Finding Order in our World:

The Primacy of the Concrete in Neural Representations and the Role of Invariance in Substance Reidentification (Extended Version)

Commentary on Ruth Garrett Millikan's "A Common Structure for Concepts of Individuals, Stuffs, and Real Kinds: More Mama, More Milk and More Mouse"*

Bruce J. MacLennan
Computer Science Department,
University of Tennessee, Knoxville TN 37996-1301.

MacLennan@cs.utk.edu; http:www.cs.utk.edu/~mclennan

Abstract: We discuss neuroscientific and phenomenological arguments in support of Millikan's thesis. Then we consider invariance as a unifying theme in perceptual and conceptual tracking, and how invariants may be extracted from the environment. Finally we consider some wider implications of Millikan's nondescriptionist approach to language, with especial application to color terms.

Since I am in substantial agreement with Millikan's thesis, my commentary will explore connections between it and neural network theories of knowledge representation. First I will discuss neuroscientific and phenomenological arguments in support of the thesis. Then I will consider *invariance* as a unifying theme in perceptual and conceptual tracking, and how invariants may be extracted from the environment. Finally I will consider some wider implications of Millikan's nondescriptionist approach to language, with especial application to color terms.

The primacy of the concrete. In his Heideggerian critique of traditional ("symbolic") cognitive science and artificial intelligence, Dreyfus (1979; 1982; 1991, pp. 115-121) pointed out the futility of trying to represent our skillful coping in the world in terms of atomic, abstract, context-free predicates. Even Husserl acknowledged the "huge concreteness" of this hypothetical abstract structure and called its phenomenological reduction an "infinite task" (see citations in Dreyfus 1982, p. 20). We may refer to this observation as the *primacy of the concrete*; that is, the world as ordinarily experienced

---

* To appear in *Behavioral and Brain Sciences*.

is primarily concrete, historical, contextual, meaningful and, in mathematical terms, effectively infinite-dimensional. Conversely, analysis into abstract, context-free, objective, low-dimensional predicates is a comparatively rare activity that we undertake in extraordinary circumstances (e.g., during "breakdowns" in skillful coping, or during scientific analysis); such analyses are always and necessarily incomplete. In Millikan's terms (section 4), our understanding of the world is primarily nondescriptionist, proceeding mostly by reidentification of relevantly similar substances; abstract descriptions are ancillary.

The primacy of the concrete is also supported by developments in neuroscience. Although sensory systems are often explained in terms of abstract "feature detectors," this terminology is inaccurate in a number of respects. Certainly, to a first approximation, neurons in early sensory areas appear to be tuned to simple abstract properties: small segments of edges and lines, patches of color, tones, etc. (Although even here we must recognize that some of the simplicity is an artifact of the simple stimuli used in such studies.) However, more detailed investigation reveals that most sensory neurons respond to complex combinations of stimulus features. For example, visual cells that respond to oriented edges may also respond to color, motion and stereo disparity (Pribram 1991, pp. 79-81). Further, it is not uncommon to find neurons in *visual* cortex that are tuned to *acoustic* frequencies (Pribram 1991 p. 81, citing Bridgeman 1982; Pribram, Spinelli & Kamback 1967). Conversely, it has been recently reported (Calvert et al. 1997) that our understanding of face-to-face communication is aided by the response of *auditory* neurons to *visual* stimuli; specifically, cells in auditory cortex are strongly activated by watching speech-like facial movements. Finally, it is worth noting that top-down signals in sensory systems can alter the receptive fields of sensory neurons, that is, their response is context-sensitive (Pribram 1991, pp. 257-258). Thus, instead of considering a sensory neuron to be a context-free feature detector, it is more accurate to view its response as an interaction between a complex combination of activities in the sensory receptors, and activity in nonsensory areas (representing context, expectations etc.).

Much of the persistence of talk about feature detectors in neuroscience can be attributed to the same descriptionist assumptions that pervade philosophy and cognitive science. If we believe that "the only game in town" is the assembly of atomic, context-free features into abstract descriptions, then that is what we will look for in the brain, and to a large extent that is what we will find.

One unfortunate consequence of this discriptivist bias is the "binding problem," which afflicts theories of neural-net knowledge representation: How are context-free features bound together to represent objects (so that, for example, perception of a red square and a green circle is different from perception of a red circle and a green square)? But the brain does not have to solve a binding problem because neurons respond to complex combinations of features, that is, to features that are already bound. (For example, there are neurons that respond to the co-occurrence of redness and aspects of circularity but not to the co-occurrence of greenness and circularity, to which other neurons respond.) Therefore the joint activity of a population of neurons can represent a unique complex macroscopic constellation of microproperties. In effect, the activity of each neuron represents a small bundle of conjoined microproperties, and the joint activity of a group of neurons represents a co-occurrence of a large number of overlapping bundles.

Substance concepts as invariants. Millikan's analogy between perceptual tracking and conceptual tracking (sections 3, 4) reveals an important idea underlying both: invariance under transformation. Invariance is well known from the psychology of perception (e.g., size and color constancy, invariance of melody under change of pitch). Invariance is of course a central concept in mathematics (especially in abstract algebra), but we must be careful applying mathematical concepts to psychology since, in particular, psychological invariants are always approximate and limited in range (MacLennan 1994).

Invariants typically arise because various aspects of a stimulus vary coherently; think of how the spatial location of an object's parts vary when the object moves or rotates. Because of this coherent variation we can have knowledge of the variation of aspects that are not being perceived. For example, when we view a rotating die, we know what the back side is doing and can predict its reappearance.

In the case of the conceptual tracking of substances we are interested in aspects that are approximately invariant over successive encounters with the substance. These are the aspects that cohere in the concept and about which it provides information. Indeed, Millikan's Aristotelian treatment of substances (section 1) views them much like objects: bundles of generally cohering properties through which they have their identity. (A view, incidentally, which supports Aristotle's similar treatment of individuals and classes — i.e. primary and secondary substances — as subjects of predication, as opposed to set-theoretic treatments, which make "Aristotle is mortal" and "man is mortal" different kinds of propositions.)

Although some invariants are "wired" into the nervous system, others — including many involved in conceptual tracking — are learned. Invariants can be detected in the coherent variation (i.e. covariation or contravariation) of multiple aspects of the stimulus. Synapses extract this information by responding to correlated activity between neurons in such a way as to strengthen strong correlations (positive or negative) and to "damp out" weak correlations (Singer 1995). Therefore, after learning, variations in certain aspects of a stimulus will lead to neural activity that mimics or primes the response to variation in other aspects that have been correlated with them in the past. Invariants become a means for generating expectations and filling in missing information. (In this way also we may simultaneously track Fido, dog, fur and bone; cf. section 5.)

The "damping out" of weak correlations causes uncorrelated aspects to be eliminated from the representation, in effect projecting the concrete stimulus from the high-dimensional space of sensory-receptor activity, in which it is given, into a lower dimensional subspace. The extreme cases, in which a stimulus is projected into a very low-dimensional subspace, produce something approximating a context-free feature detector, but such abstract features are comparatively rare and secondary to the processing of concrete microcorrelations, upon which reidentification depends. Descriptionist theories make context-free features the elementary constituents of substance concepts, but Millikan's thesis and neural network theory together show how approximately context-free features are secondary derivatives of concrete substances (that is, they show how invariants are *abstracted* — drawn out — from concrete experiences). Thus, Millikan's two continua along which the richness of real kinds can differ (section 2) can be understood as follows: the multiplicity of supported inferences results from the number of synaptic connections, and the "reliability" of the inferences from the connections' strength (synaptic efficacy), that is, the number and strength of the correlations.

The primacy of the concrete is also apparent in the context sensitivity of features. That is, the projection into lower-dimensional subspaces is dependent on some behavioral context; different features are salient depending on whether the animal is hunting prey, seeking a mate, avoiding a predator, etc. Meaning and relevance are primary; abstractions and features may follow as a consequence. Context-sensitive projections of this kind can be produced by using the neural representations of behavioral contexts to selectively activate or deactivate various sorts of microcorrelations; that is, from the complex combinations of properties to which a neuron responds, we select those relevant to the problem at hand. (MacLennan, in press, presents possible mechanisms for abstracting context-sensitive invariants and using them to control the salience of relevant aspects of the stimulus.)

The primacy of metaphor. Millikan's nondescriptionist theory of knowledge suggests a parallel development in the history of language: we should expect words to begin as context-dependent condensations out of complex clouds of pragmatic intentions. The reduction of their meaning to simple, low-dimensional predicates is secondary, and in part a consequence of descriptionist presuppositions and values. In conventional terminology, the metaphorical, concrete and context-sensitive connotation is prior to the abstract, context-free denotation (see also Lakoff & Johnson 1980). This observation even applies to such apparently abstract predicates as color terms, and part of our difficulty in understanding the use of such terms is a consequence of descriptionist assumptions. For example, ancient Greek *chlôros*, nominally translated "green," is applied to many things that are not green in color, such as dew, tears and blood (Gage 1993, p. 272n7; Zajonc 1993, p. 15). This usage is explicable when we realize that *chlôros*, like the English word "green," may refer to things that are fresh, living or moist (e.g. green wood, green riders). Also, many color terms began as univalent material-substance concepts (e.g., names for minerals or dye stuffs), but appear to be polyvalent when supposed to refer to optical color (Gage 193, pp. 34-35). So, some Medieval scarlets are black, blue, green or white in color (since scarlet was primarily to a kind of fabric); purple (originally a kind of silk) may be white, yellow, blue, black or green; *sinople* can be red or green (perhaps because these colors both derive from copper oxide coloring of glass); *glaucus* and *ceruleus* can be blue or yellow, both colors of woad leaves (Gage 1993, pp. 80, 90). The historical reduction of color to a one-dimensional predicate — wavelength — is partly a consequence of the scientific understanding of light, which began with Newton (and so offended Goethe), but we must not let this blind us to the fact that colors are primarily substances emergent from their complex meaning in our lives.

References

Bridgeman, B. (1982) Multiplexing in single cells of the alert monkey's visual cortex during brightness discrimination. *Neuropsychologia* **20**:33-42.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, P. W., Iversen, S. D. & David, A. S. (1997) Activation of auditory cortex during silent lipreading. *Science* **276**:593-596.

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

Dreyfus, H. L. (1982) Introduction, In: *Husserl, intentionality and cognitive science*, ed. H. L. Dreyfus with H. Hall. MIT Press.

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

Gage, J. (1993) *Color and culture: Practice and meaning from antiquity to abstraction.* Little, Brown & Company.

Lakoff, G., & Johnson, M. (1980) *Metaphors we live by.* University of Chicago Press.

MacLennan, B. J. (1994) Continuous computation and the emergence of the discrete. In: *Origins: Brain & self-organization*, ed. K. H. Pribram. Lawrence Erlbaum Associates.

MacLennan, B. J. (in press) Mixing memory and desire: Want and will in neural modeling. In: *The brain and values: Proceedings of the Fifth Appalachian Conference on Behavioral Neurodynamics* (tentative title), ed. J. King & K. H. Pribram. Lawrence Erlbaum Associates.

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

Pribram, K. H., Spinelli, D. N., & Kamback, M. C. (1967) Electrocortical correlates of stimulus response and reinforcement. *Science* **157**:94-96.

Singer, W. (1995) Development and plasticity of cortical processing architectures. *Science* **270**:758-764.

Zajonc, A. (1993) *Catching the light: The entwined history of light and mind.* Bantam Books.