24 Constraints on Sensory Processing

RACHEL I. WILSON

ABSTRACT The evolution of sensory systems is driven by the need to survive and reproduce in a naturalistic sensory environment. However, neural architectures are also shaped by their own internal constraints. These constraints include the evolutionary-developmental constraints that restrict the possible effects of mutations, as well as stochastic developmental noise. Neural systems are further limited by metabolic costs, particularly the cost of maintaining ionic gradients. Finally, neural systems are limited by the properties of their own electronics—namely, the existence of intrinsic electrical noise, the limited speed of signal propagation, the restricted dynamic range of synapses and firing rates, and the quasi-linear nature of synaptic integration. Importantly, many of these internal constraints are fundamentally in conflict with each other, insofar as they pressure evolution in opposite directions. Thus, neural systems are shaped by the need to find a satisfactory trade-off between competing factors. Many such trade-offs are common to many sensory modalities, and they are shared by organisms as different as flies and humans. As such, they can help explain why some features of neural systems are also shared.

"The enemy of art is the absence of limitations."
—Orson Welles

The ability of neural systems to process sensory information is subject to internal constraints. A constraint (as defined here) is a limitation on the performance of a sensory system that arises from the intrinsic properties of the neural system itself. This is distinct from limitations arising from the sensory environment, or the nature of the behavioral task.

Any given internal constraint may not directly limit performance, because it may be possible to compensate for the constraint. However, compensation may push the nervous system toward a particular architecture, rather than other architectures that might hypothetically produce better performance in the absence of this constraint. Thus, constraints often limit performance indirectly, by pressuring neural systems toward compromise architectures. A compromise may be optimal, or it may simply be good enough to permit the organism to survive and reproduce (Marder & Goaillard, 2006; Rieke, 2014).

The constraints discussed here are common to all sensory systems. Moreover, they are common to very different organisms. As such, they can potentially explain why certain neural architectures occur again and again: they are good compromises.

The design of artificial systems is sometimes inspired by biological systems. This represents an additional motivation to seriously consider the constraints on biological systems. Because artificial systems are not subject to many of the constraints discussed here, an implementation that works well in biology may be suboptimal in an artificial system, and thus engineers must be careful not to draw the wrong lessons from biological systems (Stafford, 2010).

This chapter reviews the key constraints on neural systems and their implications for sensory processing. Organisms have evolved intriguing ways to cope with some of these constraints. Such cases are generally clearer in sensory neuroscience than in other branches of neuroscience, simply because the function of sensory systems is particularly transparent. For this reason, sensory systems are a good setting for investigating how neural systems evolve in the context of constraints.

Evolutionary inheritance

Neural systems in related species have a similar organization. This is true even in species that inhabit very different ecological niches. For example, the relative size of major brain divisions is remarkably constant across mammals. A meta-analysis of 131 species showed that the volumes of all major brain divisions (including the medulla, hippocampus, cerebellum, striatum, and neocortex) were highly systematically related to total brain volume, except the olfactory bulb. Different brain divisions showed different relationships to brain volume—in particular, neocortex grew particularly steeply with increasing volume—but for all brain divisions there was a systematic dependence on brain volume that extended to species with widely varying body sizes and lifestyles. Notably, this analysis included species as diverse as simians, prosimians, insectivores, and bats. This result argues that the expansion of brain volume resulting from natural selection for any behavioral trait is constrained to be a coordinated growth of the entire nonolfactory brain (Finlay & Darlington, 1995). This finding raises a provocative question: might there be excess signaling capacity in some brain divisions, as a by-product of a strong pressure to expand other divisions?

The Cognitive Neurosciences, 5th ed., 2014, MIT Press, edited by Michael S. Gazzaniga and George R. Mangun

Whereas the size of an entire brain division (e.g., neocortex) appears to be highly constrained, there is relatively more flexibility in the regionalization of brain divisions (e.g., the division of neocortex into sensory regions). Even so, there is evidence that regionalization is also subject to constraints. For example, all mammals share a common set of primary and secondary sensory cortical regions. This includes visual areas V1 and V2, somatosensory areas S1 and S2, and auditory area A1, as defined by cytoarchitectonic landmarks and afferent/efferent connections. Even the relative positions of these regions are grossly conserved in all mammals. Because mammals inhabit a wide diversity of ecological niches, from treetops to fields to oceans, this conserved pattern of regionalization suggests that the neocortex is evolutionarily constrained to a particular architecture (Krubitzer & Kahn, 2003).

More evidence of constraints upon regionalization comes from studies of animals that have completely lost one sensory modality, and yet still preserve a vestige of the corresponding regions of neocortex and thalamus. For example, the subterranean mole rat *Spalax ehrenberghi* is completely blind: its eyes are entirely covered with skin and fur, and recordings from cortex show no evidence of visually evoked potentials. The only function for the retina in this species is to entrain its circadian clock, which occurs via projections to the superchiasmatic nucleus. Nevertheless, the retina still sends sparse projections to all the visual areas that normally process form and motion in other mammals, including the superior colliculus and the lateral geniculate nucleus (LGN) of the thalamus. Moreover, the LGN still sends a topographic projection to occipital cortex, where V1 is normally located (Cooper, Herbin, & Nevo, 1993). Although these regions are severely reduced in size, their persistence in this species argues that the regionalization of thalamus and cortex is constrained by evolutionary inheritance.

Despite sharing conserved features, homologous neural structures can dramatically switch functions. This idea is illustrated by the case of the blind mole rat, where large parts of the LGN and occipital cortex are taken over by auditory inputs (Bronchti et al., 2002; Heil, Bronchti, Wollberg, & Scheich, 1991). This illustrates the principle that inherited constraints are typically incorporated into functional neural systems.

Developmental programs

CANALIZATION Because certain developmental programs are robust to genetic variation, they tend to persist across evolutionary time, and they channel neural systems into stereotyped architectures. These programs can be likened to a canal that channels the progress of a waterway along a stereotyped route; thus, this phenomenon has been termed *canalization.* The idea that neural systems are constrained by these developmental canals is closely linked to the idea of inherited evolutionary constraints (see above). Indeed, there is no real distinction between a developmental constraint of this sort and an evolutionary constraint: developmental programs are the means by which evolutionary constraints are imprinted on an individual organism.

An example of a developmental canal is the sequence of neurogenesis in different divisions of the mammalian brain. This sequence is stereotyped across species, suggesting it is difficult to alter by genetic variation, perhaps due to some robustness of the master regulatory genetic networks that control it. The sequence of neurogenesis is important because it affects the relative volume of different brain divisions. The later the onset of neurogenesis in a particular brain division, the larger the potential pool of neural precursors in that division. Delaying the onset of neurogenesis in one brain division should therefore have cascading effects on the volume of all later-developing brain divisions. Indeed, as we might expect, the brain divisions where neurogenesis occurs last are those that have enlarged disproportionately in large-brained species. The implication is that disproportionate enlargement is constrained to occur preferentially in these brain divisions, as compared to other ones (Finlay & Darlington, 1995; Finlay, Darlington, & Nicastro, 2001). If so, then the disproportionately large size of the human neocortex (and its associated sensory regions) may have arisen initially as a by-product of a constraint on brain development.

DEVELOPMENTAL NOISE In addition to being limited by developmental canals, neural systems are also limited by developmental noise. A clear example is the rodent olfactory bulb. This structure is divided into ~1,000 discrete neuropil compartments called glomeruli. Each glomerulus is uniquely associated with a single olfactory receptor neuron type, corresponding to a single odorant receptor (figure 24.1). On a coarse spatial scale, the relatively spatial location of each glomerulus is completely stereotyped. On a fine spatial scale, however, there is notable imprecision: the relative positions of adjacent glomeruli are often swapped. Notably, imprecision along the anterior-posterior axis is significantly larger than along the medial-lateral axis (Soucy, Albeanu, Fantana, Murthy, & Meister, 2009). This is particularly interesting because anterior-posterior position is specified by an unusual axon guidance mechanism that depends on the intrinsic properties of odorant receptors themselves (Imai, Suzuki, & Sakano, 2006). This suggests that

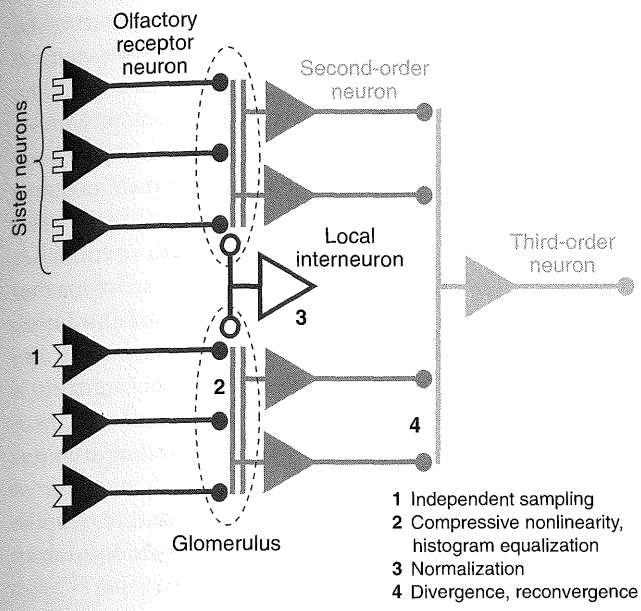

FIGURE 24.1 Circuit organization of the olfactory system. The architecture of the olfactory system provides several examples of how neural systems can respond to internal constraints. Each olfactory receptor neuron (ORN) generally expresses a single odorant receptor, and all the ORNs that express the same receptor converge on the same compartment of neuropil ("glomerulus," dashed line) in the olfactory bulb (in vertebrates) or antennal lobe (in insects). This schematic depicts only two glomeruli, but in reality there are on the order of 1,000 glomeruli (in vertebrates) and 50 in *Drosophila*. The spatial placement of each glomerulus is coarsely stereotyped but can vary on a fine spatial scale, likely reflecting stochastic developmental noise. There are on the order of 10,000 ORNs per glomerulus in vertebrates, and about 40 ORNs per glomerulus in *Drosophila*. In *Drosophila*, all sister ORNs synapse precisely on every second-order neuron, with about 3 second-order neurons per glomerulus. (Vertebrates have 10–100 second-order neurons per glomerulus.) Studies in *Drosophila* have shown that several features of this circuit can be viewed as a response to a constraint (see text): (1) sister ORNs provide independent samples of the olfactory world; (2) within a glomerulus, the transfer function from ORNs to second-order neurons takes the form of a compressive nonlinearity; (3) inhibitory local interneurons mediate lateral inhibition between glomeruli, thereby normalizing the amplitude of population activity across odor stimuli; (4) sister second-order neurons carry largely redundant signals, and they are thought to reconverge onto the same third-order neurons.

variation in glomerular position is due mainly to limitations in the precision of the hardwired developmental mechanisms that specify glomerular position, rather than variations across animals in odor-evoked neural activity. Consistent with this idea, a mutation that eliminates odor-evoked neural activity has relatively little effect on the glomerular map (Lin et al., 2000; Zheng, Feinstein, Bozza, Rodriguez, & Mombaerts, 2000).

As this example illustrates, developmental mechanisms can be noisy. The origin of developmental noise is stochasticity in signal transduction and gene expression. Stochasticity at the single-cell level reflects the low copy number of some proteins within cells (McAdams & Arkin, 1997). This type of noise places limits on the theoretical maximum rate of information transmission in biochemical signaling networks, including the biochemical signals that instruct neural development. Recent studies using information-theoretic analyses have formalized this intuition (Cheong, Rhee, Wang, Nemenman, & Levchenko, 2011).

The effect of developmental noise can be mitigated by using multiple mechanisms having partly redundant functions. For example, retinal ganglion cell axons must project in an orderly fashion to retino-recipient brain regions, forming retinotopic maps where axonal position is systematically related to retinal position. These retinotopic maps are specified by multiple ligand-receptor systems, and some of these have partly redundant functions, such that multiple mechanisms must be genetically disrupted in order to reveal any substantial phenotypic defect (Feldheim et al., 2000).

In spite of such compensatory strategies, developmental noise is still likely to limit the function of neural systems. For example, intrinsic imprecision in glomerular targeting may limit the pattern of horizontal connectivity between glomeruli, which could limit the computations performed in the olfactory bulb (Murthy, 2011). It will be interesting in future to learn whether some features of neural circuit architecture might be adaptive responses to stochasticity in single-cell developmental programs.

Metabolic constraints

ORIGINS OF METABOLIC COSTS For many species, neural systems represent a major metabolic burden. In humans, about 20% of the resting metabolic rate is consumed by the brain (Rolfe & Brown, 1997). Neurons consume more energy when they are active, but even inactive neurons impose a substantial energy burden. When neurons run out of energy, the consequences are swift: a human subject falls unconscious only 7 seconds after circulation to the neck is blocked (Ames, 2000). For these reasons, metabolic demands strongly constrain the architecture of neural systems. This idea has been explored in several comprehensive reviews (Laughlin, 2001; Niven & Laughlin, 2008), which also serve as primers on how neural systems respond to competing constraints.

The highest metabolic costs are imposed by the need to maintain steep ionic gradients across the plasma

membrane. This alone accounts for about half of the energy consumed by neural systems (Ames, 2000). Ionic gradients are dissipated by transmembrane currents, and so there is a price associated with the channels that carry these currents. Among all transmembrane currents, the most costly are the currents associated with action potentials, followed by synaptic currents and leak currents. By comparison, the cost of recycling synaptic vesicles is low (Attwell & Laughlin, 2001).

ARCHITECTURES THAT MINIMIZE METABOLIC COSTS Metabolic constraints shape neural architecture in several ways. To begin with, they create an incentive to match the intrinsic properties of neurons to the signals they must carry. A nice example is provided by the potassium conductances of fly photoreceptors. Visual signals fluctuate rapidly, and, to capture these fluctuations, photoreceptors should have fast membrane time constants. However, there is a metabolic cost to the high leak conductances that would be required to create a fast membrane (Niven, Anderson, & Laughlin, 2007; Niven & Laughlin, 2008). The solution is to match the temporal bandwidth of the membrane to the characteristics of input signals. Phototransduction is slow at low light levels and fast at higher light levels; accordingly, at rest the membrane acts as a low-pass filter, but when depolarized it acts as a high-pass filter. Moreover, fast-flying flies have faster membrane time constants, whereas slow-flying flies have slower membrane time constants (Laughlin, 1994).

Moreover, metabolic costs favor architectures where connected neurons are located near each other in space. For a typical neuron, most of the metabolic cost associated with a single spike is incurred by axonal currents, with a smaller contribution from dendritic and somatic currents (Attwell & Laughlin, 2001). The cost of axonal currents grows with axon length, and so axons should be as short as possible. This means that there is a strong pressure for connected neurons to be spatially close, and indeed most connectivity in neural systems is local rather than long range. This is sometimes called the *wiring economy principle* (Chklovskii & Koulakov, 2004).

In addition, metabolic costs create an incentive to keep redundancy low (Barlow, 1961, 2001). Indeed, metabolic costs would argue that redundancy should be minimized both in space and in time (i.e., both across neurons and within neurons). There are several ways that neural systems can do this. First, redundancy can be minimized by using an array of sensors that under-samples the sensory world. This is exemplified by the photoreceptor array in both vertebrates and invertebrates, which under-samples the optical image; this may related to the very high metabolic costs incurred by these cells (Laughlin, 1994; Snyder, Bossomaier, & Hughes, 1986). Second, redundancy can also be reduced by cell-intrinsic mechanisms of adaptation or gain control, which tend to reduce redundancy over time. Cells possess a variety of these intrinsic negative feedback mechanisms (Shapley & Enroth-Cugell, 1984; Wark, Lundstrom, & Fairhall, 2007). Third, redundancy can be reduced by *lateral inhibition* among neurons that have correlated activity. This type of lateral inhibition also represents a form of negative feedback. The key feature here is that information in other neurons is used to make a prediction about what level of gain is needed. This has been called a *predictive coding* architecture. Finally, redundancy is reduced simply by virtue of the fact that most neurons have a nonlinear spike threshold, because any nonlinearity tends to reduce linear correlations (Pitkow & Meister, 2012).

Electrical noise

ORIGINS OF NOISE As electrical signals propagate through neural systems, they are continually contaminated by noise that arises in neurons themselves. This represents a major intrinsic constraint on sensory processing. The origin of neural noise lies in stochastic microscopic processes. Chief among these are ion-channel gating and synaptic vesicle release. The noise created by these stochastic microscopic events can then be amplified by the nonlinear properties of neurons (Faisal, Selen, & Wolpert, 2008).

Intuitively, one might think that channel noise is not a major problem for most neurons, because noise should average out across many channels. Surprisingly, this is not true—even for neurons that contain relatively large numbers of channels (White, Rubinstein, & Kay, 2000). There are three reasons for this result. First, the signal-to-noise ratio (SNR) of total conductance grows only slowly with increasing channel number, because it is proportional to the square root of the number of channels (N). Given the metabolic costs of increasing N, this limits the ability of a cell to overcome noise by increasing N. Second, for a channel that is gated by depolarization, the probability of opening is low at hyperpolarized potentials, and so the SNR of total conductance can be relatively poor. Third, individual channels are not independent: a stochastic opening of one voltage-gated Na^+ channel will tend to depolarize the cell, thereby increasing the probability that another Na^+ channel will open.

Synaptic noise arises primarily from the fact that the release of synaptic vesicles is stochastic (del Castillo & Katz, 1954). At many synapses, the mean number of

The Cognitive Neurosciences, 5th ed., 2014, MIT Press, edited by Michael S. Gazzaniga and George R. Mangun

released vesicles is low, and so trial-to-trial variability in the number of released vesicles is quite high. This, together with variability in the amount of neurotransmitter per vesicle, can create large trial-to-trial fluctuations in the size of the postsynaptic response to a single presynaptic spike (Bekkers & Clements, 1999; Sargent, Saviane, Nielsen, DiGregorio, & Silver, 2005).

Together, channel noise and synaptic noise inject a substantial stochastic element into neural activity. This places a limitation on the precision of stimulus encoding, because it causes identical presentations of the same stimulus to elicit different neural responses. Moreover, the mean membrane potential of many neurons *in vivo* is just below their spike threshold; neural noise sources tend to push these neurons above their threshold, so that they fire spikes even in the absence of a stimulus. Finally, it should be kept in mind that neural noise is injected at every layer of a sensory processing circuit, and so each central neuron inherits noise from previous layers.

ARCHITECTURES THAT MINIMIZE THE EFFECTS OF NOISE Given these considerations, neural systems are under pressure to minimize the negative effects of noise. One strategy is to pool redundant signals from independent sensors. This is exemplified by the first relay in the olfactory system (figure 24.1), where each second-order neuron pools input from many olfactory receptor neurons, all of which express the same odorant receptor and all of which project their axons to the same glomerulus. In the fruit fly *Drosophila melanogaster*, all olfactory receptor neurons that express the same odorant receptor (called "sister ORNs") are known to synapse quite precisely onto each and every second-order neuron in their target glomerulus, and to make synapses of a rather uniform strength. All sister ORNs have the same noise level, and their noise is independent, so this architecture ought to maximize the SNR of second-order neurons. Indeed, the SNR of second-order neurons is better than that of their cognate ORNs, despite the fact that principal neurons are subject to additional synaptic and channel noise (Bhandawat, Olsen, Schlief, Gouwens, & Wilson, 2007; Kazama & Wilson, 2008, 2009). There are several second-order neurons in each glomerulus, and because they pool input from exactly the same sister ORNs, they carry highly redundant signals. Interestingly, there is evidence that they synapse onto some of the same third-order neurons (Marin, Jefferis, Komiyama, Zhu, & Luo, 2002; Wong, Wang, & Axel, 2002). This arrangement should allow third-order neurons to average out some of the noise that arises *de novo* in second-order neurons. Note that there is substantial redundancy in this circuit at two successive layers, in spite of its metabolic costs. Note also the peculiar architecture of this circuit (figure 24.1): signals first converge (all sister ORNs synapse onto each sister second-order neuron), then diverge (in the form of redundant sister second-order neurons), and then reconverge (as sister second-order neurons wire onto the same third-order neuron). This architecture suggests the system is under strong pressure to minimize the maladaptive effects of noise. A similar architecture has been proposed for the transmission of visual information between retina, thalamus, and visual cortex (Alonso, Usrey, & Reid, 2001).

Another strategy is to impose a filter that selectively discards noise, retaining the signal. This is possible only if the properties of signal and noise are distinctively different. For example, the phototransduction cascade downstream from rhodopsin is spontaneously active, which generates continuous voltage noise in photoreceptors. Absorption of a photon generates a predictable discrete "bump" of activity in the phototransduction cascade, which is distinctively different from continuous noise. Accordingly, the synapse between rod photoreceptors and bipolar cells is configured to impose a threshold on rod output, such that continuous noise cannot pass, but (many) single-photon responses can pass (Field & Rieke, 2002).

Yet another strategy is to distribute signals as uniformly as possible within the available coding space, ensuring that all codes are used with equal frequency. This strategy is sometimes known as *histogram equalization*, because it produces a flat histogram of response probabilities. The classic example of this phenomenon is the contrast-response function of second-order fly visual neurons, which is nicely matched to the distribution of contrasts in natural visual scenes. As a consequence, these neurons use each response level with equal probability (Laughlin, 1981). Importantly, histogram equalization cannot help combat existing noise, but it helps immunize signals from noise that is added later: when signals are well separated in coding space, adding noise has a minimal effect on their discriminability. Another example of histogram equalization occurs in the *Drosophila* olfactory system. Most odor responses of olfactory receptor neurons fall within the lower part of the dynamic range of these neurons. This might reflect a metabolic constraint on average firing rate, especially as ORNs are numerous, outnumbering second-order neurons by ~10:1. Weak ORN responses are then preferentially boosted as they are transmitted to second-order neurons. As a consequence, second-order neurons use each response level with roughly equal probability. Because second-order neurons are less numerous, the relative pressure of metabolic

constraints and noise constraints may be different in these neurons.

Constraints on neural electronics

In principle, single neurons can perform a vast array of operations on their synaptic inputs. Biophysicists tend to emphasize this viewpoint. The fanciest single-neuron operations rely on complex dendritic morphologies, the specific spatial placement of synaptic inputs onto dendritic trees, and well-tuned voltage-gated conductances in dendrites. There is ample evidence that neurons can achieve these things, although their roles *in vivo* are necessarily difficult to demonstrate (London & Hausser, 2005; Silver, 2010).

Nonetheless, the operations that a neuron can perform are also constrained by the nature of cellular electronics. Many of these constraints can be mitigated, but sometimes the solution would be costly and so is not worth the price. In short, single-neuron operations are not arbitrarily flexible. In some cases, this creates incentives for neural circuits to evolve architectures that can compensate for the limitations of single neurons.

LIMITATIONS ON SPEED Several factors limit the speed of neural processing. As we have seen, fast membrane time constants are metabolically costly, and this may be why many neurons have relatively slow membrane time constants. In addition, dendritic cable filtering tends to slow synaptic potentials as they travel to the spike initiation zone. In particular, axonal conduction delays can be as large as 100 milliseconds in long axons, much longer than the typical delay involved in synaptic transmission (150–400 microseconds; Sabatini & Regehr, 1999). Axonal conduction speed can be increased by increasing axon diameter, but because volume grows with the square of the diameter, this strategy consumes valuable space (Swadlow, 2000).

Notably, many organisms have evolved neural subsystems with unusual cellular specializations for speed. In invertebrates, these subsystems are characterized by large-diameter axons and electrical synapses, and they mediate escape reflexes (Allen, Godenschwege, Tanouye, & Phelan, 2006; Faulkes, 2008). In the brains of many mammals, the distribution of axon diameters is right-skewed, and the largest axons are always myelinated, which further increases speed. This subpopulation of particularly large axons has been proposed to serve brain functions that require fast conduction speed (Perge, Koch, Miller, Sterling, & Balasubramanian, 2009; Wang et al., 2008). The fact that these specializations for speed are only found in a small fraction of neurons is consistent with the idea that they come at a high price.

LIMITATIONS ON DYNAMIC RANGE Spike rates and vesicular release rates cannot be negative. This poses a constraint, because many sensory systems must encode fluctuations above and below some mean ambient level of a stimulus (e.g., light). In principle, neural systems might respond to this constraint by setting basal firing rates high, allowing the same neuron to encode both increases and decreases about the mean. However, high basal firing rates are metabolically costly. An alternative is to create opponent populations of neurons having opposite stimulus preferences. Classic cases are the ON-OFF neurons and color opponent neurons of the retina, but opponent neurons can also be found in mechanosensory, auditory, and thermosensory systems (Jacobs, Miller, & Aldworth, 2008; Ma, 2010; Yorozu et al., 2009).

Firing rates also cannot be arbitrarily large. Moreover, it seems that there can be strong constraints on the firing rate of a neuron averaged over long time scales. The evidence for this idea comes from the observation that firing rate distributions are exponentially distributed in visual cortical neurons. This is notable, because information theory predicts that firing-rate distributions should actually be flat, since this should maximize the rate of information transmission (see above). Exponentially distributed firing rates are consistent with the existence of a constraint on mean firing rates, together with a pressure to maximize the rate of information transmission within that constraint (Baddeley et al., 1997). Given the energy budget of the human brain, it has been estimated that the average neuron is constrained to fire at rates around 1–7 Hz (Wang et al., 2008). A constraint on firing rates will constrain the number of distinguishable messages that a neuron can send, because noise limits the number of different firing rates that can be reliably distinguished from each other.

Finally, synapses cannot be arbitrarily strong. In a passive dendrite, increasing the conductance of a synapse brings diminishing marginal returns on the postsynaptic voltage response, because a large localized synaptic conductance will simply shunt synaptic currents. Dendritic voltage-gated conductances can of course amplify synaptic potentials, but this also has a cost. Specifically, if a dendrite is endowed with active conductances, this will increase not only the efficacy of the synapse in question, but also the barrage of ongoing noise arising from spontaneous presynaptic spikes at other synapses. As a consequence, there may be no net increase in the ability of the synapse in question to

control the postsynaptic cell (London, Schreibman, Hausser, Larkum, & Segev, 2002).

All these considerations limit the dynamic range of any neuron's output. These limits create an incentive for sensory neurons to implement some form of adaptation or gain control. This allows a neuron to continually adjust its sensitivity based on the current characteristics of the sensory environment, so that both large and small inputs fall within its dynamic range (Wark et al., 2007). These mechanisms are ubiquitous, but they also come at a cost: because the signals that control adaptation or gain will tend to be noisy, any feedback mechanism based on these signals will tend to amplify noise (Dunn & Rieke, 2008). Moreover, because gain control changes the relationship between input and output, it necessarily creates ambiguity about the stimulus (Fairhall, Lewen, Bialek, & de Ruyter Van Steveninck, 2001).

Limitations Arising from Linear Summation At least to a first approximation, the dendritic trees of most neurons perform a rather simple operation on their synaptic inputs. Inputs are weighted and summed quasi-linearly, and then the neuron fires roughly in proportion to the sum above some threshold. This simplistic description, of course, does not capture all that single neurons are capable of. However, it encapsulates the basic hardware that all neurons possess.

Insofar as many neurons integrate their inputs in this manner, this situation creates a constraint on what categorizations these neurons can perform. Imagine we would like to create a neuron that responds only to a particular subset of stimuli. This neuron can receive N synaptic inputs. The free parameters we have are the weights associated with each input synapse, as well as the threshold of that neuron. It is possible to find a set of parameter values that generates the desired solution if and only if one can draw a hyperplane in N-dimensional space that separates target stimuli from off-target stimuli (Rosenblatt, 1958). In other words, stimulus representations must be *linearly separable*. An idealized binary linear classifier of this type is called a *perceptron*.

Real neurons are even more constrained than classical perceptrons. This is because synaptic weights and thresholds cannot be arbitrarily large. In addition, the synaptic inputs that can be tuned to confer the desired selectivity may be constrained to be nonnegative, because they are synaptic inputs from excitatory neurons.

There are several ways that neural systems can improve linear separability under these strong constraints. First, it is helpful to begin with a high-dimensional stimulus representation. This is because the likelihood that there is a separating hyperplane increases with the number of synaptic inputs N (Cover, 1965). For example, it has been pointed out that high dimensionality is a characteristic property of olfactory encoding (Itskov & Abbott, 2008). Odors are encoded combinatorially by the activity of many odorant receptors having diverse odor tuning, where each receptor binds multiple odors, and each odor binds multiple receptors. As the number of odorant receptors (N) grows, so does the potential selectivity of higher-order olfactory neurons that linearly sum the activity of different receptors.

Second, there is a special case of the linear separation problem where *normalization* and a *compressive nonlinearity* can be helpful. This special case can be termed sparse recoding. Imagine that we want to create an array of neurons where each neuron responds to only one stimulus out of many, and each stimulus activates at least one neuron. To achieve this, each neuron needs synaptic input weights and a threshold that confers specificity for a single stimulus. The problem of finding these values amounts to the problem of drawing a line (or in high-dimensional space, a hyperplane) that separates each individual stimulus from all other stimuli (figure 24.2). Sparse codes are typical of many sensory brain regions, so this is a biologically relevant special case of the linear separability problem.

In this situation, separability can be improved by *normalization*. Normalization involves dividing activity in individual neurons by the summed activity of many neurons in the same brain region. This operation occurs in a wide range of sensory modalities and organisms (Carandini & Heeger, 2012). Normalization facilitates sparse recoding, because it tends to equalize the total population firing rates evoked by different stimuli. In geometric terms, if we imagine each stimulus as a point in N-dimensional space, then normalization tends to move all points toward the surface of a hypersphere in that space, making it is easier to find a line or a hyperplane that separates each stimulus from the rest (figure 24.2). An example of normalization has been described in the *Drosophila* olfactory system, where input from individual olfactory receptor neurons is divided by the total activity of all olfactory receptor neurons. This is accomplished via lateral inhibition from local interneurons (figure 24.1). In simulations that were tightly constrained by data, this operation makes it easier to construct linear classifiers that respond sparsely and selectively to single odors (Luo, Axel, & Abbott, 2010; Olsen, Bhandawat, & Wilson, 2010). Normalization has been proposed to serve a similar function in visual cortical areas (DiCarlo & Cox, 2007; Olshausen & Field, 2005).

The Cognitive Neurosciences, 5th ed., 2014, MIT Press, edited by Michael S. Gazzaniga and George R. Mangun

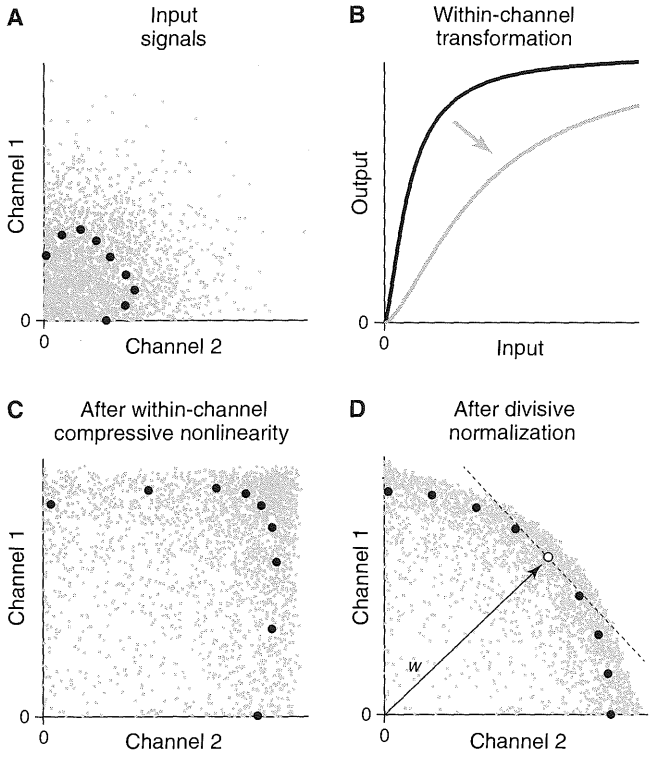

FIGURE 24.2 Compression, normalization, and linear separability. (A) Consider a sensory representation in multiple dimensions, with each dimension corresponding to a coding channel (a neuron or a chain of neurons). For concreteness, we might imagine the two-dimensional space in this schematic corresponds to two different types of olfactory receptor neurons. Every odor stimulus is a point in this space. A subset of points are highlighted for comparison across panels (black circles). (B) A compressive nonlinearity transforms signals within each channel (black curve). Divisive normalization adjusts the steepness of this nonlinearity according to the total input to both channels. Namely, when total activity is high, the nonlinearity is less steep (gray curve; Carandini & Heeger, 2012). Different stimuli will therefore fall on different curves, because they produce different amounts of total input to both channels. Curves in this schematic are fit to data recorded from *Drosophila* olfactory neurons in vivo (Olsen et al., 2010). (C) A steep compressive nonlinearity (black curve in B) preferentially boosts weak input signals (compare to A). (D) Normalization tends to equalize the distance of all representations from the origin. As a consequence, stimuli tend to lie near the surface of a circle. The vector **w** corresponds to the weights on the two input channels that define a linear separation (dashed line) between one highlighted symbol and the other symbols. Note that, in this example, we can separate each highlighted symbol from the rest using only nonnegative weights. In the *Drosophila* olfactory system, there are 50 coding channels (glomeruli), as compared to the two channels in this schematic. This system has a large capacity for generating sparse and selective representations, and compression and normalization together improve this capacity (Luo et al., 2010; Olsen et al., 2010).

In addition, sparse recoding is also facilitated by a *compressive nonlinearity* in feedforward excitation. The intuition here is similar. A compressive nonlinearity tends to make a responsive neuron fire at a fixed rate (i.e., its saturated response level). This operation produces total population firing rates that are relatively equal for different stimuli, as compared with a scenario where there is no compressive nonlinearity. Again drawing an example from the *Drosophila* olfactory system, it is notable that there is a compressive nonlinearity in the relationship between the odor responses of olfactory receptor neurons and postsynaptic second-order neurons. Implementing this nonlinearity in data-driven simulations makes it easier to construct linear classifiers that respond sparsely and selectively to single odors (figure 24.2; Luo et al., 2010; Olsen et al., 2010). It is perhaps nonintuitive that equalizing firing rates (via compressive nonlinearity and normalization) can improve separability. The key is that these operations occur in a high-dimensional coding space.

Conclusion

To be a film director, in the words of Orson Welles, is simply to "preside over accidents." This could stand as a description of how a nervous system evolves and develops—accident by accident. Neural systems are subject to constraints, and this influences which random variations are passed on to successive generations. As we have seen, the history of which accidents survive becomes another constraint on neural system architecture, in the form of the organism's evolutionary inheritance and developmental programs.

The central argument of this chapter is that internal constraints leave their imprint on the architecture of neural systems. Because these constraints are ubiquitous, they can potentially explain why some architectures are so common. Understanding how this might occur will require a comparative approach that embraces a variety of sensory modalities and organisms.

Thinking about constraints is more important than ever before, because the field of experimental neuroscience is undergoing a revolution in techniques. New techniques allow us to precisely perturb neural activity, and to test how this affects perception and behavior. In practice, one often begins with a specific element of a neural system (e.g., a cell type) and one searches for behaviors that fall apart when this element is perturbed. This search assumes we will understand the function of each element by identifying the behaviors that rely on it. However, the nature of the behavioral task is only one pressure that drives the evolution of neural systems: equally important is the pressure to

cope with internal constraints. Understanding these constraints may inspire more sophisticated experiments and interpretations.

ACKNOWLEDGMENTS Larry Abbott, James Jeanne, Andreas Liu, and Wendy Liu provided helpful comments on drafts of the manuscript. The author's research is supported by the National Institutes of Health (R01 DC008174) and the Howard Hughes Medical Institute.

REFERENCES

ALLEN, M. J., GODENSCHWEGE, T. A., TANOUYE, M. A., & PHELAN, P. (2006). Making an escape: Development and function of the *Drosophila* giant fibre system. *Semin Cell Dev Biol, 17*(1), 31–41.

ALONSO, J. M., USREY, W. M., & REID, R. C. (2001). Rules of connectivity between geniculate cells and simple cells in cat primary visual cortex. *J Neurosci, 21*(11), 4002–4015.

AMES, A., III. (2000). CNS energy metabolism as related to function. *Brain Res Rev, 34*(1–2), 42–68.

ATTWELL, D., & LAUGHLIN, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab, 21*(10), 1133–1145.

BADDELEY, R., ABBOTT, L. F., BOOTH, M. C., SENGPIEL, F., FREEMAN, T., WAKEMAN, E. A., & ROLLS, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc Biol Sci, 264*(1389), 1775–1783.

BARLOW, H. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

BARLOW, H. (2001). Redundancy reduction revisited. *Network, 12*(3), 241–253.

BEKKERS, J. M., & CLEMENTS, J. D. (1999). Quantal amplitude and quantal variance of strontium-induced asynchronous EPSCs in rat dentate granule neurons. *J Physiol, 516*, 227–248.

BHANDAWAT, V., OLSEN, S. R., SCHLIEF, M. L., GOUWENS, N. W., & WILSON, R. I. (2007). Sensory processing in the *Drosophila* antennal lobe increases the reliability and separability of ensemble odor representations. *Nat Neurosci, 10*(11), 1474–1482.

BRONCHTI, G., HEIL, P., SADKA, R., HESS, A., SCHEICH, H., & WOLLBERG, Z. (2002). Auditory activation of "visual" cortical areas in the blind mole rat (*Spalax ehrenbergi*). *Eur J Neurosci, 16*(2), 311–329.

CARANDINI, M., & HEEGER, D. J. (2012). Normalization as a canonical neural computation. *Nat Rev Neurosci, 13*(1), 51–62.

CHEONG, R., RHEE, A., WANG, C. J., NEMENMAN, I., & LEVCHENKO, A. (2011). Information transduction capacity of noisy biochemical signaling networks. *Science, 334*(6054), 354–358.

CHKLOVSKII, D. B., & KOULAKOV, A. A. (2004). Maps in the brain: What can we learn from them? *Annu Rev Neurosci, 27*, 369–392.

COOPER, H. M., HERBIN, M., & NEVO, E. (1993). Visual system of a naturally microphthalmic mammal: The blind mole rat, *Spalax ehrenbergi*. *J Comp Neurol, 328*(3), 313–350.

COVER, T. M. (1965). Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans Comput, 14*(3), 326–334.

DEL CASTILLO, J., & KATZ, B. (1954). Quantal components of the end-plate potential. *J Physiol, 124*(3), 560–573.

DICARLO, J. J., & COX, D. D. (2007). Untangling invariant object recognition. *Trends Cogn Sci, 11*(8), 333–341.

DUNN, F. A., & RIEKE, F. (2008). Single-photon absorptions evoke synaptic depression in the retina to extend the operational range of rod vision. *Neuron, 57*(6), 894–904.

FAIRHALL, A. L., LEWEN, G. D., BIALEK, W., & DE RUYTER VAN STEVENINCK, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature, 412*(6849), 787–792.

FAISAL, A. A., SELEN, L. P., & WOLPERT, D. M. (2008). Noise in the nervous system. *Nat Rev Neurosci, 9*(4), 292–303.

FAULKES, Z. (2008). Turning loss into opportunity: The key deletion of an escape circuit in decapod crustaceans. *Brain Behav Evolut, 72*, 251–261.

FELDHEIM, D. A., KIM, Y. I., BERGEMANN, A. D., FRISEN, J., BARBACID, M., & FLANAGAN, J. G. (2000). Genetic analysis of ephrin-A2 and ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping. *Neuron, 25*(3), 563–574.

FIELD, G. D., & RIEKE, F. (2002). Nonlinear signal transfer from mouse rods to bipolar cells and implications for visual sensitivity. *Neuron, 34*(5), 773–785.

FINLAY, B. L., & DARLINGTON, R. B. (1995). Linked regularities in the development and evolution of mammalian brains. *Science, 268*(5217), 1578–1584.

FINLAY, B. L., DARLINGTON, R. B., & NICASTRO, N. (2001). Developmental structure in brain evolution. *Behav Brain Sci, 24*(2), 263–308.

HEIL, P., BRONCHTI, G., WOLLBERG, Z., & SCHEICH, H. (1991). Invasion of visual cortex by the auditory system in the naturally blind mole rat. *NeuroReport, 2*(12), 735–738.

IMAI, T., SUZUKI, M., & SAKANO, H. (2006). Odorant receptor-derived cAMP signals direct axonal targeting. *Science, 314*(5799), 657–661.

ITSKOV, V., & ABBOTT, L. F. (2008). Pattern capacity of a perceptron for sparse discrimination. *Phys Rev Lett, 101*(1), 018101.

JACOBS, G. A., MILLER, J. P., & ALDWORTH, Z. (2008). Computational mechanisms of mechanosensory processing in the cricket. *J Exp Biol, 211*(Pt 11), 1819–1828.

KAZAMA, H., & WILSON, R. I. (2008). Homeostatic matching and nonlinear amplification at genetically identified central synapses. *Neuron, 58*, 401–413.

KAZAMA, H., & WILSON, R. I. (2009). Origins of correlated activity in an olfactory circuit. *Nat Neurosci, 12*(9), 1136–1144.

KRUBITZER, L., & KAHN, D. M. (2003). Nature versus nurture revisited: An old idea with a new twist. *Prog Neurobiol, 70*(1), 33–52.

LAUGHLIN, S. B. (1981). A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch, 36c*(9–10), 910–912.

LAUGHLIN, S. B. (1994). Matching coding, circuits, cells, and molecules to signals: General principles of retinal design in the fly's eye. *Prog Retin Eye Res, 13*, 165–196.

LAUGHLIN, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Curr Opin Neurobiol, 11*(4), 475–480.

Lin, D. M., Wang, F., Lowe, G., Gold, G. H., Axel, R., Ngai, J., & Brunet, L. (2000). Formation of precise connections in the olfactory bulb occurs in the absence of odorant-evoked neuronal activity. *Neuron, 26*(1), 69–80.

London, M., & Hausser, M. (2005). Dendritic computation. *Annu Rev Neurosci, 28*, 503–532.

London, M., Schreibman, A., Hausser, M., Larkum, M. E., & Segev, I. (2002). The information efficacy of a synapse. *Nat Neurosci, 5*(4), 332–340.

Luo, S. X., Axel, R., & Abbott, L. F. (2010). Generating sparse and selective third-order responses in the olfactory system of the fly. *Proc Natl Acad Sci USA, 107*(23), 10713–10718.

Ma, Q. (2010). Labeled lines meet and talk: Population coding of somatic sensations. *J Clin Invest, 120*(11), 3773–3778.

Marder, E., & Goaillard, J. M. (2006). Variability, compensation and homeostasis in neuron and network function. *Nat Rev Neurosci, 7*(7), 563–574.

Marin, E. C., Jefferis, G. S., Komiyama, T., Zhu, H., & Luo, L. (2002). Representation of the glomerular olfactory map in the *Drosophila* brain. *Cell, 109*(2), 243–255.

McAdams, H. H., & Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proc Natl Acad Sci USA, 94*(3), 814–819.

Murthy, V. N. (2011). Olfactory maps in the brain. *Annu Rev Neurosci, 34*, 233–258.

Niven, J. E., Anderson, J. C., & Laughlin, S. B. (2007). Fly photoreceptors demonstrate energy-information trade-offs in neural coding. *PLoS Biol, 5*(4), e116.

Niven, J. E., & Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *J Exp Biol, 211*(Pt 11), 1792–1804.

Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2010). Divisive normalization in olfactory population codes. *Neuron, 66*, 287–299.

Olshausen, B. A., & Field, D. J. (2005). How close are we to understanding V1? *Neural Comput, 17*(8), 1665–1699.

Perge, J. A., Koch, K., Miller, R., Sterling, P., & Balasubramanian, V. (2009). How the optic nerve allocates space, energy capacity, and information. *J Neurosci, 29*(24), 7917–7928.

Pitkow, X., & Meister, M. (2012). Decorrelation and efficient coding by retinal ganglion cells. *Nat Neurosci, 15*(4), 628–635.

Rieke, F. (2014). Are neural circuits the best they can be? In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (5th ed.). Cambridge, MA: MIT Press.

Rolfe, D. F., & Brown, G. C. (1997). Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol Rev, 77*(3), 731–758.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol Rev, 65*(6), 386–408.

Sabatini, B. L., & Regehr, W. G. (1999). Timing of synaptic transmission. *Annu Rev Physiol, 61*, 521–542.

Sargent, P. B., Saviane, C., Nielsen, T. A., DiGregorio, D. A., & Silver, R. A. (2005). Rapid vesicular release, quantal variability, and spillover contribute to the precision and reliability of transmission at a glomerular synapse. *J Neurosci, 25*(36), 8173–8187.

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Prog Retin Res, 3*, 263–346.

Silver, R. A. (2010). Neuronal arithmetic. *Nat Rev Neurosci, 11*(7), 474–489.

Snyder, A. W., Bossomaier, T. R., & Hughes, A. (1986). Optical image quality and the cone mosaic. *Science, 231*(4737), 499–501.

Soucy, E. R., Albeanu, D. F., Fantana, A. L., Murthy, V. N., & Meister, M. (2009). Precision and diversity in an odor map on the olfactory bulb. *Nat Neurosci, 12*(2), 210–220.

Stafford, R. (2010). Constraints of biological neural networks and their consideration in AI applications. *Lect Notes Artif Int, 2010*, 1–6.

Swadlow, H. A. (2000). Information flow along neocortical axons. In R. Miller (Ed.), *Time and the brain* (pp. 150–180). Amsterdam: Harwood Academic.

Wang, S. S., Shultz, J. R., Burish, M. J., Harrison, K. H., Hof, P. R., Towns, L. C., & Wyatt, K. D. (2008). Functional trade-offs in white matter axonal scaling. *J Neurosci, 28*(15), 4047–4056.

Wark, B., Lundstrom, B. N., & Fairhall, A. (2007). Sensory adaptation. *Curr Opin Neurobiol, 17*(4), 423–429.

White, J. A., Rubinstein, J. T., & Kay, A. R. (2000). Channel noise in neurons. *Trends Neurosci, 23*(3), 131–137.

Wong, A. M., Wang, J. W., & Axel, R. (2002). Spatial representation of the glomerular map in the *Drosophila protocerebrum*. *Cell, 109*(2), 229–241.

Yorozu, S., Wong, A., Fischer, B. J., Dankert, H., Kernan, M. J., Kamikouchi, A., & Anderson, D. J. (2009). Distinct sensory representations of wind and near-field sound in the *Drosophila* brain. *Nature, 458*(7235), 201–205.

Zheng, C., Feinstein, P., Bozza, T., Rodriguez, I., & Mombaerts, P. (2000). Peripheral olfactory projections are differentially affected in mice deficient in a cyclic nucleotide-gated channel subunit. *Neuron, 26*(1), 81–91.