

Chapter 15

Creating New Informational Primitives in Minds and Machines

Peter Cariani

Abstract Creativity involves the generation of useful novelty. Two modes of creating novelty are proposed: via new combinations of pre-existing primitives (combinatoric emergence) and via creation of fundamentally new primitives (creative emergence). The two modes of creativity can be distinguished by whether the changes still fit into an existing framework of possibility, or whether new dimensions in an expanded interpretive framework are needed. Although computers are well suited to generating new combinations, it is argued that computations within a framework cannot produce new primitives for that framework, such that non-computational constructive processes must be utilised to expand the frame. Mechanisms for combinatoric and creative novelty generation are considered in the context of adaptively self-steering and self-constructing goal-seeking percept-action devices. When such systems can adaptively choose their own sensors and effectors, they attain a degree of epistemic autonomy that allows them to construct their own meanings. A view of the brain as a system that creates new neuronal signal primitives that are associated with new semantic and pragmatic meanings is outlined.

15.1 Introduction

Open-endedness is an important goal for designing systems that can autonomously find new and unexpected solutions to combinatorically-complex and ill-defined problems. Classically, issues of open-ended generation of novelty in the universe have come under the rubric of the problem of emergence.

In this discussion we distinguish two general modes of creating novelty: *combinatoric emergence* and *creative emergence*. In combinatoric emergence new combinations of existing primitives are constructed, whereas in creative emergence entirely new primitives are created anew. Although combinatoric systems may differ in numbers of possible combinations, their set of possibilities is closed. Creative systems, on the other hand, have open-sets of possibilities because of the partial or

P. Cariani (✉)

Department of Otology & Laryngology, Harvard Medical School, Boston, MA, USA
e-mail: cariani@mac.com

ill-defined nature of the space of possible primitives. The dual, complementary conceptions provide two modes for describing and understanding change and creativity: as the unfolding consequences of fixed combinatorial rules on bounded sets of predefined primitives or as the effects of new covert processes and interactions that come into play over time to provide new effective dimensional degrees of freedom.

We face several related problems. We want to know how to recognise creative novelty when it occurs (the *methodological problem*). We also want to understand the creative process in humans and other systems (the *scientific problem*) such that creativity in human-machine collaborations can be enhanced and semi-autonomous, creative devices can be built (the *design problem*).

The methodological problem can be solved by the “emergence-relative-to-a-model” approach in which an observer forms a model of the behaviour of a system (Sect. 15.4). Novelty and creativity are inherently in the eye of the observer, i.e. relative to some model that specifies expected behaviours amongst possible alternatives. If the behaviour changes, but it can still be predicted or tracked in terms of the basic categories or state set of the model, one has rearrangement of trajectories of existing states (combinatorial creativity). If behaviour changes, but in a manner that requires new categories, observables, or states for the observer to regain predictability, then one has the creation of new primitives (emergent creativity).

Solution of the scientific problem of creativity requires a clear description of what creativity entails in terms of underlying generative and selective processes. Creativity exists in the natural world on many levels, from physical creation (particles, elements, stars, galaxies) through the origins and evolution of life (multicellularity, differentiated tissues, circulatory, nervous, and immune systems) to concept formation in brains and new modes of social organisation. What facilitating conditions and organisations lead to such creativity? In biological evolutionary contexts the main underlying mechanisms are Darwinian processes of genetic inheritance with variation/recombination, genetically-steered phenotypic construction, and selection by differential survival and reproduction. On the other hand, in neural contexts that support creative learning processes the mechanisms appear to involve more directed, Hebbian stabilisations of effective neural connectivities and signal productions.

Ultimately we seek to build artificial systems that can enhance human creativity and autonomously create new ideas that we ourselves unaided by machines would never have discovered. This will entail designing mechanisms for combinatorial generation and for creation of new primitives. Essentially all adaptive, trainable machines harness the power of combinatorial spaces by finding ever better combinations of parameters for classification, control, or pattern-generation. On the contemporary scene, a prime example is the genetic algorithm (Holland 1975; 1998), which is a general evolutionary programming strategy (Fogel et al. 1966) that permits adaptive searching of high-dimensional, nonparametric combinatorial spaces.

Unfortunately, very few examples of artificial systems capable of emergent creativity are yet to be found. For the most part, this is due to the relative ease and economy with which we humans, as opposed to machines, can create qualitatively new solutions. We humans remain the pre-eminent generators of emergent creativity

on our planet. It is also due in part to the primary reasons that we create machines—to carry out pre-specified actions reliably and efficiently. We usually prefer our devices to act predictably, to carry out actions we specify, rather than to surprise us in some fundamental way. In contrast, we expect our artists, designers, and scientists to continually surprise us.

Nonetheless, creatively emergent artificial systems are possible and even desirable in some contexts. In Sect. 15.3 we consider the problem of creativity in the context of adaptive goal-seeking percept-action systems that encapsulate the functional organisations of animals and robots (Figs. 15.2, 15.3, 15.6). Such systems carry out operations of measurement (via sensors), action (via effectors), internal coordination (via computational mappings, memory), steering (via embedded goals), and self-construction (via mechanisms for plastic modification). We then discuss the semiotics of these operations in terms of syntactics (relations between internal informational sign-states), semantics (relations between sign-states and the external world), and pragmatics (relations between sign-states and internal goals). New primitive relations can be created in any of these realms (Table 15.1).

We are already quite adept at creating ever more powerful computational engines, and we can also construct robotic devices with sensors, effectors, and goal-directed steering mechanisms that provide them with fixed, pre-specified semantics and pragmatics. The next step is to design machines that can create new meanings for themselves. What is needed are strategies for creating new primitive semantic and pragmatic linkages to existing internal symbol states.

Three basic strategies for using artificial devices to create new meanings and purposes present themselves:

1. *via new human-machine interactions* (mixed human-machine systems in which machines provoke novel insights in humans who then provide new interpretations for machine symbols),
2. *via new sensors and effectors on an external world* (epistemically-autonomous evolutionary robots that create their own external semantics), and
3. *via evolving internal analog dynamics* (adaptive self-organisation in mixed analog-digital devices or biological brains in which new internal linkages are created between internal analog representations that are coupled to the external world and goal-directed internal decision states).

The first strategy uses machines to enhance human creative powers, and arguably, most current applications of computers to creativity in the arts and sciences involve these kinds of human-machine collaborations. But the processes underlying human thought and creativity in such contexts are complex and ill-defined, and therefore difficult to study by observing overt human behaviour.

The second and third strategies focus on building systems that are capable of emergent creativity in their own right. In Sect. 15.3 and Sect. 15.5 respectively, we outline a basic accounts of how new primitives might arise in adaptive percept-action systems of animals and robots (emulating emergence in biological evolution) and how new neural signal primitives might arise in brains (emulating creative processes in individual humans and animals). Combinatoric and creative emergence is

first considered in the framework of a taxonomy of adaptive, self-constructing cybernetic robotic percept-action systems. One can then also consider what such an open-ended functional framework might mean for envisioning new kinds of neural networks that are capable of forming novel internal informational primitives. In this framework, adaptively-tuned neuronal assemblies function as self-constructed internal sensors and signal generators, such that new signal types associated with new concepts can be produced. The new signals then serve as internal semantic tags that function as annotative additions to the input signals that evoked their production. Emergence of new signal types in such a system increases the effective dimensionality of internal signal spaces over time, thus bringing new conceptual primitives into action within the system.

15.2 Emergence and Creativity

Emergence concerns the means by which novelty arises in the world. Intuitively, emergence is the process by which new, more complex order arises from a simpler or more predictable preceding situation. As such, images of birth, development, and evolution infuse our notions of emergence. These images provide intuitive explanations for how novelty, spontaneity, and creativity are possible and how complex organisations arise and become further elaborated.

All around us we see the complex organisations that are the emergent products of biological, psychological and social processes, and as a result, our current discourses on emergence encompass a wide range of phenomena. Novelty appears in the form of new material structures (thermodynamic emergence), formal structures (computational emergence), biological structures and functions (emergent evolution), scientific theories (emergence vs. reduction), modelling relations in observers, percepts, ideas, notational systems, and economic and social relations. Novelty and innovation are integral processes in natural and social worlds, and are coming to play ever-larger roles in artificial worlds as well.

Two fundamental kinds of emergent novelty can be distinguished, which we can call *combinatoric emergence* and *creative emergence*. Lloyd Morgan (1931) in his book “Emergent Evolution” made a similar distinction, labelling new combinations “resultants” and new primitives “emergents”. This distinction became central to my work on epistemology and evolutionary robotics, which developed an operational systems-theoretic methodology for distinguishing one process from the other (Cariani 1989). Some of my earliest inspirations came from considering the nature of novelty in biological evolution, where creation of new combinations of existing genetic alternatives and refinements of existing functions (“microevolution”) can be contrasted with creation of entirely new genes, species, morphologies, and functions (“macroevolution”). The combinatoric/creative distinction also parallels Margaret Boden’s division of explorative vs. transformational creativity (Boden 1990a; 1994; 1994b; 2006).

The two kinds of combinatoric and creative novelty reflect different deeply divergent conceptions of order and its origins, “order-from-order” vs. “order-from-noise”

(Piatelli-Palmarini 1980), that are associated with different organising paradigms (Maruyama 1977), and “world hypotheses” (Pepper 1942). Where order comes from order, novelty is but a preformationist unfolding of latent possibility or recombination of existing parts; where order arises from noise, chaos, formlessness, or ambiguity, novelty entails *de novo* formation of new realms of possibility vis-à-vis existing observational and interpretive frameworks.

My purpose in considering emergence and this combinatoric-creative distinction is and has always been primarily pragmatic. For this reason, we focus here primarily on developing heuristics for generating useful novelty rather than in engaging in philosophical debates over the status of emergent novelty vis-a-vis various postulated ontological frameworks. For useful general introductions to the philosophical problems and their implications for mind-body relations, free will, and ontological emergence, see Kim (2008) and Clayton (2004). For similar reasons we will almost entirely sidestep the literature on complexity and emergence. Complexity in and of itself does not necessarily produce anything useful, nor does it necessarily provide insights into how to do so. On the other hand, variety is the mother of invention, and increased structural complexity does provide variety in the form of more accessible states and effective degrees of freedom. Processes of “complication” (von Neumann 1951) thus serve as fodder for new functions.

15.2.1 *What Constitutes a New Primitive?*

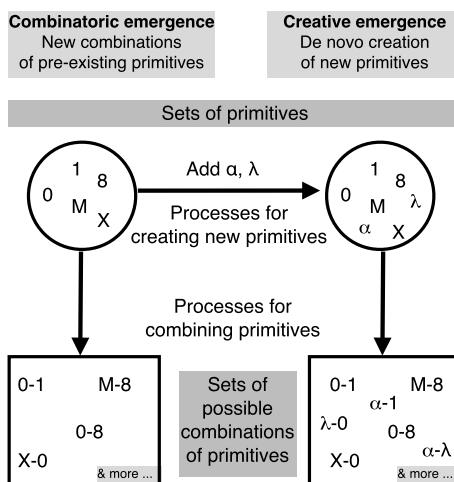
Both kinds of emergence, combinatoric and creative, entail recognition of basic sets of possibilities that constitute the most basic building blocks of the order, i.e. its atomic parts or “primitives”.

By a “primitive”, we mean an indivisible, unitary entity, atom, or element in a system that has no internal parts or structure of its own in terms of its functional role in that system. Individual symbols are the primitives of symbol string systems, binary distinctions are the primitives of flip-flop-based digital computers, and machine states are the primitives of finite state automata. To paraphrase Gregory Bateson, a primitive is a unitary “difference that makes a difference”.

Emergence then entails either the appearance of new combinations of previously existing primitives or the formation of entirely new ones (Fig. 15.1). The primitives in question depend upon the discourse; they can be structural, material “atoms”; they can be formal “symbols” or “states”; they can be functionalities or operations; they can be primitive assumptions of a theory; they can be primitive sensations and/or ideas; they can be the basic parts of an observer’s model.

Most commonly, the primitives are assumed to be structural, the parts that are put together in various combinations to make aggregate structures. Reductionist biology in effect assumes that everything that one would want to say about biological organisms can be expressed in terms of molecular parts. For many contexts and purposes, such as molecular biology and pharmaceutical development, where structure is key, this is an appropriate and effective framework. For other pursuits,

Fig. 15.1 Combinatoric and creative emergence



additional organisational and functional primitives are needed. If one wants to understand how an organism functions as a coherent, self-sustaining whole, one needs more than reductive parts-lists and local mechanisms. One needs concepts related to organisation and function, and knowledge of the natural history of how these have arisen. Living systems are distinct from nonliving ones because they embody particular organisations of material processes that enable organisational regeneration through self-production (Maturana and Varela 1973). Biological organisations also lend themselves to functional accounts that describe how goal-states can be embedded in their organisation and how goals can reliably realised by particular arrangements of processes. Full molecular descriptions of organisms do not lead to these relational concepts. Similarly, full molecular descriptions of brains and electronic computers, though useful, will not tell us how these systems work as information processing engines. If artificial systems are to be designed and built along the same lines as organisms and brains, new kinds of primitives appropriate for describing regenerative organisation and informational process are required.

15.2.2 Primitives and Interpretive Frames

Once one has defined what the primitives are or how they are recognised, then one has constructed a frame for considering a particular system. To say that an entity is “primitive” relative to other objects or functions means it cannot be constructed from combinations of the other entities of that frame, i.e. its properties cannot be logically deduced from those of the other entities. Although it may be possible, in reductionist fashion, to find a set of lower level primitives or observables from which the higher level primitives can be deduced, to do so requires a change of frame—one is then changing the definition of the system under consideration.

In the example of Fig. 15.1, the individual Roman and Greek letters and numerals are the primitives of a symbol-string system. Although concrete letters and numerals themselves do indeed have internal structure, in terms of strokes, arcs, and straight lines, these parts play no functional role in the system beyond supporting the distinction and recognition of their unitary symbol types. Once the classification of the type of symbol is made, the internal structure of the lines and curves become irrelevant. Were we to suddenly to adopt a frame in which the lines and curves are the primitives, then the appearance of the new symbols on the right, the Greek letters alpha and lambda, would not surprise us because these can be formed through combinations of the lower level strokes.

The combinatoric-creative distinction parallels ontological vs. epistemological modes of explanation. The debate that occurred in 1976 in France between Piaget, Chomsky, and Fodor over the origins of new ideas is illuminating. As organiser-participant Piatelli-Palmarini (1980) so elegantly pointed out, this was really a debate over the existence and nature of emergent novelty in the world. The two poles of the debate were held by Fodor (1980) and Piaget (1980). Fodor argued an extreme preformationist view in which all learning is belief-fixation, i.e. selection from a fixed repertoire of possible beliefs, such that entirely new ideas are not possible. Piaget presented an emergentist view in which qualitatively novel, irreducible concepts in mathematics have been created anew over the course of its history.

All that is possible in traditional ontological frameworks is recombination of existing possible constituents, whereas in epistemological frameworks, novelty can reflect surprise on the part of a limited observer. Another way of putting this is that ontologically-oriented perspectives adopt fixed, universal frames, whereas epistemologically-oriented ones are interested in which kinds of systems cause the limited observer to change frames and also what changes occur in the limited observer when frames are changed.

Second-order cybernetics (von Foerster 2003), systems theory (Kampis 1991), pragmatist theories of science (van Fraassen 1980), and constructivist epistemologies (von Glasersfeld 2007) are all concerned with “observing systems” that construct their own observational and interpretative frames. In Piaget’s words “Intelligence organises itself to organise the world” (von Glasersfeld 1992). We examine different kinds of conceivable self-constructing observing-acting systems in Sect. 15.3. When these systems change their frames, they behave in novel ways that cause those observing them to alter their own frames (Sect. 15.4).

15.2.3 *Novel Combinations of Closed Sets of Primitives*

Combinatoric emergence engages a fixed set of primitives that are combined in new ways to form emergent structures. In biological evolution the genetic primitives are DNA nucleotide sequences. On shorter evolutionary timescales microevolutionary processes select amongst combinations of existing genetic sequences, whereas on longer timescales macroevolutionary processes entail selection amongst entirely

new genes that are formed through novel sequences. On higher levels of biological organisation, emergent structures and functions can similarly arise from novel combinations of previously existing molecular, cellular, and organismic constituents. In psychology, associationist theories hold that emergent mental states arise from novel combinations of pre-existing primitive sensations and ideas. Whether cast in terms of platonic forms, material atoms, or mental states, combinatoric emergence is compatible with reductionist programs for explaining macroscopic structure through microscopic interactions (Holland 1998).

This strategy for generating structural and functional variety from a relatively small set of primitive parts is a powerful one that is firmly embedded in many of our most advanced informational systems. In the analytic-deductive mode of exploration and understanding, one first adopts some set of axiomatic, primitive assumptions, and then explores the manifold, logically-necessary consequences of those assumptions. In the realm of logic and mathematics, the primitives are axioms and their consequences are deduced by means of logical operations on the axioms. Digital computers are ideally suited for this task of generating combinations of symbol-primitives and logical operations on them that can then be evaluated for useful, interesting, and/or unforeseen formal properties. In the field of symbolic artificial intelligence (AI) these kinds of symbolic search strategies have been refined to a high degree. Correspondingly, in the realm of adaptive, trainable machines, directed searches use evaluative feedback to improve mappings between features and classification decisions. Ultimately these decisions specify appropriate physical actions that are taken. In virtually all trainable classifiers, the feature primitives are fixed and pre-specified by the designer, contingent on the nature of the classification problem at hand. What formally distinguishes different kinds of trainable machines is the structure of the combination-space being traversed, the nature of the evaluative feedback, and the rules that steer the search processes.

15.2.4 *Limits on Computations on Existing Primitives*

Combinatoric novelty is a dynamic, creative strategy insofar as it constantly brings into being new combinations of elements. However, such combinatoric realms are inherently limited by their fixed, closed sets of primitive elements. Consider the set of the digits 0–9 vs. a set of 10 arbitrarily distinguished objects. The first set is well-defined, has 10 actual members, and is closed, while the latter set is ill-defined, has an indefinite number of potential members, and is open.¹

All that can happen within well-defined universes are recombinations of existing, pre-specified symbols—there is no means by which new primitive symbols can be created by simply recombining existing ones. It seems obvious enough that one

¹ A Platonist could claim that all sets are open because they can include null sets and sets of sets *ad infinitum*, but we are only considering here sets whose members are collections of concrete individual elements, much in the same spirit as Goodman (1972).

does not create new alphabetical letter types by stringing together more and more existing letters—new types must be introduced from outside the system. This is typically carried out by an external agent. Likewise, in our computer simulations, we set up a space of variables and their possible states, but the simulation cannot add new variables and states simply by traversing the simulation-states that we have previously provided.

These ideas bear directly on fundamental questions of computational creativity. What are the creative possibilities and limitations of pure computations? Exactly how one defines “computation” is critical here. In its more widely used sense, the term refers to any kind of information-processing operation. Most often, the issue of what allows one to distinguish a computation from a non-computational process in a real-world material system is completely sidestepped, and the term is left loose and undefined. However, in its more precise, foundations-of-mathematics sense, the term refers to concrete formal procedures that involve unambiguous recognitions and reliable manipulations of strings of meaningless symbols. It is this latter, more restrictive, sense of computation as formal procedure that we will use here. For practical considerations, we are interested in computations that can be carried out in the real world, such as by digital electronic computer, and not imagined operations in infinite and potentially-infinite realms.²

In these terms, pure computation by itself can generate new combinations of symbol primitives, e.g. new strings of existing symbols, but not new primitive symbols themselves. In order for new symbol primitives to be produced, processes other than operations on existing symbols must be involved—new material dynamics must be harnessed that produce new degrees of freedom and new attractor basins that can support additional symbol types. To put it another way, merely running programs on a computer cannot increase the number of total machine states that are enabled by the hardware. In order to expand the number of total machine states that are available at any given time, we must engage in physical construction, such as fabricating and wiring in more memory.

There was a point in the history of computing devices at which self-augmenting and self-organising physical computational devices were considered. In the early 1960s, when electronic components were still expensive, growing electronic logical components “by the pound” was contemplated:

²Popular definitions of computation have evolved over the history of modern computing (Boden 2006). For the purposes of assessing the capabilities and limitations of physically-realizable computations, we adopt a very conservative, operationalist definition in which we are justified in calling an observed natural process a *computation* only in those cases where we can place the observable states of a natural system and its state transitions in a one-to-one correspondence with those of some specified deterministic finite state automaton. This definition has the advantage of defining computation in a manner that is physically-realizable and empirically-verifiable. It results in classifications of computational systems that include both real world digital computers and natural systems, such as the motions of planets, whose observable states can be used for reliable calculation. This finitistic, verificationist conception of computation also avoids conceptual ambiguities associated with Gödel’s Undecidability theorems, whose impotency principles only apply to infinite and potentially-infinite logic systems that are inherently not realizable physically.

We believe that if the “complexity barrier” is to be broken, a major revolution in production and programming techniques is required, the major heresies of which would mean weakening of machine structural specificity in every possible way. We may as well start with the notion that with 10 billion parts per cubic foot (approximately equal to the number and density of neurons in the human brain), there will be no circuit diagram possible, no parts list (except possibly for the container and the peripheral equipment), not even an exact parts count, and certainly no free and complete access with tools or electrical probes to the “innards” of our machine or for possible later repair. . . We would manufacture ‘logic by the pound’, using techniques more like those of a bakery than of an electronics factory. (Stewart 1969)

Such ideas persist today in visions of self-replicating nanobot nanotechnologies (now with the accompanying spectre of “grey goo” ecological disaster). At various times there have also existed notions of universal self-organising analog computers (see discussion of the Russian *Gutenmacher project* in Carello et al. (1984)). Such computational systems that physically grow their own hardware would be desirable, but the practical need for such self-expansions has been obviated by human ingenuity and creativity in the form of fast-evolving Moore’s-Law manufacturing efficiencies. It is simply easier to design and build a yet larger or faster machine than one that organically grows to become bigger or faster. Today’s large, scalable server arrays that permit simple addition of new modules are perhaps the closest systems we have to such growing automata.

In any case, these various means of making additional, new material dynamics accessible to the device lie outside the realm of symbol manipulation—they are non-computational. Non-computational breakout strategies are therefore required for computational systems to transcend their own initial primitive symbol sets.

15.2.5 *Creation of New Primitives*

Classically, “emergence” has concerned those processes that create new primitives, i.e. properties, behaviours, or functions that are not logical consequences of pre-existing ones (Broad 1925, Morgan 1931, Bergson 1911, Alexander 1927, Clayton 2004). How to create such fundamental novelty is the central issue for creative and transcendent emergence.

The most extreme example of emergence concerns the relationship of conscious awareness to underlying material process (Kim 1998, Clayton 2004, Rose 2006). All evidence from introspection, behavioural observation, and neurophysiology suggests that awareness and its specific contents is a concomitant of particular organised patterns of neuronal activity (Koch 2004, Rose 2006). If all experienced, phenomenal states supervene on brain states that are organisations of material processes, and these states in turn depend on nervous systems that themselves evolved, then it follows that there was some point in evolutionary history when conscious awareness did not exist.

This state-of-affairs produces philosophical conundrums. One can deny the existence of awareness entirely on behaviouristic grounds, because it can only be

observed privately, but this contradicts our introspective judgement that waking awareness is qualitatively different from sleep or anaesthesia. One can admit the temporal, evolution-enabled appearance of a fundamentally new primitive aspect of the world, a creative emergent view (Alexander 1927, Broad 1925, Morgan 1931), but this is difficult to incorporate within ontological frameworks that posit timeless sets of stable constituents. Or one can adopt a panpsychist view, with Spinoza and Leibniz, that the evolved nervous systems combine in novel ways simple distinctions that are inherent in the basic constituents of matter (Skrbina 2005). Accordingly, we could further divide creative emergence into the appearance of new structural and functional primitives that require epistemological, but not ontological reframing, and appearance of new, transcendent aspects of the world, such as the evolutionary appearance of consciousness, which require both.

Attempting to produce emergent awareness in some artificially constructed system is a highly uncertain prospect, because awareness is accessible only through private observables. One has no means, apart from indirect structural-functional analogy, of assessing success, i.e. whether any awareness has been brought into being. This is why even conscious awareness in animals, which have nervous systems extremely similar to ours, is a matter of lively debate.

More practical than *de novo* creation of new forms of being is the creation of new functions, which are both verifiable and useful to us—creativity as useful novelty. To my mind, the most salient examples of functional emergence involve the evolution of new sensory capabilities in biological organisms. Where previously there may have been no means of distinguishing odours, sounds, visual forms or colours, eventually these sensory capacities evolve in biological lineages. Each new distinction becomes a relative primitive in an organism's life-world, its sensorimotor repertoire.

Combinations of existing sensory distinctions do not create new primitive distinctions. We cannot directly perceive x-rays using our evolution-given senses, no matter how we combine their distinctions. In Sect. 15.3 we outline how evolutionary robotic devices could adaptively evolve their own sensors and effectors, thereby creating new primitives for sensorimotor repertoires.

Over the arc of evolution, the sensorimotor life-worlds of organisms have dramatically expanded. When a new sensory distinction or primitive action appears, the dimensionality of the sensorimotor combinatorial repertoire space increases. In an evolutionary landscape, the effective dimensionality of the fitness surfaces increases as life-worlds become richer and there are more means through which organisms can interact. Theoretical biologist Michael Conrad called this process “extradimensional bypass” (Cariani 2002, Chen and Conrad 1994, Conrad 1998).

The evolution of a new sensorimotor distinction and its dimensional increase can actually simplify problems of classification and decision-making. For gradient-ascending, hill-climbing optimisers, local maxima traps may become saddle points in higher dimensional spaces that open up entirely new avenues for further ascent. In the last decade, workers developing self-organising semantic webs for automated computer search have proliferated features to produce sparse, high-dimensional relational spaces (Kanerva 1988) whose partitioning becomes tractable via regularisation and linear classification techniques.

The senses of animals perform the same fundamental operations as the measurements that provide the observables of scientific models (Pattee 1996, Cariani 2011) and artificial robotic systems. Outside of artificially restricted domains it is not feasible to outline the space of possible sensory distinctions because this space is relational and ill-defined. It is analogous to trying to outline the space of possible measurements that could ever be made by scientists, past and future. Emergent creativity can be said to take place when new structures, functions, and behaviours appear that cannot be accounted for in terms of the previous expectations of the observer. For combinatorial creativity, the observer can see that the novel structures and functions are explicable in terms of previous ones, but for emergent creativity, the observer must enlarge the explanatory frame in order to account for the change. More will be said about emergent creativity and the observer in Sect. 15.4.

In this epistemological view of emergence, surprise is in the eye of the beholder. Because the observer has a severely limited model of the underlying material system, there are processes that can go on within the system that are hidden to direct observation that can qualitatively alter overt behaviour. In biological, psychological, and social systems, internal self-organising, self-complexifying processes can create novel structures and functions that in turn can produce very surprising behaviours. Because the epistemological approach is based on a limited set of macroscopic observables that do not claim any special ontological status, there is no necessary conflict with physical causality or reduction to microscopic variables (where possible). No new or mysterious physical processes or emergent, top-down causalities need to be invoked to explain how more complex organisations arise in physical terms or why they can cause fundamental surprise in limited observers. The novelty that is generated is partially due to internal changes in the system and partially due to the limited observer's incomplete model of the system, such that the changes that occur cause surprise.

15.2.6 Combinatoric and Creative Emergence in Aesthetic Contexts

A first strategy for computational creativity is to use artificial devices to cause creative responses in human participants. In aesthetic realms distinguishing between combinatoric and emergent creativity is made difficult by indefinite spaces of generative possibilities, as well as ambiguities in human interpretation and expectation. Often many prior expectations of individual human observers and audiences may be implicit and subliminal and therefore not even amenable to conscious analysis by the human participants themselves. Nonetheless, to the extent that cultural conventions exist, then it is possible to delineate what conforms to those expectations and what doesn't.

One rule of thumb is that combinatorial creative works operate within a set of stylistic or generative rules that explore new forms within an existing framework. An audience implicitly understands the contextual parameters and constraints of the medium, and the interest is in the play of particular new combinations, motifs, or

plot wrinkles. If the element recombinations are trivial, then a piece is perceived as predictable and clichéd. Emergent creative works break conventional, stylistic rules and may violate basic expectations related to the nature of the aesthetic experience itself. One thinks of the Dadaists and the world's reception of Duchamp's urinal as a found-art object.

Usually, the more creatively emergent a production, the fewer the number of people who will immediately understand it, because understanding a new art form or approach requires constructing new conceptual observables and interpretive frames in order to follow radical shifts of meaning. There is stress associated with the uncertainties of orientation and interpretation. For high degrees of novelty, the "shock of the new" causes high degrees of arousal that are in turn experienced as unpleasant.

The relation between arousal, pleasure, and aesthetics was studied by 19th century psychologists (Machotka 1980). The bell-shaped, Wundt curve plots empirical psychological data related to the relation between arousal (novelty) and experienced pleasure. Low novelty produces boredom, low arousal, and low pleasure, while extremely high novelty produces high arousal that is experienced as unpleasant. Between these two extremes is an optimal level of novelty that engages to produce moderate levels of arousal that are experienced positively. The degree to which a new piece shocks (and its unpleasantness enrages) its audiences is an indication of how many expectations have been violated. An individual's response tells us something about the novelty of the piece in relation to his or her own Wundt curve.

15.3 Creativity in Self-constructing Cybernetic Percept-Action Systems

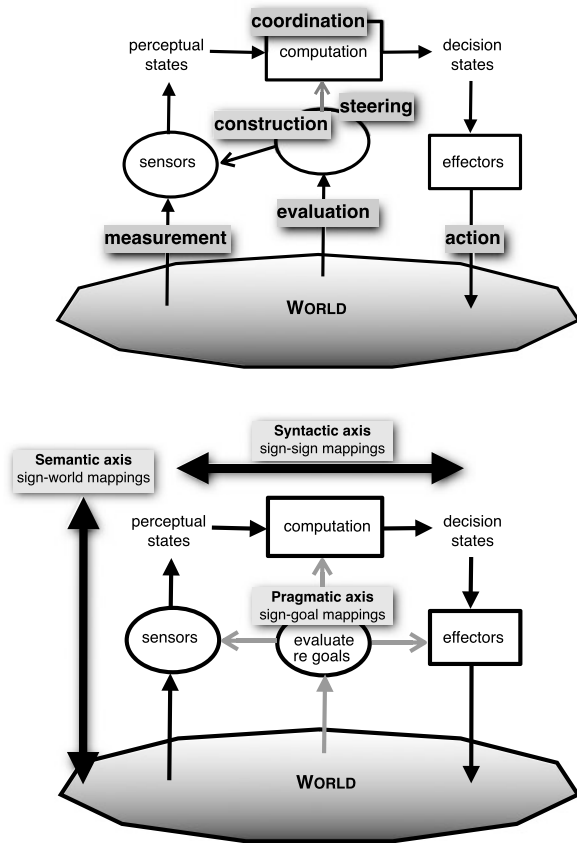
A second strategy for computational creativity involves expansions of the informational realms in the artificial devices themselves. In this section we consider artificial devices that create their own percept and action primitives, and argue that self-construction guided by internal goals and evaluative faculties is necessary to provide the structural autonomy and implicit dynamics needed to create new useful semantic linkages.

15.3.1 *A Taxonomy of Adaptive Devices*

The most straightforward way of tackling the problem of how such devices might be designed and built is to consider the different possible kinds of devices that can be conceivably constructed using a set of basic functionalities. A few taxonomies of possible mixed analog-digital adaptive and self-constructing cybernetic devices have been proposed (Cariani 1989; 1992; 1998, de Latil 1956, Pask 1961).

Here we present our own taxonomy of devices in which some functionalities are fixed, while others are adaptively modified or constructed (Figs. 15.2, 15.3, 15.4). It

Fig. 15.2 The functional organisation of cybernetic devices. *Top*: basic functionalities and their associated operations. *Bottom*: relation of semiotic dimensions to functionalities and operations



then becomes possible to consider the general capabilities and limitations of various classes of devices that possess varying abilities to adapt and evolve. Although the more structural autonomy or constructive license given to a device, the greater the potential creativity that is permitted, it should be remembered that greater degrees of autonomy and creativity come at the expense of greater complexity and longer periods of adaptive construction and evaluative testing.

The basic functionalities that constitute the functional organisation of adaptive self-constructing cybernetic devices in this taxonomy are *coordination*, *measurement*, *action*, *evaluation*, *steering*, and *construction* (Fig. 15.2, top). Computational operations here entail coordinative linking of output states with input states, and include memory mechanisms for recording and reading out past inputs. Measurement operations are carried out by an array of sensors that produce symbolic outputs whose values are contingent on the interaction of the sensors with their environs. Actions are carried out by effectors that influence the external world. Effectors produce actions contingent upon internal decisions and commands that are the output of the coordinative part. Steering mechanisms alter particular device states or state-transitions without altering the device's set of accessible states or

state-transitions. Construction processes involve adding new coordinations (computations, states, state-transitions), measurements (sensors, observables), actions (effectors), goal states and evaluative criteria, and new construction mechanisms as well. One can think of steering as switching between software alternatives, and construction as the physical construction of new hardware. Steering (switching) processes do not change the effective dimensionality of the system, whereas construction does. Many of these operations could be realised by analog, digital, or mixed analog-digital processes.

These basic functionalities arguably account for the basic operational structure of the observer-actor. There is the cycling of signals from sensors to coordinative elements to effectors (outer loop in the diagram) and “feedback to structure” (inner loops) in which evaluative mechanisms steer the modification and/or construction of hardware (sensors, computational, coordinative structures, effectors).

15.3.2 Semiotics of Adaptive Devices

It is useful to discuss such devices and their creative capabilities in terms of the semiotic triad of Charles Morris, which consists of syntactic, semantic, and pragmatic aspects (Morris 1946, Nöth 1990). Syntactics describes rule-governed linkages between signs; semantics, the relation of signs to the external world; and pragmatics, the relation of signs to purposes (goal states). These different semiotic relations are superimposed on the functional schematic of cybernetic percept-action devices in the bottom panel of Fig. 15.2. The syntactic axis runs horizontally, from sign-states related to sensory inputs to those related to coordinative transformations, and finally to decision states that ultimately lead to actions. The semantic axis runs vertically between the sign-states and the external world, where sensory organs determine world-sign causalities and effectors determine sign-world causalities. The pragmatic axis in the centre covers adaptive relationships between sign states and embedded goals. These are implemented by evaluative and adjustment processes that steer the percept-action linkages that govern behaviour and guide the construction of the device itself.

Some devices have fixed functionalities (stable systems), some can autonomously switch amongst existing alternative states to engage in combinatorial search (combinatoric systems), and some can add functional possibilities to creating new primitives (creative systems). Table 15.1 lists the effects of stable, combinatoric, and creative change for different semiotic relations. Creative emergence in the syntactic realm involves creation of new internal sign-states (or computational states) that enable entirely new mappings between states. Creative emergence in the semantic realm involves creating new observables and actions (e.g. sensors, effectors) that contingently link the outer world with internal states. Creative emergence in the pragmatic realm involves creating new goals and evaluative criteria. Table 15.1 and Fig. 15.3 schematise different classes of devices with respect to their creative capabilities.

Table 15.1 Modes of creativity with respect to semiotic dimensions

Aspect	Primitives	Type of system		
		Stable <i>Maintain structure</i>	Combinatoric <i>Search existing possibilities</i>	Creative <i>Add possibilities</i>
Syntactic	Sign-states & computations	Deterministic finite-state automata	Adaptive changes in state-transition rules <i>trainable machines</i>	Evolve new states & rules <i>growing automata</i>
Semantic	Measurements & actions	Fixed sensors & effectors (fixed robots)	Adaptive search for optimal combinations of existing sensors & effectors	Evolve new observables, actions <i>epistemic autonomy</i>
Pragmatic	Goals	Fixed goals	Search combinations of existing goals <i>adaptive priorities</i>	Evolve new goals <i>creative self-direction motivational autonomy</i>

15.3.3 Capabilities and Limitations of Adaptive Devices

One can consider the capabilities and limitations of devices with computational coordinative parts, sensors, effectors, and goal-directed mechanisms for adaptive steering and self-construction (Fig. 15.3). For the sake of simplicity, we will think of these systems as robotic devices with sensors and effectors whose moment-to-moment behaviour is controlled by a computational part that maps sensory inputs to action decisions and motor outputs. In biological nervous systems these coordinative functions are carried out by analog and mixed analog-digital neural mechanisms.

Purely computational devices (top left) deterministically map symbolic, input states to output states, i.e. they are formally equivalent to deterministic finite state automata. As they have no non-arbitrary linkages to the external world, their internal states have no external semantics save those that their human programmer-users assign to them. Because their computational part is fixed and functionally stable, such devices are completely reliable. However, they are not creative in that they cannot autonomously generate either new combinations (input-output mappings) or new primitives (sign states).

Some of the functional limitations of formal systems and computational devices are due to their purely syntactic nature, that the sign-states lack intrinsic semantics or pragmatics. The signs and operations are meaningless and purposeless, aside from any meanings or purposes that might be imposed on them by their users. Other limitations arise from their fixed nature, that pure computations do not receive contingent inputs from outside the sign-system, and therefore have no means of adaptively adjusting their internal operations—they do not learn.

One might retort that we have all sorts of computers that are constantly receiving updates from external sources and adjust their behaviour accordingly, but the moment a machine acts in manner that depends not only on its initial state and state-transition rules, its behaviour is no longer a pure computation—it is no longer

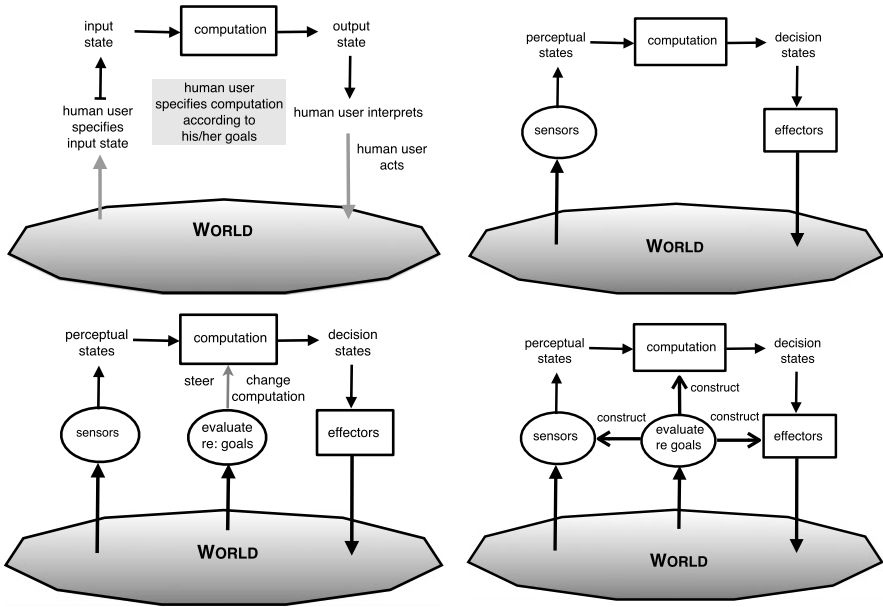


Fig. 15.3 A taxonomy of cybernetic devices. *Top left:* a fixed computational device. *Top right:* a fixed robotic device. *Bottom left:* an adaptive robotic device that modifies its computational input-output mapping contingent on its evaluated performance. *Bottom right:* a robotic device that adaptively constructs its sensing, effecting, and computational hardware contingent on its evaluated performance

emulating the behaviour of a formal system. It is as if one were to perform a calculation, say of the thousandth digit of π , but midway in the calculation the result depends partially on fine variations of the temperature in the room. Only rarely will two such procedures produce the same result, and one now has a process that is the antithesis of a formal procedure. When coupled this way such devices, in formal terms, become machines with inputs from oracles, where the internal workings of the oracle are left ill-defined (Turing 1939). Coupling a deterministic finite state automaton to a sensor that makes measurements converts the composite device to a finite state oracle machine, a decidedly different kind of beast (Hodges 2008).

Adding measurements are useful for some purposes, such as responding appropriately to changes in an external environment, but highly detrimental to others, such as performing reliable, determinate calculations, where one is interested in the purely logical consequences of the application of specified rules on initial inputs. For these reasons, our physical computing devices have been designed and built, as much as possible, to operate in a manner that is independent of their environs.

Accordingly, one can add fixed sensors and effectors to purely computational devices to create robotic devices (Fig. 15.3, top right) that have behaviours that are qualitatively different from those of formal systems. These kinds of systems, which include animals and artificial robots, have specific perception and action linkages to the external world, thereby endowing their internal states with external semantics.

Here the output productions are actions rather than symbols per se, but these devices are also not creative in that they cannot autonomously generate new behaviours.

One can then add evaluative sensors and steering mechanisms that switch the behaviour of the computational part to produce adaptive computational machines (Fig. 15.3, bottom left). This is the basic high-level operational structure of virtually all contemporary trainable machines that use supervised learning feedback mechanisms (adaptive classifiers and controllers, genetic algorithms, neural networks, etc.). Here the internal states and their external semantics are fixed, such that the evaluative-steering mechanism merely switches input-output (percept-action, feature-decision) mappings using the same set of possible states. This is a form of combinatorial creativity, because the machine searches through percept-action combinations to find more optimal ones.

Consider the case, however, where the evaluation mechanism guides the construction of the hardware of the device rather than simply switching input-output mappings (Fig. 15.3, bottom right). If sensors are adaptively constructed contingent on how well they perform a particular function, then the external semantics of the internal states of the device are now under the device's adaptive control. When a device has the ability to construct itself, and therefore to choose its sensors—which aspects of the world it can detect—it attains a partial degree of *epistemic autonomy*. Such a device can adaptively create its own meanings vis-à-vis the external world. A system is purposive to the extent that it can act autonomously to steer its behavior in pursuit of embedded goals. When it is able to modify its evaluative operations, thereby modifying its goals, it achieves a degree of *motivational autonomy*. Such autonomies depend in turn on structural autonomy, a capacity for adaptive self-construction of hardware.

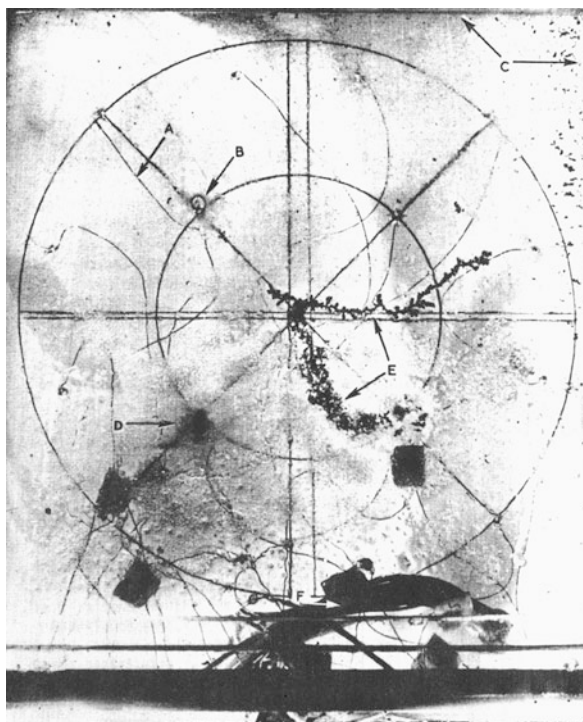
To summarise, combinatoric creativity in percept-action systems entails an ability to switch between existing internal states (e.g. “software”), whereas creative emergence requires the ability to physically modify material structures (e.g. “hardware”) that create entirely new states and state-transitions, sensors, effectors, and/or goals.

15.3.4 Pask's “Organic Analogues to the Growth of a Concept”

The most striking example of a creative emergent device is an adaptive self-constructing electrochemical assemblage that was conceived and fabricated by the brilliant and eccentric British cybernetician Gordon Pask in the late 1950s (Cariani 1989; 1993, Pask 1960; 1961, Bird and Di Paolo 2008, Pickering 2010). Pask demonstrated his device at the *Mechanisation of Thought Processes* conference in London in 1958 and described it in a paper provocatively entitled “Organic Analogues to the Growth of a Concept.” (Pask 1959)

The device's purpose was to show how a machine could evolve its own “relevance criteria”, i.e. its own external semantic meanings. Current was passed through an array of platinum electrodes immersed in an aqueous ferrous sulphate/sulphuric acid

Fig. 15.4 Gordon Pask's creative emergent electrochemical assemblage, from (Pask 1959, p. 919). The photograph looks down on a glass tank containing an aqueous solution of ferrous sulphate and sulphuric acid. Original caption labels: A. Connecting wires for electrodes. B. Platinum pillar electrodes. C. Edges of glass tank containing ferrous sulphate. D. Chemical reaction in progress. E. "Tree" threads being formed. F. Connecting cables



medium, such that iron filaments grew outwards to form bridges between the electrodes (Fig. 15.4). Here the electrodes that extend down into the medium are perpendicular to the plane of the photograph. Iron threads whose conductivity co-varied in some way with an environmental perturbation were rewarded with electric current that caused them to grow and persist in the acidic milieu. Through the contingent allocation of current, the construction of structures could be adaptively steered to improve their sensitivity. The assemblage acquired the ability to sense the presence of sound vibrations and then to distinguish between two different frequencies.

We have made an ear and we have made a magnetic receptor. The ear can discriminate two frequencies, one of the order of fifty cycles per second and the other on the order of one hundred cycles per second. The "training" procedure takes approximately half a day and once having got the ability to recognise sound at all, the ability to recognise and discriminate two sounds comes more rapidly. I can't give anything more detailed than this qualitative assertion. The ear, incidentally, looks rather like an ear. It is a gap in the thread structure in which you have fibrils which resonate at the excitation frequency." (Pask 1960, p. 261)

In effect, the device had evolved an ear for itself, creating a set of sensory distinctions that it did not previously have. Albeit, in a very limited way, the artificial device automated the creation of new sensory primitives, thereby providing an existence proof that creative emergence is possible in adaptive devices. As Pask explicitly pointed out, one could physically implement an analog perceptron with

such an assemblage: the conductances between electrodes in the electrochemical array correspond to connection weights in a connectionist neural network. His intent, however, was to show how a device could produce emergent functionality. Instead of switching inter-electrode connectivities, the thread structures could be steered and selected to become sensitive to other kinds of perturbations, such that they could be tuned with the appropriate rewards. By rewarding conductance changes associated with a particular kind of environmental disturbance, the assemblage could evolve its own sensitivities to the external world.

In the preface to Pask's book *An Approach to Cybernetics* (Pask 1961), Warren McCulloch declared: "With this ability to make or select proper filters on its inputs, such a device explains the central problem of epistemology. The riddles of stimulus equivalence or of local circuit action in the brain remain only as parochial problems." The central, most fundamental problem in epistemology is how to obtain the right observables needed to solve a particular problem. Once these are found, everything else is a matter of searching through the possibilities that these observables afford.

15.3.5 Organisational Closure and Epistemic Autonomy

Creativity and learning both require some degree of autonomy on the part of the system in question. The system needs to be free to generate its own novel, experimental combinations and modifications independent of pre-specification by a designer. The more autonomy given the system, the greater the potential for novelty and surprise on the part of the designer. The less autonomy given, the more reliable and unsurprising the system's behaviour.

When a device gains the ability to construct its own sensors, or in McCulloch's words "this ability to make or select proper filters on its inputs", it becomes *organisationally closed*. The device then controls the distinctions it makes on its external environment, the perceptual categories which it will use. On the action side, once a device acquires the ability to construct its own effectors, it thereby gains control over the kinds of actions it has available to influence the world. The self-construction of sensors and effectors thus leads to attainment of greater *epistemic autonomy* and *enactive autonomy*, where the organism or device itself can become the major determinant of the nature of its relations with the world at large. Structural autonomy and organisational closure guided by open-ended adaptive mechanisms lead to functional autonomy.

These ideas, involving adaptive self-construction and self-production link with many of the core concepts of theoretical biology and cybernetics, such as semantic closure (Pattee 1982; 2008, Stewart 2000), autopoiesis and self-production (Maturana and Varela 1973, Maturana 1981, Varela 1979, Rosen 1991, Mingers 1995), self-modifying systems (Kampis 1991), regenerative signalling systems (Cariani 2000), and self-reproducing automata (von Neumann 1951). Life entails autonomous self-construction that regenerates parts and organisations.

15.4 Recognising Different Types of Creativity

How does one distinguish combinatoric from emergent creativity in practice? This is the methodological problem. The distinction is of practical interest if one wants to build systems that generate fundamental novelty—one needs a clear means of evaluating whether the goal of creating new primitives has been attained.

15.4.1 *Emergence-Relative-to-a-Model*

Theoretical biologist Robert Rosen (1985) proposed a systems-theoretic, epistemological definition of emergence as the deviation of the behaviour of a material system from the behaviour predicted by a model of that system. At some point the behaviour of a material system will deviate from its predicted behaviour because of processes in the material world that are unrepresented in the model.

This concept can be put into concrete practice by formulating an operational definition. Like the description of an experimental method, an operational definition specifies the procedures by which different observers can reliably make the same classifications, e.g. is a given behaviour emergent, and if so, how? In this case an emergent event is one that violates the expectations of an observer's predictive model. However, simple violations, such as when the engine of one's car fails, are not so interesting or particularly useful, because they can be outward signs of the breakdown of internal mechanisms as much as signs of the evolution of new functions. Instead, we are interested in deviations from expected behaviour that are due to adaptive construction of new functions and qualitatively new behaviours of the system under study.

In the late 1980s, in conjunction with the adaptive systems taxonomy, we developed a systems-theoretic methodology for how one could go about recognising the operations of measurements, computations, and actions from the observed state-transitions of a natural system (Cariani 1989; 1992; 2011). In the process operational definitions were formulated for how these functions can be distinguished from each other, and how changes in a given functionality can be recognised. The method partitions the state-transition structure of the system into regions of state-determined transitions that resemble computations and regions of indeterminate, contingent transitions that resemble measurements and actions.

15.4.2 *Tracking Emergent Functions in a Device*

Consider the case of an observer following the behaviour of a device (Fig. 15.5, top panel). The observer has a set of observables on the device that allow him/her/it to observe the device's internal functional states and their transitions. Essentially if one were observing a robotic device consisting of sensors, a computational coordinative

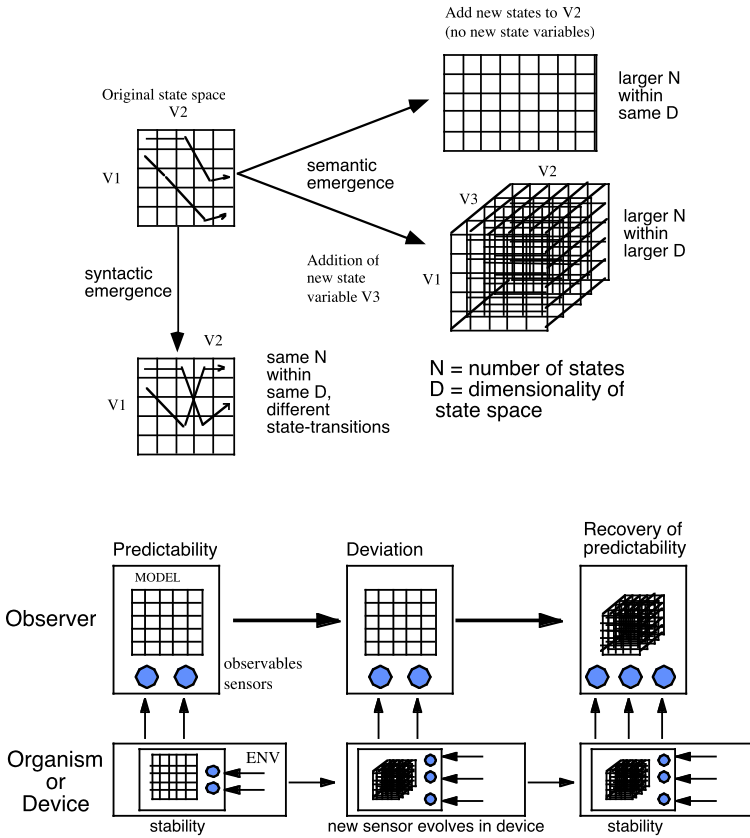


Fig. 15.5 Emergence relative-to-a-model. What changes need to be adopted by an observer in order to continue to predictively track the behaviour of an evolving, complexifying system?

part, and effectors, the computational part of the device that mapped sensory inputs into motor commands would have state-determined transitions.

One can determine if the input-output mapping of the computational part has changed by observing its state-transition structure (Fig. 15.5, top panel). If the computational part is a fixed program, this sensorimotor mapping will remain invariant. If the computational part is switched by some adaptive process, as in a trainable machine, then the sensorimotor mapping will change with training, and a new determinate input-output state transition behaviour will then ensue. From an observer's perspective, the predictive model will fail every time training alters the computational sensorimotor mapping. In order to recover predictability, the observer would have to change the state-transition rules of his or her predictive model. Thus an observer can determine whether the device under observation is performing fixed computations or whether these are being adjusted in some way over time.

Similarly, if the device evolves a new sensor, such that its behaviour becomes dependent on factors that are not registered in the observer's set of measurements,

then the observer will also lose predictability. In order to regain predictability, the observer would need to add an extra observable that was roughly correlated with the output of the device's new sensor. Thus if the observer needs to add a sensor to continue to track the device, then it can be inferred that the device itself has effectively evolved a new sensor.

The general principle involves what modifications the observer needs to make in his or her modelling framework to maintain the ability to track the behaviour of the system. If this involves rearrangement of existing states, then the system under observation appears to be combinatorically-emergent. If it requires increasing the dimensionality of his or her observational frame, then the system under observation appears to be creatively emergent. The new dimensions in the observer's complexifying modelling frame coevolve with the creation of new primitives in the observed system.

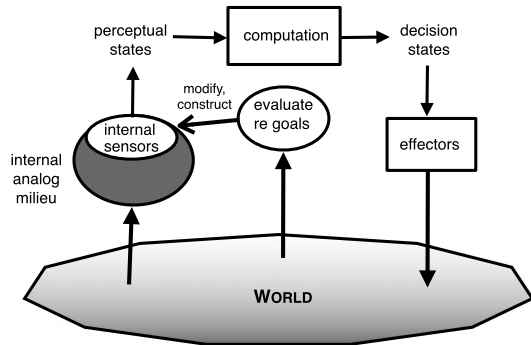
15.5 New Signal Primitives in Neural Systems

A third, and last strategy for creative emergence is to attempt to understand and emulate the creative processes inside our brains. We humans are the most formidable sources of creativity in our world. Our minds are constantly recombining existing concepts and meanings and also creating entirely new ones. The most obvious process of the construction of new concepts is language acquisition, where children reportedly add 10–15 new word meanings per day to their cognitive repertoires, with the vast majority of these being added without explicit instruction. It seems likely that this “mundane creativity” in children's brains operates through adaptive neural processes that are driven by sensorimotor and cognitively-mediated interactions with the external world (Barsalou and Prinz 1997). Although most of these processes may well fall under the rubric of the combinatorics of syntactic, semantic, and pragmatically grounded inference engines, there are rarer occasions when we experience epiphanies associated with genuinely new ways of looking at the world.

One can contemplate what creation of new signal primitives would mean for neural networks and brains (Cariani 1997). Essentially we want an account of how combinatoric productivity is not only possible, but is so readily and effortlessly achieved in everyday life. We also want explication of how new concepts might be formed that are not simply combinations of previous ones, i.e. how the dimensionality of a conceptual system might increase with experience. How might these creative generativities be implemented in neuronal systems?

We have to grapple with the problem of the primitives at the outset. Even if the brain is mostly a combinatorically creative system, the conceptual primitives need to be created by interactive, self-organising sensorimotor integration processes, albeit constrained by genetically mediated predispositions.

Fig. 15.6 Creation of new semantic primitives by means of internal sensors. Neural assemblies play the role of sensors on an internal milieu of neural activity patterns



15.5.1 New Primitives in Signalling Networks

I came to think about how neural networks might create new primitives from considering how a signalling network might increase its effective dimensionality. The simplest way of conceiving this is to assume that each element in a network is capable of producing and receiving specific signals that are in some way independent of one another, as with signals consisting of tones of different frequencies. A new communications link is established whenever a tone frequency emitted by one element can be detected by another. The effective dimensionality of such a network is related to the number of operating independent communications links. If the elements can be adaptively tuned to send and receive new frequencies that are not already in the network, then new signal primitives with new frequencies can appear over time and with them, new communications links. The dimensionality of the signalling network has thus increased.

One can conceive of the brain as a large signalling network that consists of a large number of neural assemblies of many neurons. If each neural assembly is capable of adaptively producing and detecting specific spatial and temporal patterns of action potential pulses, then new patterns can potentially arise within the system that constitute new signal primitives. In the brain we can think of Hebb's neural assemblies (Hebb 1949, Orbach 1998) as ensembles of neurons that act as internal sensors on an analog internal milieu (Fig. 15.6). The creation of a new neural assembly through an activity-dependent modification of neuronal synapses and axons can be conceived as the equivalent to adding a new internal observable on the system. Here a new concept is a new means of parsing the internal activity patterns within the nervous system. If an adaptively-tuned neural ensemble produces a characteristic pattern of activity that is distinguishable from stereotyped patterns that are already found in the system, then the neural network has created a new signal primitive that can become a marker for the activity of the ensemble and some complex combination of conditions that activates it.

The remainder of the chapter presents an outline of how a neural system might utilise this kind of dimensionally open-ended functional organisation. The nature of the central neural code in the brain is still one of science's biggest unsolved

mysteries, and the present situation in the neurosciences is not unlike biology before DNA-based mechanisms of inheritance were understood.

Despite resurgent interest in neuronal temporal synchronies and oscillations, mainstream opinion in the neurosciences still heavily favours neural firing rate codes and their connectionist architectures over temporal codes and timing architectures. For introductory overviews of how connectionist networks operate, see (Arbib 1989; 2003, Horgan and Tienson 1996, Churchland and Sejnowski 1992, Anderson et al. 1988, Boden 2006, Marcus 2001, Rose 2006). Although strictly connectionist schemes can be shown to work in principle for simple tasks, there are still few concrete, neurally-grounded demonstrations of how connectionist networks in real brains might flexibly and reliably function to carry out complex tasks, such as the parsing of visual and auditory scenes, or to integrate novel, multimodal information. We have yet to develop robust machine vision and listening systems that can perform in real world contexts on par with many animals.

In the late 1980s a “connectionism-computationalism” debate ensued about whether connectionist networks are at least theoretically capable of the kinds of combinatorial creativities we humans produce when we form novel, meaningful sentences out of pre-existing lexical and conceptual primitives (Marcus 2001, Horgan and Tienson 1996, Boden 2006, Rose 2006). Proponents of computationalism argued that the discrete symbols and explicit computations of classical logics are needed in order to flexibly handle arbitrary combinations of primitives. On the other hand, the brain appears to operate as a distributed network that functions through the mass statistics of ensembles of adaptive neuronal elements, where the discrete symbols and computational operations of classical logics are nowhere yet to be seen. But difficulties arise when one attempts to use subsymbolic processing in connectionist nets to implement simple conceptual operations that any child can do. It’s not necessarily impossible to get some of these operations to work in modified connectionist nets, but the implementations generally do not appear to be robust, flexible, scalable, or neurally plausible (Marcus 2001). It is possible, however, that fundamentally different kinds of neural networks with different types of signals and informational topologies can support classical logics using distributed elements and operations.

Because we have the strong and persistent feeling that we do not yet understand even the basics of how the brain operates, the alternative view of the brain outlined here, which instead is based on multidimensional temporal codes, should be regarded as highly provisional and speculative in nature, more rudimentary heuristic than refined model.

15.5.2 Brains as Networks of Adaptive Pattern-Resonances

Brains are simultaneously communications networks, anticipatory correlation machines, and purposive, semantic engines that analyse their sensory inputs in light of previous experience to organise, direct, and coordinate effective action. Animals

with nervous systems are cybernetic, goal-seeking percept-action systems (Arbib 1989, Powers 1973, Sommerhoff 1974, Cariani 2011, Boden 2006, Pickering 2010, McCulloch 1965, Rose 2006). Nervous systems evolved in motile animals in order to better coordinate effective action in rapidly changing situations. Lineages of animals whose nervous systems enhanced survival to reproduction persisted, whereas those with less effective steering mechanisms tended to perish. Like the adaptive devices discussed above in Sect. 15.3, animals have sensory systems that register interactions with the external world, and motor systems that influence events in the world. They have coordinative sensorimotor linkages with varying degrees of complexity, from automatic reflexes to heavily deliberated actions. Brains have embedded goal systems that enforce drive states that steer both perception and action. Embedded goals and steering mechanisms are what make intentional, purposive action possible. Finally brains have reward systems that reconfigure internal circuits to change steering functions and to build neuronal assemblies that facilitate new sensory analyses, cognitive representations, affective and motivational responses, and motor sequencing programs.

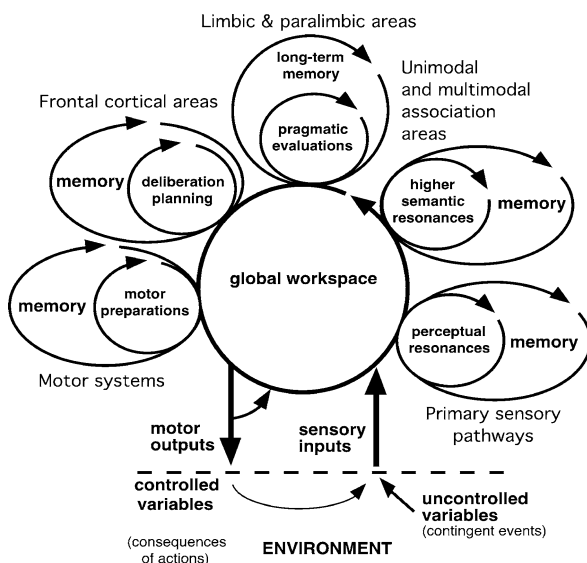
Thus globally, the brain is a goal-directed system in which the most primal goal structures, which motivate actions such as breathing, drinking, eating, fleeing, fighting, and mating, are embedded in limbic structures and dopamine-mediated reward systems. Neural goal-seeking mechanisms steer sensory and motor thalamocortical systems to act on system-goals by means of the *basal ganglia*, a large set of brain structures that mediate connections of sensory, cognitive, and motor areas of the cerebral cortex with other limbic goal and reward circuits (Redgrave 2007). The basal ganglia facilitate amplification of goal-relevant sensory information and motor programs by releasing the relevant neuronal subpopulations from inhibition, i.e. “disinhibition”.

This arrangement for steering can be thought of as a control brake on a slave braking system; when the control brake is applied, the slave brake is released. If the normal state of recurrent circuits with the inhibitory, slave brakes on is attenuating, but becomes weakly amplifying when the inhibitory slave brakes themselves are inhibited (by the basal ganglia control inputs), then those patterns of sensory, cognitive, and motor activity that are relevant to the system-goals that are currently most pressing will be the ones that will be amplified and regenerated. Neural signal regeneration in these weakly amplifying loops will then continue until some other set of competing goals becomes dominant. The behavioural steering system is reminiscent of Kilmer and McCulloch’s reticular formation action selection model (Kilmer and McCulloch 1969) and of Rodney Brooks’ subsumption architectures (Brooks 1999) in that the different goals are in constant competition for control over the organism’s mode of behaviour.

15.5.3 Regenerative Loops

The basic functionalities involved in perception, coordination, and action may be implemented by global architectures that consist of many reciprocally connected

Fig. 15.7 High level functional schematic of the brain as an overlapping and interconnected set of pattern-resonance loops each of which supports a different functionality



neuronal populations. The loops support pattern amplification cycles of neuronal signals that allow signals to be dynamically regenerated within them. The loops might be recurrent connectionist networks (Carpenter and Grossberg 2003, McCulloch 1965) or alternatively, they could be closed circuits that propagate temporal patterns of neuronal activity (Thatcher and John 1977).

That a system has the ability to regenerate alternative sets of neural signals—neural “pattern resonances”—means that the system can support alternative persistent informational states. The system can then keep a set of neuronal signals circulating indefinitely as long as they remain relevant to ongoing goals. This ability to hold signals dynamically is critical for short-term memory, informational integration in global workspaces, and may itself be a requisite for conscious awareness (Baars 1988, Cariani 2000). In dynamical systems terms these mental states are complex signal attractor states that are the stable eigen-behaviour modes of the neural system (Rocha 1996, von Foerster 2003, Cariani 2001b). In analogy with autocatalytic networks that produce the components of living cells, one could think of mutually-supporting signal-regenerations in the loops as an autopoiesis of neural signal productions (Maturana and Varela 1973).

One can sketch out a loose functional organisation based on regenerative processing loops (Fig. 15.7). Sensory information comes into the system through a number of modality-specific sensory pathways. Neural sensory representations are built up in each of the pathways through bottom-up and top-down loops that integrate information in time to form stable perceptual images. When subsequent sensory patterns are similar to previous ones, these patterns are amplified and stabilised; when they diverge, new dynamically-created images are formed from the difference between expectation and input. Such divergences are seen in evoked brain gross electrical

potentials as “mismatch negativities” and are modelled as adaptive resonances (Carpenter and Grossberg 2003, Grossberg 1988, Rose 2006).

As successive neural populations respond to the stabilised perceptual images of lower loops, higher semantic resonances are created as objects are recognised and other sets of neural assemblies that are sensitive to their implications are activated. As ensembles of neural assemblies that function as semantic “cognitive nodes” are concurrently excited, evaluative and steering circuits are also being activated and signals associated with hedonic valences become added to the circulating mixture of signals. Both short term and long term memory processes play modulatory roles in the pattern amplifications that go on in the loops, either facilitating or retarding the amplification and buildup of particular circulating patterns. In this conception, some sets of signals are limited to only one or two circuits that reciprocally connect pairs of brain regions, while others of more general relevance are circulated more broadly, through a global workspace. Sets of circuits related to evaluation of the various consequences of action and its planning are activated and finally motor execution programs and pattern generators are engaged to move muscles that act on the external world.

15.5.4 Multidimensional Signals

Although modern neuroscience has identified specific neuronal populations and circuits that subserve all these diverse functions, there is much poorer understanding of how these different kinds of informational considerations might be coherently integrated. Although, most information processing appears to be carried out in local brain regions by neural populations, a given region might integrate several different kinds of signals. Traditional theories of neural networks assume very specific neuronal interconnectivities and synaptic weightings, both for local and long-distance connections. However, flexibly combining different kinds of information from different brain regions poses enormous implementational problems. On the other hand, if different types of information can have their own recognisable signal types, then this coordination problem is drastically simplified. If different signal types can be nondestructively combined to form multidimensional vectors, then combinatorial representation systems are much easier to implement. Communications problems are further simplified if the multiple types of information can be sent concurrently over the same transmission lines without a great deal of destructive interference.

15.5.5 Temporal Coding and Signal Multiplexing

Multiplexing of signals permits them to be combined nondestructively, broadcast, and then demultiplexed by local assemblies that are tuned to receive them. Temporal coding of information in patterns of spikes lends itself to multidimensional signalling, multiplexed transmission, and broadcast strategies for long-distance neural

coordinations. The brain can thus be reconceptualised, from the connectionist image of a massive switchboard or telegraph network to something more like a radio broadcast network or even an internet (John 1972).

Neurophysiological evidence exists for temporal coding in virtually every sensory system, and in many diverse parts of the brain (Cariani 1995; 2001c, Miller 2000, Perrell and Bullock 1968, Mountcastle 1967), and at many time scales (Thatcher and John 1977). We have investigated temporal codes for pitch in the early stages of the auditory system (Cariani and Delgutte 1996, Ando and Cariani 2009, Cariani 1999). The neural representation that best accounts for pitch perception, including the missing fundamental and many other assorted pitch-related phenomena, is based on interspike intervals, which are the time durations between spikes in a spike train. Periodic sounds impress their repeating time structure on the timings of spikes, such that distributions of the interspike intervals produced in auditory neurons reflect stimulus periodicities. Peaks in the global distribution of interspike intervals amongst the tens of thousands of neurons that make up the auditory nerve robustly and precisely predict the pitches that will be heard. In this kind of code, timing is everything, and it is irrelevant which particular neurons are activated the most. The existence of such population-based statistical, and purely temporal representations begs the question of whether information in other parts of the brain could be represented this way as well (Cariani and Micheyl 2012).

Temporal patterns of neural spiking are said to be stimulus-driven in they reflect the time structure of the stimulus or stimulus-triggered if they produce response patterns that are unrelated to that time structure. The presence of stimulus-driven patterns of spikes convey to the rest of the system that a particular stimulus has been presented. Further, neural assemblies can be electrically conditioned to emit characteristic stimulus-triggered endogenous patterns that provide readouts that a given combination of rewarded attributes has been recognised (John 1967, Morrell 1967).

The neuronal evidence for temporal coding also provokes the question of what kinds of neuronal processing architectures might conceivably make use of information in this form. Accordingly several types of neural processing architectures capable of multiplexing temporal patterns have been conceived (Izhikevich 2006, Cariani 2004, Chung et al. 1970, Raymond and Lettvin 1978, Pratt 1990, Wasserman 1992, Emmers 1981, Singer 1999).

We have proposed neural timing nets that can separate out temporal pattern components even if they are interleaved with other patterns. They differ from neural networks that use spike synchronies amongst dedicated neural channels, which is a kind of time-division multiplexing. Instead, signal types are encoded by characteristic temporal patterns rather than by “which neurons were active when”. Neural timing nets can support multiplexing and demultiplexing of complex temporal pattern signals in much more flexible ways that do not require precise regulations of neural interconnections, synaptic efficacies, or spike arrival times (Cariani 2001a; 2004).

The potential importance of temporal-pattern-based multiplexing for neural networks is fairly obvious. If one can get beyond scalar signals (e.g. spike counts or firing rates), then what kind of information a given spike train signal contains can

be conveyed in its internal structure. The particular input line on which the signal arrives is then no longer critical to its interpretation. One now has an information processing system in which signals can be liberated from particular wires. Although there are still definite neuronal pathways and regions where particular kinds of information converge, these schemes enable processing to be carried out on the level of neuronal ensembles and populations. They obviate the need for the ultra-precise and stable point-to-point synaptic connectivities and transmission paths that purely connectionistic systems require.

15.5.6 Emergent Annotative Tags and Their Uses

Both stimulus-driven and stimulus-triggered temporal response patterns can function as higher-level annotative “tags” that are added to a signal to indicate that it has a particular cognitive attribute. Neural signal tags characteristic of a particular neural assembly would signify that it had been activated. Tags produced by sensory association cortical areas would connote sensory attributes and conjunctions; those produced by limbic circuits would indicate hedonic, motivational, and emotive valences, such that these neural signal patterns would bear pragmatic content. A neural assembly producing a characteristic triggered response pattern could potentially function as a cognitive timing node (MacKay 1987).

Neural assemblies could be adaptively tuned to emit new tag patterns that would mark novel combinations of perceptual, cognitive, conative, and mnemonic activation. New tags would constitute new symbolic neural signal primitives that are associated with new attributes and concepts. The appearance of a particular tag indicates that a particular state-of-affairs has been detected. Formation of new assemblies and associated signal tags would be means by which new, dedicated “perceptual symbols” could be formed from semantically and pragmatically meaningful iconic sensory representations (Barsalou 1999).

The global interconnectedness of cortical and subcortical structures permits widespread sharing of information that has built-up to some minimal threshold of global relevance, in effect creating a global workspace (Baars 1988, Dehaene and Naccache 2001, Rose 2006). In response to a particular triggering stimulus, say a picture of a large dog, the contents of such a global workspace would become successively elaborated over time as the signals produced by different sets of neural assemblies interacted (Fig. 15.8). Successive assemblies would add their own annotations to the circulating pattern associated with various experiences with other similar animals and these would in turn facilitate or suppress the activation of other assemblies.

Linkages between particular sensory patterns and motivational evaluations could be formed that add tags related to previous reward or punishment history, thereby adding to a sensory pattern a hedonic marker. In this way, these complex, elaborated neural signal productions could be imbued with pragmatic meanings which could be conferred on sensory representations that in turn have causal linkages with the external world. Neural signal tags with different characteristics could thus differentiate

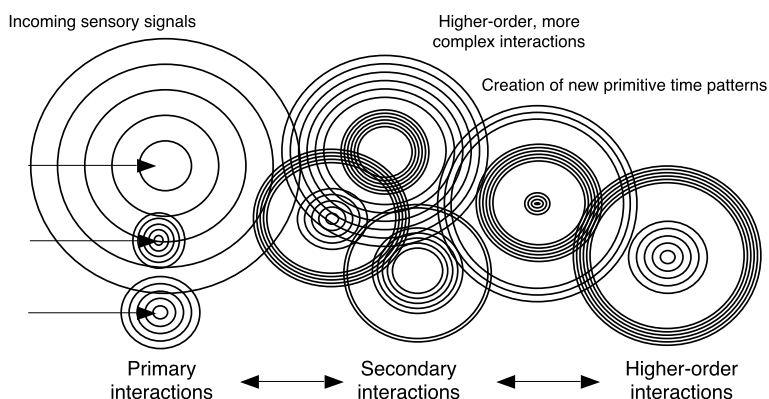


Fig. 15.8 A visual metaphor for the elaboration of evoked neural temporal pattern resonances through successive interaction. The *concentric circles* represent stimulus-driven and stimulus-triggered temporal patterns of spikes produced by neural assemblies, which interact with those of other assemblies

patterns that encode the syntactic, semantic, and pragmatic aspects of an elaborated neural activity pattern. Eventually, the signals in the global workspace would converge into a stable set of neural signals that then sets a context for subsequent events, interpretations, anticipations, and actions.

What we have outlined is an open-ended representational system in which existing primitives can be combined and new primitives formed. Combinatoric creativity is achieved in such a system by independent signal types that can be nondestructively combined in complex composite signals. The complex composites form vectors of attributes that can be individually accessed. Emergent creativity is achieved when new signal types are created through reward-driven adaptive tuning of new neural assemblies. When new signal types are created, new effective signal dimensions appear in the system.

What we do not yet know is the exact nature of the central temporal codes that would be involved, the binding mechanisms that would group attribute-related signals together into objects, the means by which relations between objects could be represented in terms of temporal tags, and how universals might be distinguished from individual instances (Marcus 2001).

Despite its incomplete and highly tentative nature, this high level schematic nevertheless does provide a basic scaffold for new thinking about the generation of novel conceptual primitives in neural networks. We want to provide encouragement and heuristics to those who seek to design mixed analog-digital self-organising artificial brains that might one day be capable of producing the kinds of combinatorial and emergent creativities that regularly arise in our own heads.

Acknowledgements I would like to gratefully thank Margaret Boden, Mark d’Inverno, and Jon McCormack and the Leibniz Center for Informatics for organising and sponsoring the Dagstuhl Seminar on Computational Creativity in July 2009 that made this present work possible.

References

- Alexander, S. (1927). *Space, time, and deity*. London: Macmillan & Co.
- Anderson, J. A., Rosenfeld, E., & Pellionisz, A. (1988). *Neurocomputing*. Cambridge: MIT Press.
- Ando, Y., & Cariani, P. G. E. (2009). *Auditory and visual sensations*. New York: Springer.
- Arbib, M. (2003). *The handbook of brain theory and neural networks*. Cambridge, MA: MIT Press.
- Arbib, M. A. (1989). *The metaphorical brain 2: neural nets and beyond*. New York: Wiley.
- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L. W., & Prinz, J. J. (1997). Mundane creativity in perceptual symbol systems. In T. Ward, S. M. Smith & J. Vaid (Eds.), *Creative thought: an investigation of conceptual structures and processes* (pp. 267–307). Washington: American Psychological Association.
- Bergson, H. (1911). *Creative evolution*. New York: Henry Holt, and Company.
- Bird, J., & Di Paolo, E. (2008). Gordon Pask and his maverick machines. In P. Husbands, O. Holland & M. Wheeler (Eds.), *The mechanical mind in history* (pp. 185–211). Cambridge: MIT Press.
- Boden, M. A. (1990a). *The creative mind*. London: George Weidenfeld and Nicolson Ltd.
- Boden, M. A. (1994). *Dimensions of creativity*. Cambridge: MIT Press.
- Boden, M. A. (1994b). What is creativity? In M. A. Boden (Ed.), *Dimensions of creativity* (pp. 75–117). Cambridge: MIT Press.
- Boden, M. A. (2006). *Mind as machine: a history of cognitive science*. Oxford: Oxford University Press.
- Broad, C. D. (1925). *The mind and its place in nature*. New York: Harcourt, Brace and Co.
- Brooks, R. A. (1999). *Cambrian intelligence: the early history of the new AI*. Cambridge: MIT Press.
- Carello, C., Turvey, M., Kugler, P. N., & Shaw, R. E. (1984). Inadequacies of the computer metaphor. In M. S. Gazzaniga (Ed.), *Handbook of cognitive neuroscience* (pp. 229–248). New York: Plenum.
- Cariani, P. (1989). *On the design of devices with emergent semantic functions*. PhD, State University of New York at Binghamton, Binghamton, New York.
- Cariani, P. (1992). Emergence and artificial life. In C. Langton, C. Taylor, J. Farmer & S. Ras-mussen (Eds.), *Santa Fe institute studies in the science of complexity: Vol. X. Artificial life II* (pp. 775–798). Redwood: Addison-Wesley.
- Cariani, P. (1993). To evolve an ear: epistemological implications of Gordon Pask's electrochemical devices. *Systems Research*, 10(3), 19–33.
- Cariani, P. (1995). As if time really mattered: temporal strategies for neural coding of sensory information. *Communication and Cognition—Artificial Intelligence (CC-AI)*, 12(1–2), 161–229. Reprinted in: K. Pribram (Ed.) (1994). *Origins: brain and self-organization* (pp. 208–252). Hillsdale: Lawrence Erlbaum.
- Cariani, P. (1997). Emergence of new signal-primitives in neural networks. *Intellectica*, 1997(2), 95–143.
- Cariani, P. (1998). Towards an evolutionary semiotics: the emergence of new sign-functions in organisms and devices. In G. Van de Vijver, S. Salthe & M. Delpos (Eds.), *Evolutionary systems* (pp. 359–377). Dordrecht: Kluwer.
- Cariani, P. (1999). Temporal coding of periodicity pitch in the auditory system: an overview. *Neural Plasticity*, 6(4), 147–172.
- Cariani, P. (2000). Regenerative process in life and mind. In J. L. R. Chandler & G. Van de Vijver (Eds.), *Annals of the New York academy of sciences: Vol. 901. Closure: emergent organizations and their dynamics*, New York (pp. 26–34).
- Cariani, P. (2001a). Neural timing nets. *Neural Networks*, 14(6–7), 737–753.
- Cariani, P. (2001b). Symbols and dynamics in the brain. *Biosystems*, 60(1–3), 59–83.
- Cariani, P. (2001c). Temporal coding of sensory information in the brain. *Acoustical Science and Technology*, 22(2), 77–84.
- Cariani, P. (2002). Extradimensional bypass. *Biosystems*, 64(1–3), 47–53.

- Cariani, P. (2004). Temporal codes and computations for sensory representation and scene analysis. *IEEE Transactions on Neural Networks, Special Issue on Temporal Coding for Neural Information Processing*, 15(5), 1100–1111.
- Cariani, P. (2011). The semiotics of cybernetic percept-action systems. *International Journal of Signs and Semiotic Systems*, 1(1), 1–17.
- Cariani, P. A., & Delgutte, B. (1996). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. II. Pitch shift, pitch ambiguity, phase-invariance, pitch circularity, and the dominance region for pitch. *J. Neurophysiology*, 76.
- Cariani, P., & Micheyl, C. (2012). Towards a theory of information processing in the auditory cortex. In D. Poeppel, T. Overath & A. Popper (Eds.), *Human auditory cortex: Springer handbook of auditory research* (pp. 351–390). New York: Springer.
- Carpenter, G., & Grossberg, S. (2003). Adaptive resonance theory. In M. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 87–90). Cambridge: MIT Press.
- Chen, J.-C., & Conrad, M. (1994). A multilevel neuromolecular architecture that uses the extradi-dimensional bypass principle to facilitate evolutionary learning. *Physica D*, 75, 417–437.
- Chung, S., Raymond, S., & Lettvin, J. (1970). Multiple meaning in single visual units. *Brain, Behavior and Evolution*, 3, 72–101.
- Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. Cambridge: MIT Press.
- Clayton, P. (2004). *Mind and emergence: from quantum to consciousness*. Oxford: Oxford University Press.
- Conrad, M. (1998). Towards high evolvability dynamics. In G. Van de Vijver, S. Salthe & M. Delpos (Eds.), *Evolutionary systems* (pp. 33–43). Dordrecht: Kluwer.
- de Latil, P. (1956). *Thinking by machine*. Boston: Houghton Mifflin.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1–2), 1–37.
- Emmers, R. (1981). *Pain: a spike-interval coded message in the brain*. New York: Raven Press.
- Fodor, J. (1980). On the impossibility of acquiring “more powerful” structures: fixation of belief and knowledge acquisition. In M. Piattelli-Palmarini (Ed.), *Language and learning: the debate between Jean Piaget and Noam Chomsky* (pp. 142–162). Cambridge: Harvard.
- Fogel, L. J., Owens, A. J., & Walsh, M. J. (1966). *Artificial intelligence through simulated evolution*. New York: Wiley.
- Goodman, N. (1972). A world of individuals. In N. Goodman (Ed.), *Problems and projects* (pp. 155–172). Indianapolis: Bobbs-Merrill. Originally appeared in *The Problem of Universals*, Notre Dame Press, 1956.
- Grossberg, S. (1988). *The adaptive brain, Vols I. and II.* New York: Elsevier.
- Hebb, D. O. (1949). *The organization of behavior*. New York: Simon and Schuster.
- Hodges, A. (2008). What did Alan Turing mean by “machine”? In P. Husbands, O. Holland & M. Wheeler (Eds.), *The mechanical mind in history* (pp. 75–90). Cambridge: MIT Press.
- Holland, J. (1998). *Emergence*. Reading: Addison-Wesley.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor: University of Michigan Press.
- Horgan, T., & Tienson, J. (1996). *Connectionism and the philosophy of psychology*. Cambridge: MIT Press.
- Izhikevich, E. M. (2006). Polychronization: computation with spikes. *Neural Computation*, 18(2), 245–282.
- John, E. R. (1967). Electrophysiological studies of conditioning. In G. C. Quarten, T. Melnechuk & F. O. Schmitt (Eds.), *The neurosciences: a study program* (pp. 690–704). New York: Rockefeller University Press.
- John, E. R. (1972). Switchboard vs. statistical theories of learning and memory. *Science*, 177, 850–864.
- Kampis, G. (1991). *Self-modifying systems in biology and cognitive science*. Oxford: Pergamon Press.

- Kanerva, P. (1988). *Sparse distributed memory*. Cambridge: MIT Press.
- Kilmer, W., & McCulloch, W. (1969). The reticular formation command and control system. In K. Leibovic (Ed.), *Information processing in the nervous system* (pp. 297–307). New York: Springer.
- Kim, J. (1998). *Mind in a physical world: an essay on the mind-body problem and mental causation*. Cambridge: MIT Press.
- Kim, J. (2008). Making sense of emergence. In M. Bedau & P. Humphreys (Eds.), *Emergence: contemporary readings in philosophy and science* (pp. 127–153). Cambridge: MIT Press.
- Koch, C. (2004). *The quest for consciousness: a neurobiological approach*. Denver: Roberts and Co.
- Machotka, P. (1980). Daniel Berlyne's contributions to empirical aesthetics. *Motivation and Emotion*, 4, 113–121.
- MacKay, D. G. (1987). *The organization of perception and action*. New York: Springer.
- Marcus, G. F. (2001). *The algebraic mind*. Cambridge: MIT Press.
- Maruyama, M. (1977). Heterogenistics: an epistemological restructuring of biological and social sciences. *Cybernetica*, 20, 69–86.
- Maturana, H. R. (1981). Autopoiesis. In M. Zeleny (Ed.), *Autopoiesis: a theory of the living*. New York: North Holland.
- Maturana, H., & Varela, F. (1973). Autopoiesis: the organization of the living. In H. Maturana & F. Varela (Eds.), *Boston studies in the philosophy of science: Vol. 42. Autopoiesis and cognition* (1980). Dordrecht: Reidel.
- McCulloch, W. S. (1965). *Embodiments of mind*. Cambridge: MIT Press.
- Miller, R. (2000). *Time and the brain, conceptual advances in brain research*. Australia: Harwood Academic Publishers/Gordon and Breach.
- Mingers, J. (1995). *Self-producing systems*. New York: Plenum Press.
- Morgan, L. (1931). *Emergent evolution* (3rd ed.). New York: Henry Holt.
- Morrell, F. (1967). Electrical signs of sensory coding. In G. Quarton, T. Melnechuk, & F. Schmitt (Eds.), *The neurosciences: a study program* (pp. 452–469). New York: Rockefeller University Press.
- Morris, C. (1946). *Signs, language, and behavior*. New York: George Braziller.
- Mountcastle, V. (1967). The problem of sensing and the neural coding of sensory events. In G. Quarton, T. Melnechuk & F. Schmitt (Eds.), *The neurosciences: a study program*. New York: Rockefeller University Press.
- Nöth, W. (1990). *Handbook of semiotics*. Indianapolis: Indiana University Press.
- Orbach, J. (1998). *The neuropsychological theories of Lashley and Hebb*. Lanham: University Press of America.
- Pask, G. (1959). *Physical analogues to the growth of a concept* (pp. 765–794). London: H.M.S.O.
- Pask, G. (1960). The natural history of networks. In M. Yovits & S. Cameron (Eds.), *Self-Organizing systems. Proceedings of an interdisciplinary conference*, May 5–6, 1959 (pp. 232–263). New York: Pergamon Press.
- Pask, G. (1961). *An approach to cybernetics. Science today series*. New York: Harper and Brothers.
- Pattee, H. H. (1982). Cell psychology: an evolutionary view of the symbol-matter problem. *Cognition and Brain Theory*, 5, 325–341.
- Pattee, H. H. (1996). The problem of observables in models of biological organizations. In E. L. Khalil & K. E. Boulding (Eds.), *Evolution, order, and complexity* (pp. 249–264). London: Routledge.
- Pattee, H. H. (2008). The necessity of biosemiotics: matter-symbol complementarity. In M. Barbi-eri (Ed.), *Introduction to biosemiotics* (pp. 115–132). Dordrecht: Springer.
- Pepper, S. C. (1942). *World hypotheses, a study in evidence*. Berkeley: University of California Press.
- Perkell, D., & Bullock, T. (1968). Neural coding. *Neurosciences Research Program Bulletin*, 6(3), 221–348.

- Piaget, J. (1980). The psychogenesis of knowledge and its epistemological significance. In M. Piatelli-Palmarini (Ed.), *Language and learning. The debate between Jean Piaget and Noam Chomsky* (pp. 23–34). Cambridge: Harvard University Press.
- Piatelli-Palmarini, M. (1980). How hard is the hard core of a scientific paradigm? In M. Piatelli-Palmarini (Ed.), *Language and learning. The debate between Jean Piaget and Noam Chomsky*. Cambridge: Harvard University Press.
- Pickering, A. (2010). *The cybernetic brain: sketches of another future*. Chicago: University of Chicago Press.
- Powers, W. (1973). *Behavior: the control of perception*, New York: Aldine.
- Pratt, G. (1990). *Pulse computation*. PhD, M.I.T.
- Raymond, S., & Lettvin, J. (1978). Aftereffects of activity in peripheral axons as a clue to nervous coding. In S. Waxman (Ed.), *Physiology and pathobiology of axons*. New York: Raven Press.
- Redgrave, P. (2007). Basal ganglia. *Scholarpedia*, 2(6), 1825.
- Rocha, L. (1996). Eigen-states and symbols. *Systems Research*, 13(3), 371–384.
- Rose, D. (2006). *Consciousness. Philosophical, psychological, and neural theories*. Oxford: Oxford University Press.
- Rosen, R. (1985). *Anticipatory systems*. Oxford: Pergamon Press.
- Rosen, R. (1991). *Life itself*. New York: Columbia University Press.
- Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron*, 24(1), 49–65. 111–125.
- Skrbina, D. (2005). *Panpsychism in the west*. Cambridge: MIT Press.
- Sommerhoff, G. (1974). *Logic of the living brain*. London: Wiley.
- Stewart, J. (2000). From autopoiesis to semantic closure. *Annals of the New York Academy of Sciences*, 901, 155–162.
- Stewart, R. M. (1969). Electrochemically active field-trainable pattern recognition systems. *IEEE Transactions on Systems Science and Cybernetics*, SSC-5(3), 230–237.
- Thatcher, R. W., & John, E. R. (1977). *Functional neuroscience, Vol. I. Foundations of cognitive processes*. Hillsdale: Lawrence Erlbaum.
- Turing, A. (1939). Systems of logic based on ordinals. *Proceedings of the London Mathematical Society*, 45(2), 161–228.
- van Fraassen, B. C. (1980). *The scientific image*. Oxford: Oxford University Press.
- Varela, F. (1979). *Principles of biological autonomy*. New York: North Holland.
- von Foerster, H. (2003). *Understanding understanding: essays on cybernetics and cognition*. New York: Springer.
- von Glasersfeld, E. (1992). Aspects of constructivism: Vico, Berkely, Piaget. In M. Ceruti (Ed.), *Evoluzione e conoscenza*, Lubrina, Bergamo, Italy (pp. 421–432). Reprinted in von Glasersfeld, *Key works of radical constructivism* (pp. 421–429).
- von Glasersfeld, E. (2007). Cybernetics and the theory of knowledge. In M. Larochelle (Ed.), *Key works in radical constructivism* (pp. 153–169). Rotterdam: Sense Publishers.
- von Neumann, J. (1951). The general and logical theory of automata, in. In L. A. Jeffress (Ed.), *Cerebral mechanisms of behavior (the Hixon symposium)* (pp. 1–41). New York: Wiley.
- Wasserman, G. S. (1992). Isomorphism, task dependence, and the multiple meaning theory of neural coding. *Biological Signals*, 1, 117–142.