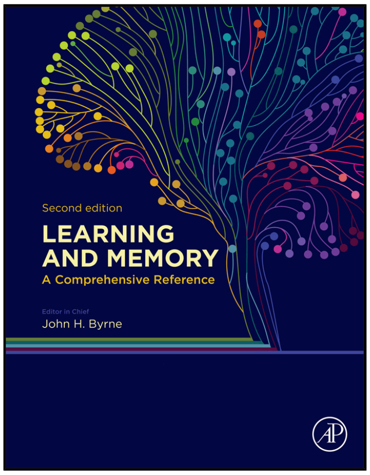

From Gallistel, R. (2017) Learning and Representation. In: Menzel, R. (ed.), *Learning Theory and Behavior*, Vol. 1 of *Learning and Memory: A Comprehensive Reference, 2nd edition*, Byrne, J.H. (ed.). pp. 141–154. Oxford: Academic Press.
http://dx.doi.org/10.1016/B978-0-12-809324-5.21009-2
ISBN: 9780128051597

Academic Press

1.08 Learning and Representation☆

Randy Gallistel, Rutgers University, Piscataway, NJ, United States

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 and Kandel, 1984; Hull, 1952; Kozma and Freeman, 2015; Pavlov, 1928; Rumelhart and 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 and King, 2010), 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., Moustafa et al., 2009), which is why they predominate in neurobiological thinking about learning.

The results of behavioral experiments on nonhuman animals have increasingly implied that much learned behavior is informed by enduring temporal and spatial representations (e.g., Balsam et al., 2010; Cheeseman et al., 2014; Menzel et al., 2005, 2011, 2012), as some prominent advocates of associative theories acknowledge (Clayton et al., 2006a; Dickinson, 2012). Moreover, direct electrophysiological observation of neural activity has shown that the nervous system represents where the animal is and has been within the environment (Aime et al., 2014; Derdikman and Moser, 2010; Kraus et al., 2015; Krupic et al., 2015; Rich et al., 2014; Stensola et al., 2012; Wills et al., 2010), how it is oriented (Bjerknes et al., 2015; Giocomo et al., 2014; Jeffery, 2014; Peyrache et al., 2015; Rubin et al., 2014; Tan et al., 2015), where objects are in relation to it (Fetsch et al., 2014; Song and McPeek, 2015; Yasuda and Hikosaka, 2015), and the timing of repetitive or predictable events (Bermudez and Schultz, 2012; Eichenbaum, 2013; Finnerty et al., 2015). Thus, there is a conceptual tension between the behavioral and electrophysiological 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 chapter: dead reckoning in insect navigation, the learning of the solar ephemeris, 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 equibility 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

---

☆*Change History*: February 2016. C.R. Gallistel updated the text and references to this chapter.

This is an update of C.R. Gallistel, 1.12 – Learning and Representation, In Learning and Memory: A Comprehensive Reference, edited by John H. Byrne, Academic Press, Oxford, 2008, Pages 227–242.

Learning and Memory: A Comprehensive Reference, Second Edition, 2017, 141–154

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 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.

1.08.1 Representations: Definition and Explication

From the perspective of cognitive science, the brain is an organ of computation (Marr, 1982; Newell and 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, Holst and Mittelstaedt, 1950).

A representation consists of signals, symbols, and the operations on them (see **Fig. 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 and the probabilities it associates with those states cannot extract from information-bearing signals the information that they carry. All it can do is generate a response of some kind.

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 currently know the physical realization of symbols in the nervous system. One school of thought doubts that they exist (Rumelhart and McClelland, 1986; Hay and Baayen, 2005; Bennett and Hacker, 2008; Proctor and Capaldi, 2012; Jones et al., 2015).

The essential features of a physically realized symbol are that it encodes information about something else (to which it refers) and it enters 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 and repressors. The binding and unbinding of translation factors (themselves usually proteins), to one another, and to promoters and repressors, control the timing and amounts of protein synthesis. The promoter–repressor components of a gene are closely analogous to the address components of a random access memory register in a conventional computer. The address portion of a memory register enables the machine to access the

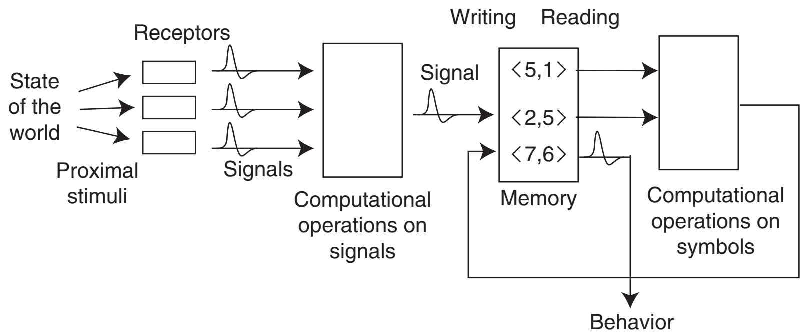

Figure 1 Schematic representation of the flow of information in a neurobiologically realized representational system. Proximal stimuli deriving from a state-of-the world (distal stimulus) act on sensory receptors to generate sensory signals, from which a perceptual signal specifying that state of the 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 gives form to behavior.

information stored in the coding portion. Similarly, the promoter–repressor portions of genes enable the protein-building machinery of the cell to access the information stored in the coding portion of the gene.

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. These signals are often 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, and then written into bit patterns in enduring memory registers. This chain of causes and effects causally connects the bit pattern in a memory register of a process-control computer, controlling, for concrete example, the temperature of a chemical bath, to the actual temperature of that bath. The function of the memory register in the process-control computers is to carry that bit pattern (hence, the information about bath temperature that it encodes) forward indefinitely in time, for use in later computations, for example, computations that determine the rate at which the temperature is changing. 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 that operate on the encoded information. The results of the computation are written back to memory for subsequent further use. This simple scheme makes possible the unbounded composition of functions.

In the nervous system, sensory stimuli 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. The action-potential signals come and go, but they leave behind some physical change that encodes the information that they carried. We do not as of yet know what that physical change is.

Because a symbolic memory mechanism has not so far been identified, it is often assumed not to exist. It is assumed that "memory elements [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 and Hepp, 2006, p. 612). 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 a pattern of associations could in and of itself specify the duration of an experienced interval or the location of an experienced boundary. Generally speaking, for associative theories of learning, this is not a problem. Because they are nonrepresentational 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 and King, 2010). 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. For new information to be combined with old information, there must be a mechanism that carries the old information forward in time in a computational 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; Gallistel and King, 2010). A homomorphism between two systems is a (partial) sameness in their abstract, mathematical form. Symbolic processes and relations in the representing system are formally parallel with nonsymbolic 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.

1.08.2 Behavioral Evidence for Representations in Learning

1.08.2.1 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 (a 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 (Grah et al., 2005; Collett and Graham, 2004; Etienne and Jeffery, 2004; Wehner and Srinivasan, 2003; Gallistel, 1990; Gould and Gould, 2012).

Dead reckoning is particularly a 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) descriptions of the processes going on in the two systems. These mappings between categories—things and processes in the world in one category, signals, symbols, and computational operations in the brain in the other category—are also called functors. 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 its position in the world 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 (change in position) to the sum of 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; that is, 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.

1.08.2.2 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 Chapter 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. That is why, mathematically speaking, a direction is an infinite set of lines that intersect at a point at infinity. Extraterrestrial landmarks such as 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; Gould and Gould, 2012). However, they have one drawback: As the earth's rotation is 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 **Fig. 2**) varies greatly depending on which hemisphere the observer is in (northern or southern), its latitude

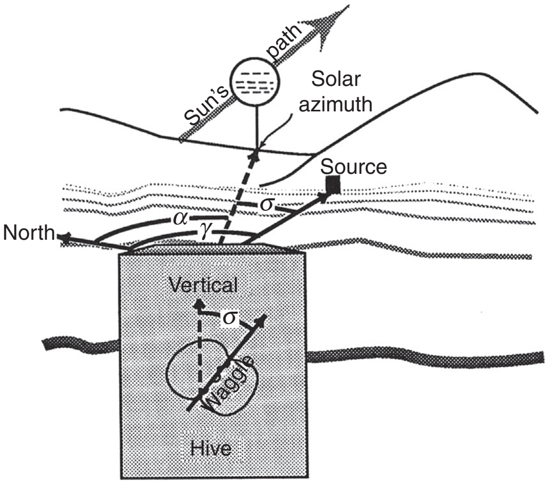

Figure 2 In the waggle dance, the returned forager circles in alternate directions from one waggle run to the next, forming a figure 8. The angle of the repeated waggle runs, relative to vertical, indicates the solar bearing of the source from which the forager has returned. The number of waggles indicates 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, C.R., 1998. Symbolic processes in the brain: the case of insect navigation. In: Scarborough, D., Sternberg, S. (Eds.), Conceptual and Methodological Foundations. Invitation to Cognitive Science, second ed., vol. 4. MIT Press, Cambridge, MA, pp. 1–51.

(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., Dyer, 2002; Budzynski et al., 2000; Kemfort and Towne, 2013; Towne, 2008). 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 **Fig. 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 (Mouritsen and Frost, 2002; Giunchi et al., 2003; Homberg, 2004; Sauman et al., 2005; Cheeseman et al., 2012). 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 (Takahashi et al., 1984; Bertolucci and Foa, 2004; Foster and Belingham, 2004). The phase of the circadian clock may also be shifted by a general anesthesia (Cheeseman et al., 2012).

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 and Dyer, 1996; Dyer and Dickinson, 1996, 1994). Built into the learning mechanism is a parameterized dynamic process – the physical realization of an equation (function) specifying the relation between 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 and 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 spatiotemporal 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 (**Fig. 2**). The direction 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 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 (Menzel and Eckoldt, 2016).

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 (von Frisch, 1967). This implies either (1) 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 (2) 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.

1.08.2.3 The Cognitive Map

A *map* is a record of multiple locations in a common coordinate framework. It gives access to other (nonpositional) 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 (Gallistel, 1990; Collett et al., 1999; McNaughton et al., 2006; Eilam, 2014).

A cognitive map enables the brain to compute the approximate range and bearing of any location from any other location, when both locations are on the map, that is, when both are in familiar territory. The technical term for this property in the representational theory of mind is *productivity* (Fodor and Pylyshyn, 1988): the machinery that enables the computation of any one range and bearing enables the computation of arbitrarily many other ranges and bearings. The productivity of an animal's cognitive map is most salient when it takes shortcuts between locations that it has previously visited only in round about ways and when it orients toward home or a food location following a large externally imposed displacement, as when a bee is displaced by a storm gust or by an experimenter (cf. Cheeseman et al., 2014).

An impressive demonstrations of the productivity of the bee's cognitive map comes from an experiment by Menzel and collaborators in which labeled and radar-tracked foraging bees learned about one location by direct experience—they repeatedly flew between it and the hive—and a second location only by witnessing the dance in the hive of other bees that were foraging at that second location (Menzel et al., 2011). On a test day, when they arrived at the location where they had been foraging, there was no artificial nectar there. After flying around that site for a while, they then departed directly for the location that they knew about only second hand. However, they did so only if the angle subtended by the bearings of the two locations from the hive was below a critical value. The more acute this angle is, the greater the savings are from taking the shortcut between the two locations, as opposed to returning first to the hive before proceeding on to the other location.

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 and 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 such as 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 (Cheeseman et al., 2014; Fukushi and Wehner, 2004; Jin et al., 2014; Menzel et al., 2005).

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 (Durier et al., 2003; Graham et al., 2004; Schultheiss and Cheng, 2011). 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 (Buzsaki and Moser, 2013; Derdikman and Moser, 2010). 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 m south and 22 m 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 degrees left and 2 m 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 and 168ff). That is, for a navigator to recognize a landmark, the navigator's estimate of its location must be within his/her uncertainty about where he/she is in relation to it on his/her cognitive map (Dale et al., 2005; Collett and Collett, 2002; Cartwright and Collett, 1983). A sailor on Long Island Sound seeing to her/his east something that looks exactly like Mount Vesuvius will not conclude that she/he is off the Bay of Naples. She/he will ponder either why she/he 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 and 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 and Wehner, 2005; Schultheiss and Cheng, 2011; Wystrach et al., 2014, 2011).

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; Menzel and Greggers, 2015).

To recognize a landmark, the brain must previously have made a record of what that landmark looks like (a "snapshot," cf. Graham et al., 2004; Jin et al., 2014). 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 selectively 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, so the "neuron" for that picture is activated. Because the image uniquely matches a built-in template, the camera (or the brain) "recognizes" the picture that you are taking.

Clearly this will not do; an infinite set of *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 vanishingly small fraction of all the possible pictures! A digital camera decomposes images into pixels. Each pixel registers the number of photons captured at one tiny location within the image. The camera uses symbolic memory to record the binary numbers from each pixel.

The pixels from images are not all random; they have a great deal of statistical structure, which is to say that knowledge of the value of one pixel or of an ensemble of pixels enables one to predict with high probability nearby and even remote pixels. Sophisticated compression algorithms based on the known statistics of natural images enable computers to compress pictures, so that they take up orders of magnitude less room in memory than they would if they were stored pixel by pixel. It is likely that even the insect brain does similar image compression (Jin et al., 2014). In finding the statistical structure in images that makes it possible to encode them more compactly, the brain also finds the structure that enables it to make effective use of the images, an application of the minimum description length principle (Grünwald, 2007; Wilkes and Gallistel, 2017).

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 encoding of a snapshot (a visual scene) for later computational use. That information must be available when the animal encounters that landmark again, 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, cf. Collett, 2010; Zeil, 2011). 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 (the pixel, which is a memory register that records the photon capture number at that location in the image during the moment when the picture was taken).

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 is not 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 is not).

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 a likelihood function based on current information is Bayesian inference or Bayesian updating.

This last point is important because it relates to findings demonstrating (1) that the cognitive map gives access to other information about the world and (2) that even insects engage in what looks like Bayesian inference, combining probabilistic information acquired at different times from different sources. In these experiments (Gould and Gould, 1988; Tautz et al., 2004; see, however, Wray et al., 2008), 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 prior 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.

1.08.2.4 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 is 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. 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 30,000 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 and Dickinson, 1998, 1999; Emery and Clayton, 2001; Clayton et al., 2003a,b; Emery et al., 2004; Clayton et al., 2006a; Dally et al., 2005a,b; de Kort et al., 2005). 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 ("3 days ago, I hid meal worms there, there and there, and 5 days ago, I hid peanuts there, there and there") and declarative memories ("meal worms rot in 2 days; peanuts do not 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 nonverbal 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 nonverbal ancestors had no high-level symbolic representation of the world prior to the emergence of language.

1.08.3 Implications for Neurobiology

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, as is evident from the following quotation (Koch, 1997, p. 210):

> And what of memory? It is everywhere (but can't be randomly accessed). It resides in the concentration of free calcium in dendrites and the cell body; in the presynaptic terminal; in the density and exact voltage-dependency of the various ionic conductances; and in the density and configuration of specific proteins in the postsynaptic terminals.—p. 210

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 information forward in time in computationally accessible form.

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 and King, 2010). 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 or other forms of sustained firing is the mechanism for storing previously computed information in recurrent neural net models, for example, in moving-activity-bump models for dead reckoning (Song and Wang, 2005; Conklin and Eliasmith, 2005; Stringer et al., 2002; Samsonovich and McNaughton, 1997; Bernardet et al., 2008; Seelig and Jayaraman, 2015).

Another constraint on the physical realization of the symbolic memory is that it must be capable of encoding information 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 postsynaptic 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. The function of an association has always been to conduct excitation from one idea or node or neuron to another. The strength of that conductance has rarely if ever been posited to refer to some quantity in the experienced world. This is just as true in a contemporary deep learning network (LeCun et al., 2015) as it was in the philosophy of John Locke (1690) or the psychology of Clark Hull (1929).

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 (joint function) of the presynaptic signal and the synaptic conductance. The synaptic conductance is usually called the synaptic weight, because it is a multiplicative constant that weighs or scales the presynaptic signal to determine what is seen by the postsynaptic integration process. Unless the postsynaptic mechanism has independent access to the presynaptic signal (unless it has information about that signal by a pathway other than the presynaptic pathway), it cannot estimate from the postsynaptic signal what the conductance of the synapse is. Similarly, the node in a deep net has no way to determine what the weight on a given input is, because the signal it sees is the product of the weight and the incoming signal, and it has no access to the incoming signal except through the weight. 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 (the distance) 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 postsynaptic effects are nonlinear in the inputs, the problem is still worse, because there is loss of information in nonlinear 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 nonconducting, and that information about facts such as 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 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 and Hepp (2006), given previously, 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 antirepresentational 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 (1) the distinction between a variable and the value of that variable, (2) the ability to bind a value to a variable, and, therefore, (3) 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 an 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 **Fig. 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 m). 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 **Fig. 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, neither 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).

This sounds far removed from biological reality, but, as shown in **Fig. 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. The 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).

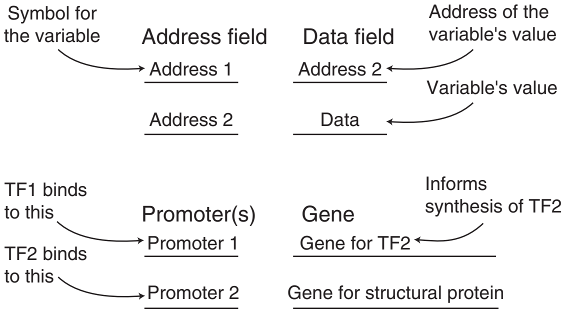

Figure 3 Indirect addressing in computers and the genome. *TF*, transcription factor. In both examples, the entry in the bottom right field ("data" or "gene for structural protein") may be yet another address (another gene for a TF). This recursive addressing logic makes possible hierarchical data structures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

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 a protein; a protein that does not 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 and read it from them as and when it becomes behaviorally relevant.

The other critical components of a machine capable of representation that is missing from our current conception of the nervous system are mechanisms 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 information conveyed by the symbols in it. (It is 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 machinery that operates 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 earlier 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 and 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 subcircuits 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 it is likely to be found at the level of molecular structure, rather than at the level of circuit structure. Implementing memory machinery and associated computational machinery at the molecular level rather than the circuit level would realize enormous savings in both space—cubic angstroms per bit stored—and in energy consumed—joules per bit computed (Gallistel, 2017; Sterling and Laughlin, 2015) as compared to the alternative of localizing memory in altered synaptic conductances and basic computing machinery in circuit-level structures dependent on high and sustained rates of neuronal firing.

Neurobiological support for this hypothesis has appeared almost simultaneously in reports from three laboratories. The Hesslow laboratory at Lund reported electrophysiological experiments with decerebrate ferrets strongly implying that the machinery that times and remembers the CS–US interval in eyeblink conditioning resides inside the cerebellar Purkinje cells (Johansson et al., 2015, 2014). The Glanzman laboratory at UCLA reported an experiment showing that conditioning of the *Aplysia* gill withdrawal persisted even when the changes in synaptic conductances heretofore assumed to mediate the conditioned response were erased (Chen et al., 2014). The Tonegawa laboratory at MIT reported that fear memory in mice engram cells persisted even when the synaptic alterations were erased (Ryan et al., 2015). These findings open a new chapter in the search for the engram, that is, the physical realization of representations in the brain. They suggest that we have been searching in the wrong place. We have been looking for enduring changes in synaptic conductances, which have never been shown to be capable of encoding simple facts such as the duration of an interval or the distance to a food source. Perhaps we should have been looking for molecular machinery inside neurons, where changes in the configurations of individual memory molecules may encode the numbers that represent our experience, that is, the symbols that refer to distances, durations, directions, rates, and the other basic quantities that constitute our experience of the world and that enter into the computations that enable us to interact effectively with that world.

See also: 1.16 Spatial Memory in Food-Hoarding Animals.

References

Aime, C., Miao, C., Jezek, K., Treves, A., Moser, E.I., Moser, M.-B., 2014. Place cells in the hippocampus: eleven maps for eleven rooms. Proc. Natl. Acad. Sci. U.S.A. 111 (52), 18428–18435.

Balsam, P.D., Drew, M.R., Gallistel, C.R., 2010. Time and associative learning. Comp. Cogn. Behav. Rev. 5, 1–22.

Bennett, M.R., Hacker, P.M.S., 2008. History of Cognitive Neuroscience. Wiley-Blackwell, Oxford.

Bermudez, M., Schultz, W., 2012. Sensitivity to temporal reward structure in amygdala neurons. Curr. Biol. 22, 1839–1844.

Bernardet, U., Bermúdez i Badia, S., Verschure, P.F., 2008. A model for the neuronal substrate of dead reckoning and memory in arthropods: a comparative computational and behavioral study. Theory Biosci. 127 (2), 163–175.

Bertolucci, C., Foa, A., 2004. Extraocular photoreception and circadian entrainment in nonmammalian vertebrates. Chronobiol. Int. 21 (4–5), 501–519.

Bjerknes, T.-L., Langston, R.-F., Kruge, I.-U., Moser, E.-I., Moser, M.B., 2015. Coherence among head direction cells before eye opening in rat pups. Curr. Biol. 25, 103–108.

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*. J. Exp. Biol. 203, 2341–2348.

Buzsaki, G., Moser, E.I., 2013. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. Nat. Neurosci. 16, 130–138.

Cartwright, B.A., Collett, T.S., 1979. How honey bees know their distance from a nearby visual landmark. J. Exp. Biol. 82, 367–372.

Cartwright, B.A., Collett, T.S., 1983. Landmark learning in bees: experiments and models. J. Comp. Physiol. A Sens. Neural Behav. Physiol. 151, 521–543.

Cartwright, B.A., Collett, T.S., 1987. Landmark maps for honey bees. Biol. Cybern. 57, 85–93.

Cheeseman, J.F., Winnebeck, E.C., Millar, C.D., Kirland, L.S., Sleigh, J., Goodwin, M., Warman, G.R., 2012. General anesthesia alters time perception by phase shifting the circadian clock. Proc. Natl. Acad. Sci. U.S.A. 109, 7061–7066.

Cheeseman, J.F., Millar, C.D., Greggers, U., Lehmann, K., Pawley, M.D.M., Gallistel, C.R., Menzel, R., 2014. Way-finding in displaced clock-shifted bees proves bees use a cognitive map. Proc. Natl. Acad. Sci. U.S.A. 111 (24), 8949–8954.

Chen, S., Cai, D., Pearce, K., Sun, P.Y.-W., Roberts, A.C., Glanzman, D.L., 2014. Reinstatement of long-term memory following erasure of its behavioral and synaptic expression in *Aplysia*. eLife. http://dx.doi.org/10.7554/eLife.03896.

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*). J. Exp. Psychol. Anim. Behav. Process. 25 (1), 82–91.

Clayton, N.S., Bussey, T.J., Dickinson, A., 2003. Can animals recall the past and plan for the future? Nat. Rev. Neurosci. 4, 685–691.

Clayton, N.S., Yu, K.S., Dickinson, A., 2003. Interacting cache memories: evidence of flexible memory use by scrub jays. J. Exp. Psychol. Anim. Behav. Process. 29, 14–22.

Clayton, N., Emery, N., Dickinson, A., 2006. The rationality of animal memory: complex caching strategies of western scrub jays. In: Nuuds, M., Hurley, S. (Eds.), Rational Animals? Oxford University Press, Oxford, pp. 197–216.

Clayton, N.S., Emery, N.J., Dickinson, A., 2006. The prospective cognition of food caching and recovery by western scrub-jays (Aphelocoma californica). Comp. Cogn. Behav. Rev. 1 (1), 1–11.

Collett, T.S., Collett, M., 2002. Memory use in insect visual navigation. Nat. Rev. Neurosci. 3, 542–552.

Collett, T.S., Graham, P., 2004. Animal navigation: path integration, visual landmarks and cognitive maps. Curr. Biol. 14 (12), R475–R477.

Collett, T.S., Cartwright, B.A., Smith, B.A., 1986. Landmark learning and visuo-spatial memories in gerbils. J. Comp. Physiol. A Sens. Neural Behav. Physiol. 158, 835–851.

Collett, M., Collett, T.S., Wehner, R., 1999. Calibration of vector navigation in desert ants. Curr. Biol. 9 (18), 1031–1034.

Collett, M., 2010. How desert ants use a visual landmark for guidance along a habitual route. Proc. Natl. Acad. Sci. U.S.A. 107 (25), 11638–11643.

Conklin, J., Eliasmith, C., 2005. A controlled attractor network model of path integration in the rat. J. Comput. Neurosci. 18, 183–203.

Dale, K., Harland, D.P., Manning-Jones, A., Collett, T.S., 2005. Weak and strong priming cues in bumblebee contextual learning. J. Exp. Biol. 208 (Pt 1), 65–74.

Dally, J.M., Emery, N.J., Clayton, N.S., 2005. Cache protection strategies by western scrub-jays, *Aphelocoma californica*: implications for social cognition. Anim. Behav. 70 (6), 1251–1263.

Dally, J.M., Emery, N.J., Clayton, N.S., 2005. The social suppression of caching in western scrub-jays (*Aphelocoma californica*). Behaviour 142 (7), 961–977.

Derdikman, D., Moser, E.I., 2010. A manifold of spatial maps in the brain. Trends Cogn. Sci. 14, 561–569.

Dickinson, J., Dyer, F., 1996. How insects learn about the sun's course: alternative modeling approaches. In: Wilson, S., et al. (Eds.), From Animals to Animats, vol. 4. MIT Press, Cambridge, MA, pp. 193–203.

Dickinson, A., 2012. Associative learning and animal cognition. Philos. Trans. R. Soc. B Biol. Sci. 367 (1603), 2733–2742.

Durier, V., Graham, P., Collett, T.S., 2003. Snapshot memories and landmark guidance in wood ants. Curr. Biol. 13 (18), 1614–1618.

Dyer, F.C., Dickinson, J.A., 1994. Development of sun compensation by honeybees: how partially experienced bees estimate the sun's course. Proc. Natl. Acad. Sci. U.S.A. 91, 4471–4474.

Dyer, F., Dickinson, J.A., 1996. Sun-compass learning in insects: representation in a simple mind. Curr. Dir. Psychol. Sci. 5 (3), 67–72.

Dyer, F.C., 2002. The biology of the dance language. Annu. Rev. Entomol. 47, 917–949.

Eichenbaum, H., 2013. The hippocampus, time and memory across scales. J. Exp. Psychol. Gen. 142 (4), 1211–1230.

Eilam, D., 2014. Of mice and men: building blocks in cognitive mapping. Neurosci. Biobehav. Rev. 47, 393–409.

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. Anim. Cogn. 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. J. Exp. Biol. 207, 1491–1508.

Fanselow, M.S., 1993. Associations and memories: the role of NMDA receptors and long-term potentiation. Curr. Dir. Psychol. Sci. 2 (5), 152–156.

Fetsch, C.R., Kiani, R., Newsome, W.T., Shadlen, M.N., 2014. Effects of cortical microstimulation on confidence in a perceptual decision. Neuron 83 (4), 797–804.

Finnerty, G., Shadlen, M., Jazayeri, M., Nobre, A., Buonomano, D., 2015. Time in cortical circuits. J. Neurosci. 35, 13912–13916.

Fodor, J., Pylyshyn, Z., 1988. Connectionism and cognitive architecture: a critical analysis. Cognition 28, 3–71.

Foster, R.G., Bellingham, J., 2004. Inner retinal photoreceptors (IRPs) in mammals and teleost fish. Photochem. Photobiol. Sci. 3, 617–627.

von Frisch, K., 1967. The Dance-Language and Orientation of Bees. Harvard Univ. Press, Cambridge, MA.

Fukushi, T., Wehner, R., 2004. Navigation in wood ants *Formica japonica*: context dependent use of landmarks. J. Exp. Biol. 207, 3431–3439.

Gallistel, C.R., King, A.P., 2010. Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience. Wiley/Blackwell, New York.

Gallistel, C.R., 1990. The Organization of Learning. Bradford Books/MIT Press, Cambridge, MA.

Gallistel, C.R., 1998. Symbolic processes in the brain: the case of insect navigation. In: Scarborough, D., Sternberg, S. (Eds.), Conceptual and Methodological Foundations. Invitation to Cognitive Science, second ed., vol. 4. MIT Press, Cambridge, MA, pp. 1–51.

Gallistel, C.R., 1999. Coordinate transformations in the genesis of directed action. In: Bly, B.M., Rumelhart, D.E. (Eds.), Cognitive Science (Handbook of Perception and Cognition), second ed., pp. 1–42.

Gallistel, C.R., 2017. The neurobiological bases for the computational theory of mind. In: Gleitman, L., de Almeida, R.G. (Eds.), Perspectives on the Work of Jerry Fodor. Oxford University Press, New York.

Giocomo, L.M., Stensola, T., Bonnevie, T., Van-áCauter, T., Moser, M.B., Moser, E.I., 2014. Topography of head direction cells in medial entorhinal cortex. Curr. Biol. 24 (3), 252–262.

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.

Gould, J., Gould, C.G., 1988. The Honey Bee. Freeman, New York.

Gould, J.L., Gould, C.G., 2012. Nature's Compass: The Mystery of Animal Navigation. Princeton University Press, Princeton, NJ.

Grah, G., Wehner, R., Ronacher, B., 2005. Path integration in a three-dimensional maze: ground distance estimation keeps desert ants *Cataglyphis fortis* on course. J. Exp. Biol. 208, 4005–4011.

Graham, P., Durier, V., Collett, T.S., 2004. The binding and recall of snapshot memories in wood ants (*Formica rufa* L.). J. Exp. Biol. 207 (Pt 3), 393–398.

Grünwald, P., 2007. The Minimum Description Length Principle. MIT Press, Cambridge, MA.

Halder, G., Callaerts, P., 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? Psychol. Rev. 91, 375–391.

Hay, J.B., Baayen, R.H., 2005. Shifting paradigms: gradient structure in morphology. Trends Cogn. Sci. 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, 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., 1929. A functional interpretation of the conditioned reflex. Psychol. Rev. 36, 498–511.

Hull, C.L., 1952. A Behavior System. Yale University Press, New Haven, CT.

Jeffery, K.J., 2014. Spatial mapping: graded precision of entorhinal head direction cells. Curr. Biol. 24, R113–R114.

Jin, N., Landgraf, T., Klein, S., Menzel, R., 2014. Walking bumblebees memorize panorama and local cues in a laboratory test of navigation. Anim. Behav. 97, 13–23.

Johansson, F., Jirenhed, D.-A., Rasmussen, A., Zucc, R., Hesslow, G., 2014. Memory trace and timing mechanism localized to cerebellar Purkinje cells. Proc. Natl. Acad. Sci. U.S.A. 111, 14930–14934. http://dx.doi.org/10.1073/pnas.14153711.

Johansson, F., Carlsson, H.A.E., Rasmussen, A., Yeo, C.H., Hesslow, G., 2015. Activation of a temporal memory in Purkinje cells by the mGluR7 receptor. Cell Rep. http://dx.doi.org/10.1016/j.celrep.2015.10.047.

Jones, M.N., et al., 2015. Models of semantic memory. In: Busemeyer, J.R., Wang, Z., Eidels, A., Townsend, J.T. (Eds.), Oxford Handbook of Computational and Mathematical Psychology. Oxford University Press, New York, pp. 232–254.

Kemfort, J., Towne, W., 2013. Honeybees can learn the relationship between the solar ephemeris and a newly experienced landscape: a confirmation. J. Exp. Biol. 216, 3767–3771.

Koch, C., Hepp, K., 2006. Quantum mechanics in the brain. Nature 440, 611–612.

Koch, C., 1997. Computation and the single neuron. Nature 385, 207–210.

Kohler, M., Wehner, R., 2005. Idiosyncratic route-based memories in desert ants, *Melophorus bagoti*: how do they interact with path-integration vectors? Neurobiol. Learn. Mem. 83 (1), 1–12.

Konorski, J., 1967. The Integrative Activity of the Brain. Chicago University Press.

de Kort, S.R., Dickinson, A., Clayton, N.S., 2005. Retrospective cognition by food-caching western scrub-jays. Learn. Motiv. 36 (2), 159–176.

Kozma, R., Freeman, W.G., 2015. Cognitive Phase Transitions in the Cerebral Cortex – Enhancing the Neuron Doctrine by Modeling Neural Fields. Springer, New York.

Kraus, B., Brandon, M., Robinson 2nd, R., Connerney, M., Hasselmo, M., Eichenbaum, H., 2015. During running in place, grid cells integrate elapsed time and distance run. Neuron 88, 578–589.

Krupic, J., Bauza, M., Burton, S., Barry, C., O'Keefe, J., 2015. Grid cell symmetry is shaped by environmental geometry. Nature 518, 232–235.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

Lindley, D., 2004. Degrees Kelvin: A Tale of Genius, Invention and Tragedy. Joseph Henry Press (National Academies Press).

Locke, J., 1690. An Essay Concerning Human Understanding.

Marr, D., 1982. Vision. W. H. Freeman, San Francisco.

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'. Nat. Rev. Neurosci. 7, 663–678.

Menzel, R., Eckoldt, M., 2016. Die Intelligenz der Bienen. Knaus, Munich.

Menzel, R., Greggers, R., 2015. The memory structure of navigation in honeybees. J. Comp. Physiol. A 201 (6), 547–561.

Menzel, R., Greggers, U., Smith, A., Berger, S., Brandt, R., Brunke, S., Watz, S., 2005. Honey bees navigate according to a map-like spatial memory. Proc. Natl. Acad. Sci. U.S.A. 102 (8), 3040–3045.

Menzel, R., Kirbach, A., Hass, W.-D., Fischer, B., Fuchs, J., Koblofsky, M., Greggers, U., 2011. A common frame of reference for learned and communicated vectors in honeybee navigation. Curr. Biol. 21, 645–650.

Menzel, R., Lehmann, K., Manz, G., Fuchs, J., Kobolofsky, M., Greggers, U., 2012. Vector integration and novel shortcutting in honeybee navigation. Apidologie 43 (5), 229–243.

Mouritsen, H., Frost, B.J., 2002. Virtual migration in tethered flying monarch butterflies reveals their orientation mechanisms. Proc. Natl. Acad. Sci. U.S.A. 99 (15), 10162–10166.

Moustafa, A.A., Myers, C.E., Gluck, M.A., 2009. A neurocomputational model of classical conditioning phenomena: a putative role for the hippocampal region in associative learning. Brain Res. 1276, 180–195.

Newell, A., Simon, H.A., 1975. Computer science as empirical inquiry: symbols and search. Commun. Assoc. Comput. Mach. 19, 113–136.

Pavlov, I.V., 1928. Lectures on Conditioned Reflexes: The Higher Nervous Activity of Animals (H. Gantt, Trans.). Lawrence & Wishart, London.

Peyrache, A., Lacroix, M.M., Petersen, P.C., Buzsaki, G., 2015. Internally organized mechanisms of the head direction sense. Nat. Neurosci. 18, 569–575.

Proctor, R.W., Capaldi, E.J. (Eds.), 2012. Psychology of Science: Implicit and Explicit Processes. Oxford University Press, Oxford.

Rich, P.D., Liaw, H.-P., Lee, A.K., 2014. Large environments reveal the statistical structure governing hippocampal representations. Science 345, 814–817.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., Bialek, W., 1997. Spikes: Exploring the Neural Code. MIT Press, Cambridge, MA.

Rubin, A., Yartsev, M.M., Ulanovsky, N., 2014. Encoding of head direction by hippocampal place cells in bats. J. Neurosci. 34, 1067–1080.

Rumelhart, D.E., McClelland, J.L. (Eds.), 1986. Parallel Distributed Processing. MIT Press, Cambridge, MA.

Ryan, T., Roy, D., Pignatelle, M., Arons, A., Tonegawa, S., 2015. Engram cells retain memory under retrograde amnesia. Science 348, 1007–1013.

Samsonovich, A., McNaughton, B.L., 1997. Path integration and cognitive mapping in a continuous attractor neural network model. J. Neurosci. 17, 5900–5920.

Sauman, I., et al., 2005. Connecting the navigational clock to sun compass input in monarch butterfly brain. Neuron 46 (3), 457–467.

Schmidt-Koenig, K., 1960. The sun azimuth compass: one factor in the orientation of homing pigeons. Science 131, 826–827.

Schultheiss, P., Cheng, K., 2011. Finding the nest: inbound searching behaviour in the Australian desert ant, *Melophorus bagoti*. Anim. Behav. 81, 1031–1038.

Seelig, J.D., Jayaraman, V., 2015. Neural dynamics for landmark orientation and angular path integration. Nature 521, 186–191.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (379–423), 623–656.

Smolensky, P., 1986. Information processing in dynamical systems: foundations of harmony theory. In: Rumelhart, D.E., McClelland, J.L. (Eds.), Parallel Distributed Processing: Foundations, vol. 1. MIT Press, Cambridge, MA, pp. 194–281.

Song, J., McPeek, R., 2015. Neural correlates of target selection for reaching movements in superior colliculus. J. Neurophysiol. 113 (5), 1414–1422.

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. J. Neurosci. 25 (4), 1002–1014.

Stensola, H., Stensola, T., Solstad, T., Frøland, K., Moser, M.-B., Moser, E.I., 2012. The entorhinal grid map is discretized. Nature 492, 72–78.

Sterling, P., Laughlin, S.B., 2015. Principles of Neural Design. MIT Press, Cambridge, MA.

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.

Tan, H.-M., Bassett, J.-P., O'Keefe, J., Cacucci, F., Wills, T.-J., 2015. The development of the head direction system before eye opening in the rat. Curr. Biol. 25, 479–483.

Tautz, J., Zhang, S.W., Spaethe, J., Brockmann, A., Si, A., Srinivasan, M.V., 2004. Honeybee odometry: performance in varying natural terrain. PLoS Biol. 2, 915–923.

Thompson, W., 1862. On the secular cooling of the earth. Proc. R. Acad. Edinburgh 5, 157–169.

Towne, W., 2008. Honeybees can learn the relationship between the solar ephemeris and a newly-experienced landscape. J. Exp. Biol. 211, 3737–3743.

Turing, A.M., 1936. On computable numbers, with an application to the Entshceidungsproblem. Proc. Lond. Math. Soc. s2-42, 230–265.

Turing, A., 1950. Computing machinery and intelligence. Mind 59, 433–460.

Vander Wall, S.B., 1990. Food Hoarding in Animals. University of Chicago Press, Chicago.

Wehner, J., Srinivasan, M.V., 2003. Path integration in insects. In: Jeffery, K.J. (Ed.), The Neurobiology of Spatial Behaviour. Oxford University Press, Oxford, pp. 9–30.

Wilkes, J.T., Gallistel, C.R., 2017. Information theory, memory, prediction, and timing in associative learning. In: Moustafa, A. (Ed.), Computational Models of Brain and Behavior. Wiley/Blackwell, New York.

Wills, T.J., Cacucci, F., Burgess, N., O'Keefe, J., 2010. Development of the hippocampal cognitive map in preweanling rats. Science 328 (5985), 1573–1576.

Wray, M.K., et al., 2008. Honeybees do not reject dances for 'implausible' locations: reconsidering the evidence for cognitive maps in insects. Anim. Behav. 76, 261–269.

Wystrach, A., Schwarz, S., Schultheiss, P., Beugnon, G., Cheng, K., 2011. Views, landmarks, and routes: how do desert ants negotiate an obstacle course? J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. 197, 167–179.

Wystrach, A., Philippides, A., Aurejac, A., Cheng, K., Graham, P., 2014. Visual scanning behaviours and their role in the navigation of the Australian desert ant *Melophorus bagoti*. J. Comp. Physiol. A 200, 615–626.

Yasuda, M., Hikosaka, O., 2015. Functional territories in primate substantia nigra pars reticulata separately signaling stable and flexible values. J. Neurophysiol. 113 (6), 1681–1696.

Zeil, J., 2011. Visual homing: an insect perspective. Curr. Opin. Neurobiol. 22, 1–9.