Riegler, A. (2001) The Cognitive Ratchet—The Ratchet Effect as a Fundamental Principle in Evolution and Cognition. *Cybernetics and Systems* 32: 411–427.

The Cognitive Ratchet

The Ratchet Effect as a Fundamental Principle in Evolution and Cognition

Alexander Riegler

CLEA, Vrije Universiteit Brussel
Krijgskundestraat 33
B-1160 Brussels, Belgium
email: ariegler@vub.ac.be

Abstract

Is the evolution and performance of cognition an asymmetric, directed process? The standard externalist definition of evolution as a mechanism of variation and selection cannot account for directed developments such as an increase in complexity of cognition. A separate cause which is responsible for complexification requires us to deviate from the usual description of cognition as ahistorical logical problem-solving: the anticipation-driven nature of behavior, and hence cognition based on a ratchet effect. On a structural level it is evident that the growth of behavioral competence, similar to biological structures, must build on previously available components, thus yielding a canalization of development. This unavoidably introduces asymmetry in the cognitive evolution. Numerous examples show the relevance of the proposed mechanisms in biology, psychology, and the artificial sciences.

Cybernetics and Systems 32: 411–427.

Running title: The Cognitive Ratchet

Introduction

Since the days of Darwin, evolution has been considered a simple, robust yet powerful mechanism of variation and selection which accounts for changes in our world. Starting from its established place in the biological sciences the use of evolutionary methods spread even to the abstract domains of genetic algorithms (Holland 1975) of genetic programming (Koza 1992) where evolutionary operators are supposed to improve the functionality of any given program-code. Similarly, cognitive capabilities could be explained as the result of the working of evolution on simple structures which undergo directed development towards higher complexity. It is believed that complexity increases in evolution, and that a direction and progress is caused by evolutionary forces.

Clearly, cognition is a dynamical process, and as such subject to influences which may affect its further development. I will argue that progression in the evolution of cognitive capabilities is not rooted in the mere interaction of variation and selection of the supposed "fittest." As Saunders & Ho (1976) already claimed, "whatever it is that drives evolution, the direction is determined not by the direction of the force but by the nature of the process." This requires us to assume that there are "two separate laws of evolution, survival of the fittest, and increase in complexity."

In this paper I will discuss what drives evolution, on which mechanisms complexification is based, and unroll its implications for cognition. Beginning with biological evidence of the anticipation-driven character of behavior and cognition, I will point out that there exists a deeper relationship between procedural anticipation and the internal structural level of biological systems responsible for complexification. The explanation of natural cognition and the design of artificial cognitive systems might therefore borrow from these insights.

Anticipation

When a snake is chasing a mouse, it uses its eyes exclusively. None of the other sensing modalities of the snake are involved (Sjölander 1995). When it comes to strike the mouse, the snake switches to olfaction. It then "assumes" that the prey is no longer able to move, and starts to swallow it by employing its touch sensors. An observer will evidently get the impression that the hunting strategy of the reptile is like a relay race in which each modality passes the courier over to the next one, trusting that the predecessor did its job correctly and provided a context in which the current modality will suffice. In other words, the phylogenetic development of the snake equipped the animal with some kind of implicit anticipatory mechanism which draws on the fact that each action will generate a new context suitable for the next piece of the action sequence. The behavior of the snake is therefore asymmetric with respect to time and order of modalities. The snake will neither be able to locate the mouse by its touch sensors nor coordinate its swallowing action by employing visual clues.

Such an anticipatory strategy is not limited to animals unable to integrate various sensory channels into one coherent picture (Riegler 1994a). Sjölander (1995) pointed out that a dog hunting a hare "does not need a full picture of a recognizable hare all the time to conduct a successful hunt. It is able to proceed anyway, guided by glimpses of parts of the hare, movements in vegetation, sounds, smell, etc. If the hare disappears behind a bush or in a ditch the dog can predict the future location of the hare by anticipating where it is going to turn up next time, basing this prediction on the direction and speed the hare had when seen last." These examples strongly suggest that animals seem to be able to anticipate future events.

Approaches to Cope with the Environment

Although the presence of anticipation as the basic element of behavior is somewhat fundamental in nature, in artificial systems, the computation of behavior is performed differently; it corresponds to the information-processing paradigm that defines the cognitive system as a system which receives data through its senses and which tries to categorize this information. A wealth of “information” is provided by the “outside” environment from which some essential features must be selected in order to decrease the enormous amount of complexity. How could simple animals ever accomplish this task?

Why the Information-Processing Approach Fails

In essence, the information-processing paradigm is an ahistorical logical approach. As such, it is prone to a variety of unsolvable problems. The most prominent are (a) the frame problem (Dennett 1984), which testifies the fruitless hope that we could formulate knowledge about the world and possible actions therein in an appropriate symbolic way; and (b) the symbol grounding problem (Harnad 1990) which addresses how the meanings of symbols can be grounded in anything but other symbols, thus rendering the entire system meaningless. Alone, these two problems leave open the question of appropriate representation of nontrivial environments (Peschl and Riegler 1999).

Dennett’s well-known analogy (1984) illustrates the shortcomings of the assumption that creatures can tackle their struggle for life in terms of such an ahistorical logical reasoning. A robot learns that its spare battery, its precious energy supply, is locked in a room with a time bomb set to go off soon. To solve this problem, the robot has to develop plans in order to rescue the battery which is located on a wagon. Equipped with a logical inference system it is able to quickly reason that pulling the wagon out of the room will also move the battery out of the supposedly dangerous room. But the robot fails because it does not pay attention to the implications of its planned actions. It did not take into consideration that the bomb is also located on the wagon and, therefore, stayed close to the battery regardless of where the robot moves the wagon. A descendent of the robot is constructed in a way that would allow it to foresee the effects of its actions. Taking possible side-effects into account, however, does not help either. As the world is very complex, an exhaustive list of all side-effects would take too long to take any action in real-time. Hence, the robot must know how to distinguish between relevant and irrelevant side-effects. But even this process of discrimination needs an enormous amount of computation, all the more as each of the possible effects must be assigned with some (quantitative) credit in order to evaluate their usefulness. Therefore, in a logical framework, too many logical implications of even the simplest actions have to be taken into account. This results in endless computations that prevent creatures from taking those actions.

The Construction of Information

Evidently, the information-processing paradigm resembles what Popper (1979) calls the *Bucket Theory of Mind*. This is the idea that “there is nothing in our mind which has not entered through our senses.” Cognition—and here cognition can be understood in its broadest sense, namely, as the capabilities necessary to successfully cope with an environment—is metaphorically seen as a bucket that is incrementally filled with knowledge through our sensory organs, like an information retrieval agent fills its database with information chunks from the internet.

Popper also offers an opposite perspective. The *Searchlight View* emphasizes that individuals actively construct knowledge in the form of a priori unjustified theories. Experience does not provide these theoretical conjectures. Rather, observations are used to select among competing theories and eventually weed out unsatisfactory ones. In other words, we first construct our world-view. This leads to anticipations which can be tested. As long as expectations match with ongoing experiences, we maintain our scaffolds of conjectures. Whenever anticipations are not fulfilled we are likely to reconstruct them. The Searchlight View, therefore, corresponds to what I initially called anticipation-driven behavior.

Of course, Popper's essay preliminarily addresses philosophy of science. But it has a wider range of applications since we can argue that scientists as well as animals are cognitive systems at different levels of sophistication. From the perspective of an observer, they all live in an environment with which they cope and from which they extract information in order to be successful. However, what is missing is that Popper isn't interested in the question of how we arrive at new conjectures, how we modify existing theories, etc. This point of view goes back to Reichenbach's distinction between the context of discovery and the context of justification (Reichenbach 1938). While the latter is subject to rigorous examinations with logical means, the context of discovery is thought to be a merely psychological issue and thus not subject to rationality. In classical neo-Darwinism, an analogical situation is prevailing. The standard explanation that the process of evolution "can be conceptualized as an interplay between variation and selection" (Heylighen 1999) shifts the emphasis to selection ("survival of the fittest") rather than paying close attention to the process of variation. The problem with standard neo-Darwinism is that it is in some respect a much too wide framework. As Ball (1999) writes, "natural selection is not entirely satisfying. Not because it is wrong, but because it says nothing about mechanism." Nor is this *externalist* perspective of pure random mutation able to explain the speed of evolutionary processes. As I will point out in this paper, *internalist* aspects are the missing link between Popper's Searchlight View and a full understanding of cognition.

Anticipation and Evolution

Examples in support of Searchlight view are abundant. The egg-retrieving behavior of geese (Lorenz & Tinbergen 1939) is yet another example for embedded anticipations in animal behavior (Riegler 1994a). If an egg falls out of the nest, the incubating goose rolls it back into the nest with its bill. If the egg is removed midway through this action pattern, the bird continues until its bill has reached the border of the nest. During this process, the goose apparently neglects environmental events: if an action pattern has been triggered, the processing of sensor information is reduced until the pattern terminates. During egg retrieving it is necessary to compensate for any sideward rolling of the egg. Obviously, from an evolutionary perspective it makes sense to expect that starting to roll back an egg will eventually get the egg back in the nest. Perturbations such as egg-stealing ethologists are the exception to the rule. Therefore, although ridiculously simple-minded at first sight, the behavior of the goose is evolutionarily successful, as it integrates the proper anticipations.

Anticipation and Perception

With regard to perceptual competence, Sacks (1995) delivers an excellent example of anticipation-driven perception. He describes the case of a man, Virgil, who had been blind since early childhood. At the age of 50 his eyelight was restored. In contrast to the general expectation, this was no help to Virgil since the way he had been living as a blind person was incompatible with the way normal sighted people perceive and organize their world. With effort and practice, he was able to interpret some of the visual data in terms of the world as he had known it through his other senses, but he had immense difficulty learning these interpretations. For instance, visually he couldn't tell his dog from his cat. For him, due to the lack of visual impressions, the temporal aspect of his world had priority. He recognized things by feeling their surface in a particular order. His didn’t get lost in his own apartment because he knew that after entering he would encounter the various pieces of furniture in a particular sequence which he perceived in a temporal order. To put it differently, he was living in a world of anticipation, of subsequent checkpoints which acted like the handing-over in a relay race. A particular cupboard was followed by a table, so once he reached the cupboard he anticipated going to the table in the next step. Apparently, we live in another world, where items are recognized “immediately” by visual perception. But are they? Lessons from artificial intelligence show that image recognition can’t simply rely on visual clues only. More likely, we also apply a certain order of recognition steps when we look at things.

Anticipation and Problem-Solving

We can extend the principle of anticipation even further towards human behavior. Classical experiments reveal the “if it ain’t broken, don’t fix it” psychology in human cognition (Riegler 1998).

Duncker (1935/1945) posed the task to support a candle on a door. The available items were matches and a box filled with tacks. Since the test subjects considered the box as a mere container they failed to empty it in order to tack it to the door where it could serve as a support for the candle. Our thinking is canalized (or fixed) with respect to the way we have learned to deal with things.

The water-jug problem, studied by Luchins (1942), provides empirical data for this assumption of “mechanization of thoughts.” He asked test subjects to measure out a specific quantity of water using a set of three jugs with known volume. The first two problems Luchins posed could be solved by applying a certain sequence of pouring water from one jug into another. Test subjects had no problems discovering this procedure. Quite the contrary. They got used to it and tried to apply it to further tasks. Like the adage says, “If it ain’t broken, don’t fix it.” What the test subjects overlooked was that much simpler procedures would have led to the same result, simply because their inductively working mind was set to the previously successful strategy.

The test subjects in these tests fail to accomplish given tasks since their minds are set in certain canalized ways of looking at things and problem-solving. Observations from every-day life support this result, as we find ourselves in situations, in which things are used in one particular context. For example, we use a hammer to drive nails into a wall, matches to light a fire. In fact, things do not seem to exist “outside” their domains of functionality because the anticipation of the result of dealing with things very much determines cognition.

Unavoidably, we arrive at the insight that observable behavior and cognition are asymmetric processes which clearly point forward in time. What can we say about the underlying structure of behavioral systems? Are their evolution and ontogenetic development asymmetric too?

Evolution

To investigate the role of asymmetry in the creation and development of behavioral structures, it is worth considering some examples.

For a long time evolutionists have been puzzled by a phenomenon known as Haeckel’s Law (1874). It states that the ontogenesis of individuals is the shortened recapitulation of phylogeny in the form of extracts such that during the ontogenesis of individuals, phylogenetically old patterns are repeated. For example, in early embryological periods, mammals develop a complete gill circulation, although gills are completely useless for mammals in their environment. In the phylogenetic tree, mammals are descendants of archaic forms of fishes. All vertebrates establish a *chorda dorsalis,* but no vertebrate has it after birth. Several examples demonstrate that old phylogenetic phases and states have to be passed through.

Another biological phenomenon is that of atavisms. Here, bygone conditions taken from phylogeny are re-established. Examples of atavisms are: the Darwin humps at the outer ear conch, the existence of reduced forms of nipples on human males, the coccygeal bone as the last tail vertebra, etc. All these parts have no selective advantage. They must be the product of not variation and selection but another mechanism.

We can best explain such phenomena by assuming some kind of "tradition" in the phylogenetic tree. No organic state exists without tribute to its ancestry, so that all building states are subsequent series of coordinations (Riedl 1977). To give another example, within rudimentation processes the chains of information can only be cut at their ends. The way in which the eye of a cave fish has been decomposed over time is the exact reverse of the evolutionary process that formed the previously fully featured eye of the cave fish' predecessor. Apparently, reversions of development can occur. But the probability for the reverse process is much smaller. We might, therefore, better speak of *asymmetry* rather than reversibility. Strictly speaking, such a reversal of biological features cannot be considered an evolutionary process, as it takes place in the absence of evolutionary pressure. The sight of the cave fish did not disappear due to selection for simplicity but because having sight or not has become irrelevant in the environment of the fish. As (Heylighen 1999) put it, "The function was lost, not because of selection for simplicity, but because of the lack of selection for sight. It is likely that sight was destroyed by a sequence of undirected, deleterious mutations, which in a different environment would have been eliminated by selection." Transferring this idea to the behavioral level, we can say that unlearning a certain behavior, for example, is not a directed developmental process but the result of not using it. If we think of the phenomenon of forgetting this becomes quite obvious.

Riedl (1977) pointed out that the tradition-principle can be interpreted as the effect of interdependencies among components, which result in system-internal *canalizations* (cf. also Waddington 1957). Due to these interdependencies and hierarchical structures among genes, the freedom of variability is enormously restricted, compared to the standard neo-Darwinist model of evolution. Old subsystems like *chorda dorsalis* are irreplaceable, and have high functional burden. Therefore, we can find it in virtually all succeeding life forms. The chances of successful mutations are very small. Exactly this is shown in Haeckel's Law. As old patterns are the basis for newer ones, they have to be repeated during ontogenesis.

Hierarchy

What can we learn from this biologically motivated example about the behavioral component of artificial systems? Although canalization was primarily proposed for morphological structures, one may apply it to knowledge structures as well (Riegler 1994b). In this view, thinking is a canalized process which builds upon previous knowledge structures. This yields the advantage of speeding up developmental processes at the expense of building a rigid system consisting of interdependent parts.

Similarly, Simon (1969) pointed out that "hierarchic systems will evolve far more quickly than non-hierarchic systems of comparable size" since the "time required for the evolution of a complex form from simple elements depends critically on the number and distribution of potential intermediate stable forms." His famous example is that of the two watchmakers Tempus and Hora. They have to build clocks consisting of 1000 parts. Unfortunately, they are interrupted in their work at random moments causing an unfinished clock to fall apart. In order to cope with the annoying interruptions, Hora divides the design of a watch into subassemblies of 10 parts each such that in the worst case only 10 components fall apart. Simple calculation shows that Hora's strategy yields a tremendous advantage with regard to the number of completed watches compared to Tempus' linear style of working. Furthermore, Hora's technique also introduces a direction. As Simon maintains, we don't need to assume any explicit teleological mechanism, no *causa finalis*, in order to account for directedness, as direction "is provided […] by the stability of the complex forms, once these come into existence."

The Hierarchical Organization of Cognition

How can we use this insight to explain cognitive capabilities? Suppose we equip an artificial creature with a knowledge system that is composed of sets of connected production rules. Such rules consist of distinctions (A, B). The distinctions are connected in the form of elementary predictions ("if A, then B") in which both sides may consist of chains of sub-ordinated distinctions. Learning takes place by modifying the elements of the rules, i.e., adding, dropping, and rearranging. This produces a potentially vast number of combinations, like we find in any natural context. By introducing interdependencies and hierarchical arrangements the evolution of these knowledge systems can be canalized in order to escape combinatorial explosion without becoming arbitrary (Riegler 1994a).

Concretely speaking, this means that we provide a structure which the snake in the introductory example obviously doesn't have. These three modalities which are employed sequentially could be combined as attributes of a central construct—for example, a mouse. If such a mouse vanishes into a mouse hole, it no longer exists for a snake, while a cat, for instance, remains in front of the hole and waits for the mouse to re-appear. Under normal circumstances, the snake need not know the concept of an entire mouse. It only perceives the mouse in slices of modalities. One could encompass its hunting behavior in three rules. R1: if see mouse then come closer. R2: if close to mouse then strike. R3: if struck then start eating. For a snake these are merely 3 subsequent rules; for us humans (as for most mammals) it forms the concept of a mouse. In this perspective, we can say that any concept is a collection of single rules and other, simpler concepts. We conclude that knowledge is a hierarchical scaffolding of rules in which rules at a higher level (i.e., concepts) are dependent on lower-level rules (Riegler 1994a). Now evolution made a snake a good hunter. That is, R1–R3 are "approved" by a sufficiently high rate of successful hunting each time the creature felt hungry. Thus there are good reasons for keeping this set of rules (i.e., this concept) rather than any other set of rules—they become "stable forms" in the words of Simon (1969). The advantage is evident. From now on, the new concept can be used as an elementary component in higher order concepts. Not only mice can be potential preys to snakes but also other small animals. While a biologist would now start a hierarchical classification of the animal kingdom, a wild animal could find this information useful to extend its menu without multiplying the same hunting strategy for different animals. (As programmers, we apply the same principle when we partition a computer program into subroutines.) In even more detail, the behavioral component of the artificial system should look separately at conditions and actions. Conditions are interpreted as the "recognition" capabilities of the system (i.e., distinctions in the above sense) whereas the "then"-part specifies possible actions. This way it is easy to assign one and the same action strategy to different perceivable objects. To go even one step further, the system may distinguish between "ontogenetic" and "phylogenetic" sets of rules. Compound rules may contain components from both sets. This way it becomes possible to build individual experience upon knowledge inherited from previous generations.

Interestingly, Latour (1987) uses the notion of "black boxes" in his philosophy of science. A black box denotes any theory that has become so well-supported that it is regarded as being unquestionably true. As such it serves as a basic entity in further discourses. This corresponds to the formation of "concepts" as described above.

Irreversibility

Is the evolution of such an (artificial) behavioral system irreversible? In biological systems, clustering of independent genomes is irreversible. This renders adaptations to changing environments more difficult. These so-called *genetic loads* (Riedl 1977) are the reason why a giraffe has the same number of neck bones as a dolphin although it could use more in order to increase its pliancy. In the above-mentioned artificial system we find a similar situation. Compounds that have been approved by their evolutionary success are treated as elementary components by compounds at higher levels. The earlier a compound appeared in the history of the system the more likely it is that many other structures are dependent on it. Deleting such an old component (say, the concept that describes a prey) will let the entire system collapse like a card house. While there is always a theoretical chance that a component can be replaced by another component that fulfills an analogous role, chances are higher that components are successfully added than removed. This finding is supported by Saunders and Ho (1976). They argue that the more complex a system is, the more organization it needs in order to survive. Consequently, it is in general much easier to add a component, which is not likely to do much harm, than to take away a component, that is likely to disturb a complex network of interdependencies.

Arthur (1993) localizes a “general law” behind this phenomenon. He writes that “complexity tends to increase as functions and modifications are added to a system to break through limitations, handle exceptional circumstances, or adapt to a world itself more complex.” Heylighen (1999) goes even a step further and speaks of an infinite jigsaw puzzle: “Every system that is selected can be seen as a piece of the puzzle that has found a place where it fits, locking in with the neighboring pieces. However, every newly added piece will add a segment to the puzzle’s outward border, where further pieces may find a place to fit. The more pieces added to the puzzle, the larger the border becomes, and the more opportunities there are for further pieces to be added.” This evokes the idea of a non-strict irreversibility (or asymmetry) again. Systems, whether natural or artificial, are driven into a continuous complexification of their structure, thus yielding asymmetry in time, caused by internalist rather than externalist mechanisms.

In the long term, changes will only take place in one, privileged direction. Evidently, this resembles a ratchet, i.e., a mechanism that consists of a wheel with asymmetrically skewed teeth and a spring-loaded pawl. Such ratchets have attracted great interest, largely theoretical, from diverse areas of science and technology. In evolutionary biology, Müller’s ratchet (1932) states that small populations are doomed to accumulate deleterious mutations and without sexual recombination, lines would eventually go extinct because this genetic load will detoriate the quality of the gene-pool. In physics, the idea of a ratchet was introduced by Feynman as a simple example of a thermal engine in order to illustrate implications of the 2nd law of thermodynamics (Feynman et al. 1966). As we have seen so far, the evolution and working of behavioral and cognitive systems resembles a ratchet as well. Accumulation of developmental components and anticipation-driven cognition let the wheel turn preferentially in one direction — an asymmetric, though not strictly irreversible mode.

The Directedness of Evolution

The implications for evolution are evident. The inherent tendency of evolution is to always explore the neighborhood but this is constrained by interdependencies among components. Any evolutionary explanation must take this into account. So must artificial learning techniques which employ evolutionary principles. In such evolutionary scenarios, complexification of structures is implicitly assumed. Gould (1996) pointed out that the identification of evolution with progress is the result of a heavily biased and hence misleading view of the course of evolution. This is based on the bias to think of evolution as a progression from simple life forms to human beings, the “Great Chain of Being” (Lovejoy 1936). It assumes human beings to be higher and therefore fitter organisms than their protozoic ancestors. Wuketits (1997) even attributes this view as pre-scientific. He writes that “the idea of progress […] is deeply rooted in philosophical preconceptions of evolution rather than in unprejudiced scientific studies.” As the successful existence of single-cell animals shows, there is no evolutionary pressure in the traditional sense towards complexification of structures, including the development of higher cognitive abilities. This is in a line with McShea (1996) who writes that “[n]atural selection does not seem to select for complexity.” Selection is only responsible for the elimination of unfit structures that are not viable enough to compensate for environmental distortions; it says nothing about the creation of new structures. Complexification must thus be explained in a different, internalist way. Clearly, behavioral and cognitive competencies have been developing over millions of years. In the absence of a corresponding evolutionary pressure, the responsible mechanism is canalization which is a result of the interdependency of subcomponents. Given this perspective, evolution has a direction and is asymmetric without being equivalent to progress. It can thus be compared with a ratchet.

Some Implications for Artificial Systems

The observations in this paper give rise to formulating some general guidelines for the design of artificially cognitive systems. Firstly, such systems should be capable of generating anticipations. This not only reverses the usual way of information processing, thus eliminating its bottlenecks, it also renders exogenous fitness functions obsolete. The success of a creature is determined by how often the anticipations are met. To this end, a rule contains checkpoints which try to match the current state of anticipations with observations of selective endo- or exogenous variables. Secondly, we can easily attach learning to such a system. A rule is deleted or modified in case the anticipations are not fulfilled at its checkpoints. Furthermore, systems that can evolve their components in a hierarchical way are faster than flat systems (cf. Simon 1969). Evolution-approved components are integrated into compounds, thus re-usable for similar tasks. Finally, as shown in (Riegler 1994a, 1994b) an easy way to take advantage of hierarchical behavioral systems is by employing a dynamical state space system controlled by an explicit set of rules. The states in this systems do not carry any pre-defined meaning, which saves the system from the symbol-grounding problem. It is alone the relational structure *among* states that provides meanings private to the system rather than the relationship between state and designer of the system (Peschl and Riegler 1999). While this means the loss of semantic transparency, the rule-based design of the system may still provide a certain degree of comprehensibility for the (human) observer (Riegler 1994a).

Conclusion

In this paper I discussed the relevance of temporal asymmetry for biological, and by extension, for artificial cognitive systems. Both performance and developmental aspects of such systems are asymmetric, i.e., irreversible in a weak sense. On the one hand, behavioral competence is grounded in the ability to generate anticipations, and thus clearly headed forward in time. On the other hand, the evolution and development of living systems follows the principle of canalization which forces irreversible complexification upon the system in question.

So regarding the question “What drives Evolution?”, we can clearly distinguish two different causes (or “forces” in the sense of Saunders & Ho 1976): (a) External selection does not yield any direction in evolution. This argument is also evident from authors like McCoy (1977) who maintains, “organic evolution is a process of divergence and wandering…” and Gould

(1996) who claims that “[e]volution is a process of constant branching and expansion.” (b) Internal selection, i.e., the interplay of mutually dependent and hierarchically arranged components, does provide this direction. Acting like a ratchet, it is responsible for the “bursts” of complexification to which we owe our existence. However, comparing this with a general notion of progress is like comparing apples with pears. Saunders and Ho’s (1976) identification of *two* separate dimensions in evolution, survival vs. increase in complexity, is more appropriate. Their argument that a more complex system needs a more extensive organization goes hand in hand with Arthur’s (1993) observation that complexity tends to increase in order to fill more and more niches of exception and special cases. The overall picture, though, is that of a ratchet the teeth of which are the generations of ancestors, and the pawl of which is the improbability necessary to mutate away all canalized developmental pathways at once. (Theoretically, one could bend the spring of the pawl wide enough in order to turn the toothed wheel backwards.)

Finally, there is psychological evidence (as shown with several experimental results) that a similar mechanism, anticipation (Riegler 1994b), also applies to cognition. The teeth of this ratchet are the cognitive encapsulation (or black-boxing) of concepts, ideas (and theories), its pawl is the limitation of the biological cognitive substrate which has to work with black-boxes in order to proceed. Among others, this view accounts for our tendency to think in paradigms, and for selective perception—cf. also Neisser (1975) who claims that the cognitive organism picks up information from its environment rather than passively filters it. Already earlier Kelly (1963) emphasized that, “[a] person’s processes are psychologically channelized by the ways in which he anticipates events.” For artificial cognitive systems it bears the potential of accelerating cognitive development and performance, and makes it possible to forgo exogenous fitness criteria. Therefore, embracing the concept of canalization as basic mechanism of asymmetric evolution and cognition is a promising candidate for explaining natural systems and engineering artificial ones.

Acknowledgments

This work was supported by the Fund for Scientific Research – Flanders (Belgium) (FWO).

References

Arthur, W. Brian “Why Do Things Become More Complex?” *Scientific American* (May 1993): 92.

Ball, Philip. *The self-made tapestry*. Oxford: Oxford University Press, 1999.

Dennett, Daniel C. “Cognitive Wheels: The Frame Problem of AI.” In *Minds, Machines, and Evolution: Philosophical Studies*, edited by Christopher Hookway, pages 129–151. London: Cambridge University Press, 1984.

Duncker, Karl. *Zur Psychologie des produktiven Denkens*. Berlin: Springer, 1935. English translation: “On Problem Solving.” *Psychological Monographs*, 58 (1945), 1–112.

Feynman, Richard P., Robert B. Leighton, and Matthew Sands. *Lectures in physics.* Reading, Mass.: Addison-Wesley, 1966.

Gould, Stephen J. *Full House: The Spread of Excellence From Plato to Darwin*. Nw York: Harmony Books, 1996.

Haeckel, Ernst. *Anthropogenie oder Entwicklungsgeschichte des Menschen*. Leipzig: Engelmann, 1874.

Harnad, Stevan “The Symbol Grounding Problem.” *Physica D,* 42 (1990): 335–346.

Heylighen, Francis. “The Growth of Structural and Functional Complexity during Evolution.” In *The Evolution of Complexity,* edited by Francis Heylighen *et al.*, pages 17–41. Dordrecht: Kluwer, 1999.

Holland, John H. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, 1975.

Kelly, George. *A Theory of Personality*. New York: Norton, 1963.

Koza, John R. *Genetic Programming: On the Programming of Computers by means of Natural Selection.* Cambridge, Mass.: MIT Press, 1992.

Latour, Bruno. *Science in Action. How to Follow Scientists and Engineers through Society.* Cambridge, Mass.: Harvard University Press, 1987.

Lorenz, Konrad Z., and Niko Tinbergen. “Taxis und Instinkthandlung in der Eirollbewegung der Graugans.” *Zeitschrift für Tierpsychologie,* 2 (1939): 1–29.

Lovejoy, Arthur O. *The Great Chain of Being*. Harvard University Press, 1936.

Luchins, Abraham S. Mechanization in Problem Solving. *Psychological Monographs,* 54 (1942), No. 248.

McCoy, J. Wynne. “Complexity in Organic Evolution.” *Journal of Theoretical Biology,* 68 (1977): 457–458.

McShea, Daniel W. “Metazoan Complexity and Evolution: Is There a Trend?” *Evolution*, 52 (1996): 477–492.

Müller, Hermann J. “Some genetic aspects of sex.” *American Naturalist*, 66 (1932): 118–138.

Neisser, Ulric. *Cognition and Reality*. San Francisco: Freeman, 1975..

Peschl, Markus and Alexander Riegler. “Does Representation Need Reality?” In *Understanding Representation in the Cognitive Sciences*, edited by Alexander Riegler *et al.*, pages 9–17, New York: Kluwer Academic/ Plenum Publishers, 1999.

Popper, Karl R. *Objective knowledge: An evolutionary approach* (rev. ed.). Oxford: Clarendon Press, 1979.

Reichenbach, Hans. *Experience and prediction.* Chicago: University of Chicago Press, 1938.

Riedl, Rupert. “A systems analytical approach to macro-evolutionary phenomena.” *Quarterly Review of Biology,* 52 (1977): 351–370.

Riegler, Alexander. “Constructivist Artificial Life.” Ph.D. diss. Vienna University of Technology, 1994a.

Riegler, Alexander. “Constructivist Artificial Life: The constructivist-anticipatory principle and functional coupling.” In *Genetic Algorithms within the Framework of Evolutionary Computation*; edited by Jörn Hopf. Saarbrücken, Germany, Max-Planck-Institut für Informatik, MPI-I-94-241, pp. 73–83, 1994b.

Riegler, Alexander. “The End of Science: Can We Overcome Cognitive Limitations?” *Evolution & Cognition,* 4 (1998): 37–50, .

Sacks, Oliver. “To See and Not to See.” In *An Anthropologist on Mars*. New York: Alfred A. Knopf, 1995.

Simon, Herbert A. “The Architecture of Complexity.” In *The Sciences of the Artificial*. MIT Press, Cambridge, 1969.

Saunders, Peter T. and Mae-Wan Ho. “On the increase in complexity in evolution.” *Journal of Theoretical Biology*, 63 (1976): 375–384.

Sjölander, Sverre. “Some cognitive breakthroughs in the evolution of cognition and consciousness, and their impact on the biology of language.” *Evolution & Cognition*, 1 (1995): 3–11.

Waddington, Conrad H. *The Strategy of the Genes*. London: Allen and Unwin, 1957.

Wuketits, Franz M. “The philosophy of evolution and the myth of progress.” *Ludus Vitalis*, vol V., no 9 (1997): 5–17.