Heylighen F. (1994) “Fitness as Default: the evolutionary basis for cognitive complexity reduction”, in: *Cybernetics and Systems '94,* R. Trappl (ed.), (World Science, Singapore), p.1595-1602.

FITNESS AS DEFAULT: the evolutionary basis for cognitive complexity reduction

Francis HEYLIGHEN*

PO, Free University of Brussels, Pleinlaan 2, B-1050 Brussels, Belgium; fheyligh@vnet3.vub.ac.be

ABSTRACT: Given that knowledge consists of finite models of an infinitely complex reality, how can we explain that it is still most of the time reliable? Survival in a variable environment requires an internal model whose complexity (variety) matches the complexity of the environment that is to be controlled. The reduction of the infinite complexity of the sensed environment to a finite map requires a strong mechanism of categorization. A measure of cognitive complexity (C) is defined, which quantifies the average amount of trial-and-error needed to find the adequate category. C can be minimized by "probability ordering" of the possible categories, where the most probable alternatives ("defaults") are explored first. The reduction of complexity by such ordering requires a low statistical entropy for the cognized environment. This entropy is automatically kept down by the natural selection of "fit" configurations. The high probability, "default" cognitive categorizations are then merely mappings of environmentally "fit" configurations.

1. INTRODUCTION

It is a recently popular trend to extend Darwin's evolutionary thinking from biology to other disciplines. For example, evolutionary economics (Saviotti & Metcalfe, 1992), evolutionary psychology (Buss, 1991), evolutionary computation, and the evolution of chemical compounds or elementary particles following the Big Bang, have all become fashionable subjects of study. One of the first of these new approaches is *evolutionary epistemology*, which was first defined by Campbell (1974). Its main thesis is that all knowledge is the product of variation and natural selection. This applies as well to the primitive knowledge stored in the genes, which allows an organism to adapt to its environment, as to the sophisticated theories of science, which undergo variation when a scientist speculates and selection when inadequate theories fail to pass empirical tests.

The Principia Cybernetica Project (Heylighen, Joslyn & Turchin, 1991) aims to carry the study of the evolutionary origin of systems to its logical end point. This means that we should not restrict our analysis to one disciplinary level (cognitive, social, biological, ...), but look at the interconnections between systems of different types, so that the development of each level can be understood as a continuation of evolution at the level below. In the limit, this should lead us to reconstruct the complete chain of variation and selection processes producing complexity, from the appearance of elementary particles to the intricate structures of present society. We have called this approach “Metasystem Transition Theory” (Heylighen, Joslyn & Turchin, in print), emphasizing the spontaneous transitions to a higher (“meta-”) level, which form the quanta of evolution.

In the present paper I wish to apply this philosophy to the origin of knowledge, however, not analysing the transitions that produced knowledge (this is being done elsewhere, Heylighen, 1991a, in print), but studying the evolutionary preconditions that make knowledge at all possible. A fundamental question every epistemologist should ask

* Postdoctoral Researcher NFWO (Belgian National Fund for Scientific Research)

is the following: given that knowledge consists of extremely simple models of an infinitely complex reality, how can we explain that knowledge is still most of the time reliable? I will try to answer that question by linking the mechanism of *defaultreasoning* to the natural selection of cognized phenomena.

2. SURVIVAL AS A CONTROL PROBLEM

Evolutionary epistemology (Campbell, 1974) assumes that the function of knowledge is to secure the survival or "*fitness*" of an organism, by helping it choose actions adequate for the given situation, while avoiding the risks of blindly trying out an action whose result may be fatal. To this analysis, cybernetics adds the premise that survival in a variable environment is a *control* problem, which requires adequate compensation of perturbations that make the system deviate from its goal (maintaining or increasing fitness). Perhaps the most famous principle of control is Ashby's *Law of Requisite Variety* (1958). It states that, in order to achieve complete control, the variety of compensatory actions a control system is capable to execute must be at least as great as the variety of perturbations that might occur. To that must be added Conant and Ashby's (1970) principle that "*every good regulator (controller) of a system must be a model of that system*". Together, the two principles imply that perfect control can only be established if there exists a one-to-one (bijective) mapping from the set of all possible perturbations to the set of all counteractions the system can execute.

This isomorphism between perturbations and potential actions reminds one of the classical reflection-correspondence theory, which sees elements of knowledge as "mirror images" of objects in the world. The obvious question in this view is: how can an infinitely extended world be isomorphically mapped onto a finite (and often extremely simple) cognitive system? The traditional answer is that the concrete mapping is not isomorphic (one-to-one) but homomorphic (many-to-one): most details or distinctions about the world are filtered out. For example, two slightly different frequencies of light will both be perceived as "red" and stored under that single heading in a person's brain.

However, under Conant and Ashby's analysis, many-to-one mappings, implying loss of information, also imply loss of control: the less information survives the mapping, the greater the variation in the outcome of the actions, and thus the larger the fluctuation around the goal state. This relation between the varieties (V) of external perturbations (E), actions (A), and outcomes (O) can be derived from Ashby's (1958) law that "only variety can destroy variety": $V(O) \geq V(E) - V(A)$.

It implies that an infinite variety of perturbations, controlled by a finite variety of actions, will still leave an infinite variety of outcomes. We must conclude that knowledge, viewed as a finite map of an infinite territory, appears like a very limited tool for achieving control. In order to explain why knowledge nevertheless seems so effective, we must note several things:

1) Ashby's law of requisite variety should not be taken as an absolute requirement. When the system is in a naturally stable state, the variety of outcomes will be *automatically* decreased or damped, even without intervening actions (see my "principle of asymmetric transitions", Heylighen, 1992).

2) unlike the reflection-correspondence theory, the present cybernetic view of knowledge does not assume a mapping between (static) *objects* and their *representations*, but between (dynamic) *perturbations* and *actions*. But what constitutes a perturbation? Apparently we only need to take into account processes that have a causal influence on the system's goal. Replacing the absolute causality of classical mechanics, where every event affects every subsequent event, by an irreversible, thermodynamic view of pro-

cesses, we may conclude that most causal signals will be dissipated or damped before they propagate to the system. In that case, what counts as "perturbations" will be a small subset of all physical events that have direct causal links to the variables defining the system's (subjective) goal.

3) even an infinite remaining variety of outcomes does not imply inadequate control, as long as the few "essential variables" distinguishing between survival and death are kept within bounds. For example, variation of the organism's horizontal position does not much affect its chances for survival. Variation of its body temperature, on the other hand, must be controlled within strict limits.

3. CATEGORIZATION AND DEFAULT REASONING

The above analysis argues that a small set of actions may still be sufficient to adapt to an infinitely complex environment. But it does not explain how an infinite variety of phenomena can be adequately *mapped* onto that small set. Every new phenomenon must somehow be put in the appropriate category, which can then be linked to an action adequate for that class of events. Perceived phenomena will activate simple sensory attributes such as "hot", "cold", "red", "heavy", "loud", etc. The function of the cognitive system is to map specific combinations of these attributes onto more abstract categories, which can then be interpreted in terms of required actions. E.g. the combination "hot", "high", "light" may elicit the concept "sun", which may trigger the action "go into the shade". In order to adequately steer the organism towards its goal of increasing fitness, a maximum number of combinations of attributes must be put into categories denoting possible dangers (e.g. fire, predators, cliffs, rivers, ...) or resources and opportunities (e.g. food, mates, water, shelters, ...). As implied by Ashby's law, the larger the variety of action-triggering categories available to the organism, the larger the control that it can achieve, and thus the better it will succeed in maintaining or improving its fitness, and the more likely that it will win the struggle for life. Evolution thus tends to increase the number of perceivable attributes and categories.

Even when the number of attributes is relatively small, the number of possible combinations will be virtually infinite. However, most of these combinations will not correspond to categories relevant to the system's survival. For example, it may be essential to distinguish a phenomenon with the attributes "moving", "small", "striped", "yellow", "black" as belonging to the category "wasp", linked to the action "avoid contact". On the other hand, combinations like "moving", "large", "green", "purple" will never be encountered, and thus need not be represented by a particular action-triggering category. Finally, a combination like "not moving", "medium", "brown" may be very common in the environment (e.g. a piece of wood or a boulder), yet be totally irrelevant for the organism's survival, and thus similarly escape categorization.

A basic mechanism for minimizing the complexity of categorization is *default reasoning*. Each time a new combination of attributes is encountered, the organism must find the appropriate category in which to fit the perceived phenomenon, in order to further infer appropriate actions. The same combination might fit several categories, or none at al. Rather than systematically test all categories (Is it a bird? Is it a plane? Is it...?), the organism will immediately pick up the "most likely" category, until it encounters evidence that another categorization is needed. It will then try out the "second most likely" category, and so on.

The classical example of default reasoning is the assumption that if something is a bird (earlier categorization or attribute), it can be expected to fly (inferred category). This is probably true in over 99% of the cases. Yet, the existence of ostriches and penguins shows that this is not a universal truth. Violations of the default expectation will be encoded in the cognitive system as exceptions: *if* it is a bird, *then* it can fly, *except* if it is penguin or an ostrich. But the awareness of the "exceptional" situation will trigger new default expectations: *if* it is an ostrich, *then* it can run; or, *if* it is a penguin, *then* it can swim. Again, there will be exceptions to these rules: if it is an ostrich, and it has a broken leg, it cannot run. But that expectation might in very unusual circumstances again be violated: perhaps an ostrich with a broken leg could still run if it was wearing some kind of artificial support...

The system behind this type of reasoning may be called a *defaulthierarchy* (Holland et al., 1986): it consists of different levels of expectations, the most likely one at the top, the less likely below. As ones goes deeper down into the exceptions and exceptions to exceptions, more attributes are added to the necessary conditions, and thus triggering conditions become more specific, and less probable. (After all, it would be quite unlikely to encounter an ostrich wearing a support around its broken leg...).

4. COMPLEXITY OF DECISION-MAKING

Let us define the *complexity of decision-making* or categorization as "the average number of alternatives (categories) that need to be explored before the appropriate one is found". This seems like a good measure for cognitive effort, amount of trial-and-error, or time spent searching. It is similar to the measure assumed by Simon (1962) in his famous paper on the "Architecture of Complexity" when arguing that *hierarchicaldecomposition* enormously decreases the complexity of problem-solving (see also Heylighen, 1991b). We will now produce a similar argument for *probability ordering*

Suppose you get information about an object with wings, and try to find out what type of object this is. Likely assumptions are that it is a bird or an insect. Less likely but still reasonable alternatives are a plane or a bat. Still more improbable is a pterodactyl, a flying lizard, or a flying fish. But you might of course also consider a harpy, a dragon, or the flying horse Pegasus... None of these possibilities can be absolutely excluded. Yet it obviously pays to first start looking whether the creature has feathers (implying that it is a bird), before you would check for hooves.

Let us consider the set of alternatives $\{a_n \mid n=1, \ldots, K\}$, each with probability $P(a_n)$. If the a_n are explored in the order of increasing n (first, alternative a_1 is tested, then a_2, then..., until a_K), the search will be successful after exactly one step with probability $P(a_1)$, after two steps with probability $P(a_2)$, and so on. The number of steps that can be expected to be necessary on average, i.e. the complexity of decision-making, is then determined by the following formula:

$$C = \sum_{n=1}^{K} P(a_n).n \qquad (1)$$

This complexity depends on two separate factors:

1) the *ordering* of the a_n: since the contribution of an alternative to complexity increases linearly with its rank (n) in the sequence, one can obviously minimize complexity by associating large n's with small $P(a_n)$'s, thus diminishing the effect of large ranks. The minimum of C for a given probability distribution will be reached if the terms are completely ordered according to decreasing probability: $P(a_i) \leq P(a_j)$, for all $1 \leq i < j \leq K$. Indeed, if for some $i<j$ we would have $P(a_i) > P(a_j)$, then C could be diminished with the positive amount $(i-j)(P(a_i) - P(a_j))$ by permutating the probability values of the i-th and j-th terms.

The strength of the decrease in complexity produced by probability ordering can be illustrated by the following example. Consider an infinite sequence where each next alternative (a_n) is a factor d ($d < 1$) less probable than the preceding one ($a_{n\text{-}1}$). Given the constraint that the sum of all probabilities must be 1, this leads us, after a few calculations, to the following complexity expression:

$$C = \sum_{n=1}^{\infty} n \frac{d^n}{\sum_{n=1}^{\infty} d^n} = \frac{1}{1-d} \tag{2}$$

For $d = 0.5$ (every next alternative half as probable as the previous one, a seemingly not very strong requirement), we get $C = 2$, i.e. the search through an infinite list of alternatives will be finished on average after just two steps! Yet, if the ordering would have been random, the search would have been infinite. For $d = 0.9$, we have $C = 10$. (If we would start from a normal (Gaussian) distribution of probabilities, C would be even smaller, as each subsequent probability factor will decrease not just with d^n but with d^{n^2}.)

2) the *distribution* of probabilities $P(a_n)$: if all probabilities would be equal, $P(a_j) = P(a_i) = 1/K$ for all $1 \leq i, j \leq K$, then there would not be any beneficial effect of ordering according to increasing P values, and C would reach a maximum value:

$$C = \frac{1}{K} \sum_{n=1}^{K} n = \frac{K+1}{2} \tag{3}$$

This value expresses the idea that if all alternatives are equiprobable (or, equivalently, if they are randomly ordered so that large probabilities are as likely to be encountered in the beginning as in the end of the sequence), we would on average need to explore half of all the alternatives in order to find the appropriate one. Since we are discussing situations where the number of possible categories (K) is virtually infinite, this means that C itself becomes infinite.

On the other hand, if, given perfect ordering, all probabilities were zero except one ($P(a_1) = 1$), the first alternative explored would always be the right one, and complexity would reach its minimum: $C_{min} = 1$. These maximum and minimum values for C as a function of the $P(a_n)$ distribution (given optimal ordering) correspond respectively to the maximum and minimum of the statistical *entropy* (H) for the distribution:

$$H = -\sum_{n=1}^{K} P(a_n).\log P(a_n) \tag{4}$$

Though it remains to be formally proven, it seems safe to conjecture that for perfect ordering C will increase monotonically with increases in H. Indeed, increases in entropy imply a more homogeneous probability distribution, i.e. smaller differences between the different $P(a_n)$, and thus larger values for the smaller $P(a_n)$ (corresponding to large n). This implies that the terms with larger n's in the complexity function will carry a larger weight (which is not compensated by the correspondingly smaller weight of the terms with small n's), increasing the sum which defines C.

In conclusion, *cognitive complexity* of choice between a given number of alternatives K will decrease with the goodness of *ordering* of the alternatives according to probability, and (given ordering) increase with the *entropy* (homogeneity) of the probability distribution. There are thus two ways to keep complexity small:

1) **Good ordering**: this is a factor which depends on the organism's capacity to learn, i.e. to store its experience as to the frequency with which a particular alternative is encountered. It is well-known that past frequency of occurrence, implying likeliness of future occurrence, is a fundamental determinant of learning. For example, associative learning in conditioning experiments or in neural network models assumes that a learned association (if "bird", then "flies") becomes stronger the more often its is activated. The "strength" of a connection determines the likeliness that the connection is later activated, and thus the (average) speed with which the alternative represented by that connection is explored.

2) **Low entropy**: this is a factor which partly depends on the organism, partly on its environment. In a high entropy environment, where all types of phenomena or situations are about equally likely, cognitive complexity would be maximal, and control through knowledge would be virtually impossible. In a "mixed entropy" environment, where some types of phenomena are equally likely, while others have strongly differentiated probabilities, cognitive complexity could still be kept within bounds by ignoring or filtering out all the high entropy categories. This is not a problem if the eliminated categories correspond to those variables which are irrelevant to the organism's fitness. For example, the Brownian motion of air molecules against one's skin is a largely entropic phenomenon, where it is practically impossible to predict the direction of the next motion given the present one. Yet the pressure exerted by this Brownian motion can be neglected as far as survival is concerned, and thus adequate cognitive modelling, enabling prediction and control, is not necessary.

5. FITNESS AS DEFAULT

Our analysis still implies that at least the variables relevant to survival should have a low entropy. This is not at all obvious, given the 2nd Law of Thermodynamics, which states that thermodynamic entropy tends to spontaneously increase. However, I have argued in an earlier paper (Heylighen, 1992) that increase of *thermodynamic* entropy can still be accompanied by decrease of *statistical* entropy (the necessary and sufficient condition for a Markov process to allow *decrease* of statistical entropy is that its transition matrix not be doubly stochastic (Koopman, 1978)). The present conclusion about cognitive complexity can be interpreted as a further, indirect argument against the common belief that the most "natural" state of the world is one of high entropy. If that were true, knowledge and control could never have developed.

The principle of asymmetric transitions (Heylighen, 1992) states that systems tend to "settle down" in their most stable (attractor) states, thus leaving the less stable states (attractor basin). (this automatically invalidates the requirement of double stochasticity necessary for entropy increase). This implies that the former states (say a_1 to a_l) will be encountered with a much higher probability than the latter ones (say a_{l+1} to a_K), a condition sufficient to allow complexity reduction by probability ordering of alternatives.

What was called "stability" when discussing asymmetric transitions, is perhaps more properly called "fitness". Fit configurations can be defined as configurations picked out by natural selection. This may happen because they are intrinsically stable, or because they are (re)produced in great quantities. Selection entails a fundamental asymmetry between fit and unfit systems: fit systems are naturally privileged, and are much more likely to be encountered than unfit ones.

In conclusion, the evolutionary principle of the "survival of the fittest" explains the existence of alternatives that are much more likely than others, playing the role of the "defaults" we introduced earlier. Thus, the most fundamental type of default assumption is that *most likely a phenomenon is fit.* This agrees with intuition. For example, lame ostriches or penguins unable to swim are clearly unfit, and will not survive very long as such: either the penguin will learn to swim, or it will die because of starvation, not being able to catch fish. More generally, given the constraint that a system is a bird (having wings and feathers), we may assume that it will be most fit when it can fly, though there are exceptional niches, occupied by ostriches and penguins, where flight is not necessary for fitness. The "defining features" of a particular category (e.g. birds) determine a limited domain of "fit" configurations. Most instances of the category will be kept within that domain by natural selection, and this explains the adequacy of default assumptions about members of that category.

Though the example of birds is typical for the predominantly biologicalinterpretation of natural selection, the proposition is much more general. A physical example: "stones are hard" is a typical default assumption. For stones, hardness is a part of fitness: soft stones will tend to pulverize or crumble under the effect of erosion, and, hence, will be quickly eliminated. Therefore, the "abstract" default assumption that a stone is fit, implies "concretely" that it is hard. Another example, "water is liquid", reminds us that fitness is relative to the environment: this assumption is valid only in the more common situations where temperatures are in between freezing and boiling. Under such temperatures, non-liquid forms of water (ice, steam) are unstable, and will be eliminated. Finally, a psychological example: "people understand language". In our present society not being able to understand language is clearly unfit: either the individual (e.g. a baby) will be taught language understanding, or he or she will get isolated (e.g. in a home for the mentally handicapped).

Not all configurations one encounters will be fit, however: variation will continuously produce fluctuations or deviations from the "normal", fit configuration. Most of these "mutations" will be quickly eliminated, except when they discover a new *niche* or fitness domain, like the ones occupied by penguins or ostriches. That is why default assumptions remain just that, they do not become "absolute truths", and for every rule there will always be an exception.

Acknowledgments: I thank C. Joslyn, V. Turchin and other contributors to the Principia Cybernetica mailing list for commenting on some of these ideas.

REFERENCES

Ashby W.R. (1958): "Requisite Variety and Implications for Control of Complex Systems", *Cybernetica* **1**, p. 83-99.

Buss D.M. (1991): "Evolutionary Personality Psychology", *Am. Rev. of Psychology* **42**, p. 459-491.

Campbell D.T. (1974): "Evolutionary Epistemology", in: *The Philosophy of Karl Popper*, Schilpp P.A. (ed.), (Open Court Publish., La Salle, Ill.), p. 413-463.

Conant R. C, and Ashby W.R.(1970): "Every Good Regulator of a System Must Be Model of that System", *Int. J. Systems Science* **1:2**, p. 89-97.

Heylighen F, Joslyn C. & Turchin V. (eds) (in print): *The Theory of Metasystem Transitions* (special issue of "World Futures").

Heylighen F. (1991a): "Cognitive Levels of Evolution", in: *The Cybernetics of Complex Systems*, F. Geyer (ed.), (Intersystems, Salinas, CA), p. 75-91.

Heylighen F. (1991b): "Coping with Complexity", *Systemica* **8**, part 1, p. 39-55.

Heylighen F. (1992): "Principles of Systems and Cybernetics: an evolutionary perspective", in: *Cybernetics and Systems '92*, R. Trappl (ed.), (World Science, Singapore), p. 3-10.

Heylighen F., Joslyn C. & Turchin V. (1991) : "A Short Introduction to the Principia Cybernetica Project", *Journal of Ideas* **2, #1** p. 26-29.

Holland J.H., Holyoak K.J., Nisbett R.E. & Thagard P.R. (1986): *Induction* (MIT Press, MA).

Koopman B. (1978): “Entropy Increase and Group Symmetry”, in: *Maximum Entropy Formalism*, M. Tribus & I. Levine (eds.), (MIT Press, Cambridge, Mass.).

Saviotti P.P. & Metcalfe J.S. (eds.) (1992): *Evolutionary Theories of Economic and Technological Change*, (Scientific and Technical Book Service, New York).

Simon H.A. (1962): “The Architecture of Complexity”, *Proc. Am. Philos. Soc.* **106**, p. 467-482.

Turchin, Valentin (1977): *The Phenomenon of Science* (Columbia University Press, New York).