Mixing Memory and Desire: Want and Will in Neural Modeling*

Bruce J. MacLennan
Computer Science Department
University of Tennessee
Knoxville, TN 37996-1301
Telephone: 423-974-5067 / Fax: 423-974-4404
E-Mail: maclennan@cs.utk.edu
Homepage: http://www.cs.utk.edu/~mclennan

Abstract

Values are critical for intelligent behavior, since values determine interests, and interests determine relevance. Therefore we address relevance and its role in intelligent behavior in animals and machines. Animals avoid exhaustive enumeration of possibilities by focusing on relevant aspects of the environment, which emerge into the (cognitive) foreground, while suppressing irrelevant aspects, which submerge into the background. Nevertheless, the background is not invisible, and aspects of it can pop into the foreground if background processing deems them potentially relevant.

This illuminates the differences between representation in natural intelligence and (traditional) artificial intelligence. Traditionally artificial intelligence has started with simple, primitive features, and attempted to construct from them a representation of the environment. If too few features are used, then the processing is imprecise and crude. However, if sufficient features are used to permit precise processing in all contexts, then the system is defeated by the combinatorial explosion of features. In natural intelligence, in contrast, we begin with a nervous system that can process in real-time the "concrete space" represented by the interface between the animal's nervous system and its environment. The separation of foreground from background then serves to increase the efficiency of this process. Instead of trying to construct the concrete world from abstract predicates, the brain projects the very high-dimensional concrete world into lower dimensional subspaces; this projection is context-sensitive and rapidly adaptable. Therefore it is not vulnerable to the combinatorial explosion.

We consider the connection between these ideas and the concepts of intentionality, as discussed by Brentano and Husserl, and information, as quantified by Shannon and Weaver. In particular, the Shannon-Weaver measure ignores relevance, which is essential to biological

* To appear in the proceedings of the Fifth Appalachian Conference on Behavioral Neurodynamics, October 18-20, 1996.

information. Further, Brentano and Husserl characterized intentionality in terms of the "directedness of consciousness," which can be explained as a decrease in the entropy (disorder) in the probability of processing, which is produced by the separation of foreground from background.

Essential to these ideas are questions of how contexts are switched, which defines cognitive/behavioral episodes, and how new contexts are created, which allows the efficiency of foreground/background processing to be extended to new behaviors and cognitive domains.

Next we consider mathematical characterizations of the foreground/background distinction, which we treat as a dynamic separation of the concrete space into (approximately) orthogonal subspaces, which are processed differently. Background processing is characterized by large receptive fields which project into a space of relatively low dimension to accomplish rough categorization of a novel stimulus and its approximate location. Such background processing is partly innate and partly learned, and we discuss possible correlational (Hebbian) learning mechanisms.

Foreground processing is characterized by small receptive fields which project into a space of comparatively high dimension to accomplish precise categorization and localization of the stimuli relevant to the context. We also consider mathematical models of valences and affordances, which are an aspect of the foreground. Cells processing foreground information have no fixed meaning (i.e., their meaning is contextual), so it is necessary to explain how the processing accomplished by foreground neurons can be made relative to the context. Thus we consider the properties of several simple mathematical models of how the contextual representation controls foreground processing.

We show how simple correlational processes accomplish the contextual separation of foreground from background on the basis of differential reinforcement. That is, these processes account for the contextual separation of the concrete space into disjoint subspaces corresponding to the foreground and background.

Since an episode may comprise the activation of several contexts (at varying levels of activity) we consider models, suggested by quantum mechanics, of foreground processing in superposition. That is, the contextual state may be a weighted superposition of several pure contexts, with a corresponding superposition of the foreground representations and the processes operating on them. This leads us to a consideration of the nature and origin of contexts. Although some contexts are innate, many are learned. We discuss a mathematical model of contexts which allows a context to split into several contexts, agglutinate from several contexts, or to constellate out of relatively acontextual processing. Finally, we consider the acontextual processing which occurs when the current context is no longer relevant, and may trigger the switch to another context or the formation of a new context. We relate this to the situation known as "breakdown" in phenomenology.

1. Relevance and Intelligence

Intelligence presupposes values. This is because a sense of *relevance* is one of the central components of intelligence. As Dreyfus (1979; Dreyfus & Dreyfus 1986) has shown in his critiques of traditional approaches to artificial intelligence, experts have a well-tuned sense of relevance that saves them from having to consider irrelevant aspects of the situation at hand. For example, unlike the chess program, which must consider every possible move, if only to reject it on the basis of heuristics, the chess master sees only the relevant possibilities. Instead of wasting time on irrelevant possibilities, as the program does, the chess master can devote greater cognitive resources to possibilities that are genuinely important.

However, a well-developed sense of relevance is not peculiar to specialized activities, such as playing chess; rather it is characteristic of expert (i.e., skillful) behavior throughout the animal kingdom. The cat stalking a bird is guided by relevance no less than the chess expert stalking an opponent's king. Relevance then is a common denominator of intelligence, since it guides effective allocation of cognitive resources.

Nevertheless, until recently (e.g. Sperber & Wilson 1986), relevance has been the abandoned child of informatics. Indeed the measure of information enshrined in Shannon's information theory (Shannon 1948; Shannon & Weaver 1949) ignores relevance. As Hamming (1980, 103) has remarked, according to Shannon's measure, the book with the most information is a book of random numbers (since it is completely unpredictable and therefore has no redundancy). This paradox arises from the divergence between our ordinary understanding of information and Shannon's formalization of it: ordinarily, for something to be informative, it must inform us of something relevant (actually or potentially).

One cannot understand relevance without understanding *context*, for aspects of the environment that are relevant in one behavioral context are not in another. Therefore, a significant component of intelligence is being able to determine the appropriate context and to switch between contexts quickly. When an aggressive dog appears nearby, the intelligent cat will switch contexts, abandon its pursuit of the bird, and flee the potential predator. When a fire alarm sounds, the intelligent chess player abandons the game and evacuates with everyone else. Different aspects of the environment, such as "Exit" signs, will become salient, while previously relevant chess configurations are ignored.

The phenomenologist philosophers, such as Husserl and Heidegger, have done much to improve our understanding of relevance and context (see, e.g., Dreyfus 1982, 17-21; Dreyfus 1991, 118-20). For example, by stressing that we are "always already in a situation" (Dreyfus 1979, 53) they remind us that perception and cognition mediated by a context-dependent sense of relevance is the norm rather than an exception. Indeed, comparatively context-free cognition occurs only in "breakdown" situations (Dreyfus 1991, 70ff), when we cannot readily determine a context appropriate to the situation.

The role of relevance and context can be understood in terms of the information-theoretic concept of entropy. If we think of a probability distribution p_x over the myriad concrete components x of the situation, or equivalently, if p_x represents the relative allocation of cognitive resources to x, then the *entropy* of this distribution, $H\{p\}$, defined,

$$H\{p\} = -\sum_x p_x \log p_x,$$

measures the "structure of relevance." The entropy is maximized (representing maximum disorder and minimum structure) when the distribution is uniform, that is, when all x get equal attention (as in context-free "breakdown"). The entropy is lower (representing greater structure) to the extent that attention is focused on some x to the exclusion of others. A lower entropy distribution gives a more definite content to consciousness, by directing it toward some aspects of the world at the expense of others (MacLennan 1988, 172-3). For the phenomenologists, such "directedness" is an essential characteristic of consciousness and intentionality (Brentano 1925; Husserl 1931, §84).

Based on the foregoing discussion we can understand the critical role played by values in intelligent behavior. The term "value" has many senses, even in philosophy (Frankena 1967). I will use *value* to refer to anything that an organism *values* (positively or negatively), that is, anything that it tends to seek or avoid. Thus, common positive values include food and sex, negative ones include pain and injury. Values define *potential interests*, which become *actual interests* in an appropriate context. Thus food is normally a potential interest until the animal becomes hungry, whereupon it becomes an actual interest. Context-sensitive processes in the animal's nervous system then direct its attention to aspects of the environment relevant to its current interests. Jung (1978, 32-3) has stressed the central role of value:

> The function of value — feeling — is an integral part of our conscious orientation and ought not to be missing in a psychological judgement of any scope, otherwise the model we are trying to build of the real process will be incomplete. Every psychic process has a value quality attached to it, namely its feeling-tone. This indicates the degree to which the subject is *affected* by the process or how much it means to him (in so far as the process reaches consciousness at all). It is through the "affect" that the whole subject becomes involved and so comes to feel the whole weight of reality.

In the following sections I will consider processes that might subserve contextual processing by identifying the appropriate context and using it to control the extraction and processing of relevant information.

2. Function of Contextual Processing

2.1. Function

In this paper I will often take a robotics perspective on the problems of relevance, context and values. That is, I will be asking the question: How *ought* contextual processing to operate in an intelligent autonomous robot? In this way, we will come to the biological functions of values and context from first principles, which will illuminate their role in natural animal intelligence. (See Kilmer, McCulloch & Blum 1969, especially §9, for an earlier parallel treatment of natural and artificial contextual processing systems.)

It appears to me that the principal role of contextual processing is the efficient allocation of neural resources (cognitive processing power). The nervous system is metabolically expensive, and so each neuron must earn its keep. Contextual processing is more efficient by devoting greater resources to relevant aspects of the world than to irrelevant aspects; allocation is *biased* toward relevant aspects. In this way, for fixed total resources, relevant aspects may be processed more extensively than would be possible with an unbiased allocation.

The disadvantage of contextual processing is a consequence of this same bias, for the implicit judgment of relevance is historical, deriving from the history of the individual and its species. Thus there is a danger that the organism or robot will be ill-equipped to handle the future, that it will be insensitive to previously irrelevant aspects that are now relevant, that it will be confined within the contextual boundaries of the past.

Therefore, if a robot or animal is to be adaptive in the face of a future that differs from the past, it cannot ignore irrelevant aspects of its environment; they must be processed, albeit with a lesser allocation of resources. Further, it must be capable of detecting situations that do not fit into known contexts, and be able to construct new contexts capturing the "structure of relevance" in such situations.

Since contextual processing focuses cognition on restricted aspects of the situation, to the partial neglect of other aspects, I will refer to the disadvantages of contextual processing as the problem of becoming "single-minded."

2.2. Means

We can identify a number of means by which contextual processing accomplishes its function. First, depending on the context, certain aspects of the situation are projected into the foreground, where more resources are devoted to them, while the remainder (or the whole situation) is projected into the background, where it is processed with a lower resource investment and in a relatively unbiased (context-free) way. That is, currently irrelevant aspects are in the background, but not "invisible." Further, the variable content of the foreground implies that the fixed neural resources devoted to foreground processing must be capable of performing different functions depending on the context.

Contextual processing, by directing attention toward some aspects, directs it away from others; therefore, management of the context becomes a critical task. First, it is important to recognize when a new context should be activated, either on the basis of foreground processing or, even more importantly, on the basis of background processing. This is, in effect, a categorization problem: classifying the current situation as a cognitive context. (This view is especially consistent with Ruth Garrett Millikan's characterization, in her Presidential Address to the American Philosophical Association, of categorization as a process of *reindentification*; see also Millikan 1984, ch. 16.)

Next, since several contexts may be demanding attention, it is important to have a representation of the urgency or importance of these competing demands, so that resources may be allocated most effectively. For example, hunger (for an animal) and low battery power (for a robot) are gradually changing conditions that we expect to make progressively more insistent demands; they can be set aside for a time, but eventually must be handled. (In psychological terms, this is the *dynamical* aspect of contextual processing.) Finally, for flexibility of behavior, it is important that both the "context triggers" and their dynamical strength be adaptable through learning.

Indeed, adaptiveness requires some means of assessing the success of contextual processing, which is provided by various kinds of reinforcement. Often this is the elimination of the conditions (e.g. hunger) that activated the context. However, just as the triggering conditions can be learned, so can the reinforcing conditions, which is important to the successful pursuit of long-term activities, for which direct (somatic) rewards may be delayed. Further, reinforcement must be contextual, otherwise it is difficult to relate it to its preconditions. This is easiest when the reinforcement occurs in the same *episode* (interval of constant context) that led to it, but often the reinforcement is delayed. In these cases the reinforcement must include information sufficient to cause reactivation of the context to which it applies. (Of course, I am not claiming these processes should be expected to work perfectly in either animal or machine; I am trying to lay down conditions that should be satisfied to have a higher probability of success.) One of primary functions of reinforcement is in determining relevance; aspects that have appeared in an episode leading to reinforcement will be *relevated* (raised in relevance, from Latin *relevare*, to lift up; Pribram 1991, 230). Beyond this, reinforcement is used to tune foreground processing to increase the probability of achievement of positive reinforcement, or avoidance of negative reinforcement.

Finally, we must "expect the unexpected"; that is, there must be some means of handling situations that do not fit into a recognized context. This could be known context arising in a new way (i.e., without its known trigger conditions), or it could be an entirely new context. Without knowing the context, which tells us how to bias processing, the situation must be processed in a relatively unbiased way, which is less efficient, since it doesn't allow resources to be diverted from irrelevant aspects. Since this *acontextual* *processing mode* is relatively inefficient, it is important to activate an appropriate context as quickly as possible. Therefore, most of the cognitive effort will be devoted to either categorizing the novel situation as a known context, or in constructing a new context (a new contextual category) to accommodate it.

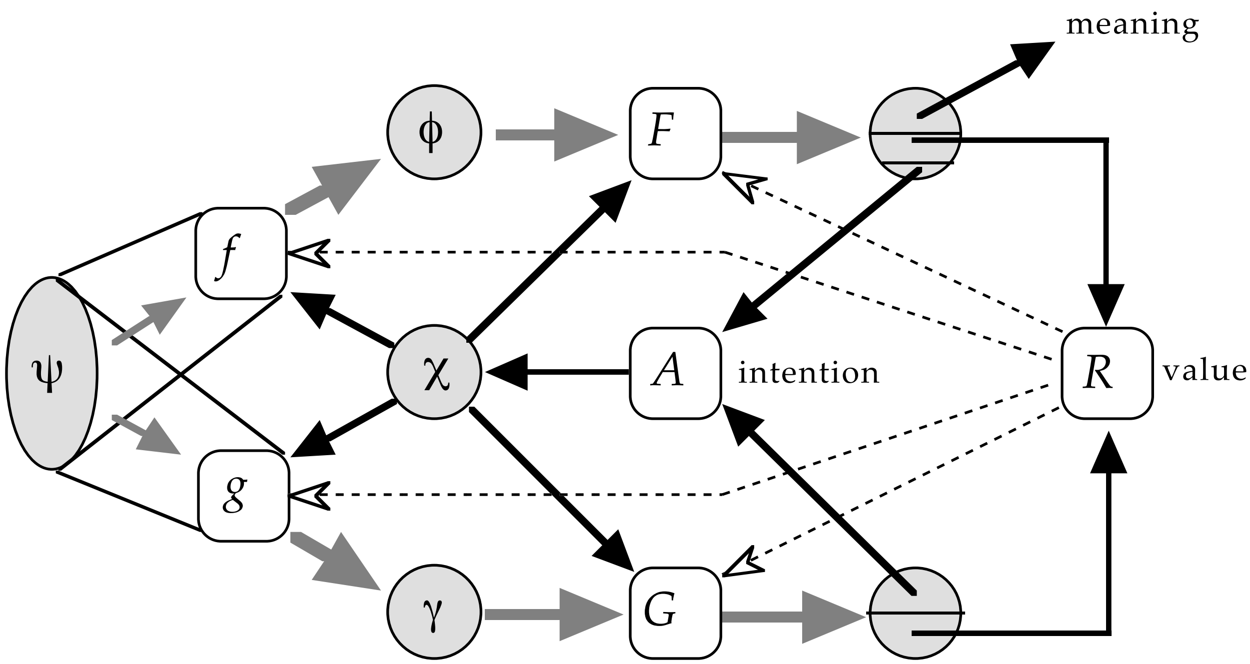

Figure 1

3. General Framework

Having outlined the general functions of contextual processing and the means to accomplish them, we can turn to a more specific model (Fig, 1).

3.1. Concretum

We begin with an abstraction, the *concretum*, which is defined as the space of all inputs to contextual processing, including all sensory and somatic inputs as well as feedback from other brain areas. More precisely, if we define contextual processing to include all the processes that are not context-free, then the concretum comprises all the neural processes on which contextual processing depends. In particular, it includes *protocritical signals* such as comfort/discomfort (Pribram 1991, 200-14). However, for the present discussion, it is not important to decide exactly what the concretum includes; it suffices to observe that it is of very high dimension.

I will use W to stand for the concretum, since it is, so far as the organism is concerned, the concrete *world*. Mathematically W is a set comprising all the possible states ψ of the concretum. The state $\psi(t)$ at a given time represents the instantaneous activity of the neural tissue corresponding to the concretum; we will generally think of this "instant" as having some finite duration, in accord with the Gabor Uncertainty Principle. That is, $\psi(t)$ is extended in time as well as space.

3.2. Primacy of the Concrete

It is necessary to say a few words about the "primacy of the concrete," a lesson we learn from phenomenology. Traditionally AI has tried to construct a representation of the world by a combination of context-free features. If too few features are used, the representation is crude and incomplete, but the use of an increasing number of features eventually leads to a combinatorial explosion of possibilities. The traditional AI program is ultimately defeated by the Scylla and Charybdis of either processing all the possibilities (context-free processing) or risking missing the unexpected (single-mindedness).

Biological intelligence does not have this shortcoming, and it is important to consider why. Biological intelligence begins with the concrete world with whatever resolution and dimensions it is represented by its nervous system. For example, the human brain has millions of sensory inputs, which suggests the dimension of our concrete world. For biological intelligence, we begin with the assumption that the organism can process its concretum in real time (for otherwise the species would not survive). Contextual processing improves the efficiency of this process by a context-dependent projection of some aspects of the concretum into a subspace where they can be processed with greater precision. Thus, instead of the *context-free features* of traditional AI, the nervous system improves its performance by means of *context-sensitive aspects*.

At root, we are really dealing with the gap between the continuous and the discrete. We can divide a continuum into smaller continua, some of which may be magnified; that is what biological intelligence accomplishes through context-sensitive projection of aspects of the concretum. On the other hand, a continuum cannot be constructed from any finite number of discrete elements; that is the impossibility that traditional AI tried to accomplish. As Zeno observed 2500 years ago in his *Dichotomy*, "There is no motion, because what moves must arrive at the middle of its course before it reaches the end" (Aris. *Phys.* $239^{b}11$). We apparently cannot get from A to B in any finite number of discrete steps, each of which goes half the remaining way. So also, no increasing number of more finely divided abstract context-free features will capture the concreteness of the world. Yet in everyday life we go from A to B all the time; continuous motion gets us from A to B and through all the points in between. A posteriori we can divide the motion into parts, but the continuous motion is primary, and its division into discrete parts secondary and approximate.

3.3. Context and the Foreground & Background Projections

First we postulate a space X of possible context representations χ; I will consider later (section 5) what they may "look like." We postulate two additional spaces Φ and Γ representing possible states of the foreground and background, and corresponding context-dependent projections $f: X \times W \rightarrow \Phi$, $g: X \times W \rightarrow \Gamma$ (i.e., *figure* and *ground*). Thus, $\phi = f(\chi, \psi)$ is the foreground representation in context χ of the concrete state ψ, and $\gamma = g(\chi, \psi)$ is the corresponding background.[1] Note that f and g are "projections" in a general sense: so g represents the state of the concretum in a manner that is unbiased with respect to context, but optimized for the detection of potential context changes. Conversely, f projects the concrete state in a manner that is specialized to the context at that time; that is, it is tuned to information relevant to that context. This implies that neurons in the foreground projection may be used differently in different contexts, which is supported by evidence of place cells in rat hippocampus (O'Keefe 1986, 82-4; Pribram 1991, 233-4).

3.4. Control of Foreground Processing

Since the primary goal of contextual processing is to allocate neural resources to different processing tasks in different contexts, we define the foreground process to be a function of context as well as of the foreground representation, $F: X \times \Phi \rightarrow \Omega_F$. The output space of foreground processing Ω_F will be left unspecified for the time being, but it must include all the outputs of attentive cognition, including those controlling motor behavior. Similarly, $G: X \times \Gamma \rightarrow \Omega_G$ is the background process with its output in Ω_G. Notice that the foreground and background processes $F[\chi, f(\chi, \psi)]$ and $G[\chi, g(\chi, \psi)]$ might be computed directly without explicit representation in the brain of the projections $\phi = f(\chi, \psi)$ and $\gamma = g(\chi, \psi)$, although it is conceptually simpler to consider them as two-step processes.

3.5. Context Activation

As noted previously, the primary function of background processing is the switching of contexts, thus avoiding "single mindedness." However, attentive foreground processing can also detect conditions for the activation of contexts. Therefore we define an "arbitration function" $A: \Omega_F \times \Omega_G \rightarrow X$ which takes the results of foreground and background processing and determines the current context. This arbitration is based on the strength with which contexts are activated, which is one of the affective components of the outputs of both foreground and background contextual categorization processes. In our model, the context representations are normalized $\|\chi\| = 1$, so that the magnitude of the

---

[1] I think it is likely that the background projection is independent of the context, but the current formulation is more general and permits, for example, subtraction of the foreground information from the background.

output $a\chi$ of the contextual categorization process reflects the strength of its activation, $\|a\chi\| = |a|$.

3.6. Reinforcement & Adaptation

Finally, there are adaptive processes that, on the basis of reinforcement, adjust the foreground and (perhaps) background projections and foreground and background processing. This is because (1) the foreground aspects are relevated on the basis of their role in achieving positive reinforcement or in avoiding negative reinforcement, (2) foreground processing must be adapted to accomplish its purpose more effectively, and (3) background processing needs to adapt to identify context switches more efficiently and reliably. This requires determining both the context to be activated and the strength of its activation.

Since they are adaptive, the foreground and background projections and processes are all repositories of "memory." This memory mediates the activation of contexts, in particular, the strength of their activation and so their potential for commanding the animal's behavior; it also mediates reinforcement, much of which is learned, though it be based on a protocritical core. For example, increasing hunger corresponds to increasing activation of a food-seeking context which causes restaurants, refrigerators, etc. to become more salient. Nevertheless, less tangible learned reinforcements, both current and future expected, may raise the activation of the doing-mathematics context, sufficient to keep us working on the derivation rather than heading for the refrigerator.

In effect, the input to the reinforcement process corresponds to the evaluative dimension (pleasant/unpleasant, etc., thus what it tends to seek or avoid), whereas the input to arbitrator corresponds to the intentional dimension (important/unimportant, thus what grabs the animal's attention). In terms adapted from Miller, Galanter & Pribram (1960, 62), we can say that memory — in the broad sense, both individual and species (cf. Jung 1978, 27-9) — is used to evaluate the current image, which in turn activates a context by which current value may be improved. *Motive* results from a combination of *value* (implicit in the evaluation) and *intention* (implicit in the context).

4. Contextual Processing

There are several ways that context can control foreground processing, thereby allowing neural resources to be used differently in different contexts. I will describe two, both based on bilinear (product forming) synapses.

4.1. Inner-product Model

Consider a single neuron in foreground processing and the inner product its dendritic net forms with the foreground representation ϕ (including here the results computed by other foreground processing neurons) before it is passed through any nonlinear activation function. For contextual processing, we would like the possibility of a different inner product $\langle \phi, \mathbf{p}_k \rangle = \phi^{\mathrm{T}} \mathbf{p}_k$ in each context χ_k. (Recall that fields such as ϕ and χ_k. are taken to be temporally as well as spatially extended, therefore the *characteristic pattern* $\mathbf{p}_k$ represents the spatiotemporal pattern to which the neuron is tuned by the context. Further, apparent matrix expressions such as $\phi^{\mathrm{T}} \mathbf{p}_k$ involve spatiotemporal inner products.) This context-dependent process is accomplished by the bilinear foreground computation $F(\chi, \phi) = \phi^{\mathrm{T}} \mathrm{M} \chi$ where $\mathrm{M} = \sum_j \mathbf{p}_j \chi_j^{\mathrm{T}}$ (a linear superposition of spatiotemporal outer products). The inner product will be exact if contextual representations are orthonormal; it will be inexact to the extent they are not. Reinforcement can lead to this kind of contextual processing through simple correlational learning, for example,

$$\Delta \mathrm{M} \propto \pm \phi \chi^{\mathrm{T}}.$$

(The sign of the change is + or – depending on whether the reinforcement is positive or negative.)

4.2. Correlation Model

Pribram (1991, 233-8) reviews O'Keefe's (1986, 82-4) observation that hippocampal function resembles holographic reconstruction with the theta-rhythm serving as a reference beam. Therefore I will briefly consider how the context can control foreground processing through a holographic (i.e. convolutional or correlational) mechanism.

To begin, we define the correlation $\chi * \psi$ of two spatiotemporally extended fields, χ and ψ. The j-th spatial component of the correlation at time t is given by a superposition of temporal correlations:

$$(\chi * \psi)_j(t) = \sum_k (\chi_{k-j} * \psi_k)(t).$$

These correlations can be computed by linear dendritic nets that have impulse responses $h_k(t) = \psi_k(-t)$. Expanding the temporal correlations by integration over the temporal extension T, we get:

$$(\chi * \psi)_j(t) = \sum_k \int_T \chi_{k-j}(s-t) \; \psi_k(s) \, \mathrm{d}s.$$

Again, we would like, in the ideal case, the linear input to a foreground neuron to be $\langle \phi, \mathbf{p}_j \rangle$ in context χ_j. If the "hologram" $\eta = \sum_k \chi_k * \mathbf{p}_k$ is a superposition of correlations between contexts and patterns, and if $\rho = \chi_j * \phi$ is the correlation of context χ_j with input ϕ, then we can show that

$$\langle \eta, \rho \rangle = \langle \alpha_j \otimes \mathbf{p}_j, \phi \rangle + \textstyle\sum_{j \neq k} \langle \kappa_{jk} \otimes \mathbf{p}_k, \phi \rangle,$$

where $\otimes$ stands for convolution, $\alpha_j = \chi_j * \chi_j$ is the autocorrelation of context χ_j, and $\kappa_{jk} = \chi_j * \chi_k$ is the cross-correlation of contexts χ_j and χ_k. The effect of the first convolution is to "blur" the characteristic pattern $\mathbf{p}_j$ by the autocorrelation field α_j. In the ideal case, $\alpha_j = \delta$ (a Dirac delta function), $\kappa_{jk} = 0$, and we have exact selection of the foreground computation, $\langle \eta, \rho \rangle = \langle \phi, \mathbf{p}_j \rangle$. The inner product can be computed in two different ways:

$$\langle \eta, \rho \rangle = \langle \eta, \chi * \phi \rangle = \langle \chi * \eta, \phi \rangle.$$

In the first case, $\langle \eta, \chi * \phi \rangle$, the context and input are allowed to interfere spatiotemporally, and an inner product is formed with the "hologram" η. In the second case, $\langle \chi * \eta, \phi \rangle$, the context is allowed to interfere spatiotemporally with the "hologram," and an inner product is formed between the result and the input. The "hologram" results from positive or negative reinforcement of the interference pattern (correlation) between the current context and the reinforced input,

$$\Delta\eta \propto \pm \chi * \phi.$$

4.3. The Foreground Projection

Pribram (1991, Apps. C – F) defines a context as a complete orthonormal system (CONS) defined by the eigenvectors of the dendritic microprocesses defined by the neural wave equation; they are the stable states that constitute an episode. In effect, the context, as CONS, defines the projection of the state onto a set of axes represented by the eigenvectors. Our model includes Pribram's as a special case by abstracting away from the details of the neural wave equation (and by taking what amounts to time-windowed or wavelet view).

In our model the foreground projection, $\phi = f(\chi, \psi)$, is controlled in the same ways as foreground processing. If we take projection to be linear process, then $\phi_i = \psi^{\mathrm{T}} \mathrm{M}_i \chi$ for the inner-product encoding, and $\phi_i = \langle \eta_i, \chi * \psi \rangle = \langle \chi * \eta_i, \psi \rangle$ for the correlation encoding. Based on these equations we can define the impulse response or characteristic pattern of the dendritic net of a foreground neuron, $h_i = \mathrm{M}_i \chi$ or $h_i = \chi * \eta_i$, respectively. Thus, keyed by the context, each foreground neuron ϕ_i receives some linear combination of spatiotemporal "microfeatures" (wavelets) from the concretum. If the neuron is unused in that context (i.e. is not a part of a foreground aspect in that context) then its characteristic pattern is the zero field; that is, $h_i = \mathbf{0}$.

Adaptation also takes place in the same ways as in foreground processing, $\Delta \mathrm{M}_i \propto \pm \psi \chi^{\mathrm{T}}$ or $\Delta \eta_i \propto \pm \chi * \psi$. In this case positive or negative reinforcement is strengthening a particular projection into foreground dimension i, up to some maximum. This strength, which corresponds to the relevance of a contextual aspect, can be measured by the norm of the characteristic pattern, $\|h_i\|$. The entropy of a context can be defined in terms of the norm of the total of the characteristic patterns of all foreground neurons, $h^* = \|\Sigma_i h_i\|$. Thus, $h^* = \|(\Sigma_i \mathrm{M}_i)\chi\|$ or $h^* = \|\chi * (\Sigma_i \eta_i)\|$, respectively. The entropy of context χ is

$$H_\chi = -(\Sigma_k h^*_k \log h^*_k) \,/\, (\Sigma_k h^*_k)$$

It seems likely to me that relevation is primarily a one-way process; that is, once reinforcement has made an aspect relevant in some context, there is little if anything that can make it irrelevant. That is, *delevation* (from Latin *delevare*, to smooth out) does not take place (or takes place extremely slowly). (From a robotics standpoint, this is what we would want, because relevant aspects need not occur in a every instance of a context.) Delevation is different from an animal learning to ignore an unreinforced foreground aspect (which is still relevant, in spite of being unreinforced). Relevation is adaptation of the foreground projection f; learning to ignore an unreinforced (yet relevant) stimulus occurs through adaptation of the foreground process F.[2]

5. Representation of Contexts

The preceding discussion shows that contextual representations work best if (in the inner-product case) they are orthonormal, or if (in the correlational case) they are strongly autocorrelated (i.e. they are aperiodic) and weakly cross-correlated. In this section I will consider in more detail the implications of orthonormality and its lack.

5.1. Approximate Discreteness

To a first approximation, contexts form a discrete set. For example, it has been estimated that a vertebrate animal has a couple dozen "behavior modes" (contexts), which are discrete and mutually incompatible (Kilmer & al. 1969). Typical behavior modes include eating, drinking, hunting, exploring, fighting, fleeing, mating, nest building, grooming, sleeping, and so forth. There are several reasons, however, for considering this discreteness to be only approximate.

First, we must say what it means, in mathematical terms, for a set to be discrete. A set is discrete (or has the discrete topology) if there are only two possible distances between the elements of the set: they have a distance of 0 if they are the same element, and they have a distance of 1 if they are not the same. That is,

---

[2]Delevation probably takes place during the creation of a context from the unbiased "breakdown" context, since unreinforced aspects must be suppressed and left to background processing.

$$d(x, x) = 0, \text{ and } d(x, y) = 1 \text{ if } x \neq y.$$

It is easy to show that a set of orthonormal vectors forms a discrete set under the inner-product metric (the metric defined by the L_2 norm):

$$\begin{aligned} d(u, v) &= \| u - v \| / \sqrt{2} \\ &= \sqrt{[(u - v)^2 / 2]} \\ &= \sqrt{[\, \|u\|^2 + \|v\|^2 - \langle u, v \rangle \,]} \\ &= \sqrt{[\, 1 - \langle u, v \rangle \,]} \\ &= 0 \text{ if } u = v, \text{ or } = 1 \text{ if } u \neq v. \end{aligned}$$

Therefore, the set of contexts will be discrete to the extent that the context representations χ are orthonormal. This is biologically unrealistic, so in practice we expect the contexts to form "clusters" or "constellations" with some overlap between them.

5.2. Innate and Learned Contexts

Although some vertebrates may have a fixed set of 25 or so behavioral modes, this is certainly not the case for human beings; we acquire many new contexts, such as playing chess, doing mathematics, and delivering papers at conferences. Like innate contexts, learned contexts define a space of relevant aspects of the world. We have already considered how, given an orthonormal representation for a context, the corresponding foreground projection and process can be learned, but we have not addressed the creation of a contextual representation.

5.3. Creation of Contexts

Representations for (approximately discrete) contexts can be created by a variety of orthogonalization processes. The existing context representations χ_1, ..., χ_n span a space, and a representation for a new context χ_{n+1} can be chosen from the orthogonal complement of that space. Recalling that contextual categorization is a problem of reidentifying a relevance structure, we can see that the creation of a contextual representation requires correlating the "trigger conditions" that will activate the context (as determined by background processing or prior foreground processing) with elements of the concretum that have emerged as potentially relevant. Since the trigger conditions are a priori relevant, they form a kernel around which a relevance structure can develop. Furthermore, new contexts often develop out of "breakdown" situations in which processing is acontextual, that is, attentive but unbiased. Reinforcement in this mode allows a tentative separation of the relevant from the irrelevant.

Suppose therefore that in acontextual processing, the activity of certain foreground neurons ϕ_{r_1}, ..., ϕ_{r_m} has been reinforced, while others ϕ_{q_1}, ..., ϕ_{q_n}

have not. We want to create a contextual representation χ that relevates $r_1, \ldots, r_n$ but delevates $q_1, \ldots, q_n$. That is, we want $\psi^T M_{r_i} \chi = 1$ and $\psi^T M_{q_j} \chi = 0$. This can be accomplished by picking a (normalized) χ that is in the subspace spanned by $M_{r_i}{}^T\psi$ but orthogonal to the subspace spanned by the $M_{q_j}{}^T\psi$. Further relevation can then take place by adjusting the M_i on the basis of reinforced foreground processing in context χ.

5.4. Superposition of Contexts

Introspection suggests that at any given time we can be operating in a limited number of compatible contexts. For example, we can do a mathematical derivation or plan a lecture while we are driving; we can look for a restaurant or a mailbox at the same time. The model of contextual processing that we have presented is easily extended to accommodate "parallel processing" in multiple contexts. Much as a quantum-mechanical system can be in a linear superposition of pure states, so an organism can be in a superposition of contexts. If we let a_k represent the strength of activation of context k, then the effective context is a linear superposition $\chi = \Sigma_k a_k \chi_k$.

Since the foreground projection, as we have defined it, is linear, the foregrounds will also exist in superposition. Of course, they could interfere with each other, but the lack of such interference is part of what makes contexts compatible and therefore amenable to parallel processing in superposition. For example, much of the visual processing involved in the looking-for-restaurant context is compatible with that in the looking-for-mailbox context.

Similarly, since the dendritic processes are linear, a superposition of contexts will create a corresponding superposition of dendritic states in foreground processing. Again, to the extent that the processes do not interfere, they can proceed in parallel. This will depend on the extent to which they have compatible, or at least relatively orthogonal, foreground projections.

It seems likely that the interference must limit the extent of parallel processing in superposition. Given a certain level of primary activation by "trigger conditions," competitive or other nonlinear processes may ensure that one context is strongly activated, and at most a couple others are at most weakly activated. In that way, the "noise" contributed by less important contexts would be kept at a manageable level. Furthermore, overt behavior would, for the most part, be under the control of a single context, thus giving the illusion of "mutually incompatible modes of vertebrate behavior" (Kilmer & al. 1969, 279).

References

Brentano, F. (1925). *Psychologie von empirischen Standpunkt.* Meiner.

Dreyfus, H. L. (1979). *What computers can't do*. New York: Harper & Row.

Dreyfus, H. L. (Ed.) (1982). *Husserl, intentionality and cognitive science*. Cambridge: MIT Press.

Dreyfus, H. L. (1991). *Being-in-the-world: A commentary on Heidegger's* Being and Time, *division I*. Cambridge: MIT Press.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: The Free Press.

Frankena, W. K. (1967). Value and valuation. In P. Edwards (Ed.), *Encyclopedia of Philosophy* (Vol. 8, pp. 229-32). New York: Macmillan Publ. Co.

Hamming, R. W. (1980). *Coding and information theory.* Englewood-Cliffs: Prentice-Hall.

Husserl, E. (1931). *Ideas* (transl. by W. R. B. Gibson of I*deen zu einer reinen Phänomenologie und phänomenologischen Philosophie*, 1913). Allen & Unwin, Macmillan.

Jung, C. G. (1978). *Aion: Researches into the phenomenology of the self*; Collected Works, Vol. 9, Pt. II, 2nd ed. (transl. R. F. C. Hull). Princeton: Princeton Univ. Press.

Kilmer, W. L., McCulloch, W. S., & Blum, J. (1969). A model of the vertebrate central command system. *International Journ. Man-Machine Studies* **1**: 279-309.

MacLennan, B. J. (1988). Logic for the new AI. In James H. Fetzer (Ed.), *Aspects of artificial intelligence* (pp. 163-92). Kluwer Academic Publishers.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.

Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism.* Cambridge: MIT Press.

O'Keefe, J. (1986). Is consciousness the gateway to the hippocampal cognitive map? A speculative essay on the neural basis of mind. In David A. Oakley (Ed.), *Brain and mind* (pp. 59-98). London & New York: Methuen.

Pribram, K. H. (1991). *Brain and perception: Holonomy and structure in figural processing.* Hillsdale: Lawrence Erlbaum.

Shannon, C. E. (1948). The mathematical theory of communication. *The Bell System Technical Journal* **27**: 379-423.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: Univ. of Illinois Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge: Harvard Univ. Press.