

To appear in Learning and memory: A comprehensive reference. R. Menzel (Vol Editor)
& J. Byrne (Editor). New York: Elsevier

Learning and Representation

C. R. Gallistel

Rutgers University

Address correspondence to:

C.R. Gallistel
Rutgers Center for Cognitive Science
Piscataway, NJ 08854-8020
email: galliste@ruccs.rutgers.edu
Phone: 1(732)445-2973

Synopsis

Behavioral evidence implies that even insect brains compute representations of behaviorally relevant aspects of the world. For example, they compute the animal's position in the world by integrating its velocity with respect to time. Other examples are the learning of the solar ephemeris, the construction of a cognitive map, and episodic memory in food caching. Representations require a symbolic read-write memory that carries information extracted from experience forward in time in a computationally accessible form. The close analogy between the architecture required to access memory in a computer and the architecture of the molecular machinery that accesses genetic information suggests the sort of memory structure to be looked for in the nervous system.

Keywords

Symbols Signals Information Read-write memory
Cognitive map Path integration Episodic memory
Sun-compass

In a representational theory of learning, the brain computes a representation of the experienced world and behavior is informed by that representation. By contrast, in associative theories of learning, which dominate neurobiological thinking, experience causes a plastic brain to rewire itself so as to make behavior better adapted to the experienced world, without the brain's computing a representation of that world (Hawkins & Kandel, 1984; Hull, 1952; Pavlov, 1928; Rumelhart & McClelland, 1986; Smolensky, 1986). The computation of a representation seems to require a functional architecture that is not transparently consistent with our current understanding of neurobiological mechanisms (Gallistel, 2006), which is why representational theories of learning have not found favor among neurobiologists. Associative theories, by contrast, have been strongly influenced by neurobiological considerations for more than a century. For them, consistency with the current understanding of neurobiology is a major consideration (e.g., Gluck *et al.*, 2005), which is why they predominate in neurobiological thinking about learning.

The results of behavioral experiments on non-human animals have increasingly implied that much learned behavior is informed by enduring temporal and spatial representations (e.g., Menzel *et al.*, 2005), as even some prominent advocates of associative theories have recently acknowledged (Clayton *et al.*, 2006a). Moreover, direct electrophysiological observation of neural activity has shown that the nervous system represents where the animal is and has been within the environment (Foster & Wilson, 2006; Hafting *et al.*, 2005; Lee *et al.*, 2004; Muller, 1996; O'Keefe & Nadel, 1978), how it is oriented (Golob *et al.*, 2001; Ranck, 1984; Sargolini *et al.*, 2006; Yoganarasimha *et al.*, 2006), where objects are in relation to it (Campos *et al.*, 2006; Goossens & Van Opstal, 2006; Muller *et al.*, 2005), and the timing of repetitive or predictable events (Ivry & Spencer, 2004; Leon & Shadlen, 2003; Meck, 2003; Penney, 2004; Schultz & Dickinson, 2000). Thus, there is a conceptual tension between the behavioral and electro-physiological findings that seem to imply a computational-representational architecture, on the one hand, and, on the other hand, current conceptions of neural

structure and mechanism, in which seemingly essential elements of the requisite functional architecture appear to be absent.

The extent to which one believes that consistency with *currently* understood neurobiological mechanisms should be a constraint on models of learning depends on whether one believes that those mechanisms provide or could provide satisfactory accounts of the behaviorally well documented phenomena that are the focus of this review: dead reckoning in insect navigation, the capacity of insects to record landmark "snapshots" and the capacity of food-caching jays to remember and make versatile use of a large number of episode-specific facts about each cache. They also depend on whether one believes that computation in the brain must be consistent with the principles that computer scientists believe constrain physically realizable computations. Perhaps, as many neuroscientists believe, the brain escapes the limitations and requirements that computer scientists believe are imposed by mathematics, logic and physics.

There have been instructive instances in the history of science where findings and analyses at a higher level of inquiry seemed to require mechanisms at a more basic level for which there was then no explanation. Throughout the latter part of the 19th century, the eminent physicists, William Thompson (Lord Kelvin) and P.G. Tait, argued that Darwin and the geologists must be grossly in error in their estimates of the age of the earth, because no heat-generation process known to physics was consistent with a solar age of more than 100 million years (Tait thought the upper limit imposed by physical considerations was 10 million years—see Lindley (2004) for an account.) Thompson (1862) wrote, for example, "It is impossible that hypotheses assuming an equibilty of sun and storms for 1,000,000 years can be wholly true." Importantly, Kelvin did not have a satisfactory theory of where the sun's heat came from—he worked on the problem off and on throughout his career—but he was confident that a satisfactory explanation could be based on physical principles and phenomena that were then understood. Radioactivity was not discovered until 1896, and it was only in 1903 that the Curies showed that it was accompanied by the liberation of heat. The following contemporary quote gives an idea of just how revolutionary this discovery was: "[this phenomenon] can barely be

distinguished from the discovery of perpetual motion, which it is an axiom of science to call impossible, [and] has left every chemist and physicist in a state of bewilderment." (Lindley, 2004, p. 302) This discovery of something that Kelvin and Tait literally could not imagine bore tellingly on their argument with Darwin and the geologists.

Whether we are in such a situation now remains, of course, to be seen. However, I argue in the following review that the behavioral evidence for representation and computation, together with basic insights in computer science about the centrality of a read-write memory mechanism in physically realized computation, implies the existence of a neurobiological read-write memory mechanism. Given the centrality of such a mechanism to computation, as computer scientists understand it, the discovery of such a mechanism may someday have an impact on neuroscience comparable to the impact of the discovery of radioactivity on physics.

Representations: Definition and Explication

From the perspective of cognitive science, the brain is an organ of computation (Marr, 1982; Newell & Simon, 1975). What it computes are representations of selected aspects of the world and the animal's relations to it (the distal stimuli). It computes these representations from the signals engendered in sensory organs by the stimuli that impinge on them (the proximal stimuli) and from signals generated by the motor system, which carry information about how the animal is moving (efference copy signals Holst & Mittelstaedt, 1950).

A representation consists of signals, symbols and the operations on them (see Figure 1). The signals and symbols carry information about properties of the experienced world. The operations on those signals and symbols enable the brain to compute explicit representations from implicit ones and to anticipate behaviorally relevant states of the world.

Information is an abstract quantity carried by any signal or symbol that can reduce the brain's uncertainty about the present state of some aspect of the world (Rieke *et al.*, 1997; Shannon, 1948). Information cannot be communicated to a device that has no representation of the world, because the measure of the amount of

information communicated is the reduction in the receiver's uncertainty about the state of the world. A receiver that has no representation of states of the world cannot have uncertainty about those states; more technically, it cannot have a probability distribution defined on those states. Thus, a receiver incapable of representing at least some states of the world is not something to which information can be communicated; it cannot extract from information-bearing signals the information that they carry.

Signals carry information from one place to another (from one spatial location to another). In a computer, the signals are current pulses. In the nervous system, they are action potentials, synaptic transmitters, and hormones. Symbols carry information forward in time (from one temporal location to a later one). In a computer, the symbols are bit patterns in a memory register. We do not know the physical realization of symbols in the nervous system. One school of thought doubts that they exist (Hay & Baayen, 2005; Rumelhart & McClelland, 1986; Wallace & Fountain, 2003)

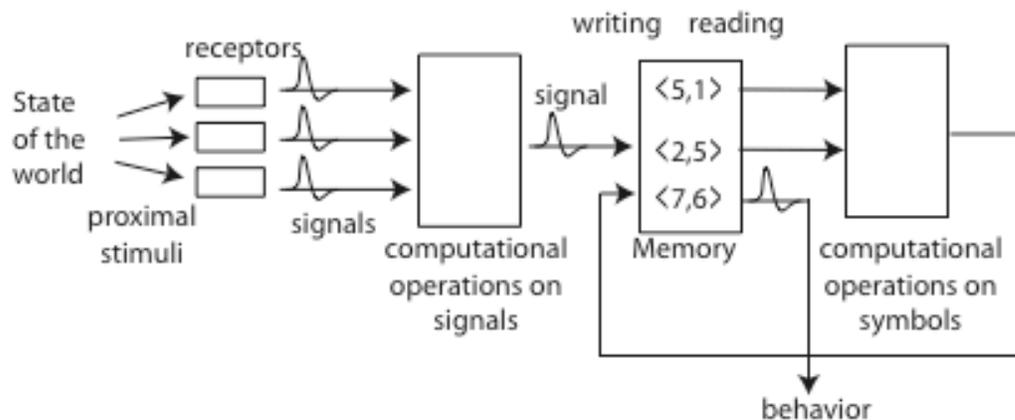


Figure 1. Schematic representation of the flow of information in a neurobiologically realized representational system. Proximal stimuli deriving from a state of the world (distal stimulus) act on sensory receptors to generate sensory signals, from which a perceptual signal specifying that state of the world is computed. The perceptual signal conveys the information to memory, where it is written into a symbol, which carries the information forward in time. Computational operations combine that symbol with other symbols to create further symbols and symbol strings in memory. The

information contained in a symbol is read from memory and converted to signals in the motor system that give form to behavior.

The essential features of a physically realized symbol are that it encode information about something else (to which it refers) and that it enter into symbol processing operations appropriate to the information that it encodes. Base-pair sequences in DNA are biological examples of physically realized symbols. The sequence of exon codons (base-pair triplets) between a start codon and a stop codon encodes the sequence of amino acids that compose a protein. These codon sequences carry forward in time evolutionarily accumulated information about functional amino acid sequences. They enter into combinatorial operations that recreate the sequences they encode. Base-pair sequences also encode promoters. The binding and unbinding of translation factors (themselves usually proteins), to one another, and to promoters, control the timing and amounts of protein synthesis. The base-pair sequences in promoters, carry forward in time evolutionarily accumulated information about functional patterns of protein synthesis (when and where it is functional to synthesize which proteins). The two kinds of information carried by base-pair sequences in DNA—exon information and promoter information—are roughly analogous to the two kinds of information carried in computer memory—data and program information.

In order to refer, the signals and symbols in a representational system must be causally connected to the things or states of the world to which they refer. In a process-control computer, the causal connection is effected by means of the transducers that generate signals proportional to critical variables such as temperature, torque, concentration, velocity, force, etc. Often these signals are analog signals (voltages), but these are usually converted almost immediately to symbols (bit patterns stored in memory buffers) by analog-to-digital converters. The bit patterns are then converted to digital signals (current pulses) that are transmitted over signal lines to the input registers (memory buffers) of the computer, where they write the bit patterns into memory registers. This chain of causes and effects causally connects the bit pattern in a memory register that represents the current temperature to the temperature to which it refers. The function of the memory register is to carry that bit

pattern (hence, the information about temperature that it encodes) forward indefinitely in time, for use in later computations. When it is to be used in a computation, the pattern is read from the register. In short, a symbolic memory register is written to by impinging information-conveying signals and read from by computational processes.

In the nervous system inputs acting on, for example, the retina of the eye or the basilar membrane of the ear, are converted first to analog signals (receptor potentials) and then into digital action potentials, which carry the information into the central nervous system. The behavioral evidence implies that the nervous system possesses a read-write memory mechanism that performs the same essential function performed by the memory registers in a computer, but we do not yet know what that mechanism is.

Because a symbolic memory mechanism has not so far been identified, it is often assumed not to exist. It is assumed that "memory element [in the brain take the] form of modifiable interconnections within the computational substrate" so that "no separate 'fetch' [read] and 'store' [write] cycles are necessary." (Koch & Hepp, 2006). These modifiable interconnections (synaptic conductances) are thought to be the physical realization of the associations in associative learning theory (Fanselow, 1993). There are, however, no proposals about how either associations in the abstract or experientially modified synaptic conductances (physically realized associations) can encode acquired information in a computationally accessible form. There are, for example, no proposals about how associations could specify the coordinates of a remembered location. Generally speaking, for associative theories of learning, this is not a problem. Because they are non-representational theories, so they do not require a symbolic memory mechanism.

For representational theories of learning, however, the absence of a symbolic memory mechanism is a problem, because a mechanism functionally equivalent to the tape in Turing's abstract conception of a general purpose computing machine (Turing, 1950, 1936) is essential to computation and representation (Gallistel, 2006). Representations are computed by combining information that arrives in dribs and drabs spread out in time, as will be illustrated shortly with the example of the dead reckoning process.

In order for new information to be combined with old information, there must be a mechanism that carries the old information forward in time in a computationally accessible form.

Symbols and the processes that operate on them create functioning homomorphisms between the symbol system and the aspects of the world to which the symbols refer (Gallistel, 1990). A homomorphism between two systems is a (partial) sameness in their abstract, mathematical form. Symbolic processes and relations in the representing system formally parallel non-symbolic processes and relations in the represented system. A functioning homomorphism is one in which the representing system exploits this parallelism to inform its interactions with the represented system.

Behavioral Evidence for Representations in Learning

Dead Reckoning

The position of an animal in its environment, as a function of time, is the integral of its velocity with respect to time. This is a mathematical fact about the relation between these vector variables (velocity and position). In mobile animals, a representation of their position relative to places of behavioral importance (its nest or resting area, food sources, hiding places, landmarks with reference to which these other places can be located, and so on) informs their behavior in many fundamental ways. The animal brain computes a representation of the animal's position by integrating with respect to time signals that convey information about its velocity. This symbolic integration process is called path integration in mathematical work and dead reckoning in traditional texts on marine navigation. It is a foundation of animal navigation (Collett & Graham, 2004; Etienne & Jeffery, 2004; Gallistel, 1990; Grah *et al.*, 2005; Kimchi *et al.*, 2004; Mittelstaedt & Mittelstaedt, 1973; Wehner & Srinivasan, 2003; Wehner & Srinivasan, 1981).

Dead reckoning is a particularly simple example of representational learning—if by 'learning' we understand the process of acquiring knowledge from experience. The animal acquires knowledge of its current position (and of past positions of behavioral interest) by means of a neurobiological process that

integrates experienced velocity, as conveyed by signals from sensory and motor mechanisms sensitive to correlates of velocity. The process of neurobiological integration in the brain, which is the representing system, parallels the physical integration of velocity that occurs as the animal moves within its environment. The animal's location in its environment is the represented system. In technical jargon, the symbolic processes in the brain are homomorphic to the displacement process in the system that the brain is representing, because there is a partial correspondence in the abstract (mathematical) form of the processes going on in the two systems. The resulting symbolic specification of position enables the brain to, for example, set the animal on the course for home. Insofar as the courses the brain sets actually bring the animal to its home, the symbolic representation of position informs functional behavior. This makes the neurobiological integration of velocity signals to obtain position symbols an example of a functioning homomorphism, that is, a representation.

Dead reckoning provides a particularly clear and simple example of the need for a symbolic memory in computation. The essence of dead reckoning is the adding of the current displacement to the sum of the previous displacements. For this to be possible, there must be a mechanism that carries the sum of the previous displacements forward in time in a form that makes it possible to add to it. The mechanism that carries the sum forward in time must not leak; the sum must not get smaller simply from the passage of time. It must not get smaller, because it is the physical realization of the symbol that specifies the subject's displacement from the origin of the reference frame, typically, the nest. If the sum gets smaller simply with the passage of time, then the brain represents the animal as getting closer to home just by sitting still.

Learning the Solar Ephemeris

The integration of the velocity vector in dead reckoning is meaningless unless the velocity vector is represented in a stable geocentric coordinate framework. The velocity vector specifies the subject's rate of displacement in two orthogonal directions (for example, latitudinal displacement and longitudinal displacement or radial displacement and tangential displacement—see Gallistel, 1990 Chap. 4 for detailed explication of path integration computations in

different coordinate frameworks). Thus, a consistent directional reference must be maintained. In navigational jargon, this is called carrying the parallel, because lines of direction at different locations (e.g., lines running north-south) are parallel. The line from the observer to any nearby landmark changes its direction as the observer moves. This is called parallax. The farther away, the landmark is, the less the parallax, hence, the better it functions as a directional referent. Extraterrestrial landmarks like the sun and the stars are, for practical purposes, infinitely far away, so they have negligible parallax. These extraterrestrial landmarks are strongly preferred as directional referents by animals of many different species (Gallistel, 1990). However, they have one drawback: because of the earth's rotation about its own axis, their terrestrial direction changes continuously.

The solar ephemeris is the direction of the sun in a terrestrial frame of reference, as a function of the time of day (that is, as a function of the angular position of the earth in its daily rotational cycle). To use the sun as a directional referent in dead reckoning, the brain must learn the azimuthal component of the solar ephemeris (Schmidt-Koenig, 1960). This function (solar compass direction vs time of day, see Figure 2) varies greatly depending on which hemisphere the observer is in (northern or southern), its latitude (angular distance from the equator) and the season of the year. Thus, it is a highly contingent fact about the animal's environment, the sort of thing that must be learned from experience.

There is an experimental literature on the learning of the solar ephemeris by homing pigeons and foraging bees (e.g., Budzynski *et al.*, 2000; e.g., Dyer, 2002; Towne *et al.*, 2005). To learn the solar ephemeris is to learn, for some point of view (e.g., the hive entrance or the nest), what parts of the local horizon profile the sun is over as a function of the time of day (see Figure 2). To learn that, the brain must represent the horizon profile, the position of the sun relative to this horizon profile, and the time of day. The time-of-day signal is provided by the animal's circadian clock (Giunchi *et al.*, 2003; Homberg, 2004; Mouritsen & Frost, 2002; Stalleicken *et al.*, 2005). The universally valid parameter of the circadian clock, its period, is genetically specified. The clock parameter that must be

adjusted based on local experience is the phase. The entrainment mechanism adjusts the phase of the clock in response to the rapid changes in photon flux that occur at dawn and dusk (Bertolucci & Foa, 2004; R. G. Foster & Bellingham, 2004; Foster *et al.*, 2003; Takahashi *et al.*, 1984).

The learning of the relation between the time-of-day signal provided by the circadian clock and the position of the sun relative to the horizon profile at the chosen point of view appears to be a curve-fitting process (Dickinson & Dyer, 1996; Dyer & Dickinson, 1996, 1994). Built into the learning mechanism is a parameterized dynamic process—the physical realization of an equation (function) specifying the relation two circular variables, time of day and position on the horizon. This built in equation specifies what is universally true about the solar ephemeris, namely, that the sun is somewhere to the east all morning and somewhere to the west all afternoon. Incubator-raised bees that have seen the sun only in the late afternoon when it is declining in the west, nonetheless represent it as being due east all morning long, stepping abruptly to due west around noon (Dyer & Dickinson, 1994). This default form for the ephemeris function is in fact valid only near the equator.

Notice that the information about where the sun is in the morning cannot have come from these incubator-raised bees' experience, because they never saw it in the morning. It is carried forward in time from the remote evolutionary past by the genes that code for the built in dynamic process relating the horizon position cycle to the circadian clock's cycle, just as the information specifying the period of the circadian clock is carried forward in time by the genes that specify its molecular structure. In experienced bees, which have observed the sun at the hive at several different times of day, the parameters of the built-in function are adjusted so that the function specifies locally appropriate horizon positions for the sun throughout the day. In the summer at mid-latitudes in the northern hemisphere, where the Dyer and Dickinson (1994) experiments were conducted, the sun rises north of east, moves continuously along the horizon in the clockwise direction to reach due south at noon and then on through west to north of west in the late afternoon. (By contrast, in winter at mid-southern latitudes, it rises north of east, and moves

continuously counterclockwise along the horizon to due north at noon and then on to north of west in the late afternoon.)

What bees learn from their experience of the sun's position over the horizon at different times of day is the locally appropriate values for the parameters of their built-in solar ephemeris function. These parameter settings are not *explicitly* specified by the bee's experiences of the sun's position above the horizon profile. There is no first-order sensory signal generated by any aspect of the bee's experience that directly specifies appropriate parameter values, in the way in which, for example, the signal from a photoreceptor explicitly specifies the number of photons captured by that receptor. Even the positional data—where the sun is over the profile—are only *implicitly* present in the spatio-temporal distribution of action potentials from the retina. The sun's profile-position coordinate must be computed from this pattern. And even then, the appropriate values for the parameters of the ephemeris equation are only implicit in data specifying several different profile positions at several different times of day. Thus, the explicit parameter values, the actual settings of the parameters, must be computed from these positional data. Moreover, the positional data are not given all at once. Experience provides different data points at different times of day. When it has extracted from the incoming retinal signals an explicit representation of the sun's profile position, the brain requires a symbolic memory to carry that information forward in time, so that it may be computationally combined with the positional data obtained later in the day to determine the appropriate values of the ephemeris parameters.

The symbolic information provided by the learned ephemeris informs not only the dead-reckoning process but other behaviors as well. When a foraging bee returns from a rich source, it does a waggle dance on the vertical surface of the hive, which symbolically specifies the direction and distance of the source (Figure 2). The direction is specified by the angle of the waggle run relative to vertical is the solar bearing of the source, the angle at which an outbound forager must hold the sun in order to fly toward the source. The solar bearing communicated by the dance is not the bearing that the returning forager has just flown; rather it is the inverse of that bearing. If the returning forager flew with the sun at

its back to reach the hive, then its dance tells the other bees to fly toward the sun. This information about the relation between the direction of the source from the hive and the direction of the sun from the hive was acquired many minutes earlier. A symbolic memory is required to carry this direction vector forward in time until it is used to inform the waggle dance (cf Menzel et al., 2006).

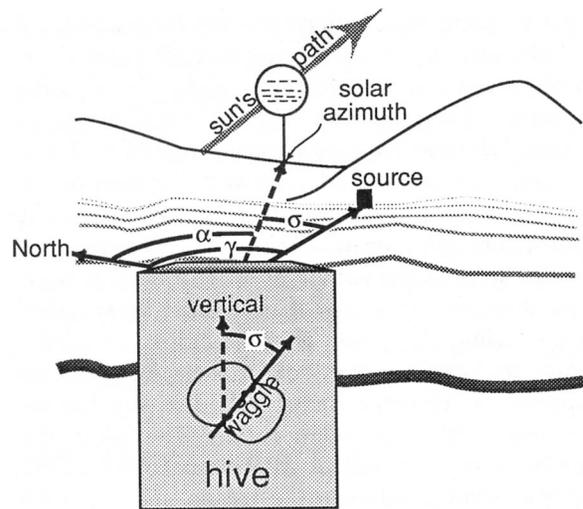


Figure 2. *In the waggle dance, the returned forager circles in alternate directions from one waggle run to the next, forming a figure 8. The angle of the repeated waggle runs, relative to vertical, indicates the solar bearing of the source from which the forager has returned. The number of waggles indicates its distance. α = the solar azimuth, its compass direction; σ = the solar bearing of the source, its direction relative to the direction of the sun; γ = the compass direction of the source. Note that the compass direction of the source is the sum of the solar azimuth and the solar bearing: $\gamma = \alpha + \sigma$. (Reproduced by permission of the author and publisher from Gallistel, 1998b)*

The information communicated by the dance is carried forward in the nervous systems of both the dancer and the observers of the dance in a computationally accessible form. If the hive is closed for several hours and then opened, outward bound foragers do not fly the solar bearing indicated by the dance they witnessed hours earlier; rather they fly a time-compensated solar bearing, a bearing that takes into account the intervening change in the sun's compass direction (Frisch, 1967). This implies either: i) A

recruited bee remembers the danced solar bearing and uses its solar ephemeris to compute the expected change in solar bearing over the interval since the dance. Or, ii) The remembered solar ephemeris was used at the time of the dance to compute a compass bearing or the indicated location on the cognitive map (map coordinates). The remembered compass bearing (or remembered map coordinates) must then be combined with the solar ephemeris function at the time of the outward bound flight to compute an appropriate solar bearing (how to fly relative to the sun so as to fly a remembered direction relative to the earth).

The learning and use of the solar ephemeris illustrates the manner in which different bits of information, gathered from experience at widely different times in the past, are computationally combined to obtain the explicit information (the current solar bearing of the food source) that informs current behavior.

The Cognitive Map

A *map* is a record of two or more locations in a common coordinate framework. It gives access to other (non-positional) information about those locations (for example, information that enables the navigator to recognize surrounding landmarks and information about the food to be found there or the dangers to be wary of). A *cognitive map* is a map of experienced terrain computed by a brain and recorded in its (presumed) symbolic memory. Dead reckoning and the construction and use of a cognitive map are intimately interrelated (M. Collett *et al.*, 1999; Gallistel, 1990, 1998; McNaughton *et al.*, 2006).

The information that dead reckoning provides about the animal's current position becomes more valuable as the number of locations on the cognitive map increases. It also becomes more valuable as the brain records more information that will enable it to recognize those places on subsequent occasions (for example, "snapshots" of surrounding landmarks, together with their direction and distance from the point of interest (Cartwright & Collett, 1979, 1983, 1987; Collett *et al.*, 1986). The information about current position becomes more valuable as the map becomes richer in information, because a major function of the cognitive map is to enable the brain to compute the range (distance from the animal)

and bearing (direction from the animal) of points of behavioral importance like the food and the nest (Collett et al., 1999). The directional information extracted by this vector computation is critical to the setting of a course toward or away from those points. The distance information is critical to the making of decisions based on estimates of the time it will take to reach a point.

Often, dead reckoning itself provides the estimate of the animal's position on its map, but sometimes either natural circumstances (gusts of wind) or experimental intervention renders the dead reckoning position useless, in which cases the brain must rely on landmark recognition to reestablish its position on its cognitive map and compute the course back to home or to the destination it was on before it was displaced (Fukushi & Wehner, 2004; Menzel et al., 2005; Menzel, et al, 2006).

Conversely, the better the positional information supplied by the dead-reckoning process, the more accurate a cognitive map becomes and the more easy it is to recognize relevant landmarks (Cartwright & Collett, 1979, 1983, 1987; Durier *et al.*, 2003; Graham *et al.*, 2004). The map becomes more accurate as dead reckoning improves, because it is likely that dead reckoning is the principal determinant of the position vectors recorded on the map. When the animal finds a point of behavioral interest, such as a food source, the position vector that represents that location on the brain's cognitive map is the dead reckoning position vector (the symbol that specifies where it is in the frame of reference used for dead reckoning, for example, 50 meters south and 22 meters east of the nest) plus the egocentric position vector for the point of interest (the symbol that specifies the location in the egocentric frame of reference, for example, 30° left and 2 meters away). Combining these two symbols is an exercise in coordinate transformation. The computation is spelled out in Gallistel (1990, p. 106ff; see also Gallistel, 1999). Thus, errors in the dead reckoning become errors in recorded locations. This is a major reason for the inaccuracies in early marine charts. In those, charts, too, estimates of longitude were largely based on dead reckoning.

It becomes easier to recognize landmarks as dead reckoning becomes more accurate, because estimates of one's location play a major role in the process. To recognize a landmark is to identify

what is now perceived with something recorded on the chart. In both human and animal navigation, position confers identity (Gallistel, 1990, p 140 & 168ff). That is, for a navigator to recognize a landmark, the navigator's estimate of its location must be within his uncertainty about where he is in relation to it on his cognitive map (Cartwright & Collett, 1983; M. Collett *et al.*, 2002; Dale *et al.*, 2005). A sailor on Long Island Sound seeing to her east something that looks exactly like Mount Vesuvius will not conclude that she is in Italy. She will ponder either why she never knew there was any part of Long Island that looked anything like Mount Vesuvius or, more likely, how remarkable that a cloud formation could look exactly like Mount Vesuvius. Thus, contrary to the common assumption, a major factor in the recognition of a landmark is that it be roughly where it is supposed to be. Being where it is supposed to be is an essential aspect of its identity. If it is not where it is supposed to be, it is not that landmark.

In short, dead reckoning provides major input to the computations that determine the recorded locations (position vectors) of points on the cognitive map. The recognition of landmarks along familiar routes (these are called waypoints in traditional navigation) is a substantial aid to dead reckoning (Kohler & Wehner, 2005). However, the animal's estimate of its position and orientation on the cognitive map is a major determinant of whether it looks for a given landmark and of whether it accepts a possible landmark as the one it is looking for. Finding a looked-for landmark helps to correct error in the dead reckoning (Etienne *et al.*, 2004) and get the animal back onto a familiar route (Kohler & Wehner, 2005).

All of this—all of our understanding of animal navigation—presupposes a symbolic memory, a memory capable of encoding information about the world and the animal's relation to it and carrying that information forward in time in a computationally accessible form. The physical realization of the geocentric position vector in the dead reckoning machinery must be such that it can be computationally combined with the physical realization of the egocentric position vector in the machinery that assigns egocentric positions to perceived landmarks (see Gallistel, 1999 for a review of the neurobiology of coordinate transformations). The resulting

vector symbol (the physical realization of the geocentric coordinates of perceived landmark) must preserve the information about the location of the landmark on the cognitive map. It must do so in a form that makes it possible to add and subtract that position vector from other position vectors, because that is what is required in computing the range and bearing of one location from another location (cf Collett et al., 1999).

In order to recognize a landmark, the brain must previously have made a record of what that landmark looks like (a "snapshot," cf Collett & Collett, 2002; Graham *et al.*, 2004). The record must contain enough information to distinguish between that landmark and other landmarks the animal has encountered in that vicinity. The number of distinguishable snapshots that the brain of an ant or bee might make is for all practical purposes infinite--larger than the number of elementary particles in the universe. Thus, there cannot be for every *possible* snapshot a genetically specified neural circuit excited by that snapshot and only that snapshot. There cannot be a "gnostic neuron" (Konorski, 1967) for every possible visual scene the ant or bee may need to record. This is derisively known as the grandmother-neuron theory of perceptual encoding.

When applied to, for example, a digital camera, the grandmother-neuron theory is that the factory builds into each camera all possible pictures that might ever be taken. When you press the shutter release, the camera detects the correspondence between the image on the camera's retina and one of the images built into it at that factory, so the "neuron" for that picture is activated. In this way, the camera (or the brain) "recognizes" the picture that you are taking.

Clearly this will not do; the possible snapshots cannot be genetically specified and prewired into the nervous system of the ant or the bee. There must be some decompositional encoding scheme that permits finite representational resources to encode the finite number of actual snapshots taken in any one lifetime. No matter how enthusiastic a photographer you are, the pictures you actually take are a negligible fraction of all the possible pictures! A digital camera decomposes images into pixels. It uses symbolic memory to record the image values for each pixel. Thus, it encodes

in memory only the pictures actually encountered, not all the pictures that might ever be encountered.

There are much more efficient encodings, as we learn when we compress image files. The use of basis functions, as in Fourier decomposition, is an example of a less obvious and potentially much more efficient encodings. Efficient encodings reduce complex inputs to vectors (in the loose sense of ordered strings of numbers) that specify parameter values for the functions that form a basis of the encoded space. Another example of the use of finite resources to encode an infinite range of possibilities is the encoding of to-be printed pages of drawings and text using Postscript.TM The number of distinguishable pages that may be printed is infinite. However, the file for any actual page is surprisingly short, which is a tribute to the efficiency of the encoding scheme.

This approach to understanding how the decidedly finite nervous system of an ant might preserve the information in a snapshot presupposes that there is a symbolic memory mechanism capable of preserving the decomposition of a snapshot (a visual scene) for later computational use. The decomposition made when the landmark was first encountered must be available when the animal encounters that landmark again. It must be available in a form that allows it to be compared to the decomposition of the currently viewed landmark. The computational architecture should allow for the generation of a measure of similarity between the two encodings being compared (the encoding of the old image in memory and the encoding of the image now on the retina). Moreover, the making of a "snapshot" as the term implies, cannot require repeated presentation of the image to be recorded; it must be a one-off process, like the process in a digital camera, which makes fundamental use of a symbolic memory mechanism.

Recognition is, of course, what attractor networks do, and they do not have a symbolic memory. (The lack of a symbolic memory is what distinguishes a neural net from a conventional computing machine.) However, because it lacks a symbolic memory, a neural net cannot compute measures of similarity or relative probability. Thus, a neural net recognizer cannot provide information about two different possible matches at the same time. Activity in the net must migrate to one attractor state or another. The net cannot be in two

different activity states at the same time. When it isn't in one attractor state or another, its activity state does not specify anything about the input. Activity states intermediate between two attractor states do not specify, for example, the relative probabilities of two different possible matches (two different attractor states that the net might be in but isn't).

The net's inability to specify a probability distribution over possible states of the world (possible matches between the current image and previous images) makes it difficult to combine the evidence from different computational procedures. For example, it makes it difficult to combine the evidence from image comparison (the probability distribution on the possible matching images, considering only the information in the images themselves) with the evidence from dead reckoning (the prior probability distribution over the possible matching images in memory, given the probability distribution on the animal's position and orientation on its cognitive map). The process of combining a prior probability distribution with

This last point is important because it relates to recent findings demonstrating: i) that the cognitive map gives access to other information about the world, and ii) that even insects engage in what looks like Bayesian inference, combining probabilistic information acquired at different times from different sources. In these experiments (Gould & Gould, 1988; Tautz *et al.*, 2004), foraging bees return from a food source on a rowboat in the middle of a pond or small lake. The returning foragers do the waggle dance, indicating the range and bearing of the food source, but the dance fails to recruit other foragers. This implies: 1) What the dance really communicates is not flying instructions (range and solar bearing) but rather map coordinates. and 2) The bees that observe the dance consult their cognitive maps before deciding whether to act on it. If their own past experience, as recorded on their map of where they have found food, indicates that no food is to be found anywhere near the location indicated by the dance, they decide not to go. Most interestingly, when the boat with the food source is moved close to an island, then the dance of the returning foragers does recruit new foragers. However the recruits do not come to the boat but rather to the shore near the boat (Tautz *et al.*, 2004). This implies that they have combined the probability distribution for the

source location indicated by the dance (the approximate location of the source implied by the dance) with the a priori probability distribution (based on their experience, as recorded on their map), to arrive at a posterior distribution, whose modal point is the nearby shore (the most probable location, all considered).

These findings suggest how grossly we may have underestimated the representational and computational capacities of even very small brains. Our gross underestimation of the representational capacity of brains as small as the head of a pin has allowed us to suppose that brains can get along without the symbolic memory mechanism that makes representation possible.

The Representation of Past Episodes

Our memory for episodes is another example of one-off memory for specific occurrences, where the number of possible episodes is infinite. Episodes by definition happen only once. We often recall them, however, long after they were encoded and committed to memory (albeit not, as accurately as we like to think). More importantly, we combine the information from one episode with the information from another to draw conclusions not justified by any single episode: the fourth time you see a man dining with an attractive woman not his wife, you begin to think he might be a philanderer. When this conclusion becomes a fixed belief, it is an example of a declarative memory ("The man's a philanderer") inferred from a sequence of episodic memories. This does not, of course, imply that the episodic memories are forgotten. Indeed, we may call on them to justify our inference to others. Recent ingeniously constructed experiments with food-caching jays show that we are not alone in our ability to remember episodes, nor to draw behaviorally informing inferences and declarative memories from them.

In times of plenty, many birds gather food and store it in caches. Western Scrub Jays are particularly remarkable cachers. They make more than thirty thousand different caches spread over square miles of the landscape (Vander Wall, 1990). Weeks and months later, during the winter when food is scarce, they retrieve food from these caches, one by one. This is another illustration of the vast demands on memory made by a cognitive map. It also

emphasizes the critical importance of a computational architecture that can effect the same computation (vector subtraction) on many different symbol combinations (pairs of vectors) without having a different neural circuit for each possible symbol combination. If we suppose that the jay can set a course toward the remembered location of any one cache while at the location of any other, then, given 30,000 locations, there are on the order of 10^{10} possible course computations (vector subtractions with distinct pairs of vectors formed from the 30,000 vectors in memory). Whatever the computational architecture of the nervous system, it cannot be such as to require 10^{10} different genetically specified neural circuits to effect these 10^{10} different computations.

A long series of ingenious experiments by Clayton and Dickinson and their collaborators have shown that jays remember much more than simply the locations of their caches (Clayton *et al.*, 2003a; Clayton *et al.*, 2003b; Clayton & Dickinson, 1998; Nicola S. Clayton & Dickinson, 1999; Clayton *et al.*, 2006b; Joanne M. Dally *et al.*, 2005a; Joanna M. Dally *et al.*, 2005b; de Kort *et al.*, 2005; Emery & Clayton, 2001; Emery *et al.*, 2004). They remember what kind of prey they hid in each cache, when they made each cache, which other jay if any was watching when they made that cache, and whether they have subsequently emptied that cache. They also remember whether they themselves have pilfered the caches of another bird. They remember the intervals that have elapsed in the past between the hiding and retrieval of a given kind of food and whether the food, when retrieved had or had not rotted. All of this remembered information combines to determine the order in which they will visit caches they have made. The information drawn from memory that is combined to inform current behavior comes from a mixture of episodic memories ("Three days ago, I hid meal worms there, there and there, and five days ago, I hid peanuts there, there and there") and declarative memories ("Meal worms rot in two days; peanuts don't rot").

These experiments demonstrate a rich representation of the jay's past experience and the ability to compute with the symbols that carry the information gained from that experience forward in time. For example, the bird's compute from the current date-time and the remembered date-time at which they made a particular

cache the time elapsed since they made that cache. They compare the time elapsed since that cache was made with the remembered time that it takes the contents of that cache to rot. If the computed time elapsed is greater than the remembered rotting time, they visit first the caches where they put the peanuts, even though they prefer (unrotted) meal worms to peanuts, and even though they made the meal worm caches more recently.

When their caching is observed by another jay, they combine the memory of which particular jay it was with a memory of the own much earlier thieving (from which they appear to infer the existence of evil) and their memory for the social status of that jay. If it was a dominant jay and if their experience of their own behavior (their own pilfering) has made them believe in the evil nature of jays, then they are likely to return to the cache when no one is looking, retrieve the food, and hide it elsewhere.

The rich representation of past episodes implied by these results again implies a decompositional scheme of some kind, because the number of possible caching episodes is infinite. Indeed, it suggests the kind of representation of our experience that appears in our linguistic descriptions of episodes (agents, actions, objects of actions, locations, markers of temporal position relative to the present and to other events, and so on). Perhaps, we should not be surprised to find evidence for such representations in non-verbal animals. Our ability to use language to communicate to others the information we have extracted from our own experience of the world is itself astonishing. It would be far more astonishing if we assumed that our non-verbal ancestors had no high-level decompositional representation of the world prior to the emergence of language.

Implications for Neurobiology

As mentioned at the outset, the hypothesis that learning depends on the brain's computing representations of selected aspects of the experienced world is controversial because we do not know what the neurobiological realization of key components of the requisite machinery is. Indeed, we do not even have a clear notion of what this realization might be, let alone persuasive evidence that an hypothesized realization is in fact the realization. The keystone in

any symbolic computation is a symbolic read-write memory mechanism. That is a mechanism to which information conveyed in signals can be written, and which will carry that information forward into the more or less indefinite future, until such time as it is needed in further computation. From a neurobiological perspective (that is from a material reduction perspective), symbols are the enduring physical changes in the nervous system that carry the information forward in time.

The behavioral evidence implies that a vast amount of acquired information is carried forward by this mechanism, even in the brains of birds and insects. This constrains physically plausible mechanisms (Gallistel, 2002). Whatever the mechanism is, it must realize a very high density of information storage; it must pack gigabytes into cubic microns. It must be thermodynamically stable, so that the preservation of the information has little or no metabolic cost. One cannot pack a great deal of information into a very small volume using a mechanism that depends on the expenditure of metabolic energy to preserve the information, because 1) there is no way to supply the requisite amount of energy within the small volume and 2) if there were, there would be no way to dissipate the heat generated.

In short, we must remember that the laws of thermodynamics apply to whatever symbolic memory mechanism we may imagine. These thermodynamic considerations make reverberating circuits an implausible physical realization of a mechanism whose essential function is to store large amounts of information for long periods of time. This is important because reverberating activity is the mechanism for storing previously computed information in recurrent neural net models, for example, in moving-activity-bump models for dead reckoning (Conklin & Eliasmith, 2005; Samsonovich & McNaughton, 1997; Song & Wang, 2005; Stringer *et al.*, 2002).

Another constraint on the physical realization of the symbolic memory is that it must be capable of encoding information and in a readable form. That is, it must be possible to envision how the enduring physical change hypothesized to be the physical realization of a symbol could in principle specify a fact about the world and how that fact could be recovered (read) from it. It is this consideration that makes Hebbian synapses implausible as the

physical realization of symbolic memory. Hebbian synapses are synapses that enduringly change their conductance as the result of the pairing of pre- and post-synaptic activity. The conductance of a synapse is the amplitude and duration of the transient change in the postsynaptic membrane potential when a spike arrives at the presynaptic terminal.

The first thing to be noted in this connection is that there are few if any published suggestions that it is the synaptic conductances themselves that encode information about the world (as opposed to the gnostic neurons that are connected through those conductances). As already noted, changes in synaptic conductances are the hypothesized physical realization of changes in the strengths of associative bonds. Associative theories of learning have traditionally not been representational theories, for the simple reason that it is hard to make a symbol out of an association.

Traditionally, associative strengths have been assumed to change slowly and in a manner dependent on many different aspects of a repeated experience: how close two stimuli are in time, how strong each of them is, how often they have occurred in close temporal conjunction, and so on. Thus, the strength of an association, that is, the conductance of a modifiable synapse, is the product of many different aspects of experience. This means that it cannot encode any one of those aspects. Mathematically speaking, the mapping from experience to an associative strength is a many-one function, and many-one functions are not invertible; you cannot recover the many from the one.

Even if the process of modifying a synaptic conductance were somehow constrained in such a manner that the conductance of a "memory synapse" could be made, for example, proportional to a to-be-remembered distance, the architecture of the nervous system, *as currently understood*, would not permit the conductance of that synapse to be read. The postsynaptic signal is a product (or joint function) of the pre-synaptic signal and the synaptic conductance. The synaptic conductance is usually called the synaptic weight, because it is a multiplicative constant that weights or scales the presynaptic signal to determine what is seen by the postsynaptic integration process. Unless the postsynaptic mechanism has independent access to the pre-synaptic signal (unless it has

information about that signal by a pathway other than the pre-synaptic pathway), it cannot estimate from the postsynaptic signal what the conductance of the synapse is. Thus, even if experience had made the strength of a synaptic conductance proportional to the distance to the food, it is hard to see how that information could be recovered by the postsynaptic integration process.

The mathematical impossibility of recovering the conductance of a synapse from the postsynaptic effect of an unknown presynaptic signal does not go away when one considers instead the problem of recovering the synaptic conductances (weights) in a neural network from the activity in its output neurons. If the postsynaptic effects are linear in the presynaptic signal strengths, then the activities of the output neurons may be regarded as the known values in a linear algebra problem, in which the input signal strengths and the intervening synaptic conductances are the unknown values. Unless there are as many knowns (output signals) as there are unknowns (input signals and synaptic conductances), it is a basic algebraic truth that the values of the unknowns cannot be recovered.

If the network is a richly connected one, with more synapses than inputs and outputs combined, then the conductances of the synapses cannot be recovered even when both the output and the input signals are known. If post-synaptic effects are non-linear in the inputs, the problem is still worse, because there is loss of information in non-linear operations. Intuitively, this is because when there are thresholds, one cannot estimate from the signal on the other side of the threshold what is going on below the threshold. Thus, recognizing that the synapses whose alteration is supposed to carry information forward in time are embedded in complex networks does not on the face of it make the problem of readability less of a problem; it makes it more of a problem.

This is not to say that some particular network architecture may not solve the problem. A particular architecture dictated by the logic of the problem is just what we see in the read-write memory of a computer. One might imagine, for example, that modifiable synapses only had two states, conducting and non-conducting, and that information about facts like distance and direction was encoded in banks of such synapses, using the same binary code by which

distance and direction information is encoded in the memory registers of conventional computers. However, this, too, would require both revising our conception of the functional architecture of the nervous system so as to enable write and read operations. And, as indicated by the quote from Koch (2006 see earlier p.), the current understanding of the nervous system has led to the conclusion that this is precisely what its architecture does not support. Moreover, as the Koch quote indicates, this is taken to be an advantage of the architecture, not a deficiency. Thus, *within our current understanding of the functional architecture of the nervous system*, there does not appear to be a way to make modifiable synaptic conductances be the mechanism of symbolic memory. This accounts in some measure for the strong strain of anti-representational theorizing in neurobiologically inspired models of learning.

The presence of a read-write memory mechanism has far reaching architectural implications. There is no point in writing information to memory if you cannot find it later. Thus, the existence of a read-write mechanism implies or presupposes that the architecture of the system supports memory addressing. It makes no functional sense to have a read-write memory in a machine whose architecture does not support memory addressing.

It is memory addressing that makes possible: i) the distinction between a variable and the value of that variable, ii) the ability to bind a value to a variable, and, therefore, iii) the creation of symbolic structures (data structures). Memory addressing makes all this possible, because the address in memory at which a given piece of information (a symbol) is to be found is itself information that may be written to memory at another address. This leads to indirect addressing in which the value to be operated on is specified not by its address, but rather by the address where its address is stored. The bit pattern specifying the first address (Address 1 in Figure 3) is the symbol for the variable itself (e.g., the distance to the food source), not for the value that the variable happens to have at any one time (e.g, 10 meters). The symbol for the variable gives access to the value of the variable by means of the bit pattern stored in the data field at the address specified by the symbol for the variable. The "data" stored in that field is the address of the value (Address 2

in Figure 3). The bit pattern (symbol) that specifies the value is in the data field at the second address. To change the value of the variable, the computer changes the bit pattern stored at the second address, not the bit pattern for the first address (the symbol for the variable) nor the bit pattern stored at the first address (the address of the variable's value).

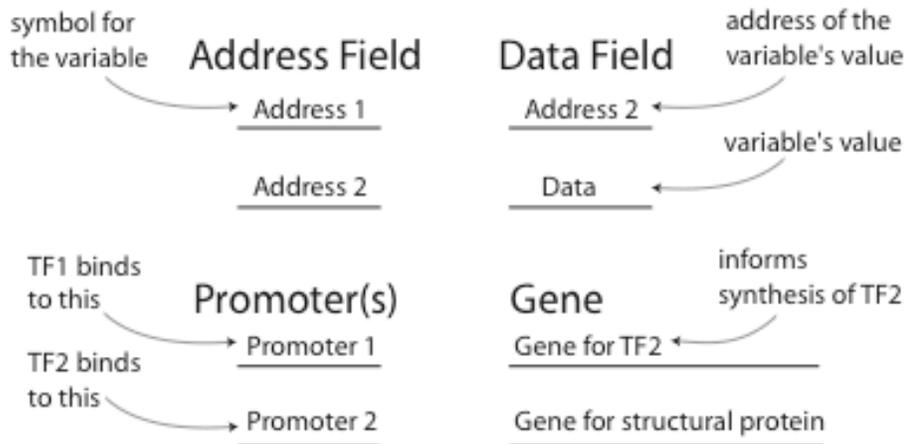


Figure 3 *Indirect addressing in computers and the genome. TF = transcription factor. In both examples, the entry in the bottom right field ("Data" or "Gene for structural protein") may be yet another address (another gene for a TF). This recursive addressing logic makes possible hierarchical data structures with a (potentially) unlimited number of levels.*

This sounds far removed from biological reality, but, as shown in Figure 3, the information stored in DNA is retrieved by same functional architecture, implemented now at the molecular level, which is to say more compactly than in current computer memory. The address at which a datum (the codon sequence specifying a protein) is stored is the promoter for that gene. Transcription factors initiate the reading of that information (transcription) by binding to the promoter. When a transcription factor binds to the promoter for a gene, the amino acid sequence coded for by that gene is used by the molecular machinery that synthesizes proteins to control the synthesis of the protein defined by that sequence. These protein thereby synthesized may be either a structural

protein (a building block of cellular structure, analogous to the actual value of variable) or transcription factor (an address). Addressing the promoter of a transcription factor (synthesizing a protein that binds to its address) gives access to the addresses (promoters) of the genes to which that transcription factor binds. As that transcription factor is synthesized, it binds to those promoters, leading to the synthesis of the proteins coded by their genes, many of which proteins may themselves be yet further transcription factors. This indirect addressing makes possible the hierarchical structure of the genome. It makes it possible to have an "eye" gene that, when activated (addressed by the transcription factor for its promoter), leads to the development of an entire eye (Halder *et al.*, 1995). The eye gene codes only for one protein. That protein does not itself appear anywhere in the structure of the eye. It is a transcription factor. It gives access to the addresses of other transcription factors and, eventually, through them, to the addresses of the proteins from which the special tissues of the eye are built and to the transcription factors whose concentration gradients govern how those tissues are arranged to make an organ.

What is needed for a brain to have learned representations is a similar architecture for accessing the information acquired from experience, an architecture in which it is possible to write information into biochemically realized symbols as well as read it from them.

The other critical components of a machine capable of representation that is missing from our current conception of the nervous system are the machines that operate on symbols. Needed are mechanisms that perform unary operations, such as negation, and binary operations, such as the arithmetic operations, string building operations (e.g., concatenation), and logical operations (e.g., AND and OR). The existence of mechanisms operating on symbols would be implied by the existence of a symbolic memory mechanism, because it does not make functional sense to have a read-write memory if there are no mechanisms for operating on the symbols in it. (It's like having chromosomes without ribosomes, which are the machines that put the amino acids together to form a protein.) What makes a representation are the symbols together with the operations on them. The operations enable the extraction

of explicit representations from information that is only implicitly present in the symbols already computed. A representation, that is, a functioning homomorphism, exists only when a machine can operate on its symbols to construct new symbols, symbol strings, and data structures.

A constraint on the mechanisms that operate on symbols is that there not be as many different mechanisms as there are distinct variables. As noted above in connection with the 30,000 cache locations that a jay can remember, it is not plausible to imagine an architecture in which there is a separate neural circuit dedicated a priori to each different vector subtraction that might ever have to be performed. The reason is once again the infinitude of the possible: there are too many possibilities.

This last constraint may seem almost too obvious to mention. However, in many neural net models of even simple computations, such as those involved in dead reckoning, the combinatorial operations are implemented by a table-look-up architecture (e.g., Samsonovich & McNaughton, 1997). There is a separate look-up table for each instance of a given kind of operation--a different table for each different case in which two variables must be added. A given table can effect the multiplication of all possible values of two variables, but it can operate only on those two variables. Moreover, each such table is composed of tens of thousands of neurons. because each different combination of values for the two variables is effected by the neurons that compose the corresponding cell in the table. In short, there are as many different look-up-tables as there are pairs of variables whose values may have to be combined, and within each such table, there are as many different neural sub-circuits as there are pairs of values for the two variables. Such an architecture is prodigally wasteful of material resources. It is nakedly exposed to combinatorial explosions that lurk behind every tree in the computational forest. That appears to be the price that must be paid for doing without a symbolic memory.

These considerations suggest that the nervous system may in fact contain a yet-to-be discovered read-write memory mechanism. They suggest, moreover, that the mechanism is likely to be found at the level of molecular structure, rather than at the level of cellular structure. Finally, they suggest that the discovery of such a

mechanism and the functional architecture that is required to make it effective would change our conception of how the brain works in fundamental ways.

References

- Bertolucci, C., & Foa, A. (2004). Extraocular photoreception and circadian entrainment in nonmammalian vertebrates. *Chronobiol Int*, *21*(4-5), 501-519.
- Budzynski, C. A., Dyer, F. C., & Bingman, V. P. (2000). Partial experience with the arc of the sun is sufficient for all-day sun compass orientation in homing pigeons, *Columba livia*. *Journal of Experimental Biology*, *203*, 2341-2348.
- Campos, M., Cherian, A., & Segraves, M. A. (2006). Effects of eye position upon activity of neurons in macaque superior colliculus. *Journal of Neurophysiology*, *95*(1), 505-526.
- Cartwright, B. A., & Collett, T. S. (1979). How honey bees know their distance from a nearby visual landmark. *Journal of Experimental Biology*, *82*, 367-372.
- Cartwright, B. A., & Collett, T. S. (1983). Landmark learning in bees: Experiments and models. *Journal of Comparative Physiology*, *151*, 521-543.
- Cartwright, B. A., & Collett, T. S. (1987). Landmark maps for honey bees. *Biological Cybernetics*, *57*, 85-93.
- Clayton, N.S., Emery, N., & Dickinson, A. (2006a). The rationality of animal memory: Complex caching strategies of western scrub jays. In Hurley (Ed.) *Book title to be specified*. (pp. 197-207).
- Clayton, N. S., Emery, N. J., & Dickinson, A. (2006b). The prospective cognition of food caching and recovery by western scrub-jays (*aphelocoma californica*). *Comparative Cognition & Behavior Reviews*, *1*(1), 1-11.
- Clayton, N.S., Yu, K. S., & Dickinson, A. (2003a). Interacting cache memories: Evidence of flexible memory use by scrub jays. *Journal of Experimental Psychology: Animal Behavior Processes*, *29*, 14-22.

- Clayton, N. S., Bussey, T. J., & Dickinson, A. (2003b). Can animals recall the past and plan for the future? *Nature reviews Neurosciences*, *4*, 685-691.
- Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*, 272-274.
- Clayton, N. S., & Dickinson, A. (1999). Memory for the content of caches by scrub jays (*Aphelocoma coerulescens*). *Journal of Experimental Psychology: Animal Behavior Processes*, *25*(1), 82-91.
- Collett, M., Collett, T. S., & Wehner, R. (1999). Calibration of vector navigation in desert ants. *Current Biology*, *9*(18), 1031-1034.
- Collett, M., Harland, D., & Collett, T. S. (2002). The use of landmarks and panoramic context in the performance of local vectors by navigating bees. *Journal of Experimental Biology*, *205*, 807-814.
- Collett, T. S., Cartwright, B. A., & Smith, B. A. (1986). Landmark learning and visuo-spatial memories in gerbils. *Journal of Comparative Physiology A*, *158*, 835-851.
- Collett, T. S., & Collett, M. (2002). Memory use in insect visual navigation. *Nature Reviews of Neuroscience*, *3*, 542-552.
- Collett, T. S., & Graham, P. (2004). Animal navigation: Path integration, visual landmarks and cognitive maps. *Current Biology*, *14*(12), R475-477.
- Conklin, J., & Eliasmith, C. (2005). A controlled attractor network model of path integration in the rat. *J Comput Neurosci*, *18*(2), 183-203.
- Dale, K., Harland, D. P., Manning-Jones, A., & Collett, T. S. (2005). Weak and strong priming cues in bumblebee contextual learning. *Journal of Experimental Biology*, *208*(Pt 1), 65-74.
- Dally, J. M., Emery, N. J., & Clayton, N. S. (2005a). Cache protection strategies by western scrub-jays, *aphelocoma californica*: Implications for social cognition. *Animal Behaviour*, *70*(6), 1251-1263.

- Dally, J. M., Emery, N. J., & Clayton, N. S. (2005b). The social suppression of caching in western scrub-jays (*aphelocoma californica*). *Behaviour*, *142*(7), 961-977.
- de Kort, S. R., Dickinson, A., & Clayton, N. S. (2005). Retrospective cognition by food-caching western scrub-jays. *Learning and Motivation*, *36*(2), 159-176.
- Dickinson, J., & Dyer, F. (1996). How insects learn about the sun's course: Alternative modeling approaches. In S. Wilson & e. al. (Eds.), *From animals to animats. Vol. 4* (pp. 193-203). Cambridge, MA: MIT Press.
- Durier, V., Graham, P., & Collett, T. S. (2003). Snapshot memories and landmark guidance in wood ants. *Current Biology Vol 13*(18) Sep 2003, 1614-1618.
- Dyer, F., & Dickinson, J. A. (1996). Sun-compass learning in insects: Representation in a simple mind. *Current Direction in Psychological Science*, *5*(3), 67-72.
- Dyer, F. C. (2002). The biology of the dance language. *Annual Review of Entomology*, *47*, 917-949.
- Dyer, F. C., & Dickinson, J. A. (1994). Development of sun compensation by honeybees: How partially experienced bees estimate the sun's course. *Proceedings of the National Academy of Sciences, USA*, *91*, 4471-4474.
- Emery, N. J., & Clayton, N. S. (2001). It takes a thief to know a thief: Effects of social context on prospective caching strategies in scrub jays. *Nature*, *414*, 443-446.
- Emery, N. J., Dally, J., & Clayton, N. S. (2004). Western scrub-jays (*Aphelocoma californica*) use cognitive strategies to protect their caches from thieving conspecifics. *Animal Cognition*, *7*, 37-43.
- Etienne, A. S., & Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, *14*(2), 180-192.

- Etienne, A. S., Maurer, R., Boulens, V., Levy, A., & Rowe, T. (2004). Resetting the path integrator: A basic condition for route-based navigation. *Journal of Experimental Biology*, 207, 1491-1508.
- Fanselow, M. S. (1993). Associations and memories: The role of nmda receptors and long-term potentiation. *Current Directions in Psychological Science*, 2(5), 152-156.
- Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084), 680-683.
- Foster, R. G., & Bellingham, J. (2004). Inner retinal photoreceptors (irps) in mammals and teleost fish. *Photochemical Photobiological Science*, 3(6), 617-627.
- Foster, R. G., Hankins, M., Lucas, R. J., Jenkins, A., Munoz, M., Thompson, S., et al. (2003). Non-rod, non-cone photoreception in rodents and teleost fish. *Novartis Foundation Symposium*, 253, 3-23; discussion 23-30, 52-25, 102-109.
- Frisch, K. v. (1967). *The dance-language and orientation of bees*. Cambridge, MA: Harvard Univ. Press.
- Fukushi, T., & Wehner, R. (2004). Navigation in wood ants *formica japonica*: Context dependent use of landmarks. *Journal of Experimental Biology*, 207, 3431-3439.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.
- Gallistel, C. R. (1998). Symbolic processes in the brain: The case of insect navigation. In D. Scarborough & S. Sternberg (Eds.), *Invitation to cognitive science* (Vol. 4). Cambridge, MA: MIT Press.
- Gallistel, C. R. (1999). Coordinate transformations in the genesis of directed action. In Bly, B. M. & Rumelhart, D. (Eds.) *Cognitive science*. (a volume in *Handbook of perception and cognition*. 2nd ed). San Diego: Academic (pp 1-42).

- Gallistel, C. R. (2002). The principle of adaptive specialization as it applies to learning and memory. In R. H. Kluwe, G. Lüer & F. Rösler (Eds.), *Principles of human learning and memory* (pp. 250-280). Berlin: Birkenaeuser.
- Gallistel, C. R. (2006). The nature of learning and the functional architecture of the brain. In Q. Jing & e. al. (Eds.), *Psychological science around the world: Proceedings of the 28th international congress of psychology* (Vol. 1, pp. 63-71). Sussex: Psychology Press.
- Giunchi, D., Mongini, E., Pollonara, E., & Baldaccini, N. E. (2003). The effect of clock-shift on the initial orientation of wild rock doves (*Columba l. livia*). *Naturwissenschaften*, *90*(6), 261-264.
- Gluck, M. A., Myers, C., & Meeter, M. (2005). Cortico-hippocampal interaction and adaptive stimulus representation: A neurocomputational theory of associative learning and memory. *Neural Networks*, *18*(9), 1265-1279.
- Golob, E. J., Stackman, R. W., Wong, A. C., & Taube, J. S. (2001). On the behavioral significance of head direction cells: Neural and behavioral dynamics during spatial memory tasks. *Behavioral Neuroscience*, *115*, 285-304.
- Goossens, H. H., & Van Opstal, A. J. (2006). Dynamic ensemble coding of saccades in the monkey superior colliculus. *Journal of Neurophysiology*, *95*(4), 2326-2341.
- Gould, J., & Gould, C. G. (1988). *The honey bee*. New York: Freeman.
- Grah, G., Wehner, R., & Ronacher, B. (2005). Path integration in a three-dimensional maze: Ground distance estimation keeps desert ants *cataglyphis fortis* on course. *Journal of Experimental Biology*, *208*, 4005-4011.
- Graham, P., Durier, V., & Collett, T. S. (2004). The binding and recall of snapshot memories in wood ants (*Formica rufa l.*). *Journal of Experimental Biology*, *207*(Pt 3), 393-398.

- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*, 802-806.
- Halder, G., P. Callaerts, & Gehring, W. J. (1995). Induction of ectopic eyes by target expression of the *eyeless* gene in *Drosophila*. *Science*, *267*, 1788-1792.
- Hawkins, R. D., & Kandel, E. R. (1984). Is there a cell-biological alphabet for simple forms of learning? *Psychological Review*, *91*, 375-391.
- Hay, J. B., & Baayen, R. H. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Science*, *9*(7), 342-348.
- Holst, E. v., & Mittelstaedt, H. (1950). The reafference principle (interaction between the central nervous system and the periphery). English translation in *The behavioral physiology of animals and man. Selected papers of E. von Holst. Vol. 1*. Coral Gables, FL: University of Miami Press. 1973
- Homberg, U. (2004). In search of the sky compass in the insect brain. *Naturwissenschaften*, *91*(5), 199-208.
- Hull, C. L. (1952). *A behavior system*. New Haven, CT: Yale University Press.
- Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current Opinion in Neurobiology*, *14*, 225-232.
- Kimchi, T., Etienne, A. S., & Terkel, J. (2004). A subterranean mammal uses the magnetic compass for path integration. *Proceedings of the National Academy of Sciences U S A*, *101*(4), 1105-1109.
- Koch, C., & Hepp, K. (2006). Quantum mechanics in the brain. *Nature*, *440*, 611-612.
- Kohler, M., & Wehner, R. (2005). Idiosyncratic route-based memories in desert ants, *melophorus bagoti*: How do they

- interact with path-integration vectors? *Neurobiology of Learning and Memory*, 83(1), 1-12.
- Konorski, J. (1967). *Integrative activity of the brain; an interdisciplinary approach*. Chicago: Chicago University Press.
- Lee, I., Yoganarasimha, D., Rao, G., & Knierim, J. J. (2004). Comparison of population coherence of place cells in hippocampal subfields Ca1 and Ca3. *Nature*, 430(6998), 456-460.
- Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, 38(2), 317-327.
- Lindley, D. (2004) *Degrees Kelvin: A Tale of Genius, Invention and Tragedy*. Washington, DC: Joseph Henry Press.
- Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.
- McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M.-B. (2006). Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, 7, 663-678.
- Meck, W. H. (Ed.). (2003). *Functional and neural mechanisms of interval timing*. New York: CRC Press LLC.
- Menzel, R., DeMarco, R.J. & Greggers, U. (2006) Spatial memory, navigation and dance behaviour in *Apis mellifera*. *Journal of Comparative Physiology*, 192, 889-903.
- Menzel, R., Greggers, U., Smith, A., Berger, S., Brandt, R., Brunke, S., et al. (2005). Honey bees navigate according to a map-like spatial memory. *Proceedings of the National Academy of Sciences USA*, 102(8), 3040-3045.
- Mittelstaedt, H., & Mittelstaedt, M.L. (1973) Mechanismen der Orientierung ohne richtende Aussenreize. *Fortschritte der Zoologie*, 21, 46-58.
- Mouritsen, H., & Frost, B. J. (2002). Virtual migration in tethered flying monarch butterflies reveals their orientation

mechanisms. *Proceedings of the National Academy of Sciences USA*, 99(15), 10162-10166.

- Muller, J. R., Philiastides, M. G., & Newsome, W. T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences USA*, 102(3), 524-529.
- Muller, R. U. (1996). A quarter of a century of place cells. *Neuron*, 17, 813-822.
- Newell, A., & Simon, H. A. (1975). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery*, 19, 113-136.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.
- Pavlov, I. V. (1928). *Lectures on conditioned reflexes: The higher nervous activity of animals* (H. Gantt, Trans.). London: Lawrence & Wishart.
- Penney, T. B. (2004). Electrophysiological correlates of interval timing in the stop RT task. *Cognitive Brain Research*, 21, 234-249.
- Ranck, J. B. J. (1984). Head direction cells in the deep layer of dorsal presubiculum in freely moving rats. *10(1)*, 599.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.
- Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.
- Samsonovich, A., & McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience*, 17, 5900-5920.
- Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M. B., et al. (2006). Conjunctive representation of

- position, direction, and velocity in entorhinal cortex. *Science*, 312(5774), 758-762.
- Schmidt-Koenig, K. (1960). The sun azimuth compass: One factor in the orientation of homing pigeons. *Science*, 131, 826-827.
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, 23, 473-501.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, 27, 379-423, 623-656.
- Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Foundations* (Vol. 1, pp. 194-281). Cambridge, MA: MIT Press.
- Song, P., & Wang, X. J. (2005). Angular path integration by moving "hill of activity": A spiking neuron model without recurrent excitation of the head-direction system. *Journal of Neuroscience*, 25(4), 1002-1014.
- Stalleicken, J., Mukhida, M., Labhart, T., Wehner, R., Frost, B., & Mouritsen, H. (2005). Do monarch butterflies use polarized skylight for migratory orientation? *Journal of Experimental Biology*, 208(Pt 12), 2399-2408.
- Stringer, S. M., Rolls, E. T., Trappenberg, T. P., & de Araujo, I. E. (2002). Self-organizing continuous attractor networks and path integration: Two-dimensional models of place cells. *Network*, 13(4), 429-446.
- Takahashi, J. S., De Coursey, P. J., Bauman, L., & Menaker, M. (1984). Spectral sensitivity of a novel photoreceptive system mediating entrainment of mammalian circadian rhythms. *Nature*, 308, 186-188.
- Tautz, J., Zhang, S. W., Spaethe, J., Brockmann, A., Si, A., & Srinivasan, M. V. (2004). Honeybee odometry: Performance in varying natural terrain. *PLoS Biology*, 2, 915-923.

- Towne, W. F., Baer, C. M., Fabiny, S. J., & Shinn, L. M. (2005). Does swarming cause honey bees to update their solar ephemerides? *Journal of Experimental Biology*, *208*(21), 4049-4061.
- Turing, A.M. (1950). Computing machinery and intelligence. *Mind*, *59*, 433-460.
- Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society 2nd ser.*, *42*, 230-265.
- Vander Wall, S. B. (1990). *Food hoarding in animals*. Chicago: University of Chicago Press.
- Wallace, D. G., & Fountain, S. B. (2003). An associative model of rat serial pattern learning in three-element sequences. *Quarterly Journal of Experimental Psychology B*, *56*(3), 301-320.
- Wehner, J., & Srinivasan, M. V. (2003). Path integration in insects. In K. J. Jeffery (Ed.), *The neurobiology of spatial behaviour* (pp. 9-30). Oxford: Oxford University Press.
- Wehner, R., & Srinivasan, M. V. (1981). Searching behavior of desert ants, genus *Cataglyphis* (*formicidae*, *hymenoptera*). *Journal of Comparative Physiology*, *142*, 315-338.
- Yoganarasimha, D., Yu, X., & Knierim, J. J. (2006). Head direction cell representations maintain internal coherence during conflicting proximal and distal cue rotations: Comparison with hippocampal place cells. *Journal of Neuroscience*, *26*(2), 622-631.