ (1997) The neural basis of cognitive development: a constructivist manifesto. *Behav Brain Sci* 20: 537–596
- Reich DS, Mechler F, Victor JD (2001) Independent and redundant information in nearby cortical neurons. *Science* 294: 2566–2568
- Richmond BJ, Gawne TJ, Jin GX (1997) Neuronal codes: reading them and learning how their structure influences network organization. *BioSystems* 40: 149–157
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL, PDP Research Group (eds) *Parallel distributed processing: explorations in the microstructure of cognition*, vol 1. MIT Press, Cambridge, MA, pp 318–362
- Ryder D (2004) SINBAD neurosemantics: a theory of mental representation. *Mind Lang* 19: 211–240
- Ryder D, Favorov OV (2001) The new associationism: a neural explanation for the predictive powers of cerebral cortex. *Brain Mind* 2: 161–194
- Segev I, London M (2000) Untangling dendrites with quantitative models. *Science* 290: 744–750
- Segev I, Fleshman JW, Burke RE (1989) Compartmental models of complex neurons. In: Koch C, Segev I (eds) *Methods in neuronal modeling*. MIT Press, Cambridge, MA, pp 63–96
- Singer W (1995) Development and plasticity of cortical processing architectures. *Science* 270: 758–764
- Stone J (1996) Learning perceptually salient visual parameters using spatiotemporal smoothness constraints. *Neural Comput* 8: 1463–1492
- Stuart GJ, Häusser M (2001) Dendritic coincidence detection of EPSPs and action potentials. *Nat Neurosci* 4: 63–71
- Stuart G, Spruston N, Sakmann B, Häusser M (1997) Action potential initiation and backpropagation in neurons of the mammalian CNS. *Trends Neurosci* 20: 125–131
- Vetter P, Roth A, Häusser M (2001) Propagation of action potentials in dendrites depends on dendritic morphology. *J Neurophysiol* 85: 926–937
- Vinje WE, Gallant JL (2000) Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287: 1273–1276
- Von der Malsburg C, Singer W (1988) Principles of cortical network organization. In: Rakic P, Singer W (eds) *Neurobiology of neocortex*, Wiley, New York, pp 69–99
- Wang Z, Xu NL, Wu CP, Duan S, Poo MM (2003) Bidirectional changes in spatial dendritic integration accompanying long-term synaptic modifications. *Neuron* 37: 463–472
- Williams SR, Stuart GJ (2000) Backpropagation of physiological spike trains in neocortical pyramidal neurons: implications for temporal coding in dendrites. *J Neurosci* 20: 8238–8246
- Woolley CS (1999) Structural plasticity of dendrites. In: Stuart G, Spruston N, Häusser M (eds) *Dendrites*. Oxford University Press, Oxford, UK, pp 339–364