

A Computational Introduction to the Biological Brain-Mind

Juyang Weng

Michigan State University, USA

*corresponding author: weng@cse.msu.edu

Abstract

The brain-mind is hyphenated because of the tight integration of the brain model and the mind model. On one hand, the neuroscience literature has provided very rich data about the brain, but such data tend to mislead us to think that the brain is composed many special purpose modules (e.g., Brodmann areas) where the role of each module is largely determined by the genes (e.g., detecting edge orientation or a human face). On the other hand, traditional artificial neural networks (e.g., Self-Organization Map SOM) perform general-purpose signal processing and they learn. However, they cannot autonomously learn and develop like a brain with its body. Autonomous mental development models how a brain-like system, natural and artificial, develops autonomously through interactions with the environments. The most fundamental difference between traditional machine learning (using symbolic or neural net methods) and autonomous mental development is that a developmental program is task non-specific so that it can autonomously generate internal representations for a wide variety of simple to complex tasks. This paper first discusses why autonomous development is necessary based on a concept called task muddiness. No traditional methods can perform muddy tasks. If the electronic system that you design is meant to perform a muddy task, you need to enable it to autonomously develop its own mind. Then some basic concepts of autonomous development are explained, including the paradigm for autonomous development, brain-mental architectures, developmental algorithm, a refined classification of types of machine learning, spatial complexity and time complexity. Finally, the architecture of a brain-like spatiotemporal machine that is capable of autonomous development is described.

1. Biological Development

A human being starts to develop from the time of conception. At that time, a single cell called a zygote is formed. In biology, the term *genotype* refers to all or part of the genetic constitution of an organism. The term *phenotype* refers to all or part of the visible properties of an organism that are produced through the interaction between the genotype and the environment. In the zygote, all the genetic constitution is called genome, which mostly resides in the nucleus of a cell. At the conception of a new human life, a biological program called the *developmental program* starts to run. The code of this program is the genome, but this program needs the entire cell as well as the cell's environment to run properly.

The biological developmental program handles two types

of development, *body development* and *mental development*. The former is the development of everything in the body excluding the brain. The latter is the development of the brain (or the Central Nervous System CNS). Through the body development, a normal child grows in size and weight, along with many other physical changes. Through the mental development, a normal child develops a series of mental capabilities through interactions with the environment. Mental capabilities refer to all known brain capabilities, which include, but not limited to, perceptual, cognitive, behavioral and motivational capabilities. In this paper, the term *development* refers to mental development unless stated otherwise. The biological mental development takes place in concurrence with the body development and they are closely related. For example, if the eyes are not normally developed, the development of the visual capabilities is greatly affected. In the development of an artificial agent, the body can be designed and fixed (not autonomously developed), which helps to reduce the complexity of the autonomous mental development.

The *genomic equivalence* principle [1] is a very important biological concept for us to understand how biological development is regulated. This principle states that the set of genes in the nucleus of every cell (not only that in the zygote!) is functionally complete -- sufficient to regulate the development from a single cell into an entire adult life. This principle is dramatically demonstrated by cloning. This means that there are no genes that are devoted to more than one cell as a whole. Therefore, development guided by the genome is *cell-centered*. Carrying a complete set of genes and acting as an autonomous machine, every cell must handle its own learning while interacting with its external environment (e.g., other cells). Inside the brain, every neuron develops and learns in place. It does not need any dedicated learner outside the neuron. For example, it does not need an extra-cellular learner to compute the covariance matrix (or any other moment matrix or partial derivatives) of its input lines and store extra-cellularly. If an artificial developmental program develops every artificial neuron based on only information that is available to the neuron itself (e.g., the cellular environment such as pre-synaptic activities, the developmental program inside the cell, and other information that can be biologically stored intra-cellularly), we call this type of learning *in-place learning*.

This in-place concept is more restrictive than a common concept called “local learning.” For example, a local learning algorithm may require the computation of the covariance matrix of the pre-synaptic vector that must store extracellularly. In electronics, the in-place learning principle can greatly reduce the required electronics and storage space, in addition to the biological plausibility. For example, suppose that every biological neurons requires the partial derivative matrix of its pre-synaptic vector. As the average number of synapses of a neuron in the brain is on the order of $n = 1000$. Each neuron requires about $n^2 = 1,000,000$ storage units outside every neuron. This corresponds to about 1,000,000 of the total number of synapses (10^{14}) in the brain!

Conceptually, the fate and function of a neuron is not determined by a “hand-designed” (i.e., genome specified meaning of the external environment. This is another consequence of the genomic equivalence principle. The genome in each cell regulates the cells mitosis, differentiation, migration, branching, and connections, but it does not regulate the meaning of what the cell does when it receives signals from other connected cells. For example, we can find a V1 cell (neuron) that responds to an edge of a particular orientation. This is just a facet of many emergent properties of the cell that are consequences of the cells own biological properties and the activities of its environment. A developmental program does not need to, and should not, specify which neuron detects a pre-specified feature type (such as an edge or motion).

2. Why Autonomous Mental Development?

One can see that biological development is very “low level”, regulating only individual neurons. Then, why is it necessary to enable our complex electronic machines to develop autonomously? Why do we not design high-level concepts into the machines and enable them to carry out our high-level directives? In fact, this is exactly many symbolic methods have been doing for many years. Unfortunately, the resulting machines are brittle — they fail miserably in real world when the environment fall out of the domains that have been modeled by the programmer.

To appreciate what are faced by a machine to carry out a complex task, Weng [2] introduced a concept called *task muddiness*. The composite muddiness of a task is a multiplicative product of many individual muddiness measures. There are many possible individual muddiness measures. Those individual muddiness measures are not necessarily mutually independent or at the same level of abstraction, since such a requirement is not practical nor necessary for describing the muddiness of a task. They fall into five categories: (1) external environment, (2) input, (3) internal environment, (4) output and (5) goal, as shown in Table I. The term “external” means external with respect to the brain and “internal” means internal to the brain.

TABLE I
A LIST OF MUDDINESS FACTORS FOR A TASK

Category	Factor	Clean ←→	Muddy
External Env.	Awareness	Known	Unknown
	Complexity	Simple	Complex
	Controlledness	Controlled	Uncontrolled
	Variation	Fixed	Changing
	Foreseeability	Foreseeable	Nonforeseeable
Input	Rawness	Symbolic	Real sensor
	Size	Small	Large
	Background	None	Complex
	Variation	Simple	Complex
	Occlusion	None	Severe
	Activeness	Passive	Active
	Modality	Simple	Complex
	Multi-modality	Single	Multiple
Internal Env.	Size	Small	Large
	Representation	Given	Not given
	Observability	Observable	Unobservable
	Impossibility	Imposable	Nonimposable
	Time coverage	Simple	Complex
Output	Terminalness	Low	High
	Size	Small	Large
	Modality	Simple	Complex
	Multimodality	Single	Multiple
Goal	Richness	Low	High
	Variability	Fixed	Variable
	Availability	Given	Unknown
	Telling-mode	Text	Multimodal
	Conveying-mode	Simple	Complex

The composite muddiness of a task can be considered as a product of all individual muddiness measures. In other words, a task is extremely muddy when all the five categories have a high measure. A chess playing task with symbolic input and output is a clean problem because it is low in categories (1) through (5). A symbolic language translation problem is low in (1), (2) and (4), moderate in (3) but high in (5). A vision-guided navigation task for natural human environment is high in (1), (2), (3) and (5), but moderate in (4). A human adult handles extremely muddy tasks that are high in all the five categories.

From the muddiness table Table I we have a more detailed appreciation what a human adult deals with even in a daily task, e.g., navigating or driving in a city environment. The composite muddiness of many tasks that a human or a machine can execute is proposed by Weng [2] as a metric for measuring required intelligence.

A human infant is not able to perform those muddy tasks that a human adult performs everyday. The process of mental development is necessary to develop such a wide array of mental skills. Much evidence in developmental psychology has demonstrated that not only a process of development is necessary for human intelligence, the environment of the development is also critical for normal development.

Likewise, it is not practical for a human programmer to program a machine to successfully execute a muddy task. Computers have done very well for clean tasks, such as playing chess games. But they have done poorly in performing muddy tasks, such as visual and language understanding. Enabling a machine to autonomously develop task skills in

its real task environments is the only approach that has been proved successful for muddy tasks — no existing higher intelligence for muddy tasks is not developed autonomously.

3. The Paradigm of Autonomous Development

By definition an agent is something that senses and acts. A robot is an agent, so is a human. In the early days of artificial intelligence, smart systems that caught the general public's imagination were programmed by a set of task-specific rules. The field of artificial intelligence moved beyond that early stage when it started the trend of studying general agent methodology [3], although the agent is still a task-specific machine.

As far as we know, Cresceptron 1993 [4], [5] was the first developmental model for visual learning from complex natural backgrounds. By developmental, we mean that the internal representation fully emerges from interactions with environment, without allowing a human to manually instantiate a task-specific representation. By the mid 1990's, connectionists had started the exploration of the challenging domain of development [6], [7], [8].

Due to a lack of the breadth and depth of the multi-disciplinary knowledge in the single mind of a researcher or a reviewer, there have been various doubts from domain experts, mainly due to the widespread lack of sufficient cross-disciplinary knowledge discussed above. Examples include: (1) Artificial intelligence does not need to follow the brain's way. (2) Modeling the human mind does not need to follow the brain's way. (3) Your commitment to understanding the brain is laudable but naive; (4) Regardless a clear advance of knowledge and its importance, I need to see X.

There is a lack of bylaws, guidelines and due process that contain the negative effects of human nature that are well documented by Thomas Kuhn [9]. Such negative effects eroded "revolutionary advances" required by some federal programs. Serious overhauls and investments for the infrastructure for converging research on intelligence is urgently needed. Such infrastructure is necessary for the healthy development of science and technology in the modern time.

Not until the birth of the new AMD field marked by the NSF and DARPA funded Workshop on Development and Learning 2000 [10], [11] has the concept of the task-nonspecific developmental program caught the attention of many researchers. A hallmark difference between traditional artificial intelligence approaches and autonomous mental development [11] is the task specificity. All the existing approaches to artificial intelligence is task specific except the developmental approach. Table II lists the major differences among existing approaches to artificial intelligence. An entry marked as "avoid modeling" means that the representation is emergent from experience. See Weng 2012 [12] for a review of symbolic models and emergent models and the comparison thereof.

Traditionally, given a task to be executed by the machine, it is the human programmer who understands the task and, based on his understanding, designs a task-specific representation. Depending on different approaches, different techniques are used to produce the mapping from sensory inputs to effector outputs. The techniques used range from direct programming (knowledge-based approach), to learning the parameters (in the parametric model), to genetic search (genetic approach). Although genetic search is a powerful method, the chromosome representations used in artificial genetic search algorithms are task specific.

Using the developmental approach, the tasks that the robot (or human) ends up doing are unknown during the programming time (or conception time), as illustrated in Fig. 1. The ecological conditions that the robot will operate under must be known, so that the programmer can design the body of the robot, including sensors and effectors, suited for the ecological conditions. The programmer may guess some typical tasks that the robot will learn to perform. However, world knowledge is not modeled and only a set of simple reflexes is allowed for the developmental program. During "prenatal" development, internally generated synthetic data can be used to develop the system before birth. For example, the retina may generate spontaneous signals to be used for the prenatal development of the visual pathway. At the "birth" time, the robot's power is turned on. The robot starts to interact with its environment, including its teachers, in real time. The tasks the robot learns are in the mind of its teachers. In order for the later learning to use the skills learned in early learning, a well designed sequence of educational experience is an important practical issue.

4. Learning Types

In the machine learning literature, there have been widely accepted definitions of learning types, such as supervised, unsupervised, and reinforcement learning. However, these conventional definitions are too coarse to describe computational learning through autonomous development. For example, it is difficult to identify any type of learning that is completely unsupervised. Further, the traditional classification of animal learning models, such as classical conditioning and instrumental conditioning, is not sufficient to address computational considerations of every time instant of learning. A definition of a refined classification of learning types is necessary.

We use a variable i to indicate internal task-specific representation imposed by human programmer (called *internal-state imposed* $i = 1$) or not (called *internal-state autonomous* $i = 0$).

We use e to denote autonomy of effector. If the concerned effector is directly guided by the human teacher or other teaching mechanisms for the desired action, we call the situation *action imposed* ($e = 1$). Otherwise, the learning is effector autonomous ($e = 0$).

TABLE II
A COMPARISON OF APPROACHES TO ARTIFICIAL INTELLIGENCE

Approach	Species Architecture	World knowledge	Agent behaviors	Task specific
Knowledge-based	Model	Model	Model	Yes
Learning-based	Model	Parametrically model	Model	Yes
Behavior-based	Model	Avoid modeling	Model	Yes
Genetic	Genetic search	Parametrically model	Model	Yes
Developmental	Parametrically model	Avoid modeling	Minimize modeling	No

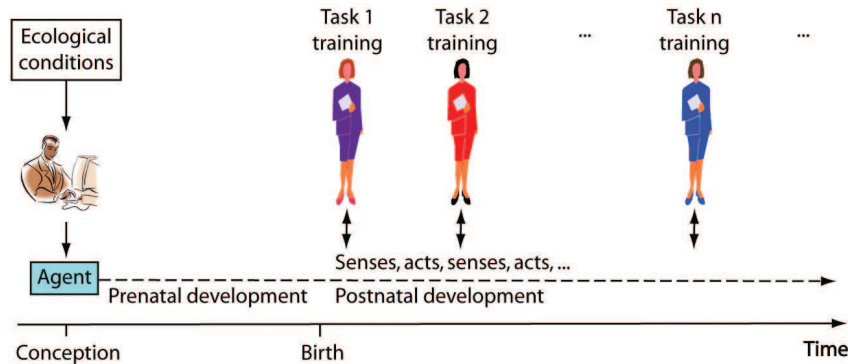


Figure 1. Illustration of the paradigm of developmental agents, inspired by human mental development. No task is given during the programming (i.e., conception) time, during which a general-purpose task-nonspecific developmental program is loaded onto the agent. Prenatal development is used for developing some initial processing pathways in the brain using spontaneous (internally generated) signals from sensors. After the birth, the agent starts to learn an open series of tasks through interactions with the physical world. The tasks that the agent learns are determined after the birth.

We need to distinguish the channels of reward (e.g., sweet and pain sensors) that is available at the birth time, and other channels of reward that are not ready to be used as reward at the birth time (e.g., auditory input “good” or “bad”) but implies a value after a certain amount of development. We define (inborn) biased sensors:

If the machine has a predefined preference pattern to the signals from a sensor at the birth time, this sensor is an (inborn) biased sensor. Otherwise, it is an (inborn) unbiased sensor.

In fact, all the sensors become biased gradually through postnatal experience — the development of the value system. For example, the image of a flower does not give a newborn baby much reward, but the same image becomes pleasant to look at (high value) after the baby has grown up.

We use the third variable b to denote whether a biased sensor is used. If any biased sensor is activated (sensed) during the learning, we called the situation reinforcement ($b = 1$). Otherwise, the learning is called communicative ($b = 0$).

Using these three key factors, any type of learning can be represented by a 3-tuple (i, e, b) , which contains three components i , e , and b , each of which can be either represented by 0 or 1. Thus, there are a total of 8 different 3-tuples, representing a total of 8 different learning types. If we consider ieb as three binary bits of the type index number of learning type, we have 8 types of learning defined in Table III. We can also name each type. For example, Type 0 is state-

TABLE III
EIGHT TYPES OF BIOLOGICAL AND ARTIFICIAL LEARNING

Type (binary)	Internal state	Effector	Biased sensor
0 (000)	Autonomous	Autonomous	Communicative
1 (001)	Autonomous	Autonomous	Reinforcement
2 (010)	Autonomous	Imposed	Communicative
3 (011)	Autonomous	Imposed	Reinforcement
4 (100)	Imposable	Autonomous	Communicative
5 (101)	Imposable	Autonomous	Reinforcement
6 (110)	Imposable	Imposed	Communicative
7 (111)	Imposable	Imposed	Reinforcement

autonomous, effector-autonomous, communicative learning. Type 7 is state-imposable, effector-imposed, reinforcement learning, but it has not been included in the traditional definition of either supervised learning or reinforcement learning. However, this learning is useful when teaching a positive or negative lesson through supervision.

Using three key features, state-imposed, effector-imposed and reinforcement, eight learning types are defined. This refined definition is necessary to understanding various modes of developmental and nondevelopmental learning.

All learning types using a non-developmental learning method corresponding to Types 7 to 4. This is because the task-specific representation is at least partially handcrafted after the task is given. Autonomous mental development uses Types 0 to 3.

5. Brain-Mind Architectures

Weng 2007 [13] proposed a SASE model through which the agent can autonomously learn to think, while the thinking behavior is manifested as internal attention. Attention is a key to emergent intelligence.

5.1. Top-down Attention is Hard

Consider a car in a complex urban street environment. Attention and recognition is a pair of dual-feedback problems. Without attention, recognition cannot do well; recognition requires attended areas (e.g., the car area) for the further processing (e.g., to recognize the car). Without recognition, attention cannot do well; attention requires recognition for guidance of the next fixation (e.g., a possible car area).

1) *Bottom-up attention*: Studies in psychology, physiology, and neuroscience provided qualitative models for bottom-up attention, i.e., attention uses different properties of sensory inputs, e.g., color, shape, and illuminance to extract saliency. Several models of bottom-up attention have been published. The first explicit computational model of bottom-up attention was proposed by Koch & Ullman in 1985 [14], in which a “saliency map” is computed to encode stimuli saliency at every location in the visual scene. More recently, Itti & Koch et al. 1998 [15] integrated color, intensity, and orientation as basic features in multiple scales for attention control. An active-vision system, called NAVIS (Neural Active Vision) by Baker et al. 2001, was proposed to conduct the visual attention selection in a dynamic visual scene [16]. Our SASE model to be discussed next indicates that saliency is not necessarily independent of learning: The top-down process in the previous time instant may affect the current bottom-up saliency.

2) *Top-down attention*: Volitional shifts of attention are also thought to be performed top-down, through spatially defined and feature-dependant controls. Olshausen et al. 1993 [17] proposed a model of how visual attention can be directed to address the position and scale invariance in object recognition, assuming that the position and size information is available from the top control. Tsotsos et al. 1995 [18] implemented a version of attention selection using a combination of a bottom-up feature extraction scheme and a top-down position selective tuning scheme. Rao et al. 2004 [19] described a pair of cooperating neural networks, to estimate object identity and object transformations, respectively. Schill et al. 2001 [20] presented a top-down, knowledge-based reasoning system with a low-level pre-processing where eye movement is to maximize the information about the scene. Deco & Rolls 2004 [21] wrote a model of object recognition that incorporates top-down attention mechanisms on a hierarchically organized set of visual cortical areas. In the above studies, the model of Deco & Rolls 2004 [21] was probably the most biologically plausible, as it incorporates bottom-up and top-down flow into individual neuronal computation, but unfortunately the

top-down connections were disabled during learning and no recognition performance data were reported.

Where-What Networks (WWN) are embodiment of a brain-mind model called Developmental Network (DN). In the Where-What Network 2 (WWN-2) experiment [22] discussed later, we found that the corresponding network that drops the L4-L2/3 laminar structure gave a recognition rate lower than 50%. In other words, a network that treats top-down connection similar to bottom-up connection (like a uniform liquid state machine [23]) is not likely to achieve an acceptable performance.

5.2. Motor Shapes Cortical Areas

On one hand, high-order (i.e., later) visual cortex of the adult brain includes functionally specific regions that preferentially respond to objects, faces, or places. For example, the fusiform face area (FFA) responds to face stimuli (Kanawisher 1997 [24], 1999 [25], Grill-Spector et al. 2004 [26]) and the parahippocampal place area (PPA) responds to place identity (O’Keefe & Dostrovsky 1971 [27], Ekstrom et al. 2003 [28], Bohbot & Corkin 2007 [29]). How does the brain accomplish this feat of localizing internal representation based on meaning? Why is such a representation necessary?

In the cerebral cortex, there is a dense web of anatomically prominent feedback (i.e., top-down) connections (Kennedy & Bullier 1985 [30], Perkel et al. 1986 [31], Felleman & Van Essen 1991, [32], Katz & Callaway 1992 [33], Salin & Bullier 1995 [34], Johnson & Burkhalter 1996 [35]). It has been reported that cortical feedback improves discrimination between figure and background and plays a role in attention and memory (Hupe et al. 1998 [36], Grossberg & Raizada 2000 [37], Sullivan & de Sa [38]). Do feedback connections perform attention? Furthermore, do feedback connections play a role in developing abstract internal representation?

The computational roles of feedback connections in developing meaning-based internal representations have not been clarified in existing studies reviewed above. The Self-Abstractive Architecture next indicates that in the cerebral cortex, each function layer (L4 and L2/3) is a state at this layer. We will show that, unlike the states in POMDP, HMM, Hopfield network and many others, the states in the Self-Abstractive Architecture integrate information from bottom-up inputs (feature inputs), lateral inputs (collaborative context) and top-down inputs (abstract contexts) into a concise continuous vector representation, without the artificial boundaries of a symbolic representation.

5.3. Brain Scale: “Where” and “What” Pathways

Since the work of Ungerleider and Mishkin 1982 [39], [40], a widely accepted description of visual cortical areas is illustrated in Fig. 2 [32], [17]. A ventral or “what” stream that runs from V1, to V2, V4, and IT areas TE0 and TE

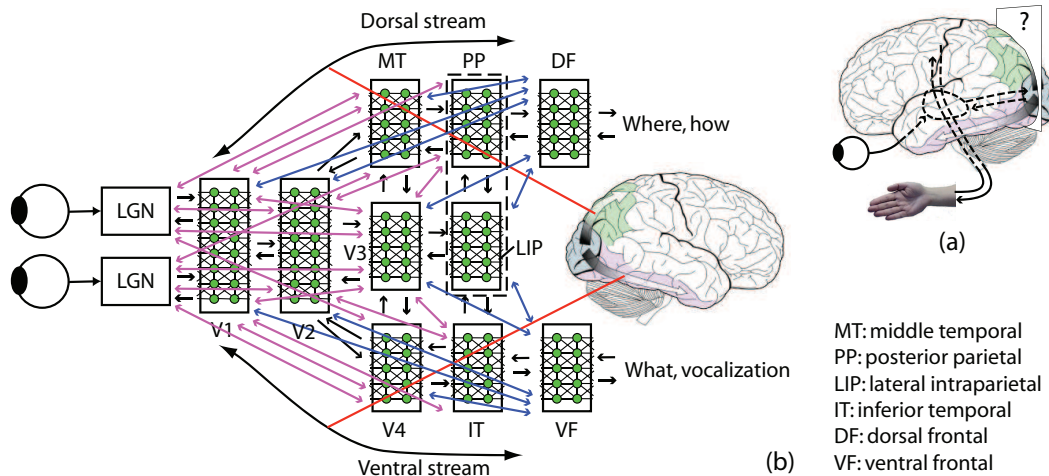


Figure 2. (a) How does the brain generate internal representation? The imaginary page slices the brain to “peek” into its internal representation. The only external sources are sensors and effectors. They are the ports for the brain to exchange information with the external environment. (b) An example of brain connections — the visuomotor streams. It consists of two major streams among others: the dorsal “where and how” stream and the ventral “what” stream. The nature of the processing along each stream is shaped by not only sensory input-output but also the motor input-output. In principle, every area needs connections with all other areas. An area pair that has only weak connection means that this pair has only weak statistical correlation. This diagram only schematically illustrates the cortical connection patterns. Every two-way arrow means two one-way arrows in opposite directions.

computes properties of object identity such as shape and color. A dorsal or “where” stream that runs from V1, to V2, V3, MT and the medial superior temporal areas MST, and on to the posterior parietal cortex (PP) computes properties of the location of the stimulus on the retina or with respect to the animal’s head. Neurons in early visual areas have small spatial receptive field (RFs) and code basic image features; neurons in later areas have large RFs and code abstract features such as behavioral relevance. Selective attention coordinates the activity of neurons to affect their competition and link distributed object representations to behaviors (e.g., see the review by Serences and Yantius 2006 [41]).

With the above rich, suggestive information from neuroscience, I propose that the development of the functions of the “where” and “what” pathways is largely due to:

- 1) Downstream motors. The motor ends of the dorsal pathway that perform position tasks (e.g., stretching an arm to reaching for an apple or a tool), and the motor ends of the ventral pathway that perform type classification and conceptual tasks (e.g., different limbic needs between a food and an enemy);
- 2) Top-down connections. The top-down connections from motor areas that shape the corresponding pathway representations.

Put in a short way, *motor is often abstract*. Any meaning that can be communicated between humans is motorized: spoken, written, hand-signed, etc. Of course, “motor is abstract” does not mean that every stage of every motor action sequence is abstract. However, the sequences of motor actions provide statistically crucial information for the development of internal abstractive representation.

5.4. System views

The system level architecture is illustrated in Fig. 3.

An agent, either biological or artificial can perform regression and classification

Regression: The agent takes a vector as input (a set of receptors). For vision, the input vector corresponds to a retinal image. The output of the network corresponds to motor signals, with multiple components to be active (firing) The brain is a very complex spatiotemporal regressor.

Classification: The agent can perform classification before it has developed sophisticated human language capability to verbally tell us the name of a class. For example, each neuron in the output layer corresponds to a different class.

1) *Two signal sources: sensor and motor:* The brain faces a major challenge as shown in Fig. 2(a). It does not have the luxury of having a human teacher to implant symbols into it, as the brain is not accessible directly to the external human teacher. Thus, it must generate internal representations from the two signal sources: the sensors and the effectors (motors). This challenging goal is accomplished by the brain’s where-what networks schematically illustrated in Fig. 4. The system has two motor areas, the where motor that indicates where the attended object is and the what motor that tells what the attended object is. This specialization of each pathway makes computation of internal representation more effective.

5.5. Pathway Scale: Bottom-up and Top-down

It is known that cortical regions are typically interconnected in both directions [32], [42], [43]. However,

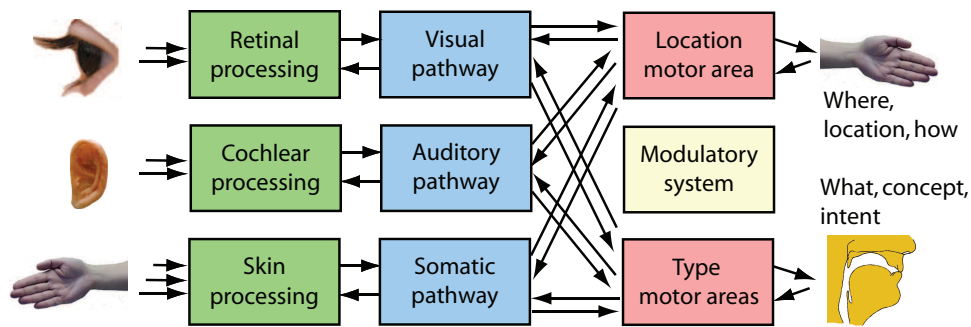


Figure 3. Schematic connections for multimodal integration. Multi-sensory and multi-effector integration is achieved through interactive learning. The modulatory (motivational) system is distributed over the entire brain through different types of neural transmitters. The block in the diagram only indicates the existence of the distributed modulatory system.

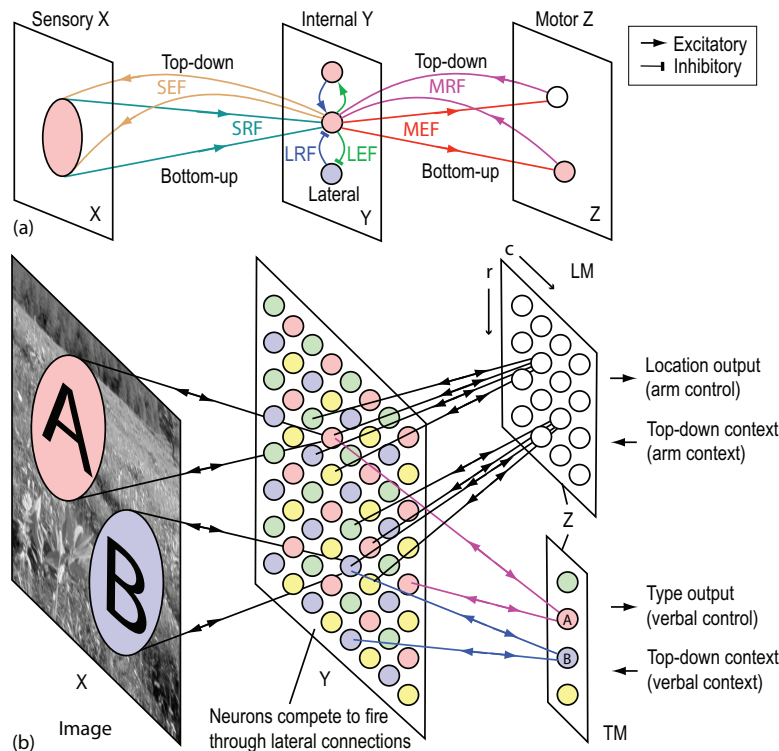


Figure 4. A simple WWN as a schematic developmental model of the brain. (a) The hextuple field for each neuron: SRF, MRF, LRF, SEF, MEF, and LEF, thus are highly recurrent. (b) A simple WWN with four areas (image as the X area, the brain as Y , and LM and TM as the Z area) and its hextuple network representation. Each wire connects if the pre-synaptic and post-synaptic neurons have co-fired. The weight is the frequency of pre-synaptic co-firing when the post-synaptic neuron fires. Within each cortical area, each neuron connects with highly correlated neurons using excitatory connections (e.g., NMDA-ergic) but connect with highly anti-correlated neurons using inhibitory connections (e.g., GABA-ergic). This forces neurons in the same area to detect different features in SRF and MRF. These developmental mechanisms result in the shown connections. Every Y neuron is *location-specific* and *type-specific*, corresponding to an object type (marked by its color) and to a location block (2×2 size each). Each LM neuron is location-specific and type-invariant (more invariance, e.g., lighting-direction invariance, in more mature WWNs). Each TM neuron is type-specific and location-invariant (more invariance in more mature WWNs). Each motor neuron pulls all applicable cases in Y as top-down context. A two-way arrow means two one-way connections. With each area, all the connections within the same area are omitted for clarity.

computational models that incorporate both bottom-up and top-down connections have resisted full analysis [44], [45], [46], [21], [47], [48], [49]. The computational model, illustrated in Fig. 5, provides further details about how each

functional level in cortex takes inputs from the bottom-up signal representation space X and top-down signal representation space Z to generate and update self-organized cortical **bridge representation** space Y . This model further

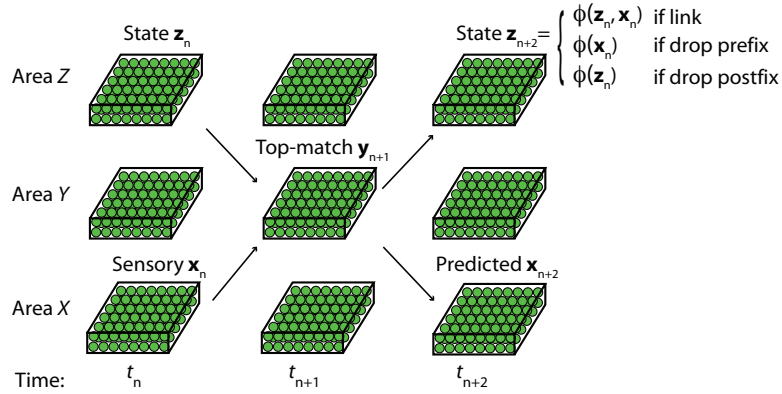


Figure 5. Cortex scale: The spatial SASE network for both spatial processing and temporal processing without dedicated temporal components. At each temporal unit shown above (two time frames), three basic operations are possible: link, drop prefix and drop postfix. After proper training, the TCM is able to attend any possible temporal context up to the temporal sampling resolution.

computationally predicts that a primary reason for the dorsal and ventral pathways to be able to deal with “where” and “what” (or achieving identity and positional invariances [49]), respectively, is that they receive top-down signals that drive their motors.

From where does the forebrain receive teaching signals that supervise its motors? Such supervised-motor signals can be generated either externally (e.g., a child passively learns writing while his teacher manually guides his hand) or internally (e.g., from the trials generated by the spinal cord or the mid brain). As illustrated in Fig. 4, the model indicates that from early to later cortical areas, the neurons gradually increase their receptive field and gradually reduce their effective field as the processing of the corresponding *bridge representations* becomes less sensory and more motoric.

5.6. Cortex Scale: Feature Layers and Assistant Layers

The cerebral cortex contains six layers: layer L1 is the superficial layer and layer L6 is the deep layer. Weng et al. 2008 [50] reasoned that L4 and L2/3 are two feature detection layers as shown in Fig. 5 with L5 assisting L2/3 and L6 assisting L4, in the sense of enabling long range lateral inhibition. Such long range inhibitions encourage different neurons to detect different features. The model illustrated in Fig. 5 was informed by the work of Felleman & Van Essen [32], Callaway and coworkers [43], [42], and others (e.g., [37]). There are no top-down connections from L2/3 to L4, indicating that L4 uses unsupervised learning (U) while L2/3 uses supervised (S) learning. Weng et al. 2008 [50] reported that such a *paired hierarchy* USUS led to better recognition rates than the unpaired SSSS alternative.

5.7. Level Scale: the Dually Optimal CCI LCA

As shown in Fig. 5, given parallel input space consisting of the bottom-up space X and the top-down input space Z , represented as $X \times Z$, the major developmental goal of each cortical level (L4 or L2/3 as two representative levels

of each area in Fig. 5) is to have different neurons in the level to detect different features, but nearby neurons should detect similar features.

Each feature level faces two pairs of conflicting criteria which are probably implicit during biological evolution: (1) The spatial pair: with its limited number of neurons, the level must learn the best internal representation from the environment while keeping a stable long-term memory. (2) The spatiotemporal pair: with its limited child time for learning, the level must not only learn the best representation but also learn quickly without forgetting important mental skills acquired long time ago. The sparse coding principle [51] is useful to address the first pair: Allowing only a few neurons (best matched) to fire and update. Other neurons in the level are long-term memory because they are not affected. In other words, in each cortical region, only closely related mental skills are replaced each time. Therefore, the role of each neuron as working memory or long-term memory is dynamic, depending on the feature match (i.e., binding) with the input, as shown in Fig. 6. However, this rough idea is not sufficient for optimality.

The cortex inspired Candid Incremental Covariance-free (CCI) Lobe Component Analysis (LCA) [52], [53] has the desired dual optimality: spatial and spatiotemporal, as illustrated in Fig. 6. CCI LCA models optimal self-organization by a cortical level with a limited resource: c neurons. The cortical level takes two parallel input spaces: the bottom-up space X and top-down space Z denoted as $P = X \times Z$ as illustrated by Fig. 5. Each input vector is then denoted as $\mathbf{p} = (\mathbf{x}, \mathbf{z})$ where $\mathbf{x} \in X$ and $\mathbf{z} \in Z$. CCI LCA computes c feature vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c$. Associated with these c feature vectors is a partition of the input space P into c disjoint regions R_1, R_2, \dots, R_c , so that the input space P is the union of all these regions. For the optimal distribution of neuronal resource, we consider that each input vector \mathbf{p} is represented by the winner feature \mathbf{v}_j which has the highest response r_j :

$$j = \arg \max_{1 \leq i \leq c} r_i$$

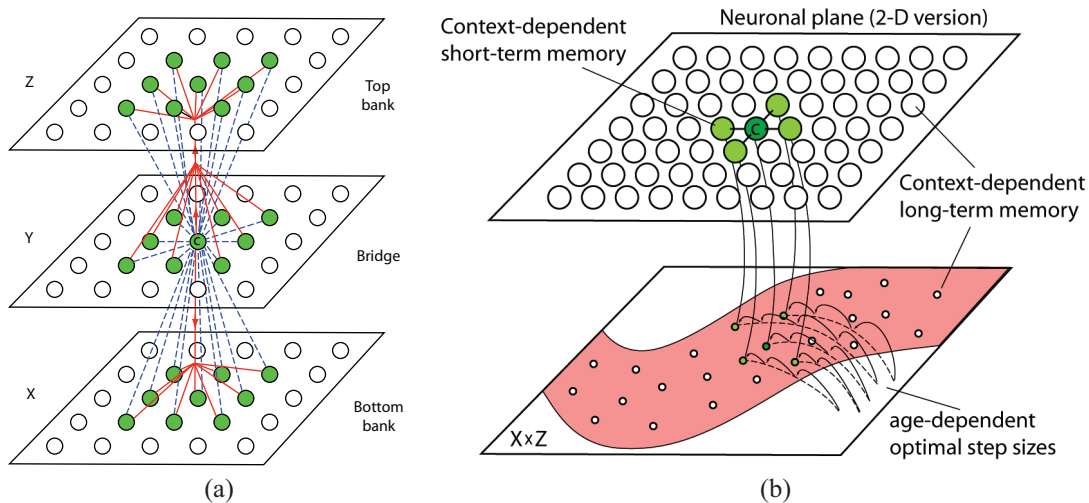


Figure 6. (a) The default connection pattern of every neuron in the brain. The area Y is a bridge for its two banks X and Z . The connections are local but two-way. Blue: neuronal input; red: axonal output. For each feature neuron (e.g., pyramidal neuron) in the brain, some near neurons (e.g., green for the center neuron) are connected to the neuron by excitatory connections (for prediction) and some far neurons (white ones) are connected to the center neuron by inhibitory connections (competition resulting in detection of different features by different neurons). Neurons that are not connected with the center neuron c are not considerably correlated or anti-correlated with it. (b) Cell-centered learning. The upper layer indicates the positions for the neurons in the same area: firin neurons are (context-dependent) working memory and those do not fir are (context dependent) long-term memory. The lower layer indicates the very high dimensional input space ($X \times Z$) of the area Y . The purple area in $X \times Z$ indicates the manifold of the input distribution. The connection curve from the upper neuron and lower small circle indicates the correspondence between the upper neuron and the feature that it detects. The neuronal weight vectors must quickly move to this manifold as the inputs are received and further the density of the neurons in the purple area should reflect the density of the input distribution. The challenge of learning and fast adaptation at various maturation stages of development is as follows: The updating trajectory of every neuron is a highly nonlinear trajectory. The statistical efficiency theory for neuronal weight update (amnesic average) results in the nearly minimum error in each age-dependent update, meaning not only the direction of each update is nearly optimal, but also every step length.

where r_i is the projection of input \mathbf{p} onto the normalized feature vector \mathbf{v}_i : $r_i = \mathbf{p} \cdot (\mathbf{v}_i / \|\mathbf{v}_i\|)$. The form of approximation of \mathbf{p} is represented by $\hat{\mathbf{p}} = r_i \mathbf{v}_i / \|\mathbf{v}_i\|$ and the error of this representation for \mathbf{p} is $e(\mathbf{p}) = \|\hat{\mathbf{p}} - \mathbf{p}\|$.

1) *Spatial optimality*: The spatial optimality requires that the spatial resource distribution in the cortical level is optimal in minimizing the representational error. For this optimality, the cortical-level developmental program modeled by CCI LCA computes the best feature vectors $V = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c)$ so that the expected square approximation error $\| \hat{\mathbf{p}}(V) - \mathbf{p} \|^2$ is statistically minimized:

$$V^* = (\mathbf{v}_1^*, \mathbf{v}_2^*, \dots, \mathbf{v}_c^*) = \arg \min_V E \| \hat{\mathbf{p}}(V) - \mathbf{p} \|^2. \quad (1)$$

where E denotes statistical expectation. The minimum error means the optimal allocation of limited neuronal resource: frequent experience is assigned with more neurons (e.g., human face recognition) but rare experience is assigned with fewer neurons (e.g., flower recognition for a nonexpert). This optimization problem must be computed incrementally, because the brain receives sensorimotor experience incrementally. As the feature vectors are incrementally updated from experience, the winner neurons for the past inputs are not necessarily the same if past inputs are fed into the brain again (e.g., parents' speech when their baby was little is heard again by the grown-up baby). However, while the

feature vectors are stabilized through extensive experience, the partition of the input space becomes also stable. Given a fixed partition, it has been proved that the best feature set V^* consists of the c local first principal component vectors, one for each region R_i . The term "local" means that the principal component vector for region R_i only considers the samples that fall into region R_i . As the partition is tracking a slowly changing environment (e.g., while the child grows up), the optimal feature set V^* tracks the slowly changing input distribution (called nonstationary random process).

Intuitively speaking, the spatial optimality means that with the same cortical size, all the children will eventually perform at the best level allowed by the cortical size. However, to reach the same mental skill level one child may require more teaching than another. The spatiotemporal optimality is deeper. It requires the best performance for every time t . That is, the child learns quickest allowed by the cortical size at every stage of his age.

2) *Temporal optimality*: The spatiotemporal optimality gives optimal step sizes of learning. Each neuron takes response weighted input $\mathbf{u}(t) = r(t)\mathbf{x}(t)$ at time t (i.e., Hebbian increment). From the mathematical theory of statistical efficiency, CCI LCA determines the optimal feature vectors $V^*(t) = (\mathbf{v}_1^*(t), \mathbf{v}_2^*(t), \dots, \mathbf{v}_c^*(t))$ for every time instant t starting from the conception time $t = 0$, so that the distance

from $V^*(t)$ to its target V^* is minimized:

$$V^*(t) = \arg \min_{V(t)} E \|V(t) - V^*\|^2. \quad (2)$$

CCI LCA aims at this deeper optimality — the smallest average error from the starting time (birth of the network) up to the current time t , among all the possible estimators, under some regularity conditions. A closed form solution was found that automatically gives the optimal retention rate and the optimal learning rate (i.e., step size) at each synaptic update [52]

In summary, the spatial optimality leads to Hebbian incremental direction: response weighted pre-synaptic activity (*rp*). The deeper spatiotemporal optimality leads to the best learning rates, automatically determined by the update age of each neuron. This is like different racers racing on a rough terrain along a self-determined trajectory toward an unknown target. The spatially optimal racers, guided by Hebbian directions, does not know step sizes. Thus, they cover other trajectories that require more steps. The spatiotemporally optimal racer, CCI LCA, correctly estimates not only the optimal direction at every step as illustrated in Fig. 6, but also the optimal step size at every step. In our experiments, CCI LCA out performed the Self-Organization Map (SOM) algorithm by an order (over 10 times) in terms of percentage distance covered from the initial estimate to the target. This work also predicts cell-age dependent plasticity schedule which needs to be verified biologically.

6. Temporal Processing

Mauk & Buonomano 2004 [54] argued that the brain uses its intrinsic mechanisms to deal with time, and it does not have explicit delay lines and does not have a global clock. Drew & Abbott 2006 [55] proposed that the gradual change in the level of membrane potential inside a neuron may record some temporal information. However, this seems also not sufficient and robust for long time dependency, as argued by Ito et al. 2008 [56]. How the cortex deals with long time context has been elusive, especially considerably beyond around 30 ms modeled by Spike Timing-Dependent Plasticity (STDP) [57], [58].

To discuss how the brain deal with time, it is beneficial to discuss the Finite Automata (FA), also called finite state machines. In the symbolic world with vocabulary Σ , a sensory sequence of a life from conception up to the current time is a string $x \in \Sigma^*$, where Σ^* denotes the set of all strings of finite lengths. If the required skill set S for an agent can be defined by a partition of Σ^* into c equivalent sets:

$$\Sigma^* = [q_1] \cup [q_2] \cup \dots \cup [q_c]$$

where q_i is the equivalent perceptive-cognitive-behavior state, plus the state transitions among these states in the form of $q \xrightarrow{\sigma} q'$.

A Developmental Network (DN) is a generalization of Where-What Network whose motor area can represent any

state: location, type, or any other cognitive state or behavior state or both. It has been established that given any FA, a DN can simulate any FA [59].

From the theory of FA [60], [61], the above conditions for state partition and state transitions are equivalent to the existence of the corresponding FA. It can be proved that a DN can simulate any FA. Further, during learning, the DN needs to learn every state transition of the FA only once.

Although the above discussion does not explicitly mention time, all the time properties are imbedded in the concept of equivalent states, such as the skills to estimate time duration, deal with time warping, conduct arbitrary temporal attention, and deal with context of any temporal length. Those properties are explicitly proved in Weng 2010 [62].

This completeness is symbolic since the environment of FA is symbolic. A major difference between a symbolic world and the real world is that the latter must deal with attention in new unobserved environments. Then, how a learned DN performs in a new environment and it can generalize depends on factors such as the similarity between a new environment and the learned environments. It is interesting here to utilize the power of the automata theory and the benefit of having mapped any arbitrary but static and symbolic FA to a general purpose but dynamic and emergent DN. Using the capability of DN to learn any Finite Automaton (FA), it can be proved that DN can abstract at least as well as FA-based symbolic models [63]. This addresses the correct criticisms by Marvin Minsky [64] and others in that traditional neural networks do not abstract well.

The DN has two types action and two types of sensing proposed by Weng 2007 [13]. Internal action is called *internal attention* here and external action is called *external behaviors*. Internal sensing is accomplished by autonomous internal wiring, *internal competition* and emergent representation. External sensing is realized by the sensors of the DN and its effectors (e.g., saccades, locomotion and limb actions which change what is sensed).

The external behaviors should include expressions about perception and cognition (e.g., speak), as well as actual actions acting on the external physical world. Although those skills are very different on surface, the DN model treats them in a unified way so that a single unified developmental program (DP) of a DN-based brain model potentially is sufficient to model complex skills. It is still unknown at this time what substantial limitations such a model has in modeling the human brain-mind.

7. Experiments

This introduction is theoretical. Experimental discussion is beyond the scope of this introduction. Some experimental results guided by this theory have published elsewhere. A visual WVN-2 [22] reached 92.5% in object recognition rate and 1.5 pixels in average position error with 75% of the area in each image filled with unknown natural

backgrounds. The WWN-3 [65] has shown a capability to deal with multiple learned objects in complex backgrounds. The user can specify either goal (location or type) and WWN-3 reports the reasoning results for other concepts (from location goal to type, or from type goal to location). WWN-4 [66] investigated internal connection constraints — it showed that deep learning as a cascade of areas between the sensory port and the motor port does not perform as well as shallow learning — multiple areas each having connections with both ports. WWH-5 [67] deals with objects with multiple scales in complex backgrounds.

WWN equipped with automatic synapse maintenance [68] demonstrated that how each neuron in WWN can segment object along its natural contour to get rid of the background in its sensory receptive field. A temporal version for visual recognition [69] has reached an almost perfect recognition rate for centered objects viewed from any of the 360° object views. The stereo version of WWN [70] has shown that pre-screening is truly necessary for the temporal mechanisms to improve the result.

A text processing version [71] has been tested for part-of-speech tagging problem (assigning the words in a sentence to the corresponding part of speech, about 99% correct); and chunking (grouping sequences of words together and classify them by syntactic labels, about 96% success rate) using text corpus from the Wall Street Journal. A version of generative DN [72] has been shown to transfer skills from words through their temporal association, such as member-to-class transfer and member-to-member transfer.

8. Summary

The material in this paper outlines a series of tightly intertwined advances we recently made in understanding and modeling how the brain-mind develops and works. The grand picture of the biological brain-mind, although controversial and subject to further refinement seems getting increasingly clear. However, much work is needed to further verify whether such a grand model can generate rich behaviors that are consistent with known biological data and produce machine capabilities that traditional agent models cannot. A critical key that many computational models about cortical signal processing is *autonomous development*. Without understanding and modeling development, such models are not only incomplete, but also misleading, incorrect, and computationally inscalable. Developmental robots and machines urgently need industrial electronics for real-time, brain scale computation and learning. This need is here and now. This raises a great challenge for the field of industrial electronics, but an exciting future for understanding natural and artificial intelligence as well.

REFERENCES

[1] W. K. Purves, D. Sadava, G. H. Orians, and H. C. Heller. *Life: The Science of Biology*. Sinauer, Sunderland, MA, 7 edition, 2004.

[2] J. Weng. Task muddiness, intelligence metrics, and the necessity of autonomous mental development. *Minds and Machines*, 19(1):93–115, 2009.

[3] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Upper Saddle River, New Jersey, 1995.

[4] J. Weng, N. Ahuja, and T. S. Huang. Learning recognition and segmentation of 3-D objects from 2-D images. In *Proc. IEEE 4th Int'l Conf. Computer Vision*, pages 121–128, May 1993.

[5] J. Weng, N. Ahuja, and T. S. Huang. Learning recognition and segmentation using the Cresceptron. *International Journal of Computer Vision*, 25(2):109–143, Nov. 1997. Cited early versions in IJCNN 1992 and ICCV 1993.

[6] J. L. McClelland. The interaction of nature and nurture in development: A parallel distributed processing perspective. In P. Bertelson, P. Eelen, and G. d'Ydewalle, editors, *International Perspectives on Psychological Science*, volume 1: Leading Themes, pages 57–88. Erlbaum, Hillsdale, New Jersey, 1994.

[7] J. L. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. *Rethinking Innateness: A connectionist perspective on development*. MIT Press, Cambridge, Massachusetts, 1997.

[8] S. Quartz and T. J. Sejnowski. The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences*, 20(4):537–596, 1997.

[9] T. S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, IL, second edition, 1970.

[10] J. Weng and I. Stockman. Autonomous mental development: Workshop on development and learning. *AI Magazine*, 23(2):95–98, 2002.

[11] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291(5504):599–600, 2001.

[12] J. Weng. Symbolic models and emergent models: A review. *IEEE Trans. Autonomous Mental Development*, 4(1):29–53, 2012.

[13] J. Weng. On developmental mental architectures. *Neurocomputing*, 70(13-15):2303–2323, 2007.

[14] C. Koch and S. Ullman. Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, 4:219–227, 1985.

[15] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, November 1998.

[16] G. Backer, B. Mertsching, and M. Bollmann. Data- and model-driven gaze control for an active-vision system. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(12):1415–1429, December 2001.

[17] B. A. Olshausen, C. H. Anderson, and D. C. Van Essen. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, 13(11):4700–4719, 1993.

[18] J. K. Tsotsos, S. M. Culhane, W. Y. K. Wai, Y. Lai, N. Davis, and F. Nuflo. Modeling visual attention via selective tuning. *Artificial Intelligence*, 78:507–545, 1995.

[19] R. P. N. Rao and D. H. Ballard. Probabilistic models of attention based on iconic representations and predictive coding. In L. Itti, G. Rees, and J. Tsotsos, editors, *Neurobiology of Attention*. Academic Press, New York, 2004.

[20] K. Schill, E. Umkehrer, S. Beinlich, G. Krieger, and C. Zetzsche. Scene analysis with saccadic eye movements: Top-down and bottom-up modeling. *Journal of Electronic Imaging*, 10(1):152–160, 2001.

[21] G. Deco and E. T. Rolls. A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research*, 40:2845–2859, 2004.

[22] Z. Ji and J. Weng. WWN-2: A biologically inspired neural network for concurrent visual attention and recognition. In *Proc. IEEE Int'l Joint Conference on Neural Networks*, pages +1–8, Barcelona, Spain, July 18-23 2010.

[23] M. Rabinovich, R. Huerta, and G. Laurent. Transient dynamics for neural processing. *Science*, 321:48–50, 2008.

[24] N. Kanwisher, J. McDermott, and M. M. Chun. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11):4302–4311, 1997.

[25] N. Kanwisher, D. Stanley, and A. Harris. The fusiform face area is selective for faces not animals. *NeuroReport*, 10(1):183–187, 1999.

- [26] K. Grill-Spector, N. Knouf, and N. Kanwisher. The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, 7(5):555–562, 2004.
- [27] J. O’Keefe and J. Dostrovsky. The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1):171–175, 1971.
- [28] A. Ekstrom, M. Kahana, J. Caplan, T. Fields, E. Isham, E. Newman, and I. Fried. Cellular networks underlying human spatial navigation. *Nature*, 425:184–188., 2003.
- [29] V. D. Bohbot and S. Corkin. Posterior parahippocampal place learning in h.m. *Hippocampus*, 17:863–872, 2007.
- [30] H. Kennedy and J. Bullier. A double-labelling investigation of the afferent connectivity to cortical areas v1 and v2 of the macaque monkey. *Journal of Neuroscience*, 5(10):2815–2830, 1985.
- [31] D. J. Perkel, J. Bullier, and H. Kennedy. Topography of the afferent connectivity of area 17 of the macaque monkey. *Journal of Computational Neuroscience*, 253(3):374–402, 1986.
- [32] D. J. Felleman and D. C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1:1–47, 1991.
- [33] L. C. Katz and E. M. Callaway. Development of local circuits in mammalian visual cortex. *Annual Review of Neuroscience*, 15:31–56, 1992.
- [34] P. A. Salin and J. Bullier. Corticocortical connections in the visual system: structure and function. *Physiological Review*, 75(1):107–154, 1995.
- [35] R. R. Johnson and A. Burkhalter. Microcircuitry of forward and feedback connections within rat visual cortex. *Journal of Comparative neurology*, 368(3):383–398, 1996.
- [36] J. M. Hupe, A. C. James, B. R. Payne, S. G. Lomber, P. Girard, and J. Bullier. Cortical feedback improves discrimination between figure and background by v1, v2 and v3 neurons. *Nature*, 394:784–787, Aug. 20 1998.
- [37] S. Grossberg and R. Raizada. Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research*, 40:1413–1432, 2000.
- [38] T. J. Sullivan and V. R. de Sa. A model of surround suppression through cortical feedback. *Neural Networks*, 19:564572, 2006.
- [39] L. G. Ungerleider and M. Mishkin. Two cortical visual systems. In D. J. Ingel, editor, *Analysis of visual behavior*, pages 549–586. MIT Press, Cambridge, MA, 1982.
- [40] M. Mishkin, L. G. Ungerleider, and K. A. Macko. Object vision and space vision: Two cortical pathways. *Trends in Neuroscience*, 6:414–417, 1983.
- [41] J. T. Serences and S. Yantis. Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences*, 10(1):38–45, 2006.
- [42] A. K. Wiser and E. M. Callaway. Contributions of individual layer 6 pyramidal neurons to local circuitry in macaque primary visual cortex. *Journal of neuroscience*, 16:2724–2739, 1996.
- [43] E. M. Callaway. Local circuits in primary visual cortex of the macaque monkey. *Annual Review of Neuroscience*, 21:47–74, 1998.
- [44] J. Moran and R. Desimone. Selective attention gates visual processing in the extrastriate cortex. *Science*, 229(4715):782–784, 1985.
- [45] L. Itti and C. Koch. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2:194–203, 2001.
- [46] T. S. Lee and D. Mumford. Hierarchical bayesian inference in the visual cortex. *J. Opt. Soc. Am. A*, 20(7):1434–1448, 2003.
- [47] M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc. National Academy of Sciences U S A*, 103(26):10046–10051, 2006.
- [48] T. J. Buschman and E. K. Miller. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315:1860–1862, 2007.
- [49] Z. Ji, J. Weng, and D. Prokhorov. Where-what network 1: “Where” and “What” assist each other through top-down connections. In *Proc. IEEE Int’l Conference on Development and Learning*, pages 61–66, Monterey, CA, Aug. 9-12 2008.
- [50] J. Weng, T. Luwang, H. Lu, and X. Xue. Multilayer in-place learning networks for modeling functional layers in the laminar cortex. *Neural Networks*, 21:150–159, 2008.
- [51] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, June 13 1996.
- [52] J. Weng and N. Zhang. Optimal in-place learning and the lobe component analysis. In *Proc. IEEE World Congress on Computational Intelligence*, pages +1–8, Vancouver, BC, Canada, July 16-21 2006.
- [53] J. Weng and M. Luciw. Dually optimal neuronal layers: Lobe component analysis. *IEEE Trans. Autonomous Mental Development*, 1(1):68–85, 2009.
- [54] M. D. Mauk and D. V. Buonomano. The neural basis of temporal processing. *Annual Review of Neuroscience*, 27:307–340, 2004.
- [55] P. J. Drew and L. F. Abbott. Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *The National Academy of Sciences of the USA*, 103(23):8876–8881, 2006.
- [56] I. Ito, R. C. Ong, B. Raman, and M. Stopfer. Sparse odor representation and olfactory learning. *Nature Neuroscience*, 11(10):1177–1184, 2008.
- [57] G. Bi and M. Poo. Synaptic modification by correlated activity: Hebb’s postulate revisited. *Annual Review of Neuroscience*, 24:139–166, 2001.
- [58] Y. Dan and M. Poo. Spike timing-dependent plasticity: From synapses to perception. *Physiological Review*, 86:1033–1048, 2006.
- [59] J. Weng. A 5-chunk developmental brain-mind network model for multiple events in complex backgrounds. In *Proc. Int’l Joint Conf. Neural Networks*, pages 1–8, Barcelona, Spain, July 18-23 2010.
- [60] J. E. Hopcroft, R. Motwani, and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Boston, MA, 2006.
- [61] J. C. Martin. *Introduction to Languages and the Theory of Computation*. McGraw Hill, Boston, MA, 3rd edition, 2003.
- [62] J. Weng. A general purpose brain model for developmental robots: The spatial brain for any temporal lengths. In *Proc. Workshop on Bio-Inspired Self-Organizing Robotic Systems, IEEE Int’l Conference on Robotics and Automation*, pages +1–6, Anchorage, Alaska, May 3-8 2010.
- [63] J. Weng. Why have we passed “neural networks do not abstract well”? *Natural Intelligence: the INNS Magazine*, 1(1):13–22, 2011.
- [64] M. Minsky. Logical versus analogical or symbolic versus connectionist or neat versus scruffy. *AI Magazine*, 12(2):34–51, 1991.
- [65] M. Luciw and J. Weng. Where What Network 3: Developmental top-down attention with multiple meaningful foregrounds. In *Proc. IEEE Int’l Joint Conference on Neural Networks*, pages 4233–4240, Barcelona, Spain, July 18-23 2010.
- [66] M. Luciw and J. Weng. Where What Network 4: The effect of multiple internal areas. In *Proc. IEEE 9th Int’l Conference on Development and Learning*, pages 311–316, Ann Arbor., August 18-21 2010.
- [67] X. Song, W. Zhang, and J. Weng. Where-what network 5: Dealing with scales for objects in complex backgrounds. In *Proc. Int’l Joint Conference on Neural Networks*, pages 2795–2802, San Jose, CA, July 31 - August 5 2011.
- [68] Y. Wang, X. Wu, and J. Weng. Synapse maintenance in the where-what network. In *Proc. Int’l Joint Conference on Neural Networks*, pages 2823–2829, San Jose, CA, July 31 - August 5 2011.
- [69] M. Luciw, J. Weng, and S. Zeng. Motor initiated expectation through top-down connections as abstract context in a physical world. In *IEEE Int’l Conference on Development and Learning*, pages +1–6, Monterey, CA, Aug. 9-12 2008.
- [70] M. Solgi and J. Weng. Developmental stereo: Emergence of disparity preference in models of visual cortex. *IEEE Trans. Autonomous Mental Development*, 1(4):238–252, 2009.
- [71] J. Weng, Q. Zhang, M. Chi, and X. Xue. Complex text processing by the temporal context machines. In *Proc. IEEE 8th Int’l Conference on Development and Learning*, pages +1–8, Shanghai, China, June 4-7 2009.
- [72] K. Miyan and J. Weng. WWN-Text: Cortex-like language acquisition with What and Where. In *Proc. IEEE 9th Int’l Conference on Development and Learning*, pages 280–285, Ann Arbor, August 18-21 2010.