Trends in Cognitive Sciences

Opinion

The Coding Question

C.R. Gallistel[1],*

Recent electrophysiological results imply that the duration of the stimulus onset asynchrony in eyeblink conditioning is encoded by a mechanism intrinsic to the cerebellar Purkinje cell. This raises the general question – how is quantitative information (durations, distances, rates, probabilities, amounts, etc.) transmitted by spike trains and encoded into engrams? The usual assumption is that information is transmitted by firing rates. However, rate codes are energetically inefficient and computationally awkward. A combinatorial code is more plausible. If the engram consists of altered synaptic conductances (the usual assumption), then we must ask how numbers may be written to synapses. It is much easier to formulate a coding hypothesis if the engram is realized by a cell-intrinsic molecular mechanism.

Trends

Recent results suggest that some simple memories are not stored in plastic synapses.

The interval between the conditioned stimulus and the unconditioned stimulus in a simple form of Pavlovian conditioning appears to be stored in a mechanism intrinsic to the cerebellar Purkinje cell.

How spike trains encode numbers – the symbols for experienced quantities such as duration, distance, rate, probability, size, and numerosity – is among the most important questions in neuroscience.

No code has been proposed for use with plastic synapses.

When it comes to transmission and storage, information may be treated as a string of numbers irrespective of what they refer to (Shannon's insight).

Good codes are composed of a few elements (e.g., the 0 and 1 of a binary code) that convey different messages when arranged in different orders.

The polynucleotide code (DNA and RNA) is an example of a good code; it compactly and lastingly stores information and makes it accessible to computation.

A Cell-Intrinsic Quantity Memory

Recent electrophysiological results imply that individual **cerebellar Purkinje cells** (see Glossary) time and remember the interval between the onset of an artificial **conditioned stimulus** (CS) and the onset of an artificial **unconditioned stimulus** (US) [1] (Figure 1). They imply that the timing mechanism is inside the Purkinje cell, as is the **engram** that encodes the duration of the CS–US interval. They suggest that the engram lies somewhere within an intracellular signaling cascade whose beginning and end have been tentatively identified. Somewhere within that cascade we may find the mechanism that reads the engram.

At least in eyeblink conditioning, memory appears not to be materially realized by altered synaptic functional connectivity. This runs contrary to the general assumption about the material basis of memory [2]. Recent results from other labs suggest the same conclusion [3–6]: memory does not reside in altered **synaptic conductances**. Altered synaptic conductances are also known as plastic synapses, Hebbian synapses, or connection weights. At least some memories appear to reside in altered or synthesized molecular structures in the cell membrane or within the cell. I call this the cell-intrinsic memory hypothesis.

On this hypothesis, quantities extracted from experience (Box 1) are written to the hypothesized molecular engram, and synaptic inputs read what has been written to control and direct subsequent behavior. In this conceptual framework, we have not understood the engram until we have deciphered the code by which it represents behaviorally important quantities. I single out the Purkinje cell results because of the strength and clarity of the case they make for the encoding of a specified experiential quantity at a cell-intrinsic locus.

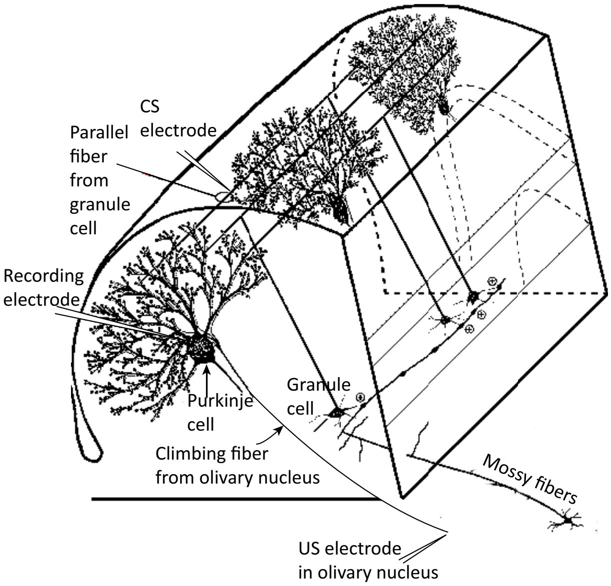
We do not yet know the cell-intrinsic molecular mechanism that encodes the duration of the CS–US interval, nor in what part of the cellular structure it resides. *A fortiori*, we do not know the physical change that encodes the interval. That an encoding of some type occurs is, however, clear, which is why we must now begin to think seriously about the coding question in neuroscience [7]. We may make inferences about the code without knowing its physical implementation. Indeed, without such inferences we will have no idea what to look for.

[1] Rutgers Center for Cognitive Science, 152 Frelinghuysen Road, Piscataway, NJ 08854-8020, USA

*Correspondence:
galliste@ruccs.rutgers.edu
(C.R. Gallistel).

http://dx.doi.org/10.1016/j.tics.2017.04.012

© 2017 Elsevier Ltd. All rights reserved.



Trends in Cognitive Sciences

Figure 1. The Preparation for Direct Pavlovian Conditioning of a Cerebellar Purkinje Cell. The experiment is conducted in the **decerebrate** ferret [1]. The conditioned stimulus (CS) electrode stimulates **parallel fibers**, which make monosynaptic connections onto the Purkinje cell. At a fixed interval after the onset of this input, the unconditioned stimulus (US) electrode directly stimulates the **climbing fiber**. The firing of the Purkinje cell is monitored by the recording electrode. It has a high endogenous firing rate (80–100 Hz). On the first several conditioning trials, the CS stimulation produces a modest increase in this firing rate or, sometimes, no noticeable change. After 30 to 200 or more conditioning trials, the CS stimulation begins to elicit a profound pause in the firing of the Purkinje cell. The pause lasts only as long as the CS–US interval, regardless of the duration and other parameters of the CS stimulation, and whether or not the US stimulation is delivered. The pause is not blocked by drugs that block the potent inhibitory effect of off-beam stimulation [1], ruling out inhibitory network inputs as the cause. The pause controls the latency of the conditioned eyeblink [50–52]. It is blocked only by agents that block the **mGlu7** receptor. The binding of **glutamate** to this **metabotropic** receptor inhibits the spontaneous firing of the cell by way of an intracellular signal cascade that terminates at the **Kir3 potassium channel** [53].

In thinking about the coding question in neuroscience, it helps to keep in mind the code by which DNA represents the linear structure of a protein and how genetically coded transcription factors specify abstract aspects of organic structure [8]. We should aspire to be able to read the engram code and infer from it properties of the animal's past experience as readily as molecular geneticists read the molecular genetic code and infer from it properties of the built animal.

Signals and Symbols

Signals carry information from place to place. Symbols, by contrast, carry information forward in time. Symbols are the information-carrying elements in a memory. The essential feature of a signal is its transience; the essential feature of a symbol is its relative permanence. In both cases, however, their function is to convey information. Through memory, the past communicates with the future.

Within cells, signals are physically realized by numerous and complex biochemical cascades. Between cells, signals are physically realized by action potential sequences (**spike** trains), neurotransmitter releases, and hormone releases.

The DNA molecule is the only well-understood biological symbol repository. The symbols in it – the coding genes – convey hereditary information from generation to generation. Within a generation, the information in DNA directs the construction of the organism. It also governs

Glossary

Address-addressable: a memory system in which stored messages may be retrieved by probe signals that specify the location (address) of the memory without specifying anything about the content of the message. Computer memory is, with few exceptions, address-addressable. The memory in auto-associative neural networks is, by contrast, content-addressable. The probe that retrieves the whole message must specify a portion of its content.

Bit: a unit of information – the amount of information conveyed by the answer to a well-posed yes/no question; hence, the amount of information encoded in the position of a two-position switch, or by specifying one element of a two-element signal (e.g., the 1 or the 0 in the binary code, or one member of a pair of complementary nucleotides).

Cerebellum: a large cortex-like layered structure that sits atop the midbrain. It contains 80–95% of all the neurons in a vertebrate brain. The Purkinje cells are located in the top layer of the cerebellar cortex.

Climbing fiber: axons that originate from cells in the olivary nuclei deep beneath the cortical layers in the cerebellum. There is only one climbing fiber for each Purkinje cell, but it embraces the cell body, making numerous synaptic contacts. Short double spike bursts in a climbing fiber produce complex (multi-humped) action potentials in the Purkinje cell.

Codes for numbers: different ways of recording a quantity such that it may be referred to in the indefinite future or transmitted to some other location: '||||||||||' is a unary code for ten of something (e.g., dollars, bushels of wheat, seconds, meters, etc.); 10 is the decimal code for that quantity; 1010 is the binary code. Any machine that manipulates quantities computationally must have a code for number.

Conditioned stimulus (CS): a sensory input that does not produce a conditioned response in a naive (untrained, unconditioned) subject but does elicit that response once the subject has been conditioned, that is, taught that that stimulus predicts the unconditioned stimulus.

Decerebrate: an animal in which the cerebrum – the part of the brain in

Box 1. Numbers in the Brain

In the brain the symbols for quantities must represent:

(i) Discrete quantities (numerosities [55–58]).
(ii) Continuous quantities (e.g., durations [1,59] and distances [60,61]).
(iii) Bernoulli probabilities (proportions between numerosities) [62–64] and relative likelihoods (the probabilities of the observed events given various hypotheses about what generated them) [57,65,66].
(iv) Rates (numerosities divided by durations) [65,67].
(v) Directed (that is, signed) quantities (as in dead-reckoning [68] and the **vectors** that compose a cognitive map [12,69,70]).
(vi) Large quantities: bees navigate over mapped environments measured in kilometers ([69], pp. 246ff); the Arctic tern makes annual round trips between the Arctic and Antarctica, covering as much as 80 000 kilometers in a year [71].
(vii) Small quantities, such as the already mentioned probabilities.
(viii) The values of physical variables such as inertia and compliance that are essential to the brain's representation of the physics of our body and the objects it must manipulate, a representation upon which even simple behaviors such as grasping a glass depend [72].
(ix) Views (snapshots, patterned photon catches) [68,73].

much of the daily functioning of an organism. The circadian clock, for example, is an intracellular biochemical cycle with a genetically specified period of approximately 24 h. The phase of this cycle (its position in time) is adjusted by the light–dark transitions at dawn and dusk. These adjustments reset the clock each day to keep it in phase with the local day–night cycle. This experience-settable intracellular molecular clock plays a fundamental role in many aspects of daily physiology and behavior [9].

A symbol must be read for the information it contains to enter into the neural computations that underlie behavior. Reading a symbol transcribes the information in it from the memory code to a signal code. We see this not only in computers but also in molecular biology – where the reading of a gene is called transcription because the information it carries is translated from a memory code into a signal code.

Whether the conveyance of information is from place to place, as in signals, or from earlier points in time to later points in time, as in symbols, is irrelevant to many aspects of the coding question.

Why Number Coding Is Fundamental

Remembered quantities, such as CS–US intervals and food locations, are essential to the execution of behaviors informed by information extracted from previous experience [10,11] (Box 1 and Figure 2). On first consideration, however, the Purkinje cell example is not representative of the general case because the Purkinje cell directly experiences the CS–US interval. The other memory-resident quantities in Box 1 cannot be directly experienced by individual neurons. There is no way to produce a presynaptic input to a neuron that is 100 m distant from another presynaptic input to that neuron. We know that the brain encodes distances not only from behavioral results (Figure 2) but also from neurobiological results: boundary-vector cells in the hippocampus signal the distance and direction of navigational boundaries (walls, cliffs) from the current location of the animal, even when the boundaries are not perceptible from that location [12].

Experienced quantities such as duration and distance are naturally represented by numbers because numbers are symbols for quantities. The beauty of a number is that it can refer to any quantity. Therefore, you can transmit any quantity simply by transmitting a number. When it comes to transmitting information, it does not matter what the information-carrying signals and symbols refer to. In an opening paragraph of the paper that created **information theory** [13], Shannon wrote:

front of the midbrain – has been severed from the midbrain, thereby putting all the neural tissue in the cerebrum out of play.

Engram: the physical substrate for a memory in a brain.

Glutamate: the neurotransmitter released onto postsynaptic Purkinje cell dendrite sites from the presynaptic endings of parallel fibers by spikes arriving at those endings. It binds to several different receptor molecules embedded in the postsynaptic membrane, one of which is the little-known mGlu7 metabotropic receptor molecule.

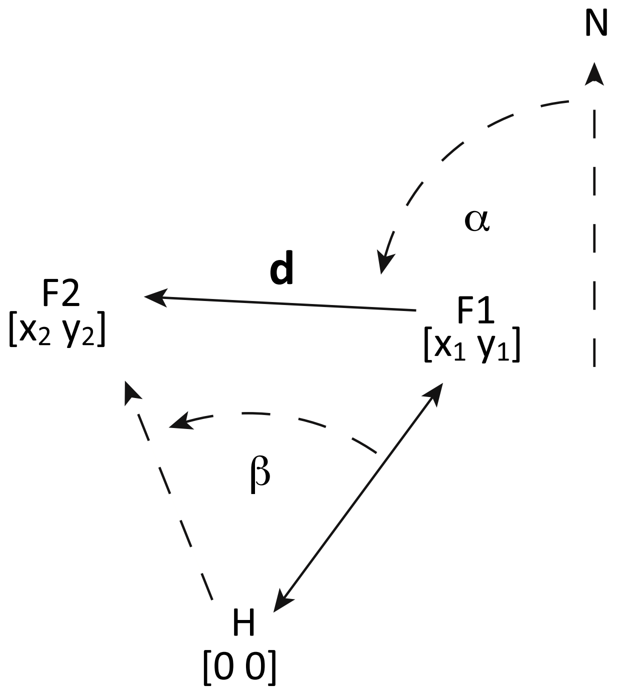
Information theory: the body of mathematical results underlying modern information technology, and the design of machines that communicate, store, and process information. It was created in 1948 by Claude Shannon, an engineer at Bell Laboratories. It also plays a fundamental role in modern physics, particularly thermodynamics. It provides a way to measure information, thereby taking information out of the realm of intuition and into the realm of measurable and physically important quantities.

Isomerization: a switch-like change in the form of a molecule, hence a possible molecular-level mechanism for information storage.

Kir3 potassium channel: a complex molecular assemblage embedded in the membrane of the Purkinje cell that behaves as a one-way gate: when the gate is opened it allows potassium ions to flow inward more easily than to flow outward. The inward flow of potassium ions hyperpolarizes the membrane, suppressing spike formation.

Metabotropic: a metabotropic postsynaptic receptor is a molecule embedded in the postsynaptic membrane that sets in motion an intracellular signaling cascade when the transmitter released by a presynaptic signal binds to it. In other words, it transcribes from an extracellular signal to an intracellular signal.

mGlu7: the name of a complex molecular assemblage, called a metabotropic postsynaptic receptor, which is embedded in the postsynaptic membranes of the synapses that parallel fibers make on Purkinje cell dendrites. The binding of glutamate to this receptor triggers an



Trends in Cognitive Sciences

Figure 2. A Behavioral Manifestation of the Computational Capacity of the Brain [54]. Bees fly back and forth between the hive (H), at the location represented by the vector [00], and a nectar source at F1 (represented by $[x_1\ y_1]$). In the hive, they observe the dances of other bees foraging at F2 (represented by $[x_2\ y_2]$). When they subsequently find no nectar at F1, they fly the vector **d** from F1 to F2, but only if β is less than a critical value. N = north; α = the compass bearing of F2 from F1. If there is a way to compute the range-and-bearing vector, **d,** from the F1 and F2 location vectors without implementing addition, subtraction, and multiplication, then someone could make an important contribution to theoretical neuroscience by explaining how to do that. Absent an alternative, this example and many like it are evidence that these arithmetic operations are implemented in the brain on stored vectors, that is, on number strings stored in an engram.

> Frequently the [conveyed] messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

What Shannon understood was that all information may be treated as numerical, irrespective of what the numbers refer to. This insight simplifies and focuses the coding question in neuroscience – how does neural tissue transmit and preserve numbers?

The number-coding question is equally foundational when the brain is approached from a computational perspective. At the foundation of a computing machine is the representation of numbers. The machinery that implements the basic computational operations – addition, subtraction, and multiplication – cannot be designed (or reverse engineered) until one has decided (or determined) how the numbers themselves are to be represented.

How Are Numbers Transmitted by Spike Trains?

The question of how numbers are written to and read from engrams is not well posed until we think about how numbers are embedded in the spike trains that convey them to and from engrams. That is one reason to begin by considering how spike trains **code for numbers**. Another reason is that the coding question for spikes has been explicitly addressed in both

intracellular signaling cascade (a sequence of biochemical reactions).

Neural net model: a model (theory) of how the brain computes predicated on the assumption that computations are performed by population rate-code signals propagating through networks of neurons in which the synaptic weights are adjusted by a learning algorithm. This computational architecture does not have a read-write memory nor an address-addressable memory.

Nucleotide: one of the four bases the two pairs of which form the rungs in the DNA double helix – also a building block of RNA molecules. Nucleotides form complementary pairs: guanine pairs with cytosine and adenine pairs with thymine (or in RNA with uracil). These complementary pairs resemble the complementary elements (0 and 1) in the binary code for number.

Parallel fibers: the parallel fibers originate from the tiny granule cells that form the granular layer of the cerebellar cortex. They rise to the top layer, where they bifurcate and run parallel to the folds in the cerebellar cortex. Approximately 200 000 parallel fibers contact each Purkinje cell.

Polynucleotide: a string of nucleotides. Because there are four possible nucleotides at any one position in the string, the nucleotide found at a given position can carry 2 bits of information.

Population rate code: a rate code that avoids implausibly long reading windows by assuming that the spikes being counted arrive more or less simultaneously over a large population of presynaptic axons.

Purkinje cells: cells in the top layer of the cerebellar cortex with an enormous flat dendritic tree oriented perpendicular to the folds of the cerebellar cortex, hence to the parallel fibers.

Source statistics: the relative frequencies with which the different possible messages from a given source are conveyed; hence the probability distribution on the messages from that source. For example, for reasons that are somewhat mysterious, the small digits (1–3) are sent and stored much more often than the large digits (7–9), and irrespective of the source (what they refer to).

experimental and theoretical work [14–17]. No work known to me has explicitly considered the engram coding question.

When you listen to a spike train fed to an audio system, it sounds like a telegraphic transmission – a sequence of clicks that vary in their frequency depending on what is happening. In the telegraphic case, we know how the information is encoded into the click sequences: three interclick intervals, chosen to be easily distinguishable by a human receiver, are the code elements. The shortest interval (the dot) establishes the timescale. The other two intervals (the dash and the word space) are three and seven times this base interval. The code is combinatorial: changing the sequence in which the code elements occur – from dot→dash to dash→dot, for example – changes the message conveyed. The code is timescale-invariant because only the relative intervals between the clicks matter. In short, telegraphic information transmission uses a timescale-invariant three-element combinatorial code in which the information is carried by the intervals between the clicks.

Spike-train signaling is often assumed to use a rate code. This assumption is not well specified until one specifies the time-window within which the receiver reads the rate ([14], chapter 7) – because a rate is a count divided by a duration. In a rate code, the sequence of interspike intervals within the reading window does not matter; only their number matters. In a combinatorial code, on the other hand, changing the sequence in which two or more interspike intervals occur changes the message ([15], chapter 1) (when the counting window is reduced to approximately the average interspike interval, the distinction between a rate code and spike-timing code becomes hard to make [14–16]).

Rate codes for quantities are the temporal equivalent of ‘hash mark’ (also known as ‘tally mark’) codes. Hash marks are the oldest codes for conveying quantitative information; they appear in the Paleolithic [18]. Rate codes, similarly to hash mark codes, are unary codes: to convey a particular number one must use as many code elements as the numerosity to which the number refers. Thus, the code instantiates the number.

Code Efficiency (Bits Per Spike)

Rate codes convey less than 1 **bit** per spike. To see this, consider the contrast between the hash mark code for twenty (‘||||||||||||||||||||’; that is, 20 tally strokes), as contrasted with the digital code (20) or the binary code (10100). The latter two codes are combinatorial. In a combinatorial code there is more than one symbol element (the 10 digits in a decimal code, the 0 and 1 in a binary code) and the ordering of the elements matters. The string ‘12’ does not convey the same message as the string ‘21’.

Combinatorial codes are vastly more efficient than unary codes. If a machine must be able to store and communicate any number in a set of one million numbers, then in a unary code one of the numbers will be encoded by one million hash marks. By contrast, it takes at most six decimal digits or 20 binary digits (bits) to convey any one of one million possible numbers (possible messages). Moreover, if the coding scheme is efficient – if it accurately reflects the **source statistics** – the average message requires many fewer code elements [19–21]. Source statistics refers to the relative frequencies with which different messages from a given source are sent. An efficient combinatorial code is source-specific: it uses short strings for the frequent messages and long strings for the infrequent messages, thereby minimizing the average number of code elements used to convey the messages from that source (a symbol string is a message-bearing sequence of code elements; for example, some number of spikes in a rate code or a sequence of interspike intervals in a timing code).

Spike: action potential in an axon; the basic element of a neural signal – a current pulse that propagates rapidly along an axon.
Synaptic conductance: the scale factor that determines the magnitude and sign of the postsynaptic effect (the amount of depolarization or hyperpolarization) produced by a presynaptic spike. In most neurocomputational models it is represented by a single number, called the weight. The weight multiplies (scales) another single number called the rate.
Thermodynamic stability: a property of molecules that are not reconfigured or broken up by the jostling that goes on at the molecular level (thermal noise).
Unconditioned stimulus (US): a motivationally important stimulus that elicits an observable response in an untrained (unconditioned) subject.
Vector: an ordered set of numbers, for example the latitude and longitude of a location. Changing the order of the numbers changes the reference of the vector and the consequences of computational operations performed on it. That is, the position of a number within a vector is syntactically important.
Weber fraction: the fraction by which a perceptible quantity, for example a given weight, duration, or numerosity, must be increased to produce a detectably greater quantity. The measure of the imprecision with which the brain represents a quantity: the bigger the Weber fraction, the more imprecisely the brain represents a given dimension of experience.

Electrophysiological experiments on sensory channels construct mappings between stimuli and the spike sequences they produce. Electrophysiological experiments on motor channels attempt to specify the mapping between spike trains in these channels and the actions that follow. However, the question how much information is conveyed by these spike trains is rarely asked. When this question has been posed and answered, the spike trains have been shown to convey between three and seven bits of information per spike [14,15]. In other words, in the best-understood sensory channels, the information about quantitative variation in an environmental variable is conveyed by the interspike intervals; reordering the intervals changes the message [22]. The precise timing of the spikes has also recently been shown to determine a motor result [23].

In a sensory channel, one may know with some confidence what the messages being transmitted are. In the central nervous system the coding question is much harder because we are uncertain what the messages are. However, the results from the minority of experiments whose data analysis has been directed by information-theoretic considerations suggest the tentative assumption that information about behaviorally important quantities such as those in Box 1 is conveyed to engrams not by the rates at which spikes arrive at synapses but rather by the intervals between their arrivals. They further suggest that different interspike intervals constitute different elements in a combinatorial code, exactly as they do in a telegraphic transmission. They suggest, in other words, that the process that writes to the engram is sensitive to the intervals that separate arriving signals. I note also that the relevant interspike intervals are often much longer than those relevant to spike-dependent plasticity [24].

Rate codes are extremely costly energetically. Cortical spike conduction together with transmitter release and reuptake requires the hydrolysis of 10^8 ATP molecules per spike [25]. If the spike-train code is combinatorial, the energy cost of communicating quantities (ATPs per bit communicated) becomes exponentially smaller. Metabolic cost has been proposed as a unifying principle in neuronal biophysics [26]. The radical reductions in metabolic costs achievable by using a combinatorial code is a further motivation for the hypothesis that spike trains convey numbers to engrams by means of a combinatorial interspike-interval code.

If the information about behaviorally important quantities arrives at sites of engram formation in the form of informative variations in a sequence of interspike intervals, then the conditioning of Purkinje cells by the interval between the onset of the artificial CS spike train and the artificial US spike train may be more representative of the general case than it at first appears. Information about non-temporal quantities such as distance and direction may arrive in the same form, because what the sequence conveys is not intervals as such but rather the numbers that specify the quantities, whatever they may be.

Auto-Scaling

The symbols in memory must be able to convey information about quantities that range over many orders of magnitude (many powers of 10). Probabilities, for example, must be represented by numbers much less than 1. On the other hand, cognitive maps of familiar terrain span distances from centimeters to thousands of kilometers [27]. Shannon's 'set of possible messages' (see the third paragraph of the section 'Why Number Coding Is Fundamental') must therefore be a very large set: distinguishable numbers that may have to be transmitted and remembered number in the millions.

Computers represent quantities ranging from Angstroms (10^{-10} m) to parsecs (3.09×10^{16} m) by dividing them into two parts. One part, the exponent, specifies the scale, that is, the magnitude ballpark; the other part, the significand, specifies an interval within that ballpark. The number of bits in the significand determines the precision with which a quantity is represented.

It determines the size of the intervals that tile the ballpark. When the significand has four bits, it can specify $2^4 = 16$ different intervals within the ballpark. The bipartite representation of a number is also seen in scientific number notation, examples of which are 1.23e46 (a very big number) and 1.5e−19 (a very small number). The digits following the ‘e’ (for exponent) specify the scale (ballpark), that is, the power of 10. The number of digits between the decimal point and the ‘e’ determines the precision with which the number is encoded, hence the interval of uncertainty regarding the quantity it represents. When there is but one digit between the decimal point and ‘e’, the quantities are represented with a precision of ± 10%, that is, the values for significand divide the ballpark into 10 quantity intervals. When there are two such digits, then the precision is ±1%, and so on; the more digits, the more finely the ballpark is subdivided.

Scientific notation and the representation of quantity in conventional computers obey Weber’s law, the first quantitative law to emerge from experimental psychology. Weber’s law is that the experienced values on a quantity dimension (e.g., weight, duration, or distance) are specified in the brain to within some fixed relative precision, that is, ± some percentage. The **Weber fraction** specifies the imprecision. It varies from dimension to dimension because the brain represents some quantities more precisely than others.

This two-part coding principle (scale and significand) is seen in the single sensory neuron that signals the yaw (unwanted lateral turning) in the flight-control system of the fly [28]. The interspike intervals in this neuron convey information about the waveform of the yaw (how it varies over time) and about its scale (how big those variations are). When the air turbulence increases or decreases, thereby changing the scale of the yawing, this neuron signals the change in scale within less than 1 s. Given that we find this level of coding sophistication in a single insect neuron, it seems likely that we will also find it in the code that mammalian brains use to preserve behaviorally important quantities for future computational use, that is, in the engram code for number.

How Do Engrams Encode Number?

Part of the conceptual problem posed by the hypothesis that the engram resides in altered synaptic conductances is that synaptic machinery, as commonly conceived, is designed for the transmission and modulation of transient signals. That is not the function of a memory mechanism. The function of a memory mechanism is to carry large amounts of information forward in time in a computationally accessible form. A memory mechanism should have high **thermodynamic stability** (last a long time) and low or negligible energy costs (not drain the batteries). It should be realizable in a maximally compact volume, using as few elements as possible (store a great many bits in very little space). It should be capable of representing a huge range of quantities (store very small and very big numbers) and be addressable on a short timescale – that is, it is quickly readable on the basis of location in memory, without reference to informational content.

The Plastic Synapse Hypothesis

In the literature on the synaptic memory hypothesis, it is difficult to find discussions of how altered synaptic conductances satisfy these desiderata. One recent experiment report does, however, address the question of how many bits can be stored at a synapse. The authors report that single axons in hippocampal neuropil make ‘two or more synaptic contacts on the same dendrites, having shared histories of presynaptic and postsynaptic activity. The spine heads and neck diameters . . . were nearly identical in size’ [29]. There was a minimum of 26 distinguishable synaptic strengths (head sizes). They suggest that a synapse can therefore store $\log_2(26) = 4.7$ bits of information.

Like the cerebellar Purkinje cell finding, however, this important finding raises many unanswered questions. The first is – how could this information be read? If one assumes a rate code in the presynaptic spikes arriving at these synapses, then the postsynaptic effects of those signals will be jointly determined by the number of spikes within the reading window and the strength of the synapse. In **neural net models**, the postsynaptic machinery has no access to the number of presysnaptic spikes that arrived during the reading window; it has access only to the product of the synaptic strength (determined by the size of the spine head) and the number of spikes that arrived. The postsynaptic cell has access only to wr, where w is synaptic strength, and r is the rate. The problem is that when I tell you only that $wr = 28$, you have no way of knowing either w or r, which means you have no way of knowing the synaptic strength.

The puzzle of how the bits in a synapse could be read highlights the fact that the ‘engrams’ in the synaptic-strength theory of memory are not information-storing entities. Ever since Sherrington, synapses have been conceived of as conductive junctions between cells: they modulate signal flow but they do not encode specifiable experiential facts such as CS–US intervals. Therefore, they are not read in the information-theoretic sense of the word. The concept of a synaptic strength representable by a single number is a physiological interpretation of the associative strengths posited first by empiricist philosophers and subsequently by behaviorist psychologists. The history of the concept explains why the literature on the neurobiology of memory rarely discusses how altered synaptic strengths could meet the above-listed desiderata for engrams. Associative bonds were never conceived of as symbols, and neither are their neurobiological proxies. In neural net models, plastic synapses are molded by experience, not inscribed with readable information.

The Cell-Intrinsic Hypothesis

It is easier to imagine how engrams could be inscribed when we entertain the cell-intrinsic memory hypothesis – because we know that there are molecules inside neurons that carry large amounts of abstract information forward in time in a computationally accessible form. This knowledge can serve as a guide to our thinking about the engram – it should look like a gene to which experience can write.

The **polynucleotide** sequences in DNA carry hereditary information forward in time using a four-element code (adenine, guanine, thymine, and cytosine). DNA carries information at remarkable volumetric densities (more bits per cubic micron than in the best currently available memory chips), with negligible energetic cost (the code can be read thousands of years after the animal has died), and with an **address-addressable** logic-based reading mechanism [7,30]. This molecularly realized intracellular memory architecture creates hierarchical data structures such as those one finds in the RAM of conventional computers [7]. This architecture makes possible a gene for an eye [31] and individual genes for anatomical abstractions such as anterior, dorsal, and distal [8].

Adding one **nucleotide** to a polynucleotide string adds 2 bits of information, and it requires the hydrolysis of only one ATP – eight orders of magnitude less energy than is required to propagate one cortical spike. It is obvious how to store numbers in polynucleotide sequences using either a binary code (a code with two elements, e.g., 0 and 1) or a quaternary code (a code with four elements, e.g., 0, 1, 2, 3). Because of the above-listed properties of polynucleotide memory, some computer engineering laboratories are exploring the use of bacterial DNA as the memory medium in a conventional computer [32].

The central dogma of molecular biology is that information flows out of DNA, not into it [33]. In addition, it is commonly thought that accessing the information in DNA is slow. Therefore, nuclear DNA itself may not be a plausible medium for the storage of acquired information.

However, cells are full of extranuclear microRNAs. These short polynucleotide strings store information in the same way as DNA. They can store 1 or 2 bits per nucleotide, depending on whether one uses a binary or a quaternary number code. Indeed, the shortness of the microRNA strings commends them for this purpose. They are known to play important roles in intracellular signaling cascades [34–36], which means that they are readable.

There are many other molecular possibilities for the hypothesized cell-intrinsic engram. Methylation patterns are one possibility. Another possibility is the flipping (**isomerization**) of membrane-bound linear arrays of molecular switches such as the thermodynamically stable opsins. The isomerization of opsins by photons mediates vision. Opsins are an interesting example of the type of molecular switch one might look for as a molecular memory mechanism because, although they are thermodynamically stable, their structure can be altered on a nanosecond timescale, and those alterations inform visually guided behavior on a sub-second timescale. We know from this example that molecules whose form can be set by sensory experience can inform behavior on the same fast timescale as memory informs behavior.

One might also imagine that quantities out in the world are represented by quantities of some protein within a cell. This would be a hash-code, with the drawbacks inherent in such codes. If the engram code is combinatorial, then the elements of the code must be embedded within a linear structure of some kind. Otherwise, there would be no way to specify their order. The ordering of the code elements is essential in a combinatorial code because element order determines string meaning.

The cell-intrinsic engram hypothesis is consonant with the argument that the computations realized in neural net models by spike trains propagating through plastic synapses could be realized instead by intracellular neurochemistry, with enormous savings in the energy required [37]: in their recent book, Sterling and Laughlin write (p. 124): ‘These advantages – compactness, energy efficiency, and ability to adapt and match – all suggest the principle compute with chemistry. It is cheaper.’

The quest for the molecular realization of the CS–US engram in the Purkinje cell will need to be carried out using the latest tools in the rapidly evolving molecular biology toolbox. It is likely to take some time before a clear story emerges. The immediate impetus this quest gives to progress in computational neuroscience is to shift our focus to the coding question. When one entertains the cell-intrinsic hypothesis and focuses on memory for quantity, the coding question comes to the fore. One may hope that giving prominence to the coding question will inspire champions of the synaptic-conductance hypothesis to address it. We cannot have a computationally relevant understanding of the hypothesized synaptic engram until we have a testable hypothesis about the code by which altered synaptic conductances specify numbers.

The Cerebellum As a Prediction Machine

If the cell-intrinsic hypothesis proves to be correct for the eyeblink engram, it is unlikely that it will apply only to the small bit of the cerebellum that mediates the conditioned eye blink [38,39]. The cerebellum contains 80–90% of the neurons in the brain [40]. In humans and apes it has evolved more rapidly than the neocortex [41]. The cerebellar microzones number in the hundreds to the thousands. The neuronal circuitry consists of a small number of neuronal elements arranged in a highly stereotyped configuration (Figure 1).

The anatomy and physiology of cerebellar circuitry scream prediction. Yogi Berra is famously (if perhaps mythologically) quoted as saying: ‘It’s difficult to make predictions, particularly about the future.’ It is difficult because a great many different predictors (potential CSs) might predict any one event of interest (any one US). Only experience can distinguish the good from the bad predictors, and only experience can tell us when to expect the US given a predictive CS. These facts about prediction seem to be directly reflected in the cerebellar circuitry and in the physiology of the Purkinje cell: 200 000 CS fibers synapse on each Purkinje cell versus only one US fiber, while the Purkinje cell at the confluence of these signals times and remembers CS–US intervals, that is, prediction latencies.

Associative learning is not about the formation of associations; it is about using the remembered past to anticipate the future. The ability to anticipate the future depends fundamentally on the ability to efficiently encode the past. From the minimum description length approach to stochastic model selection [42,43] we learn that the more efficient the encoding of the past (the data already seen), the better the prediction of the future (the data not yet seen) [44–46].

These considerations about cerebellar circuitry and the role that the efficient encoding of the past plays in the prediction of the future suggest that the lessons to be learned from further pursuit of the cell-intrinsic engram hypothesis may have far-reaching consequences for our understanding of how the brain works. The information stored in memory is the foundation of computation [7], exactly as the information stored in DNA is the foundation of life. Everything rests on and derives from those foundations.

Intuitions about Neural Plausibility

Absent an answer to the coding question, it is hard to see how we can justify intuitions about what is neurobiologically plausible in a computational theory of the mind/brain. Nonetheless, plausibility intuitions play a large role in the formulation of contemporary neurocomputational theories. Most such theories take two assumptions for granted: (i) a **population rate code**, and (ii) acquired information resides in altered synaptic conductances (connection weights) [14,47,48]. A computing machine that does computations such as those that underlie the behavior portrayed in Figure 2 cannot be built on these foundations, which is why a recent proposal adds a read/write memory stack to a conventional neural net [49]. Perhaps the brain really does compute in much the same sense as, and in much the same way as, a conventional computer. Why should the conveyance of acquired information proceed by principles fundamentally different from those that govern the conveyance of heritable information?

Outstanding Questions

How are numbers encoded into spike trains? That is, how do spike trains convey quantitative facts from one location within neural tissue to a distant location within that tissue?

How many bits are transmitted per spike?

How energy-efficient is the spike-transmission process (how many joules per bit)?

What is the material realization of the memory for numbers – the engram that conveys previously learned quantitative facts forward in time for use in later computations?

What is the process or mechanism that writes to the engram the numbers encoded in incoming spike trains?

What is the process that reads (transcribes) numbers encoded in the engram into a spike train that conveys them to other neurons?

What are the primitive computational operations in neural tissue?

Are the primitive computational operations implemented primarily by circuit-level structures, as in neural net models, or primarily by intracellular neurochemistry?

References

1. Johansson, F. *et al.* (2014) Memory trace and timing mechanism localized to cerebellar Purkinje cells. *Proc. Natl. Acad. Sci. U. S. A.* 111, 14930–14934
2. Poo, M.-M. *et al.* (2016) What is memory? The present state of the engram. *BMC Biol.* 19, 40
3. Chen, S. *et al.* (2014) Reinstatement of long-term memory following erasure of its behavioral and synaptic expression in *Aplysia*. *eLife* 3, e03896
4. Kass, M.D. *et al.* (2013) Fear learning enhances neural responses to threat-predictive sensory stimuli. *Science* 342, 1389–1392
5. Ryan, T. *et al.* (2015) Engram cells retain memory under retrograde amnesia. *Science* 348, 1007–1013
6. Queenan, B.N. *et al.* On the research of time past: the hunt for the substrate of memory. *Cogn. Neurosci.* (in press)
7. Gallistel, C.R. and King, A.P. (2010) *Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience,* Wiley/Blackwell
8. Carroll, S.B. (2005) *Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom,* Norton
9. Gallistel, C.R. (1990) *The Organization of Learning,* Bradford Books/MIT Press
10. Cordes, S. *et al.* (2007) Common representations of abstract quantities. *Curr. Dir. Psychol. Sci.* 16, 156–161
11. Gallistel, C.R. (2011) Mental magnitudes. In *Space, Time and Number in the Brain: Searching for the Foundations of Mathematical Thought* (Dehaene, S. and Brannon, L., eds), pp. 3–12, Elsevier
12. Lever, C. *et al.* (2009) Boundary vector cells in the subiculum of the hippocampal formation. *J. Neurosci.* 29, 9771–9777
13. Shannon, C.E. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423 623-656
14. Gerstner, W. *et al.* (2014) *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition,* Cambridge University Press
15. Rieke, F. *et al.* (1997) *Spikes: Exploring the Neural Code,* MIT Press
16. Theunissen, F. and Miller, J.P. (1995) Temporal encoding in nervous systems: a rigorous definition. *J. Comput. Neurosci.* 2, 149–162
17. Thorpe, S. *et al.* (1996) Speed of processing in the human visual system. *Nature* 381, 550–552
18. Flegg, G. (1983) *Numbers: Their History and Meaning,* Courier Dover Publications
19. Olshausen, B.A. and Field, D.J. (2004) Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* 14, 481–487
20. Simoncelli, E.P. and Olshausen, B. (2001) Natural image statistics and neural representations. *Annu. Rev. Neurosci.* 24, 1193–1216
21. Ganguli, D. and Simoncelli, E.P. (2014) Efficient sensory encoding and Bayesian inference with hetrogeneous neural populations. *Neural Comput.* 26, 2103–2134
22. Lazar, A.A. and Zhou, Y. (2014) Reconstructing natural visual scenes from spike times. *Proc. IEEE* 102, 1500–1519

23. Srivastava, K.H. *et al.* (2016) Motor control by precisely timed spike patterns. *Proc. Natl. Acad. Sci. U. S. A.* 114, 1171–1176
24. Hesslow, G. *et al.* (2013) Classical conditioning of motor responses: what is the learning mechanism? *Neural Netw.* 47, 81–87
25. Laughlin, S.B. (2004) The implications of metabolic energy requirements for the representation of information in neurons. In *The Cognitive Neurosciences III* (Gazzaniga, M.S., ed.), pp. 187–196, MIT Press
26. Hasenstaub, A. *et al.* (2010) Metabolic cost as a unifying principle governing neuronal biophysics. *Proc. Natl. Acad. Sci. U. S. A.* 107, 12329–12334
27. Cheeseman, J.F. *et al.* (2014) Way-finding in displaced clock-shifted bees proves bees use a cognitive map. *Proc. Natl. Acad. Sci. U. S. A.* 111, 8949–8954
28. Brenner, N. *et al.* (2000) Adaptive rescaling maximizes information transmission. *Neuron* 26, 695–702
29. Bartol, T.M. *et al.* (2015) Nanoconnectomic upper bound on the variability of synaptic plasticity. *eLife* 4, e10778
30. Gallistel, C.R. (2017) The neurobiology of the language of thought. In *Minds on Language and Thought: The Status of Cognitive Science and Its Prospects* (Almeida, R.G.D. and Gleitman, L., eds), Oxford University Press (in press)
31. Halder, G. *et al.* (1998) Eyeless initiates the expression of both sine oculis and eyes absent during *Drosophila* compound eye development. *Development* 125, 2181–2191
32. Scudellari, M. (2015) Inner workings: DNA for data storage and computing. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15771–15772
33. Crick, F. (1970) Central dogma of molecular biology. *Nature* 227, 561–563
34. Alle, H. *et al.* (2009) Energy-efficient action potentials in hippocampal mossy fibers. *Science* 325, 1405–1408
35. He, L. and Hannon, G.J. (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* 5, 522–531
36. Vienberg, S. *et al.* (2016) MicroRNAs in metabolism. *Acta Physiolologica* 219, 346–361
37. Sterling, P. and Laughlin, S.B. (2015) *Principles of Neural Design,* MIT Press
38. Longley, M. and Yeo, C.H. (2014) Distribution of neural plasticity in cerebellum-dependent motor learning. *Prog. Brain Res.* 210, 79–101
39. Mostofi, A. *et al.* (2010) Electrophysiological localization of eyeblink-related microzones in rabbit cerebellar cortex. *J. Neurosci.* 30, 8920–8934
40. Herculano-Houzel, S. (2009) The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neurosci.* 3, 31
41. Barton, R.A. and Venditti, C. (2014) Rapid evolution of the cerebellum in humans and other great apes. *Curr. Biol.* 24, 2440–2444
42. Rissanen, J. (1978) Modeling by shortest data description. *Automatica* 14, 465–471
43. Rissanen, J. (1999) Hypothesis selection and testing by the MDL principle. *Comput. J.* 42, 260–269
44. Gallistel, C.R. and Wilkes, J.T. (2016) Minimum description length model selection in associative learning. *Curr. Opin. Behav. Sci.* 11, 8–13
45. Grünwald, P. (2007) *The Minimum Description Length Principle,* MIT Press
46. Wilkes, J.T. and Gallistel, C.R. (2017) Information theory, memory, prediction, and timing in associative learning. In *Computational Models of Brain and Behavior* (Moustafa, A., ed.), Wiley/Blackwell (in press)
47. Eliasmith, C. (2013) *How to Build a Brain: A Neural Architecture for Biological Cognition,* Oxford University Press
48. Joanisse, M.F. and McClelland, J.L. (2015) Connectionist perspectives on language learning, representation and processing. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 235–247
49. Graves, A. *et al.* (2016) Hybrid computing using a neural network with dynamic external memory. *Nature* 538, 471–476
50. Johansson, F. *et al.* (2016) Mechanisms for motor timing in the cerebellar cortex. *Curr. Opin. Behav. Sci.* 8, 53–59
51. Jirenhed, D.A. and Hesslow, G. (2016) Are Purkinje cell pauses drivers of classically conditioned blink responses? *Cerebellum* 15, 526–534
52. Heiney, S.A. *et al.* (2014) Precise control of movement kinematics by optogenetic inhibition of Purkinje cell activity. *J. Neurosci.* 34, 2321–2330
53. Johansson, F. *et al.* (2015) Activation of a temporal memory in Purkinje cells by the mGluR7 receptor. *Cell Rep.* 13, 1741–1746
54. Menzel, R. *et al.* (2011) A common frame of reference for learned and communicated vectors in honeybee navigation. *Curr. Biol.* 21, 645–650
55. Dehaene, S. (1997) *The Number Sense,* Oxford University Press
56. Ditz, H.M. and Nieder, A. (2016) Sensory and working memory representations of small and large numerosities in the crow endbrain. *J. Neurosci.* 36, 12044–12052
57. Berkay, D. *et al.* (2016) Probabilistic numerical discrimination in mice. *Anim. Cogn.* 19, 251–265
58. Yang, T.I. and Chiao, C.-C. (2017) Number sense and state-dependent valuation in cuttlefish. *Proc. R. Soc. Lond. Ser. B* 283, 20161379
59. Balsam, P.D. *et al.* (2010) Time and associative learning. *Comp. Cognit. Behav. Rev.* 5, 1–22
60. Howard, L.R. *et al.* (2014) The hippocampus and entorhinal cortex encode the path and euclidean distances to goals during navigation. *Curr. Biol.* 24, 1–10
61. Teichroeb, J.A. and Aguado, W.D. (2016) Foraging vervet monkeys optimize travel distance when alone but prioritize high-reward food sites when in competition. *Anim. Behav.* 115, 1–10
62. Gallistel, C.R. *et al.* (2014) The perception of probability. *Psychol. Rev.* 121, 96–123
63. Kheifets, A. and Gallistel, C.R. (2012) Mice take calculated risks. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8776–8779
64. Tosun, T. *et al.* (2016) Mice plan decision strategies based on previously learned time intervals, locations, and probabilities. *Proc. Natl. Acad. Sci. U. S. A.* 113, 787–792
65. Gallistel, C.R. *et al.* (2001) The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *J. Exp. Psychol. Anim. Behav. Process.* 27, 354–372
66. Nassar, M.R. *et al.* (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in changing environment. *J. Neurosci* 30, 12366–12378
67. Gallistel, C.R. *et al.* (2007) Is matching innate? *J. Exp. Anal. Behav.* 87, 161–199
68. Cheng, K. and Cody, F.W. (2015) Path integration, views, search, and matched filters: the contributions of Rüdiger Wehner to the study of orientation and navigation. *J. Comp. Physiol. A* 201, 517–532
69. Menzel, R. and Eckoldt, M. (2016) *Die Intelligenz der Bienen,* Knaus
70. Taube, J.S. *et al.* (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J. Neurosci.* 10, 420–435
71. Egevang, C. *et al.* (2010) Tracking of Arctic terns *Sterna paradisaea* reveals longest animal migration. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2078–2081
72. Shadmehr, R. and Mussa-Ivaldi, S. (2012) *Biological Learning and Control: How the Brain Builds Representations, Predicts Events, and Makes Decisions,* MIT Press
73. Collett, T.S. and Graham, P. (2015) Insect navigation: do honeybees learn to follow highways? *Curr. Biol.* 25, R240–R242