FUNCTION, REDUCTION AND NORMATIVITY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

GÖKHAN AKBAY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN
THE DEPARTMENT OF PHILOSOPHY

MAY 2010

Approval of the Graduate School of Social Sciences

Prof. Dr. Sencer Ayata
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Arts.

Prof. Dr. Ahmet İnam
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Arts.

Assoc. Prof. Dr. Ayhan Sol
Supervisor

Examining Committee Members

Assoc. Prof. Dr. Ayhan Sol	(METU,PHIL)	
Asst. Prof. Dr. John Herbert Bolender	(METU, PHIL)	
Assoc. Prof. Dr. Mehmet Elgin	(Muğla Uni., PHIL)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Gökhan Akbay

Signature :

ABSTRACT

FUNCTION, REDUCTION AND NORMATIVITY

Akbay, Gökhan

M. A., Department of Philosophy

Supervisor: Assoc. Prof. Dr. Ayhan Sol

May 2010, 84 pages

Normativity of biological functions create a serious obstacle against the reduction of functional biology into molecular biology. Normativity of biological entities has two interconnected sources. One is the internal complexity and self organization demonstrated by the organism. The second source is external to the organism: Natural selection. An organism adapts to its environment by its internal autonomy. Species or populations adapt by natural selection. If these two sources of normativity can be reduced to statistical generalities achieved by theoretical models, reductionism will prevail.

Keywords: molecular biology, function, reductionism, normativity, autonomy

ÖZ

İŞLEV, İNDİRGEME VE NORMATİFLİK

Akbay, Gökhan

Yüksek Lisans, Felsefe Bölümü

Tez Yöneticisi: Doç. Dr. Ayhan Sol

Mayıs 2010, 84 sayfa

Biyolojik işlevlerin normatifliği, işlevsel biyolojinin moleküler biyolojiye indirgenmesi önünde büyük bir engeldir. biyolojik varlıkların normatifliği iki bağlantılı kaynaktan gelir. Bunlardan ilki canlının sergilediği içsel karmaşıklık ve kendini organize edebilmedir. İkinci kaynak ise canlıya dışsaldır: Doğal seçilim. Canlı çevresine içsel otonomisiyle uyarlanır. Türler veya popülasyonlar ise doğal seçilim ile uyarlanırlar. Eğer normatifliğin bu iki kaynağı, moleküler biyolojinin istatistiksel genellemelerine indirgenebilirse, indirgemecilik başarıya ulaşacaktır.

Anahtar Kelimeler: moleküler biyoloji, işlev, indirgemecilik, normatiflik, otonomi

To My Father,

ACKNOWLEDGEMENTS

I would like to thank my supervisor Ayhan Sol, who made suggestions about both technical and substantive problems in every draft I sent to him. I also would like to thank my jury members John Bolender and Mehmet Elgin for their helpful suggestions.

The people who provided both physical and psychological conditions for me to cope with the lengthy and stressful period of thesis writing also deserve mentioning. I thank my family and my aunt Zekine Kırteke for providing me with a shelter that protected me from the problems of daily life.

Lastly, I thank my love Ayşegül Özkan who helped me at every stage of the thesis writing period.

TABLE OF CONTENTS

CHAPTER 1

Introduction

Teleology, or mainly teleological explanations are considered as a back-step from modern science. Whenever we think about modern science, we have physics and chemistry in mind. Why do we consider teleological explanations as more primitive than those law-governed and purpose-free explanations of physics and chemistry? We should first state the difference and then proceed by stating that a teleological explanation refers to ends, purposes or in modern biology, functions. Those objects that we attribute ends, goals and functions are all physically realized. However, we observe on organization, sequencing and regularity in some physical phenomena that look like a means to an end. We observe these phenomena mainly in the realm of biology, psychology and engineering. An eye is for seeing, a diploma is for applying for a job, a fight is for the beloved, a cell-phone is for communication, and so on. What makes these objects and happenings so special? They are all physical in the sense that we do not postulate any non-physical forces to understand them. But they have a difference from other purposeless physical phenomena such as rain, earthquakes, stones and other stuff. We are not inclined to say that rain is for our agricultural activities, earthquakes are for punishing the infidels or stones are for sculpting. The orderliness, complexity and functionality distinguish teleological phenomena from other physical phenomena.

Functions debate in philosophy of biology originates from a pre-Darwinian period where the complex organization of living things seemed mysterious and required a better-than-chance explanation. The molars of a mammal are always in the back and canines are always in the front. There are deviations of course, but the point is that, whenever a deviation occurs, the unlucky creature is less fit than the normal ones. Another mystery was reproduction, which creates new progeny from an existing pair (or single) of creatures. The mystery of reproduction was in a sense derivative of the mystery of complex organization because the interesting question was to reproduce that organization without the (visible) aid of any intelligent being. Dolphins give rise to baby dolphins and cockroaches give rise to baby cockroaches. The continuity and fidelity of reproduction hinted at some deeper organization but nobody knew what was being transmitted before the dawn of molecular genetics. The ‘essence’ which passes from parents to the offspring was unknown. However, the visible features in the offspring resembled the parents so much that it was inevitable to conclude that the organization was somehow being transmitted and unfolded through the process of ontogeny.

Since the resemblance was powerful enough to be abstracted from the minute deviations, it was possible to attribute an essence to living things, which was being inherited through periods much longer than an individual observer’s life time. Some deviations were fatal, which means they prevented the offspring to reach adulthood. Most of the deviations, however, were slight enough to be omitted. Hence, the generations of organisms seemed to oscillate around a point of attraction in a harmonic way such that once the pendulum moves out of these boundaries, the whole structure collapses and fades into chaos. These boundaries can be named as the norms of nature.

Norms are subject to discussion mostly in ethics and philosophy of mind and language. They are tightly linked to intelligent systems. In ethics, norms or normativity get their meanings in opposition to what actually happens in nature. Norms say what ought to be done, regardless of what is the case. Thus, the phrase “norms of nature” may seem a contradiction to those who are accustomed to ethical thinking. Their intuitions might be on the right track but we should decide the question only after examining it thoroughly. The question is whether there really are norms of nature or they are just fictions of our anthropomorphic way of thinking.

Norms appear in two other philosophical areas. In the philosophy of mind, a representation is said to be about something. I will not delve into the details of intentionality debate. A short description is adequate. The “aboutness” of representations makes them normative in the sense that their relation to reality is governed by rules rather than physical laws. The nature of these rules and how they originate in the first place might have a naturalistic explanation, but that will be out of the concerns of this thesis.[1]

In language, the same relation holds. Normativity in these areas can be understood when we contrast it to deterministic causal relations. For instance, suppose that A causes B, or B supervenes on A. This means that whenever A is present (or it occurs), B occurs too. Let’s contrast this with the representing relation. Suppose A represents B. If the presence of B deterministically causes A in all possible worlds, then A will never be false of B. Suppose A is a type of representation (not a token) and B is a type of physical event or entity. If all of our representations were of this kind, then we would never have wrong ideas. Since we sometimes have wrong ideas, some of our representations must be in another relation to reality than deterministic causality. Biological functions are similar in this respect.

Biological functions are best understood in comparison to the functions of artifacts. Think of the eye. The lens of the eye does the same job as the lens of the camera. It focuses light rays coming from different angles to an area. If there is a problem in focusing, the resulting picture you experience may be blurry (as in the astigmatic). Unlike the truth relation, however, function

---

[1] Teleosemantics is such a naturalistic account of representations. According to the teleosemantic approach, the relation between representation and reality is functional (i.e. content of a representation is its function). Teleosemantic approaches inherit the selected effects account of functions and apply that to semantics. For a detailed discussion of the topic, see Macdonald, G. and Papineau, D. 2006

relation depends on how it benefits the organism. If astigmatic was a beneficial feature, we would call astigmatic lenses functional.

As I will discuss in the following chapters, function is more liberal than truth. It has to be more liberal because otherwise no new function would emerge from existing functions by accumulation of changes if the normativity of functions was an all or nothing phenomenon.

Normativity in nature might just be an illusion, like a mirage.[2] Illusions are the consequence of two factors: the physical arrangement of the outer world and our cognitive systems. For instance, in the case of mirages, the density difference of air near and above the surface and optical refraction is the "outer-world" part of the story. The other part consists of how our brains interpret the input coming from the outer world. An illusion is an illusion due to our beliefs generated by it. The mirage is an illusion because one who sees it will believe there is water nearby, expect to satisfy her thirst with it, etc. In the case of organisms, this distinction (outer-world basis and our expectations) can be made too. We may be falsely interpreting the input, but the outer-world part of the illusion stands and needs explanation. A complete project to eliminate this kind of illusion would take into account both factors, but it would be too ambitious for this thesis. Hence, I will focus selectively on the interesting common properties of organisms which provide the input for our misinterpretation that the biological world is a product of intelligent design, or that it shows genuine normativity.

Eliminating the illusion of normativity is important for biology. Biology is a special science. Special sciences are those about special objects of inquiry. Living things constitute a minute portion of the universe. Organisms exhibit at least *prima facie* differences compared to nonliving stuff of the universe. The

---

[2] Sober, E.: 1993, pp. 1-2

apparent normativity is one of those differences. Reduction of biology to physics or to an intermediate science (like molecular biology) is an ideal worth pursuing because it will enable us to unite life and non-life. The emergence and evolution of life will be explained by well-known physical processes (rather than occult forces like divine intervention).

Normativity is an obstacle against reductionism. Functional biology is said to be a historical science which lacks laws, and even *ceteris paribus* laws. It is historical because functions are the product of lengthy chains of evolution and these chains are contingent and local.[3] There are no laws of biology because function is a product of natural selection and natural selection is insensitive to structure.[4] Local conditions and contingency are so powerful that there can be no universal biology, according to anti-reductionists.

Even if molecular biology and its sub-branches (molecular developmental biology, molecular genetics, etc.) look like good candidates for being the reducing field, normativity of functions which is a consequence of historical contingencies, seem to be an obstacle against synchronic analysis. Most of the information will be lost in the evolutionary process and since synchronic analysis can only give the here and now *modus operandi* of functional structures, it will be difficult to understand the rationale behind existing structures. Hence, we will need another concept of function, in addition to the synchronic concept, to include the historical aspect of functional structures.

There is a valuable source of historical information residing in the genomes of organisms. Even if the exact details of history cannot be reproduced, some

---

[3] Rosenberg , A.: 2001

[4] Ibid.

critical changes can be inferred from the genomes of organisms. Most of the functional structures do not fossilize but genes do fossilize.[5] A fossil gene is a stretch of DNA which served in the production of proteins in the past but is now dysfunctional. Sequence comparison between different species already gave us a great deal of historical information (the period when two lineages diverged). Fossil genes and their functional cousins may give useful information on environmental changes and changes in lifestyles of a species. To combine a historical and a synchronic perspective in biology is vital for a complete understanding of biological function.

Molecular biology is a good starting point because in molecular biology, we can both observe functional structures in action and infer about environmental changes that produced these structures. However, if molecular biology is just as normative as functional biology, it will not be an appropriate reducing field. But, although molecular biology uses terms like "a dysfunctional protein", "an uncoding region of DNA", there is a sharper consciousness about the scale of structural differences that cause functional differences. In molecular biology, the boundary between function and malfunction is visible and also transgressable. An enzyme can function even if some amino acid residues are replaced by others but it might lose function when some critical residues are removed. The normativity is a gradual one such that although the catalytic activity of an enzyme is substrate specific, it is still not absolutely specific because of the weak molecular interactions. Even if there are norms in molecular biology, they are of the statistical sort. An enzyme has an "affinity" towards the substrate, an antibody can recognize (although weakly) other antigens than the one it was produced to interact with. The liberal normativity of molecules has given evolution the freedom of improvisation and it will give us the freedom to unite historical and non-historical sources of normativity.

---

[5] Carroll, S. B.: 2006, p. 36

CHAPTER 2

Reduction in Biology

Reduction has been analyzed into three different senses: ontological, epistemological and explanatory.[6] Ontological reduction is characterized by the thesis that higher level properties of an entity supervene on lower level properties of the components of that entity. Supervenience is a one-way dependence relation between higher level and lower level properties. A higher level property supervenes on a lower level property if and only if all the causal powers of a higher level property are constituted by the lower level properties of its components. In other words, if there is no difference in lower level properties of components, there will be no difference in higher level properties.

For instance, if two batteries have the same atomic configuration, their longevity will be the same. This is a one way relation because the same higher level property (e.g. battery longevity) can be achieved by different lower level configurations (e.g. atomic configurations).

This type of reductionism is straightforward and accepted by the majority of philosophers except dualists and vitalists. This idea is also called token-token reductionism. Token-token reductionism is the theses that tokens of a higher level kind are identical to their low level counterparts. For instance, even if the heart is a functional kind and is not identical with any configuration of atoms, an individual heart may well be identical with a particular configuration of atoms. Token-token identity thesis (i.e. token token reductionism) can best be understood in comparison to type-type identity thesis. According to the latter,

---

[6] Brigandt, I. and Love, A.: 2009

we can find common lower level properties for a high level property and hence find out an identity relation between a type in the higher level and a type in the lower level. In contrast, token-token identity thesis is a rejection of type-type identity thesis on the grounds that for some higher level properties (or predicates) we cannot find a definitive collection of lower level properties, even if we can identify any particular token of a type with a collection of lower level predicates.

The second type of reductionism is epistemological in the sense that it is about the relations between theories rather than kinds of things. Epistemological reductionism or theory-theory reductionism was first clarified by Ernst Nagel in his *Structure of Science.*

A theory is said to be reduced to another theory if the laws of the reduced theory can be logically deduced from the laws of the reducing theory. If the terms of the reducing theory and reduced theory are prima facie incompatible, reduction is achieved by formulating translation rules. These translation rules are biconditionals that connect the terms of two theories. They are also called bridge laws.

There are some problems with this account of reduction. One reason to doubt about this account is the fact that logical deduction is not possible if the reduced theory contains false propositions. This is usually the case in science. One theory replaces another not because it is only more general, but also the former theory makes false predictions. One can save Nagel's account by adding the condition that reduced theory should be corrected before reduction but this does not work because the false and corrected theories will not be the same. In that case, it would be elimination rather than reduction. [7]

---

[7] *Ibid.*

Another and more relevant criticism is that, some special sciences like biology do not contain laws to be deduced. There are generalizations and mathematical models in these sciences but they do not count as laws because they apply to only a very little portion of the universe (organisms, intelligent beings, etc.) Hence, we seem to need a different type of reduction to understand the relations between special sciences and lower level sciences. The third type of reduction is an answer to these problems. In biology, there are no laws like $e=mc^2$. However, there are empirical generalizations at higher or lower levels. In biology, reduction and reductionism are understood as decomposing a system into its components and explaining a higher level phenomenon (e.g. phenylketonuria) by means of lower level causes (e.g. genetic deficiencies). Both proponents and opponents of reductionism in biology accept this. Thus, this is the common ground we should step on in order to understand the problem.

Explanatory reduction is what happens when a higher level generalization, which is supposed to explain a higher level phenomenon, is grounded or corrected by lower level causes. In one sense, the higher level refers to wholes and lower level refers to parts and their interactions. The fur color of a cat is an inherited property. Hence, the higher level generalization is that “Black cats beget black kitten”. This generalization is explained and corrected by molecular genetics such that the genes that control the production of melanin are identified. Another example would be developmental deficits (in bone, cartilage and pigmentation) in zebrafish. By the (molecular) differences that make (phenotypic) differences, biologists construct causal pathways leading from molecules to phenotypes.[8]

---

[8] From now on I will use *synchronic reduction* as a subtype of *explanatory reduction* along with *diachronic reduction.*

Patricia Churchland explains reduction as an operation that takes place between theories. A theory is said to be reduced to another if “the causal powers described at the macrolevel are explained as the outcome of events and processes at the lower level”.[9] In this regard, classical thermodynamics is said to be reduced to statistical mechanics. Properties of heat in general are macrolevel properties and statistical analysis of molecular motions can explain them. The same can be said about classical optics and electromagnetic theory. Visible light’s macroproperties are explained in a more general scheme of properties of electromagnetic radiation. In life sciences, reduction of Mendelian genetics to molecular genetics is another analogue of this process. If we carefully look at history of science, we will find other examples of reductions.

Reduction also has another sense as used in science. This time we are not concerned with theories themselves but functional components of organisms, from systems to organs and to molecules. I have previously called it explanatory reduction (or synchronic reduction). This sense of the term is also recognized by Churchland, who claims that “if you want to understand how a thing works, you need to understand not only its behavioral profile, but also its basic components and how they are organized to constitute a system. If you do not have the engineering designs available for reference, you resort to reverse engineering – the tactic of taking apart a device to see how it works”.[10]

In molecular biology, we have the opportunity to observe this kind of reduction at work. In molecular biology, intentionally suppressing some natural events such as expression of a certain gene and looking for its correlation to other

---

[9] Churchland, P.S. 2005

[10] Churchland, P. S. 1997, p127

changes in metabolism is a very valuable method of analysis. Sometimes it is called "loss of function" analysis (or experimental ablation in neuropsychology). Functions of many *Drosophila* genes, for example, have been identified by intentionally causing mutations and trying to find out what macrolevel changes occur. This point is verified by the fact that gene names for *Drosophila melanogaster* refer to lack of function rather than the function itself (e.g. eyeless, flattened eye, deformed eye, distorted eye, etc). With this methodology, higher level functions of organisms can be reduced to lower level functions of genes, proteins, gene networks and signaling networks. Since the organism is given to the scientist as a whole, she tries to understand the basis of higher level functions – such as the regulation of the expression of a certain protein – and main concern is to understand here-and-now *modus operandi* of the system, synchronic reduction becomes a necessity.

Explanatory reduction resembles mechanistic explanation. Considering organisms as machines can be traced back to Descartes. Mechanistic explanations decompose, localize and reconstruct the causal determinants of observed properties. For decomposition to succeed, wholes should really consist of discrete parts. Living things are very complex. Thus, mechanical explanations are not straightforward and simple.[11] Interactions among parts are the main reason for complexity. The problem is to divide the organism into such discrete parts that your division would reflect real distinctions. In human made machines, functional parts of entities are usually isolated, which means that deficiencies in one would not affect others.[12] There are exceptions, of course, such as the central processor, power source, or the engine which affect the working of the whole system. These kinds of central parts do not refute the main claim: there is a suitable analysis of a human-made system because it was

---

[11] Bechtell, W. and Richardson, R. C.: 1993, p. 21

[12] Dennett, D. C.: 1996, p. 215

planned and constructed from distinct parts. The problem in biology is that we lack the suitable analysis of a system. Reverse engineering is always an option but in some cases it leads to seeing everything as functional. Not everything in an organism is functional. Some parts (i.e. rudiments) have lost their functions. Some parts are free riders: they are there because of their connection to a functional part. Hence, it is vital in biology to distinguish genuine functionality from pseudofunctionality, or non-functionality.

Another type of reduction is diachronic reduction, or forward engineering. This is the adaptationist viewpoint in evolution. According to this view organisms are complex structures that have adapted to their environments by means of gradual design accumulation that can also be seen as emergence of functions step by step, by means of natural selection. Dennett summarizes this kind of reduction as follows:

> Here is the most general form of the schema of Darwinian explanation. The task of getting from the early time when there wasn't any *x* to the later time when there is lots of *x* is completed by a finite series of steps in which it becomes less and less clear that "there still isn't any *x* here, not really," through a series of "debatable" steps until we eventually find ourselves on steps where it is really quite obvious that "of course there is *x*, lots of *x*." We never draw any lines.[13]

Here stands the source of biological normativity. As mentioned in Introduction, normativity and norms belong to social and ethical affairs. In social affairs, norms regulate interactions among the members of a society. They might be explicit and coded as rules or they may be implicit and do not need linguistic expression to regulate. Ethical affairs constitute a subset of social affairs and norms regulate human conduct by acting as ideals we strive to reach. In biology, norms serve the purpose of distinguishing functional, disfunctional

---

[13] *Ibid*, p. 200

and malfunctional aspects of organisms. Function, in the everyday usage of the term, implies purposes. By attributing a function to something, we express our expectations on what it is supposed to do. The expectation is grounded on our knowledge that artifacts are produced to achieve some ends. In biology, function attribution reflects our expectations about how a biological structure should contribute to the fitness of its bearer. However, this expectation cannot be grounded on the intentions of an intelligent being, assuming the naturalistic stance of contemporary biology. What might ground that expectation will be the central theme of this thesis.

An organism, or a single cell, acts on its own behalf. The flexibility of an organism to adapt to different conditions resemble following a goal and overcoming obstacles. Another aspect of organismal purpose is historical perfection of an adaptation to fit to a specific environment.

A fish that lives in cold and dark depths of the ocean adapts to this extreme environment by synthesizing an anti-freeze protein. In a sense, the purposeful looks of organisms come from two sources: the synchronic self-regulation to achieve a consistent and stable internal state, and following a diachronic trajectory to fit to a particular environment’s relatively stable regularities. Thus one source is internal and the other is external.

CHAPTER 3

The Functions Debate

3.1 Overview

Functions debate was at first a debate about teleology. The question was whether biological entities were designed by an omnipotent God or they were the outcomes of a natural process. Now, an omnipotent God has lost its significance as an alternative to naturalistic theories of evolution. However, the apparent harmony and complexity of biological world still charms many scientists, including the arch evolutionists like Richard Dawkins.[14] Functions debate turned out to be a debate internal to naturalism.

The question of function has two dimensions. I will call the first one "the explanatory question" and the second one "the substantive question". The first is related to the nature of functional explanations. What do we explain when we attribute functions to entities or processes? Given the physical description of a system, do functional explanations contribute anything to our knowledge? These are some of the explanatory questions I will trace in the history of the debate.

The second type of questions involves the nature of functional entities that include natural entities like enzymes, organs, systems, etc. They also include artifacts that humans produce and use. The question is whether we can find an *interesting* common property that will unite functional entities in both domains while excluding all nonfunctional entities. I underline the word "interesting" here because one may easily find some property that captures the extension of the term "functional" but it may not tell us anything about the nature of

---

[14] Dawkins, R.: 1998

functional entities. Are functional entities really special? This question is tightly related to the second explanatory question I mentioned.

At first, I will discuss the pioneers of functions debate. These include Larry Wright (1973) and Robert Cummins (1975). Their definitions determined the boundaries of the modern debate. Larry Wright is well known for his “selected effects”[15] account of functions and Robert Cummins is the founder of “causal role” account.[16] Before these pioneers, debate around functions was focused on goals and teleology. Their preferences mark a shift in the traditional ways of approaching functions. The second part will consist of the followers of pioneers and modern discussion of the topic. The discussion both includes proponents of the pioneers and combinatorial accounts to unite those original accounts.

3.2 Pioneers

What do we mean when we say that the function of mirrors is to reflect light? And what does a biologist mean when she says “the function of B-lymphocytes

---

[15] Proponents of the selected effects account claim that function attribution is made to explain the presence of an entity or a process by means of some of its effects. In other words, funtional explanations refer to some specific effect of an entity in order to answer the question: “Why is this entity here?” For instance, according to them, to say that heart’s function is to pump blood is to say that hearts exist because they pump blood. In other words, funtional explanations refer to some specific effect of an entity in order to answer the question: “Why is this entity here?”

[16] Beth Preston (1998) prefers “systemic functions” rather than causal role functions. Her preference depends on the observation that Cummins himself does not equate causal analysis with functional analysis in his later works. I prefer “causal role” because Cummins’ original paper deals with organized causal pathways and the method for analyzing them.

is to produce antibodies"? We can say mirrors are very good at reflecting light and we use them to reflect light. We can also appeal to the design intentions of mirror designers. In the second case, we can neither find designers nor users of B-lymphocytes. Natural functions pose more serious problems than conscious functions. In order to justify our usage of word 'function' we should find a covering definition for both types (conscious and natural).

I deliberately selected paradigmatic use of the term 'function'. In paradigmatic use, function is less context-dependent. The function of something is such and such independent of how it is used. Things may get more complicated when we take everyday use of the term function. We can easily imagine someone saying "the old house down there now functions as a playground for children" or "the sun functioned as a natural timekeeper those days". The second type of use, in opposition to paradigmatic use, is the peripheral use of the term.[17] The difference from the paradigmatic case is apparent even from the phrases used: the function of something vs. something functions as something[18].

If we are to use 'function' in a proper way, we should either show that paradigmatic and peripheral uses mean the same thing or they mean different things (certainly, an overlap is also possible). If we accept their difference, we should show what makes them different. Wright's aim is twofold: unification of the meaning of proper uses of the term, and making a very general distinction in order to exclude counter intuitive uses of the term. Wright offers some crucial distinctions to find the core notion of function, and then criticizes former attempts to give a unified definition. Lastly he provides his own alternative formulation and argues for its superiority against former

---

[17] Wright, L.: 1973, p.139

[18] *Ibid.* p.140

formulations.

The first distinction Wright makes is between goal and function. Wright claims that a function cannot be equated with a goal because goal implies behavior[19]. The chitin wall of an insect has the function of protecting its internal organs and a knife has the function of cutting even if both show no behavior.[20] We can ascribe goals to autonomous agents but we can ascribe functions to artifacts- which may or may not be autonomous. Equation even fails in the case of biological functions. We can say that the function of hemoglobin is to carry oxygen from lungs to tissues but we cannot say hemoglobin molecules have goals.

The second distinction is between "a" function and "the" function of something[21]. In some contexts, we may prefer to say "a" function of B-lymphocytes is to produce antibodies, in order to make room for further function ascriptions. B-Lymphocytes, in addition to producing antibodies, also produce cytokines.

The third distinction is between function and accident.[22] An entity may have many effects: Your heart pumps blood, makes noise, fills space, consumes oxygen and sugar, emits infrared waves, etc. Not all of those effects are its functions. Some effects have a privileged position. Pumping blood is such a

---

[19] *Ibid.*

[20] Both the chitin layer and knife may be considered to have certain dispositions under certain conditions. One might call their resistance to external effects a type of behavior but Wright here has certainly a more active type of behavior in mind.

[21] *Ibid.* p.141

[22] *Ibid.*

privileged effect. The reason of some effects' being privileged will be made clear when we discuss Wright's own definiton of function. Some other effects are merely accidental. In fact, the same reason that makes some effects privileged also makes others accidental. Accidental here doesn't mean unpredictable or chaotic. Accidental in this context means *unclassified*.[23] The explanatory theory at hand determines what counts as accidental and what counts as essential. In Wright's theory, function attribution explains the presence of an entity. Effects produced by that entity are classified regarding their capacity to explain the presence of that entity. All effects lacking such a capacity are termed accidental. If we turn back to the hemoglobin example, it is a matter of fact that hemoglobin can also bind to CO (carbon monoxide), which is a lethal toxin, but it is not the function of hemoglobin, it is an accidental side-effect of hemoglobin's structure because binding CO doesn't explain why hemoglobin is present. The same can be said for artifacts. The computer has many functional parts but it would be dubious to say 'the function of computer fans is to make noise'. I think this distinction is the central theme of Wright's article. As we will see in the following paragraphs, former attempts failed to give a satisfactory account of this intuitive distinction.

The last distinction Wright examines is the distinction between conscious and natural functions[24]. We already said a good deal about this point. To

---

[23] The case is similar in history. One might explain the founding of the Anglican Church by Henry the VIII's desire to have a child and Church's strict prohibition of divorce (his Spanish wife was infertile). However, if one's theory explains such important events by means of impersonal forces (e.g. Marxism), Henry's desires will count as accidental. If one's theory is theory of evolution by natural selection – as it is in Wright's case – the effects of a trait that drive natural selection to fixate that trait will be its functions and others will be its accidental side effects.

[24] *Ibid.* p.142

summarize our previous remarks I should underline the point that conscious functions (artifacts) are usually – but not always – design-dependent. We ascribe those functions partly regarding the design intentions behind their production and partly the way we use them. Wright does not accept ‘design intention’ as a necessary condition for function ascription[25]. This preference seems natural if we keep his aims in mind (unification and exclusion).

Wright is not the first philosopher trying to give an account of function ascriptions. As he emphasizes, his analysis depends on the strengths and weaknesses of his predecessors. Morton Beckner and John Canfield are his main targets in this paper. Beckner offers two accounts, first in 1959 and a more refined version in 1969. First account is very simple: Something has a function F in a system, iff that thing’s presence in the system is necessary and sufficient for the occurrence of F[26]. For instance, presence of a heart is necessary and sufficient for the occurrence of pumping blood.

Wright criticizes this formula from two angles. The first is that it underestimates the multiple realizibility issue. Hearts are not necessary for pumping blood; some other physical structure may have achieved it. The second point is that Beckner’s analysis does not capture function-accident distinction. By the same rule, the presence of computer can be a necessary and sufficient condition for the fan noise I hear, but it is not a function at all. Beckner’s second and refined account can be summarized in the following way: Saying that something in a system has the function F is to say that the functional thing is the part of a systematic whole where it contributes to a system-wide action and parts of the system are individuated functionally. The

---

[25] *Ibid.*

[26] *Ibid.* p.144

problem with this account is that, it cannot be applied to artifacts. Much of our artifacts are not parts of systematic wholes[27]. Of course one can arbitrarily define a system where knives contribute to the system-wide action of cutting but this seems counter-intuitive. Another point Wright raises is that this formulation cannot grasp function-accident distinction either.[28] Imagine a virus inserting into your genome and producing microbial resistance proteins such as restriction enzymes. It has become part of your genome (a systematic whole) and it contributes to the system-wide function of microbial resistance. According to Beckner's formula, we should say that the function of some parts of virus' genome is to produce antimicrobial proteins for the sake of its host. In fact, from the virus' perspective, their function is exactly to fight with microbes that will kill the host cell. However, from the perspective of the host, of which the the virus is now a part, that is just an accidental benefit.

John Canfield is the second target of Wright. His account is very similar to Beckner's first account but he adds another criterion: usefulness. According to Canfield, functions are useful to the systems where they are performed. Wright's criticisms are similar. Usefulness cannot discriminate among functions and accidents; it is also not applicable to artifacts because finding the proper containing system for which something is useful is very difficult. A property may be accidentally useful but not functional: Consider a hole in a tree that serves as a nesting ground for budgerigars. It is not the function of the hole to protect the mother and the chick but it is certainly useful to its hosts. It is also difficult to find a proper containing system for much of artifacts. For instance, what would be the proper containing system of a coffee cup such that it will benefit that system? One can certainly define an arbitrary containing

---

[27] *Ibid.* p.152

[28] *Ibid.*

system in which an artifact is useful but that will not differentiate among accidental side effects and genuine functions. Wright also claims that an artifact or part of an artifact doesn't have to be useful in order to be functional. For instance, part of a car (e.g. brake pedal) does still have a function even if nobody drives the car. According to Wright, intentions of designers give artifacts their functions, even if the intended effect is useless, nonsense or even harmful.[29]

After a lengthy discussion of previous analyses of function statements, Wright proposes his own solution. For Wright, most critical point of making function ascriptions is to explain the reasons for the presence of functional entities. [30] For example, when we claim that the function of knives is to cut, we mean that knives are present because they cut. In the case of artifacts, the point is usually clear. An artifact may have many effects, but only one effect can explain why it is there. We can use CDs for ornamenting our cars but they are produced to store information. Wright asserts that the same point applies to biological functions as well. Hemoglobin is present because it carries oxygen. It may have other effects such as binding to CO, but natural selection has maintained and spread hemoglobin for its oxygen binding capacity.

Explaining something by its causal history is called etiological explanation. Etiology is not confined to function-talk; we can properly call any explanation that depends on the causal history an etiological explanation. In the case of function, we should bring in one more constraint: the performance of the function must be a consequence of the functional thing. By that qualification, Wright claims to have made a distinction among etiologies: purely causal

---

[29] *Ibid.* p.146

[30] *Ibid.* p.154

etiologies and functional etiologies. According to Wright, a purely causal etiology will only cite proximate causes. For instance, the proximate causes of having a heart may be the genetic program, its implementation (i.e. development), etc. However, functional etiologies consist of ultimate causes such that a heart's existence is explained by means of the fitness advantages having a heart provided against heartless individuals in the same population.[31] Hence, Wright's formula is the following:

> The function of *X* is *Z* means
>
> (a) *X* is there because it does *Z*
>
> (b) *Z* is a consequence (or result) of *X*'s being there.[32]

---

[31] Wright never mentions proximate and ultimate causes. *Proximate* and *ultimate* explanations are technical terms introduced by Ernst Mayr in his "Cause and effect in biology" (1961) According to Sober (2000), proximate explanations refer to non-historical causes but ultimate explanations refer to the evolutionary rationale behind the presence of a trait. Sober exemplifies this by comparing two meanings of the question: "Why do ivy plants grow toward the sunlight?" One might be wondering either the internal mechanisms that generate the end result, or the evolutionary-ecological reasons for the prensence of such a capacity. Even though Wright never uses those terms, his classification of etiological explanations reflects such a distinction. His original example concerns the presence of oxygen in human blood. He claims there are two possible etiological explanations for the presence of oxygen in blood. The causal (i.e. proximate) explanation cites the facts that there is abundant hemoglobin in blood and hemoglobin has a high affinity for oxygen. The functional (i.e. ultimate) explanation, in contrast, cites the role of oxygen in energy producing reactions in cells. The example is rather awkward, and this I think is the primary reason why it has been overlooked, because oxygen is a very simple chemical whose existence (in blood or elsewhere) must be explained without recourse to functions.

[32] *Ibid.* p.161

According to Wright, this formula satisfies both of his aims: unification and exclusion. It unifies conscious and natural functions. For him, both artifactual and biological functions owe their presence to their specific effects (functions). Counter intuitive cases are excluded successfully. Neither accidental effects nor confusions regarding containing systems can enter into our definition. With this definition, we avoid ascribing functions to fan noise and invading viruses.

The second substantial contribution to functions debate comes from Robert Cummins' 1976 article which was an attempt to explicate the real aim of functional explanations in science. According to Cummins, functional explanations have nothing to do with the presence of some entity. In contrast, functional explanations seek to explain what contribution an entity or a process makes to the capacities of a containing system. Function attributions thus explain how a containing system produces complex outcomes by means of simple processes and entities.

Cummins diagnoses one fundamental mistake in etiological theories such that they confuse an inference to the best explanation with an explanation itself[33]. Such an inference is legitimate if we are dealing with products of conscious design. For instance, there is a sense in which the presence of a part of an artifact is best explained by referring to its function because a conscious designer must have put the part there for that function. The legitimacy of these explanations derive from the rationality of designers (e.g. if this screw is here, then the designer should have put it there for some reason). We are reverse-engineering in those cases.

However, the function of biological entities and processes do not accord this scheme. Biological entities, according to the theory of evolution, are not

---

[33] Cummins, R.: 1975, p. 748

products of conscious design. They are rather products of a lengthy chain of transformations in which different non-conscious forces (selection, drift, mutation) are acting. Hence, Cummins concludes that function ascriptions in the etiological sense depend on a superficial similarity. The difference lies in the genuine explanatory role functional claims play in science.

Cummins claims that the apparent scientific plausibility of etiological approaches originates from a misunderstanding of evolutionary theory. According to Cummins, etiological approaches falsely consider natural selection as an explanation of the presence of functional traits in organisms. For Cummins, natural selection does not determine which traits an individual has. It is determined by the genetic plan. Natural selection just explains why some traits are more frequent than others by referring to their adaptive advantages. In his own words "natural selection cannot alter a plan, but it can trim the set."[34]

Another option Cummins rejects is the view that functions are effects which are beneficial to a species. For example, one can say that even if natural selection can not explain the presence of a token of a functional entity, it can explain why wings exist by referring to the benefits of flight in the survival of some bird species. However, Cummins argues that even if flight ceases to be beneficial to birds, the relation between flight and wings still needs an explanation. A working pair of wings, even if flight is detrimental to the species, should be analyzed into its simpler capacities to explain their role in flight. [35]

Even a detrimental effect needs to be functionally analyzed, so, the point of

---

[34] *Ibid.* p.751

[35] *Ibid.* p.756

functional analysis must be different from finding out privileged effects of an entity. The point in functional analysis can be understood by contrasting it with a more familiar way of explanation. I think Cummins' (1976) central concern was to contrast functional explanation with deductive nomological form of explanation.[36]

A deductive nomological explanation seeks to deduce the explanandum from the explanans. Explanation is achieved by taking natural law(s) and precipitating events as premises (explanans) and deducing the explanandum as the conclusion.[37] In functional analysis, the organization of the system is of the essence. DN explanations can subsume all events in a system under laws of physics or higher laws.[38] However, the genuine contribution of functional analysis is to explain how certain higher order capacities are realized by means of lower order dispositions. This requires lower order dispositions to be programmed in a certain way. Dispositions and program (i.e. the sequence, distribution and causal connections among the performances of dispositions) explain the end product.

---

[36] Cummins (1975 p.757) expresses his aim in these sentences: "Such approaches have not given much attention to the characterization of the special explanatory strategy science employs in using functional language, for the problem as it was conceived in such approaches was to show that functional explanation is not really different in essentials from other kinds of scientific explanation. Once the problem is conceived in this way, one is almost certain to miss the distinctive features of functional explanation, and hence to miss the point of functional description."

[37] Churchland, P.M.: 1988, pp. 56-58

[38] Cummins himself mentions law-like regularities rather than higher level laws. But I think we mean the same thing here. (*Ibid*. p.759)

Flowchart diagrams, abstract descriptions of electronic circuits or assembly lines are best examples of Cummins' approach. Consider the computer fan example I mentioned before. The function of the fan is to cool the processor. Cooling the processor is a capacity of the computer which is achieved by the inner workings of the fan. The fan can be decomposed into its simpler parts. These simpler parts contribute to the cooling capacity of the fan, hence, one can explain how the cooling capacity of the computer is realized by means of analyzing relevant parts into simpler and simpler capacities they have.

3.3 Followers

As we saw in Wright's paper, it was an aim of conceptual analysis to distinguish functions from accidental effects. Wright had proposed a distinction between "the function of" and "functions as" to capture the different usages of the term. The function of something was the reason of its presence; it was also a consequence of that thing's presence.

This line of inquiry was elaborated by Ruth Millikan in her theory of proper functions.[39] She constructed a theoretical definition that would unite purposeful phenomena under one definition. Before examining the vices and virtues of her project we should first see the reasons for her project.

As Millikan points out, her former ideas were parallel with Wright's approach.[40] However, she saw some difficulties in his approach. Firstly, Wright's definition referred to present performance of the trait in order to explain the presence of the trait (or he does not clarify the point). Past

---

[39] Millikan, R. G. 1984, Excerpt from *Language, Thought and Other Biological Categories* in Buller, D.(Ed.): 1999

[40] Millikan, R. G.: 1995

performance of a trait cannot explain the presence of the trait itself, for a very simple reason.[41] Millikan concluded that the performance of a past trait should be the cause of present trait. Thus, the function of the present trait is what its antecedents did in the past that caused the present trait's existence.

Millikan adds a further constraint on functional entities: reproduction. By reproduction she means something like copying. Copying occurs when some properties of an original entity determine the properties of its descendants in a regular way. For instance, the copying machine maps written marks on the original paper onto the marks on a sheet of paper and produce a similar paper. In copying, descendants systematically covary with the original entity (the ancestor or the model).

Another constraint Millikan adds is selection. For Millikan, we cannot determine the function of a trait just by looking at its present benefits. The point is that we cannot distinguish fortuitous benefit from genuine function by just looking at its present performance. How beneficial should a trait be in order to be functional? Millikan identifies two extremes regarding biological cases.[42] The first extreme is the view that the trait in question should be a necessary component of the living system. Homeostatic functions (i.e. thermoregulation in haemothermal animals and such necessary dispositions of an organism) may be a candidate for this extreme view. However, even homeostatic functions were not present in some ancestors of today's organisms. So they cannot be "necessary" in the strict sense. There are also some functions of traits that are not performed in the lifetime of an individual. Mating displays of a lonely budgie would not occur unless a friend from opposite sex is placed into his cage.

---

[41] The trait should exist beforehand to contribute its preservaion or maintenance.

[42] *Ibid.* p.35

Another extreme is the view that a past token might have a function just because it has benefited only once in the reproduction of the trait. As Millikan states, it is natural selection that draws the line between these two extremes. Hence, selection is a necessary part of proper function definitions according to Millikan:

> Whether something has a function is not a matter of how often it has accidentally helped out in the movement from generation to generation. Anything whatever might occasionally have done that. It is a matter of whether it was selected to help out in this way.[43]

A crucial point of Millikan's account is the reference to historical conditions. As we observed before, it is the complexity of the system and relevant simplicity of dispositions that make a causal role analysis of functions plausible. It is the explanatory aims of the researchers that determine which capacities would deserve analysis. This may be true for machines in general. They are specified by certain rules of performance. There are rules that specify the proper input, internal states and expected output. This is why abstract characterizations like flowchart diagrams are well suited to describing machines. However, organisms have no such explicit rules of performance specifying what counts as input, output, damage, etc. According to Millikan, a Cummins type analysis is suitable in *designed* systems where there are well defined boundaries between input, output, external and internal states, etc. [44] Those boundaries in organisms, if they exist at all, are results of historical-ecological conditions under which the organisms have evolved: "But in any realm in which history plays a role, such as biology, members drift in and out

---

[43] *Ibid*. p. 38

[44] Millikan, R. G.: 1999

of lawful categories over time, leaving their boundaries ragged."[45]

A problem with Cummins' account is that it is too generous in ascribing functions. Millikan gives water cycle as an example:

> Cummins' definition it is, arguably, the function of clouds to make rain with which to fill the streams and rivers, this in the context of the water-cycle system, the end result to be explained being, say, how moisture is maintained in the soil so that vegetation can grow. Now it is quite true that, in the context of the water cycle, clouds function to produce rain, function as rain producers; that is their function in that cycle. But in another sense of "function", the clouds have no function at all-because they have no purpose.[46]

One can identify various functions according to the analytical account one has in mind. However, Cummins restricts functional analysis to complex entities whose systemic capacities can be analyzed into simpler or qualitatively different parts. This leads Millikan to assume that causal role functions coincide with proper functions in most of the cases because where there are such complex capacities there should be active selective forces (either natural or artificial).

Another problem with Cummins' account is that it cannot capture the normative aspect of function ascriptions.[47] For Millikan, since Cummins does not ascribe functions to parts that do not actually contribute to a systemic capacity, his analysis cannot account for malfunctions. According to Millikan, an entity has a proper function not because of its contemporary dispositions.

---

[45] Pinker, S.: 1997, p. 308

[46] Millikan, R. G.: 1989, pp. 294-295

[47] *Ibid.*

Having a function depends on history. The historical link between normativity and function is provided by natural selection.

There are two different senses of the term ‘selection’.[48] The first is “selection for”. If a trait increases in frequency because of an effect it has, we say that there is selection for that item. For instance, whiteness of polar bears increases the efficiency of camouflage. Since this trait is heritable and carriers of that trait reproduce better than other individuals, the frequency of whiteness increased in polar bear populations. In this instance, we say that there is selection for whiteness. This sense of selection explains the normativity of function claims. Even if a polar bear in a zoo cannot benefit from it, whiteness still has a function because of its evolutionary history. Normativity issue is clearer in the case of malfunctions. A mutant hemoglobin molecule which binds to CO_2 rather than O_2 still has the function of carrying oxygen. The other sense of selection is “selection of”. In this case, the trait in question increases in frequency by random drift or other chance events. It may or may not benefit its bearers. There is no selection for it either because there were no variants to compete with or the trait was no fitter than the alternatives.

Explaining function with reference to natural selection has been seriously criticized by many philosophers. One line of criticism is that selection is too strong a criterion for identifying functions. Recent followers of causal role functions claim that some item can have function in the absence of selection. One example given by Lewens involves a population of moths with orange wings, *without any variation*.[49] This population lives on green leaves where predation rates are high. One day the smoke coming from a factory turns leaves

---

[48] Sober, E.: 1993, p.100

[49] Lewens, T.: 2004, p. 93

into orange. Now, since predation is decreased due to camouflage, reproductive rates increase. However, there was no preceding variation in the butterfly population, and since variation is a necessary condition of selection, no selection has occurred. In this case, no selection occurred so we can not talk of functions according to Millikan's account. The wing color of moths does not have the proper function of camouflage. But intuitively there is such a function.

Another attempt to exclude selection from the definition of function comes from David Buller. Buller proposes a weaker etiological account in which to have a function, it is sufficient for a trait to increase fitness even if there is no actual selection for the trait.[50] Reconsider the orange winged moth example. In that case, since there was no variation, there was no selection.

However, having orange wings increased the fitness of organisms. But as Millikan stresses, in these cases the counterfactual reasoning undermines the relevant details of actual processes of evolution. In actual processes, we determine a trait's fitness enhancing capacity by comparing it with other competing traits. According to a textbook definition of adaptation:

> An *adaptation* is a characteristic that enhances the survival or reproduction of organisms that bear it, relative to alternative character states (especially the ancestral condition in the population in which the adaptation evolved).[51]

To say that there has been no variation in a trait is to say that it is out of the blue. Buller treats biochemical characteristics of organisms as if they did not vary anytime.[52] I think Millikan's theory does not require recent selection. The

---

[50] Buller, D.: 1998, p.513

[51] Futuyma, D.: 2005, p. 247

[52] Buller, D. :1998, p. 512

functions of complex biochemical pathways may have been fixed long ago and variants may be eliminated because tinkering with them would most probably produce offspring that will die before maturity. Showing little or no variance does not imply that no selection has occurred. It only implies that there was strong purifying selection for those traits.

Other counter examples involve swampman[53] (a molecule to molecule duplicate of a person who randomly and immediately appears), screws that accidentally fall into a machine and make a functional connection[54], brand-new antibiotic resistance gene that enters a bacterium[55], etc. All of these attempts aim to show one thing. History is not essential to function ascriptions. These counterexamples seem to be missing the point in a crucial respect. If Millikan were conducting a conceptual analysis, in other words, if she were trying to determine the criterion for true uses of the term 'function', these examples would show that including selection and even reproduction in the definition of function would lead to constant failures. However, Millikan explicitly states that she is not in search for a clarification of the older versions of function concept.[56] Rather, she was trying to find a theoretical definition of purposeful phenomena.

Although she does not explicitly say what the difference will be, I can cite one property of theoretical definitions: they may conflict with your intuitions. Consider "burning". A theoretical definition of burning would cite chemical

---

[53] Sterelny, K. and Griffiths, P.: 1999, p.222

[54] Kitcher, P. "Function and Design," in Hull, D. L. and Ruse, M. (eds.)

[55] Sterelny, K and Griffiths, P. *Ibid,* p.222

[56] Millikan, R. G.: 1989a, p.293

reactions which involve oxidation. Thus, corrosion would be burning according to the theoretical definition. However, someone who lacks that theoretical knowledge would find this characterization counterintuitive. It is not a conceptual necessity that functional things have a selection history. It is how things work on our planet.

Preston criticizes Millikan's account for overestimating the role of selection. The reason is that there are biological cases in which natural selection has not occurred yet but the traits under scrutiny are still functional. These traits are called exaptations. Exaptations are biological traits that have been designed by natural selection for some purpose and then co-opted for another use. Preston claims that if there are permanent exaptations, we should ascribe them causal role functions regardless of their selection history[57]. She gives mantling herons' behavior as an example in which wings are adapted for flight and used for mantling as an additional function. The example is taken from Gould and Vrba's important article where they first invented the term "exaptation".[58] In that article, Gould and Vrba suggest the term "exaptation" to be used for characteristics of organisms which are not selected for the effects they are currently performing but whose effects are somewhat useful to their bearers. Gould and Vrba suggest that wings have been selected for flight but not for assisting mantling behavior because heron wings haven't been optimized for mantling, compared to sister species.

Millikan criticizes Preston for confusing exaptations with causal role functions.[59] Cummins type functions, according to Millikan, cannot be equated

---

[57] Preston, B.: 1998, pp. 215-254.

[58] Gould, S. J., Vrba, E.: 1982

[59] Millikan, R. G. 1999, p.193

with exaptations because most of those functions, in biological cases are adaptations. Since ‘adaptation’ and ‘exaptation’ are mutually exclusive terms (i.e. an exaptation is a useful structure or behavior that is *not* an adaptation), causal role functions and exaptations are not identical.

I agree with Millikan’s answer because every adaptation begins as a fortuitous benefit. Biologists use the term “preadaptation” for those traits. Preadaptations are useful traits of organisms which currently serve a different function from the functions they were selected for.[60] Most complex adaptations should begin as preadaptations because otherwise we would have to suppose that some complex structures evolve by chance and then are selected as fully developed structures before the need for them arises. [61] The concept of preadaptation was

---

[60] Futuyma, D.:2005, p.261

[61] At this point, I feel the necessity to express a personal opinion. After Gould and Vrba’s article, the talk of exaptations has been very popular, along with spandrels, in philosophy and cognitive science but if they are natural kinds like adaptation, I would expect a science of exaptations (or spandrels). I would at least expect some positive definition of exaptation or spandrel. Exaptations and spandrels are defined merely in opposition to adaptations. Hence, exaptations and spandrels are conceptually parasitic on adaptations. At the genetic level, most changes in the genome have been shown to be neutral, after Motto Kimura. However, nobody has yet demonstrated, to my knowledge, a better explanation than natural selection for useful (either currently or historically) phenotypic traits. I suspect the usefulness of those concepts (exaptation and spandrels) other than warning evolutionists that not all biological traits are adaptations, following Dennett (1996). Adaptations should begin as accidentally beneficial traits. This accidental usefulness of some traits is what makes them exaptations. However, not all exaptations are preadaptations because their bearers may go extinct before the trait is improved by natural selection. Even if all exaptations are not preadaptations, all preadaptations should be exaptations because otherwise complex adaptations should have arisen before the need for them has arisen.

introduced to overcome the challenge of creationists, who argue that Darwin's explanation of adaptive complexity is false because for a design accumulation process like natural selection to succeed, every step leading to a complex adaptation (e.g. a 5% eye) should benefit the organism. To answer this challenge, evolutionists argued that steps leading to an adaptation were selected for some other function and then coopted for the present function.[62] In addition to this, Millikan's theory never searches for "the" proper function of entities. There might be more than one selective pressure acting on a trait so we can ascribe more than one proper function to an entity.

Pluralist and combination theories have been offered against Millikan's unificationist[63] account. Amundson and Lauder claim that there are areas of biology – such as functional anatomy – where causal role functions are the "only game in town".[64] They insist that causal role functions cannot be eliminated in favor of proper functions because finding evolutionary explanations for every trait is difficult (and also out of concern). Even if there is a proper function and selective history for regulatory pathways in multicellular organisms, molecular biologists can study the synchronic aspects and causal relations among the constituents of those pathways without referring to the evolutionary history of those pathways or constituent molecules.

They also argue against Millikan's claim that biological categories are

---

[62] Gould, S. J. , Vrba, E. , *Ibid.,* p. 11

[63] Millikan's unificationism differs from combinatorial approaches because she doesn't try to benefit from the virtues of both accounts (selected effects and causal role accounts). She rather prefers her own concept of "proper functions" as a unifying account of purposeful (not in the conscious sense) phenomena.

[64] Amundson, R. and Lauder,G. V. : 1998, in Hull, D. L. and Ruse, M. (eds.)

individuated functionally. Their point is that anatomy can individuate entities just by looking at their structures. For example, in comparative anatomy homologies are identified by structures rather than functions because homologous structures usually differ in functions.

I can give one more example from molecular biology that seems to support their theses of ineliminability of causal role functions. In molecular biology, interactions of signaling molecules, receptors and transcription factors are analyzed exactly in the way Cummins analyzes functions. They are so much like flowchart diagrams that someone might think that some form of molecular computation is going on. Evolutionary considerations do not enter those diagrams and the function of a molecule is defined by its contribution it makes to the specific biochemical pathway. Hence, even if proper functions can capture the usage of function ascriptions in some fields, there are other fields in which Cummins functions fit better.

Philip Kitcher argues for a unified concept of functions. He tries to achieve this by defending the role of *design* in function attributions. He claims that function of x is what x is designed to do.[65] For Kitcher, both artifacts and biological organisms are products of design processes. In the artifact case, intentional design is present. In biological organisms, natural selection is the designer. The grounding processes for design might differ in two cases but it is still possible to unify them under one title. Kitcher defends his thesis by citing some evolutionary and physiological explanations where the shared component is the presupposition of design.

Kitcher states that intentional design is considered unproblematic but biological function poses the serious problems. He claims that seeing

---

[65] Kitcher, P. 1998, "Function and Design," in David. L. Hull and Michael Ruse (eds.)

organisms as products of design is a relic of religious beliefs about creation. However, he states, even pre-Darwinian biologists could explain organismic functions by means of the organism's relation to its environment. Some basic needs of organisms (i.e. nutrition, mating) and the environment imposed some demands over the organisms. The structure of organisms reflected those demands. For instance, the big molars of herbivors reflected the demand to gain energy from plants. The same reasoning, says Kitcher, can be reformulated in the theory of evolution. The "demands of the environment" is translated into "evolutionary pressures". Hence, the environment in which the organism lives (and selection in that environment) might be seen as a designing force that shapes organismic structure in accordance with functional requirements. By such an analysis, Kitcher cliams to have unified two function concepts (SE and causal role) and also artifact functions and biological functions.[66]

Davies is the most radical proponent of causal role functions. He claims that everything that can be explained by proper functions can be explained by a modified version of causal role theory.[67] In addition, he claims that causal role functions are epistemically necessary to ascribe proper functions. His defense of this thesis depends on two auxiliary assumptions. First one is that, there are no real norms in the biological world. In other words, there is no malfunctioning biological entity. He says that the apparent normativity in biological world is a consequence of our expectations. When we see a complex and hierarchically organized entity, we expect it to continue its existence. Hence we interpret its actions in accordance with norms.

---

[66] *Ibid.*

[67] Davies, P. S.: 2003

The second reason to discard proper functions is that natural selection itself can be explained with a Cummins style analysis. In that case, the capacity to evolve would be a systemic capacity and natural selection would be one of the simpler capacities of a population that contributes to this capacity. Since evolution itself can be described in this way, proper functions account does not contribute anything to our knowledge (it is parasitic on causal role theory) and should be eliminated.

CHAPTER 4

Autonomy and Self-Maintenance

Autonomy is acting on one's own behalf. In other words, autonomy is calibrating one's actions to achieve an internal goal. An autonomous agent will manipulate both its external and internal environment to fit its goals. The term 'goal' needn't refer to conscious goals. For instance, if you assign to a bacterium's movement towards a food source the goal to feed, you don't mean it has a conscious representation of that goal. Its flagellar movements, the way it senses the glucose gradient are all assigned functions according to that goal. The concept of goal is more appropriate than need, fitness or survival because it is more general and fits cases when an organism sacrifices itself, its progeny or perform other kinds of self-destructive deeds (like apoptosis).

Autonomy is a source of normativity because actions performed by autonomous agents can be judged according to their contribution to a specific goal. Hence, it is possible to assign functions to autonomous systems. To make the point clearer, let's consider an HIV virus that enters the T-cells of our immune system.[68] Normally, T-cells present antigens to B-cells for inducing the mass production of antibodies specific for that antigen. An HIV infected T-cell loses this capacity and HIV viruses proliferate by great rates after disabling host defences. There is a latency period in which HIV viruses enter the genome of T cells and remain silent. This period may range from six months to 15 years

---

[68] The autonomy of a virus may sound surprising for most of the biologists. However, having no metabolism does not mean that viruses are not autonomous. They use the host's mechanisms on their own behalf, and autonomy can be formulated not only for individuals but lineages. See the following pages for more detail.

depending on the infected person. The first response of the immune system is powerful: Antibodies for nearly all recognizable parts of the virus are produced in great numbers. However, this doesn't completely eliminate some mutant strains of the virus. They mutate, integrate into the host's genome and replicate at very high rates to escape human immune system's defenses. They remain silent and after a lengthy period, they exit their host cells and infect new cells.[69]

The individual actions (or parts) of the virus that result in this consequence can be judged regarding their contribution to this specific capacity to destroy T-cells. Deviations from those actions can only be called "malfunctional" relative to that goal. For instance, if a virus doesn't destroy T-cells and thus the antigen is presented to B-cells and the deficient part (e.g. a mutated coat protein such as gp120) is identified, it is legitimate to call that protein malfunctional. Normativity necessitates goals and only autonomous systems have intrinsic goals.

Autonomy is best studied if we look at its simplest forms. A eucaryotic cell would not be the best starting point because its metabolism is too complicated. Let's remember that we want to find what defines autonomy and how it can be a source of biological normativity. Hence we better begin with theoretical models. This is necessary because even the simplest unicellular organism has a DNA, tools to transcribe it into proteins, a semi-permeable membrane, etc. We first want to know whether any of these are necessary components to be autonomous. Theoretical models abstract from the details of actual organisms that we doubt to be essential. Thus, they allow us to find the minimal conditions to be autonomous.

Best studied minimally autonomous systems are autocatalytic networks.

---

[69] Wood, P. J.: 2006, Chapter 14: AIDS

Catalysis is a term used in chemistry. In a chemical reaction, there is an energy barrier to be passed in order for the reaction to occur. To catalyze a reaction is to help reactants (i.e. molecules or atoms that react) to pass that barrier. This can be done in many ways. For instance, the reactants can be confined to a two dimensional surface, the medium can be heated or specific chemicals (catalysts) can be added to the medium. The function of catalysis is to speed up the reaction, which will occur very slowly otherwise. Autocatalytic systems are those systems which can catalyze the production of their own *essential* components. Essential components are those molecules that cannot be obtained from the medium.

The simplest autocatalytic network is exemplified in Kaufmann (1993) as follows:

$A \rightarrow B \rightarrow A$

Here, A catalyzes the synthesis of B and vice versa. The building blocks of A and B are omitted in the diagram for the sake of simplicity. If we add them, the system becomes something like this:

$A^1+A^1 \rightarrow A$

$B^1+B^1 \rightarrow B$

In this case, arrows represent the synthesis of A and B from the precursor molecules. The first reaction is catalyzed by B and the second is catalyzed by A.

Autocatalytic networks form the simplest kind of autonomous systems. Here, autonomy means only the capacity to reproduce essential components of a system by that system itself. Building blocks or precursor molecules are usually stable. Bigger molecules, however, tend to break down easily. A system should reproduce them if they are its means to persist. Autocatalytic networks are not independent from environment. Both energy and precursors should be provided from the environment.

Basic autonomy was defined in various ways. For instance, Kauffman defines it as autocatalysis plus the capacity to reproduce and “to perform one or more thermodynamic work cycles”[70]. I already explained what autocatalysis is. Ability to reproduce is tightly related to autocatalysis. An autocatalytic chemical network would, by definition, have the capacity to reproduce.

As in the schematic description above, provided that the system has fragments of A and B and adequate energy, it will continue to produce A and B. Since A and B are themselves the catalysts, the set will reproduce itself.

The arguments in favor of the spontaneous generation of autocatalytic reaction networks (after a certain complexity threshold is passed) are these:

1. In a random graph consisting of dots and bonds, when the ratio of bonds to dots passes the threshold value of ½, giant clusters of connected dots emerge.

2. Bonds represent reactions and dots represent reactants.

3. We can assign a random probability of catalysis to any molecule (i.e. dot) in the reaction network.

4. There are more reactions that produce constituent molecules than the molecules themselves because there are many different ways to produce a given molecule. For instance, AAABBB can be produced by

---

[70] Kaufmann, S:.2000

combining ABs, AAAs and BBBs, or cleaving AAABBBAABB to AAABBB and AABB. Hence, bonds/dots is greater than ½.

5. When the diversity of molecules increase, the number of catalyzed reactions increase even faster.

6. When a diversity threshold is passed, it is almost certain that a network in which every molecule will be produced by a reaction that is catalyzed by another molecule inside that network. Thus, an autocatalytic network will emerge.

There are some possible problems with this simple definition of autocatalysis. Energy requirements for the production of large molecules should be satisfied. Large molecules break down with a higher rate then their synthesis, unless energy is provided from outside. This is achieved by coupling endergonic (i.e. energy requiring) reactions with exergonic (i.e. energy producing) reactions. In addition, a lipid bilayer will confine the space in which the molecules interact and raise the possibility of their well-oriented interaction. Hence, the simplest form of life can be defined as an autocatalytic network of reactions bounded spatially and energized from outside by means of simple food molecules both acting as energy sources and constituents. [71]

The ability to perform work cycles is a more elaborate matter. Complex and far from equilibrium systems tend to break down easily in accordance with the second law of thermodynamics, unless energy flux into the system is used appropriately in order to repair the system. An autocatalytic set or network can be constructed out of purely spontaneous reactions. Spontaneous reactions are exergonic reactions in which energy is transmitted. They do not need energy

---

[71] Kaufmann, S.: 1995, Chapter 3: We the Expected

supplies to proceed. However, living systems are high energy systems. In thermodynamics, the second law suggests that all closed systems will lose energy and order untill it reaches a low energy stable state. Living things are open thermodynamic systems which take energy from their surroundings. They use this energy to maintain their high energy (and ordered) state. For this end, exergonic and endergonic (energy requiring) reactions are coupled in an organism's metabolism.

The critical point in work cycles is to convert energy, provided to the system from either internal exergonic reactions or external sources, into useful forms. For instance, the chlorophyl molecule in plants convert energy of photons into a more appropriate form for the plant: glucose. The energy is stored in the chemical bonds of glucose. This kind of conversion is vital because organisms should control the flow of matter and energy. If they don't, their autocatalytic network will collapse. The capacity to perform work cycles is to convert energy in order to control it and use it in appropriate amounts. It is also to use it where and when it is needed.

Organizational closure (Mossio et al 2009) and catalytic closure (Kauffman 2000) are also necessary for function attribution. Catalytic closure means that for every catalyzed reaction in an autocatalytic network, one can find a catalyst inside the system. In other words, the catalytic task space of a closed system is filled from within. A catalytic task space is simply the following. Catalysts are special molecules that speed up specific reactions. However, no catalyst is absolutely specific. The catalytic performance of a catalyst is determined by the shape of its active site (this is especially true for complex catalysts). The active site binds to the transition state produced by the substrate(s). The transition state is a transient high energy configuration of the substrate(s) before the end product is produced. A shape space of transition states can be formed just like a similarity space in three (or more) dimensions. Similar reactions, whose transition states will have similar shapes, will be catalyzed by similar catalysts and vice versa. Thus, we can identify in which part of this space a catalyst will be active. A specialist catalyst will fill a small subspace in the shape space (i.e. it will catalyze a small number of reactions). A generalist catalyst will fill a greater subspace. In this context, catalytic closure means that the shape space of the reactions in a system is filled by the catalysts in the system itself. In addition, the reactions that produce the catalysts are within the filled space.

Organizational closure is a more general concept than catalytic closure such that it implies a circular causal relation between the higher level configuration and the microstructures that constitute that configuration[72]. Higher level configuration of a system helps maintain those microstructures and vice versa. Convection cells provide a good example to this pehomenon. If a homogenous liquid is heated from below, the molecules form ordered structures after the liquid reaches a certain temperature threshold. These structures are called convection cells. The constituent molecules in a convection cell rotate on either a clockwise or an anticlockwise hexagonal trajectory. The direction of the molecules is determined by the convection cell they belong and its immediate neighbors. If in one cell all the molecules rotate clockwise, the molecules in the neighbor cell rotate anticlockwise. Hence, the rotations of single molecules maintain the overall structure and overall structure imposes constraints upon their rotation.

Organizational closure and catalytic closure are both introduced to explain why a system is autonomous. Autonomous systems act on goals but as I said before, goal concept is at least as normative as function concept. Hence, researchers try to naturalize goal states as well. An additional task for autonomous system research is to internalize those goals. To do this, they try to find which internal processes and structures contribute to some specific capacity of the system. To

---

[72] Mossio, M., Saborido, C., Moreno, A. 2009

achieve this, it is necessary to find a nonarbitrary way to distinguish specific effects from other effects produced by the system's components. In an organizationally closed system, goal states are defined as the states where the system can exist and functions are those effects that contribute to maintain or reach those states.[73] Malfunctioning of a trait will be producing such effects that push the system away from goal states. Since goal states are purely physical states, this should be considered as a step towards naturalization of norms.

Simplest autonomous systems, according to Bickhard, are those which can sustain themselves but do not have specialized processes to achieve this. [74] He gives the example of a candle flame. The candle flame is a self maintaining system. Flame vaporizes the wax and this provides fuel for the flame. It creates a pressure difference which sucks oxygen rich air and pushes carbon dioxide away. There are no differentiated (or specialized) processes in this system. Hence, it demonstrates the simplest form of autonomy. Undifferentiated processes contribute to the persistence of the system.

A more complex autonomous system would have differentiated but automatic processes. These processes would perform specific tasks in accordance with the overall organization of the system. They would lack the ability to regulate themselves (and be regulated by other components) under changing environmental conditions. In other words, their action would be totally deterministic and would be insensitive to any environmental information. This is simply because these processes do not contain switches.

The last and the most living-like systems should have, in addition to

---

[73] *Ibid.*

[74] Bickhard, M.: 1993

specialized self-maintaining processes, switches sensitive to environmental changes. These switches should control the flow of energy and matter. Different conditions trigger different reactions, by means of those switches. Then, these systems should "decide" what to do under certain circumstances. Their decisions would determine when and where to transfer energy and other resources. They would be more flexible than the simpler ones, under changing environmental conditions. They would have the capacity to respond adaptively. For instance, a bacterium swimming in a glucose gradient changes its flagellar movements when it enters in a low concentration zone. Since it has a switch that enables it to behave adaptively under changing conditions, it can control its movement in accordance with the external stimuli, which in this case is glucose gradient. This kind of systems is recursively self-maintaining. Recursivity comes from their ability to act to maintain the conditions of their self-maintenance. They maintain themselves by means of maintaning the conditions for their maintenance.[75]

Self maintenance, autonomy and normativity are closely tied. An action is called functional if it contributes to the self maintenance of a system. However, not all self maintaining systems bear functions. Some of them, the simplest ones, do not bear functions because they do not have differentiated processes. Differentiation and specialization are necessary for functional analysis. An undifferentiated process or an unstructured entity will be as suitable as anything to a task. Suppose that, instead of blood, your veins are filled with tap water. Even if tap water can carry oxygen and nutrients, its efficiency in doing that will be very low. Water can interact in many ways with importants biological molecules. Some of these interactions are ubiqutous as they may

---

[75] *Ibid.* A recursive function can take its output as its input for an indefinite number of cycles. In the context of Bickhard's article, the function is self-maintenance itself and the self maintaining system acts to maintain the property of self-maintenance.

occur anywhere and anytime. For instance, hydrogen bonds and Van der Waals forces are some of the weak interactions which may form between enzymes and water molecules surrounding them, anytime and anywhere. However, these types of non-specific interactions are difficult to control. Life processes need to be, to some extent, isolated from each other and their surroundings. Metabolic pathways need to be separated biochemically by means of interaction specificity, like the specificity of enzyme-substrate interactions. This is necessary to respond adaptively to environmental changes. Non-specialized systems can also carry out almost any function but with a poor performance. Improbability of control and poor performance imposes specificity. Thus, omnipotency is impotency regarding life processes.

Specificity is the core property that connects structure and function. In a sense, all living things are the result of interplay between specificity and flexibility. Induced fit is a good example of interplay in molecular processes. According to the induced fit hypothesis, binding sites of enzymes undergo conformational changes after the binding of the substrates. These changes give the binding site a more similar shape to that of the substrate. Before the interaction, enzyme has certain specificity for the substrate. This prior specificity may be considered as a genetically coded and selected specificity. The increased specificity after binding may be considered as a synchronic and input-driven property of the enzyme.

A better example of the interplay between coded-selected specificity and input driven and synchronic enhancement of this specificity can be observed in the working of the immune system. An antigen binds to an antibody in its variable regions. Binding increases the affinity of the antibody and increases the number of weak interactions. If the specificity was absent prior to binding and was only the consequence of binding, it would be impossible to discriminate self from non-self. If specificity was entirely coded, the immune system should have a prior stock consisting of antibodies for almost any antigen. The solution is to produce inducibly specific antibodies.[76]

Relating autonomy and self-maintenance to functions debate is an important contribution to causal role accounts of functions. Causal role accounts were usually criticized for underspecifying functional traits and being insensitive to the normative aspect of function attributions.[77] They were also criticized for being too generous in attributing functions. Constraining function attribution to self-maintaining systems seems to solve these problems. Two points are important here. Firstly, norms are ascribed only to autonomous sytems which are recursively self-maintaining. The simplest autonomous systems like candle flames do not have functions because they do not have differentiated parts that contribute to their maintenance.[78] The second point is about reproduction. Sometimes, traits that enhance reproduction may have no effect on self maintenance or can even be deleterious (e.g. altruistic traits). Hence, some traits should be functional without contributing positively to self-maintenance of the *individual*. This is one reason why natural selection is a complementary source of biological normativity.

In functional analysis, organization is important in explaining the functions performed by the whole system. For instance, the functioning of the heart necessitates an appropriate spatial organization of parts and an appropriate timing of performances. Hence, parts in a mechanism constrain each other's activities. This is one source of normativity in biological systems because parts acting out of those constraints will be rightly called to be malfunctioning. For

---

[76] Nelson, D. L., Cox, M. M.: 2000, p.231

[77] Mossio, M. et al., *Ibid.*

[78] *Ibid.*

example, a mitral valve that does not close after the contraction of the left atrium will let blood flow back to the left atrium. Such a deficit can be called a malfunction for two reasons. First, it decreases the pumping performance of the heart. Second, it does not act within the timing constraints imposed by the heart muscles and sinoatrial and atrioventricular nodes.[79]

Bringing functions back here and now is inevitable because functional analysis is used to explain how biological organisms as complex systems work. Selected effects (SE) accounts make functions epiphenomenal in the sense that current working of the system is by no means explained by function attributions. Bickhard explains this point with an example.[80] He claims that, according to SE accounts, traits are functional iff they have the right history. The causal contribution of a process to the survival and reproduction of an organism is considered functional only in the context of past selection. However, if we accept this, the accidental replica of a living organism (e.g. swampman) would bear no functions because it would not have any selective history. This seems counterintuitive because the replica will be structurally identical to the actual organism. Its parts will contribute to its self maintenance just like the original one. In addition to this, only the current states of a system can be *causally efficacious*. History or past states can exert an influence on the system only by affecting current states. If function attribution cannot be grounded on the current state of the system, it will have no explanatory power about the current working of the system. This leads to epiphenomenalism about functions because having functions will make no causal difference. Hence, function attribution will have no role in causal mechanical explanations. In order to avoid epiphenomenalism, we should find out how history contributed

---

[79] Sinoatrial and atrioventricular nodes are nerve centers that send electrical signals to heart muscles.

[80] Bickhard, M.: 1993, p. 6

to the specific details of current organisms.

Differentiation and specialization provide the prime source of intuition in search for new functions. They also provide the first clue to start a functional decomposition of an organism. For functional decomposition to succeed, our analytic account should have non-arbitrary grounds for distinguishing functional structures. Not all systems are suitable for functional decomposition. Some systems are highly integrated. Their working cannot be analyzed into the activities of simpler and lower level component parts in a non-arbitrary manner. To understand the difference, we should glimpse at a foregoing discussion in cognitive science, whose sides are the supporters of the classical paradigm and connectionists.

Classical cognitive science or the classical paradigm in cognitive science is the research program in which mind is considered as a symbol manipulating device and its operations are governed by explicit rules (like that of logical rules). According to some prominent followers of this paradigm, such as Jerry Fodor, some parts of the mind (i.e. perception and language) can be functionally decomposed into domain specific computational units. These units are called modules. The most prominent features of modules are their domain specificity (i.e. they have a very limited database), innateness, encapsulation, hardwiredness and automaticity.[81]

Brain lesion data are interpreted by this school as an evidence for functional specialization and spatial localizability of certain computational units because lesions in specific areas cause selective deficits.[82]

---

[81] Fodor, J.: 1983

[82] Pinker, S., *Ibid.,* p.19

Encapsulation, in this context, means that a module's internal workings cannot be changed by external influences. In other words, either any external influence will count as an input for the computed function or it will not affect anything at all. The module's influence on other modules will be similar in this regard. If our module is computing the function $f(x)=x^2$ where $5>x>-1$, an input of 3 will give 9 as output and 6 will give nothing. The module is sensitive to only a small subset of possible inputs, and no input can change its internal operations. For instance, no input can change the internal computation from $f(x)=x^2$ to say $f(x)=\log(x)$.

Connectionism is radically different from the classical approach. The major difference lies in their concepts of computation. In classical models, rule governed symbol manipulation is of the essence. In connectionist networks, computation is performed at a subsymbolic level such that the representation (i.e. the symbol) is distributed among units. Connectionist networks act in parallel, there are no symbols (hence their operation is understood as subsymbolic), and representations are distributed among different units of the network. Even if the output of a connectionist network looks as if there were a rule governed activity, this is not the case. Rules or rule governed activity is just an emergent property of the system.

Another important difference is that, in connectionist models, computations are not performed by individual units. These models consist of simple units. Every such simple unit is connected to many other units. Units are classified according to a hierarchy of computation: input units, hidden units and output units. These units are so simple that their operations do not explain any interesting computation carried out by the system. Units are not specialized. The only thing they do is to transmit activation (or inhibition) accross their connections. Computation is performed by the whole network of connected units.

Rather than computing a fixed internal function, connectionist networks are input driven. The network does not know what to do at the beginning.[83] It is trained by providing samples of a given type as input. There is an expected response to a prototypical instance of that type. This expected response (output) is used as a criterion for correction. By applying that criterion, the success of the real response is calculated and the connection strengths are calibrated according to the generalized delta rule.[84] Hence, the network learns how to respond *appropriately*.[85]

There are some lessons to be learned from the above comparison. First, as mentioned above, not all systems are appropriate for functional decomposition. Connectionist systems are of this kind. They are structurally decomposable but not functionally decomposable. There are units of the system which belong to different layers (input, hidden, and output). There are connections, connection

---

[83] Even if it is possible to set fixed connection weights before training, connectionist networks are interesting mainly because of their learning capacity in terms of adapting their own connection weights upon training. The above possibility is omitted in order to increase the contrast between classical models and connectionist models. For a full discussion of these ideas, see Bechtel and Richardson's *Discovering Complexity,* especially chapter 9.

[84] Churchland, P.M., *Ibid.*

[85] I wrote "appropriately" in italics to emphasize a simplification in connectionist model building. In connectionist models, the *appropriateness* of the response of a network is determined by the preferences of the model builder. For instance, in Paul Churchland's (1988, Section 7.5) example, the network is trained to respond to sea mines and distinguish mines from rocks and other sorts of stuff. The trainer forces the network, by means of the generalized delta rule, to respond to mines with the vector <1,0>. Thus, the emergence of norms is guided by the corrections made on the initial vector. The vector is adjusted to fit a prototypical example of a mine that is selected by the observer herself.

weights, etc. that can be structurally analyzed. However, units are too simple to carry out any siginificant function. The overall organization (i.e. connections, connection strengths, etc.) is a functionally distinguishable unit. In contrast, classical models are constructed out of functionally distinguishable units. It is crucial to decide whether biological systems are best explained by modular or non-modular terms. It is crucial because this will change the subject matter of function attributions (from parts to wholes).

The second point is that not all systems are crafted beforehand to perform functions. In connectionist networks, function emerges after the training. In classical models, functions are innate.

As we observe in biology, different levels of organization are best explained by different types of models. Spontaneous organization in chemical networks will resemble the emergence of functions (or norms). At higher levels, e.g. organs, there is rather a crafting of functions beforehand. 'Beforehand' implies evolution by natural selection, hence, adaptation. Divisions in biology are not as clearcut as I depict here. One might find optimal solutions (i.e. adaptations) even at the lowest level. Even biochemical pathways exhibit symptoms of optimality, rather than law governed emergence by spontaneous organization.[86] The opposite may also be true. The overall organization of a multicellular organism is determined in the process of development, and development itself may be the result of spontaneous organization.[87] Even if the boundaries are fuzzy, the distinction is still important. Spontaneous organization is one way to explain the emergence of norms. Evolution by natural selection is another. They are complementary and sometimes alternatives to each other.

---

[86] Cornish-Bowden, A.: 2004

[87] Kaufmann, S.: 1995, Chapter 5: The Mystery of Ontogeny

CHAPTER 5

Adaptation

The second source of normativity is natural selection. In fact, invoking natural selection to explain normativity was considered to be the main advantage of SE accounts. SE theorists claim that this is the only way to naturalize functions and normativity. They also claim that the general features of a selectionist explanation could be extended to account for the normativity of language and mind. In the case of language and mind, nongenetic selection has been used in addition to classical genetic selection.

Natural selection happens when there is heritable variation in fitness. Let me explain this step by step. Organisms in a population vary in some of their properties. Some are tall, some are short. Some are fast and others are slow. Some of the differences are invisible to the bare eye. For instance, some individuals may have a mutated enzyme that enhances a metabolic pathway or detoxifies hazardous materials found in the environment. Variation by itself cannot produce natural selection. Some variation may appear due to environmental factors like nutrition. This type of variation will be transient, it will not affect next generation. Those variations which are transmitted through generations are called heritable variations. In particular, orthodox biological view posits only changes that affect the genomes of germline cells (whole cells in the case of unicellular organisms) as the source of natural selection. Only these count as heritable variation because only these changed properties will appear in the progeny with such a frequency that cannot be explained by mere chance or environmental induction.

Variation, either heritable or not, cannot cause selection by itself. Genetic variation should make a difference in the survival and reproduction capacity (i.e. fitness) of its owners to cause selection. If a variant and heritable property makes such a difference, it is plausible to say there is selection for (or against) that property. Most of the genetic changes do not affect fitness. These changes are called neutral mutations. They are neutral either because they happen in noncoding regions of the DNA or they do not change the amino acid sequence of proteins. In some cases, even if the protein sequence is changed, mutation affects only a non-vital part of the protein and it does not influence the functioning of that protein.

Here is a short summary of natural selection. Organisms in a population may have *different traits*: Birds have differing wing lengths, bacteria have differing capacities of antibiotic resistance, humans have different eye colors, etc. Some of these differences are *heritable*. For instance, resistant bacteria have resistant offspring, tall parents have tall children, etc. These heritable differences affect *reproduction rates* of their owners. For instance, resistant bacteria produce more offspring, tall parents produce greater number of children, long winged birds have more offspring, etc. Having more offspring *changes the composition* of the population in favor of fitter organisms. The fitter trait increases in the population. In short, selection requires *heritable variation in reproductive success*.

According to one definition of evolution, evolution is the change in gene frequencies.[88] Gene frequencies are determined by averaging the number of different alleles of a gene per capita. There may be different sources of evolution. One source is mutation. Mutations are changes in a DNA sequence.[89]

Other sources are genetic drift, migration, immigration, and natural selection.

---

[88] Sober, E.: 2000 p. 1

[89] Futuyma, D. *Ibid*., p. 165

Natural selection changes gene frequencies by means of differing reproductive rates of variant organisms. For instance, if a mutated gene X increases the fitness of its bearers, they will reproduce better than the bearers of the nonmutated allele. This will cause the population to have a greater percentage of X individuals in the next generation. In some cases, this process is repeated untill X reaches 100%. In other cases, some factors may limit the power of natural selection and prevent X from reaching fixation.

Natural selection is usually invoked to explain design-like features of organisms. Teleology of biological traits is the main source of philosophers' interest in functions. There seems to be a means-ends relation between the function bearer and the function it performs. For instance, restriction enzymes seem to be specifically produced for cutting certain regions of nucleic acids. Ion-gated channels in neurons look as if they were designed to control ion flow through the membrane. There seems to be an end that acts as a criterion against which the performance of the trait can be judged. In addition, the function bearer seems to be perfectly fit for the task. Or more correctly, it is optimal under internal and external constraints.

Design-like features of organisms are summarized in Dawkins' concept of adaptive complexity. Dawkins (following Williams) claims that only adaptive complexity in organisms deserves an explanation by natural selection. Other traits can be explained by other evolutionary forces like mutation, migration and drift. However, adaptive complexity, or functional complexity, is best explained by natural selection. Dawkins considers natural selection as a blind watchmaker.

Natural selection designs complex and functional traits without having the foresight possessed by human designers. Design, whether conscious or not, is what makes function attribution plausible. In the case of conscious design, intentions of designers determine the means-ends relations. In naturally designed systems, a long history of interaction with the environment determines what means are to be used for what ends.

Adaptive complexity is very similar to organizational differentiation (or specialization) we discussed in the last chapter. One meaning of complexity is having qualitatively different constituents. Heterogeneity of the components of a system forms its material complexity (Mossio et al 2009).

Adaptive complexity or functional complexity differs from material complexity because in adaptively complex systems, components are both structurally and functionally differentiated. These systems have an organization that is essential to their working.

Material complexity can be likened to William Wimsatt's aggregativity. Suppose *W* is a whole and *i*'s are its constituents. For W to be an aggregate, it should satisfy the following conditions:

> 1) *W* is invariant under the rearrangement and inter-substitution of *i*'s.
>
> 2) *W* remains *qualitatively similar* with the addition and subtraction of *i*'s.
>
> 3) There is no *inhibitory or cooperative interactions* between *i*'s.[90]

Aggregates are unstructured wholes; their causal powers are the sum of the causal powers of their components. In contrast, a whole is structured if *the organization of its elements matters to its activities.* We cannot ascribe functions to the parts of aggregates. Adaptive complexity, in contrast, requires a collaborative organization among functional components. Organization and the resulting functionality are what make a system or structure adaptively complex.

---

[90] Wimsatt, W.: 1986

There are differences between adaptive complexity and organizational differentiation. Adaptivity is judged according to how good a system fits into its normal environment. In other words, adaptivity is the contribution of a component to its bearer's fitness. Adaptivity is more strongly related to reproduction than survival and self-maintenance. Organizational differentiation is more tightly related to self-maintenance and survival.

Adaptive complexity is a way to characterize the functional parts of an organism which are too complex and too optimal to be explained by mere chance, self organization and other non-historical forces. Adaptive complexity is a result of a lengthy chain of design accumulation. This lengthy process is natural selection.

Natural selection is thought as a force that accumulates useful variation and discards harmful ones. Hence, it is an external and historical force that guides the process. More truely, natural selection is invoked to explain what environmental problem is to be solved by the trait at hand. In Brandon's words "we can see that there is no way, even in principle, to carve up an organism into its functional traits apart from its selective environment."[91]

On the other hand, organizational differentiation is not sensitive to historical forces. Organizational differentiation does not make any reference to the process out of which such systems emerge. It only refers to current properties displayed by the system. It is invoked to identify what kinds of systems can be ascribed functions, just by looking at their current properties. The practitioners of the research project which has created this concept explicitly argue that "biological systems, as organizationally closed and differentiated self-maintaining systems, possess the necessary properties for adequately

---

[91] Brandon, R.: 1999,. p.173

grounding both the teleological and normative dimensions of functions in their *current organization*"[92].

According to the externalist account of Dawkins and other adaptationists, organisms carry information about their environment. However, this information is not coded via learning, experience or other such simultaneous interactions with the environment. This type of externalism supposes that environmental regularities are coded or internalized by the organisms in a lengthy chain of small evolutionary changes. Hence, this type of externalism can be named historical externalism.

The difference between historical and nonhistorical externalism is important in functions debate because in the former, organisms are supposed to have delicate internal structures to deal with specific environmental challenges. Proximate explanation of a behavior will depend heavily on these internal structures. In Bechtel and Richardson's[93] words, "the locus of control" will be the organism itself. Only the ultimate explanation – the explanation of the origin of those structures – will cite historical and external facts. In the latter case, the organism will be thought as a non-differentiated simple system whose behavior is explained by external factors such as environmental regularities. Hence, the locus of control will be the environment. To understand this difference better, one can examine the difference between radical behaviorism and evolutionary psychology.

Radical behaviorism is a research program in psychology. The distinctive ideas that ground radical behaviorism are the associationism of Hume and Spencer

---

[92] Mossio et al.: 2009, italics mine

[93] Bechtel, W. and Richardson, R. C.: 1993

and functional psychology of William James and John Dewey.[94] According to the proponents of radical behaviorism, the best strategy to study human mind is to examine overt behavior and relate it with a carefully classified set of external stimuli. In other words:

> The only behavioral changes with psychological significance are responses to environmental changes; moreover, our learning capacities are the result of environmentally induced modifications in learning, and it is in terms of these simple learning mechanisms that complex capacities must be understood.[95]

The simple learning procedure consists of basic elements and their association by means of experience. The associationism of radical behaviorists forces them to accept no special mental faculty other than the association faculty.

Evolutionary psychology is a combination of computational theory of mind and adaptationist perspective in evolution.[96] According to the computational theory of mind, the mind is a symbol manipulating device like a universal Turing machine. The mind consists of propositions, proposition parts and rules for manuplating propositions. The Neocartesian followers of the computational theory such as Chomsky and Fodor claim that there are special faculties for some mental tasks such as language acquisition. These faculties, according to the neocartesians, are innately specified bunches of propositional contents. Mind's rough structure is the structure consisting of semantic relations among propositions (e.g. implication relation).[97]

---

[94] *Ibid,* p. 42

[95] *Ibid.* pp. 42-43

[96] Pinker, S. *Ibid.*, p. 23

[97] Fodor, J. *Ibid.*

Evolutionary psychologists ground the existence of innately specified faculties in terms of natural selection. For instance, Steven Pinker claims that: “The mind is a system of organs designed by natural selection to solve the kinds of problems our ancestors faced in their foraging way of life, in particular, understanding and outmaneuvering objects, animals, plants and other people.”[98]

According to adaptationism, function is, in a sense, information about the past environments where organisms evolved. All functions are selective. Their domains are specific. For instance, a bird’s wings are adapted to fly only in a fluid which has a viscosity similar to earth’s air. All of the enzymes are highly specific to their substrates. Eyes have a specific range of wavelengths for sight. It seems as if natural selection coded relevant information into our genes. The case is also relevant to artifacts. Artifacts are also domain specific. Their specificity depends on the things on which they are used. For example a pen is specific to some kinds of surfaces, you cannot write on anything. In a sense, a pen includes information about the nature of the surface on which it is intended to leave a mark.

The specificty of functional entities is explained by referring to environmental problems they are supposed to solve. Specific problems need specific solutions. In the biological world, we see many examples of specialization to different (and sometimes extreme) environments. For instance, some fish living in the depths of oceans have evolved natural anti-freeze molecules. These help their blood flow to continue normally in extremely cold environments. This fact needs explanation. Can natural selection provide a sufficient explanation of the emergence of such traits?

---

[98] Pinker, S. *Ibid.*, p. 21

Natural selection is a force that changes gene frequencies in a population. But how then will it explain the presence of adaptive complexity? You may recall Cummins' criticism of etiological theories for misinterpreting natural selection as a creative force. He claimed that SE theories gain their plausibility from a misunderstanding of the evolutionary theory. According to Cummins, natural selection does not explain the presence of a trait. It is explained by mutation and the genetic program (developmental steps that produce the adult organism). For him, natural selection is just a negative force. It only eliminates harmful variants.

These criticisms are partially right. The textbook conception of evolution equates evolution with the change in gene frequencies. Natural selection usually conceived as a sieve that eliminates the harmful and keeps the useful. No individual organism or individual trait is modified by natural selection. Organisms may adapt to their environments in their lifetimes by subtle changes in their metabolism. For instance, a human living in a coastland adapts to a mountain climate by a change in the number of her erythrocites. However, this sense of adaptation is completely different from the evolutionary concept of adaptation. Ontogenetic adaptation consists of the changes that happen in the lifetime of an organism and benefit it. Evolutionary adaptation needs many generations to emerge. Generations of organisms with heritable fitness differences will lead to stable and adaptive traits that characterize their nature. Mutation provides the raw matter for natural selection. If it is useful, it is kept. If it is harmful, it is discarded. Selection is conceived as a negative force lacking any creative power.

This criticism is partially right because natural selection is used in two different kinds of explanations. These are distribution explanations and origin explanations.[99] The criticism is true for distribution explanations. Distribution explanations deal with the question: "Why is this trait found in such and such frequency in this population?" Or we might be asking the reasons for an increase (or decrease) in the frequency of that trait. Sieve metaphor is partially valid in this context where the relative fitness of traits is important. Alternative alleles of a gene are compared to understand the frequency differences. Origin explanations are formulated to explain why the trait exists at all. In this context, the absolute fitness of that trait is important. A mutation may increase the relative fitness of an individual without increasing the absolute fitness. For instance, in a population where the census size is limited by food resources, a mutation that doubles the reproductivity of its bearers will increase relative fitness but not absolute fitness.[100] Wild type individuals will be eliminated by the rapid increase in the frequency of mutant ones. However, since the census size is limited by an external factor, at the end, mutant will have the same reproductive rates.

Natural selection has a special role in origin explanations. In every step towards an adaptively complex trait, mutation is the sole source of innovation. However, if selection increases the absolute fitness of its bearers, the total number of those variants will increase. This will raise the probability of mutations to happen in the right places. Suppose the steps going to a complex adaptation involve point mutations (i.e. mutation in single nucleotides) and the complex trait is an assembly of 4 necessary components. Let's symbolize the trait as abcd. The letters in abcd represent the components. Suppose xbcd, axcd, abxd and abcx are precursors of that trait where x is a trait only a point

---

[99] Godfrey-Smith,P. "What Darwinism Explains", Unpublished Conference Presentation

[100] Sober, E.: 2000, p.96

mutation away from the corresponding trait. If the mutation leading to axcd increases the absolute fitness of its bearers, the total number of axcd individuals will increase. This will raise the probability of the mutation from axcd to abcd. The same rationale can be traced back to earlier precursors of the trait (e.g. axxd). Thus, invoking natural selection explains why the trait is there indirectly by explaining why useful mutations happen in the right places.

The common point in all adaptationist explanations is their insistence on considering every interesting biological phenomenon as the functional product of a gradual design accumulation process. To say that a trait has function(s) is to say that it has some features that cannot be explained by mere chance. Adaptationism puts natural selection in the place of chance and intelligent designer. According to this theory, which is also called the synthetic theory, organisms are totalities of organized functions that have been selected for their greater reproductive success. Stephen Jay Gould criticizes this research strategy for overestimating the creative power of natural selection and underestimating the effects of chance, neutral mutations and exaptations.[101]

In addition to this, graduality of evolutionary progression is enabled by mutations and recombinations, which are simple enough to be realized step by step. Every function is inherited for the advantage it provides to the organism. The emergence of new functions depends merely on a series of gradual modifications in the genome of the organism. Random mutations provide a genetic diversity in the population and the ones who produce more progeny inherit more of their genes than the unsuccessful ones.

Adaptive evolution is not an unconstrained process. The presence of a suitable variation is the most important factor that limits the power of natural selection.

---

[101] Gould, S. J.:1985

There are other constraints on the power of selection which are important in understanding the emergence of biological norms. The structure of the genome (i.e. the organisation and interactions among genes) of an organism is one such constraint.

If the traits of organisms are totally independent (i.e. absolutely modular) and if the number of traits is huge, each trait's contribution to overall fitness will be very small. Hence, selection pressures acting on those traits will be weaker than mutations. An *error catastrophe* will destroy previously gained functional traits.

If characters are totally dependent on each other (i.e. absolute integration) and their numbers are huge, optimization of one character will lower the fitness contribution made by other traits. Hence, the population will be locked in a relatively low fitness peak. A *complexity catastrophe* will prevent the population from searching for higher average fitness.[102]

The constraints on adaptive evolution can be simulated by using NK type fitness landscapes.[103] NK models simulate the activities of a number of units (i.e. "N" in NK) and their interactions in order to predict the overall statistical properties of complex systems. They are very similar to connectionist networks discussed before. For every unit, we assign a random state and a Boolean function. The Boolean function, the state of the unit and input(s) coming from connected units determine the future state of the unit.

---

[102] Kaufmann, S.: 1993, p. 36

[103] A fitness landscape (or adaptive landscape) is a graph which shows the correlation between the frequency of an allele and the mean fitness of individuals in a population. An NK type fitness landscape is such a graph in which there are N different loci (with usually two alleles for each) with K epistatic interactions.

In evolutionary scenarios, K represents the number of interactions per gene and N represents the number of genes in a genome. In other words, K is the number of genes a gene affects. Here, the only relevant effect is the effect of a gene on the fitness contributions of related genes. When K is increased, the average number of epistatic interactions will increase and the organism will become a more integrated whole. In that case, increasing the fitness contribution of one gene will most likely decrease the fitness contributions made by the related genes. Hence, the more integrated an organism is, the more difficult for a population of that organism to reach optimal design (i.e. a global fitness peak). If we follow Kaufmann's reasoning, we must conclude that *the more modular* an organism is, the better it adapts.[104] So, organisms must be nearly decomposable in Bechtel and Richardson's (1993) sense in order to be subject to gradual design accumulation.[105]

---

[104] Modularity in developmental biology is different from modularity in cognitive science. In developmental biology, module is one of the repeating structures which constitute the body plan of the organism. A module in cognitive science is domain specific unit of computation. The only important similarity is that both types of modules act inedependently from each other. However, developmental or evolutionary modules may well be integrated in the adult organism. Their indepence from each other is only regarding development. Modularity and diversification of segments (i.e. differentiation of the identities of segments) have played an important role in animal evolution. For instance, diversification of segments in the common ancestor of onycophora and arthropods is beleived to be the hallmark of their separation. Differentiation of modules gives them the ability to acquire new functions. For example, reptile teeth are uniform. However, teeth of mammals have diversed into different functional structures such as molars, canines, incisors. See Carroll, S.B. et. al's (2005) *From DNA to Diversity.*

[105] Kaufmann, S.: 1995, p.173

We can observe the same trend of modularization in protein evolution. The interactions among proteins and their strengths show negative correlation with the speed of protein evolution. The reason for such a relation is that, if the number and strengths of interactions increase, any change in one protein will negatively affect the working or regulation of many interrelated proteins. Since the most likely result of those changes will be a decrease in overall fitness, mutants will be eliminated by purifying selection.[106] This might also explain why developmental genes like *Hox* genes are preserved in animal evolution. They are connected to many other genes and changes in them will certainly affect almost all other developmental pathways.

Local optima approximate the mean fitness if N and K are both great.[107] This is because of conflicting constraints we mentioned before. Then, it is a better choice to build systems which have a few relatively independent parts with correlated subparts. Another important point is that, in addition to N and K values, the strength of correlations can be tuned to produce a variety of systems. It will be just like tuning K. Suppose we call this variable as P which is just like the connection weight in connectionist networks. It represents how much an incoming input will change the receiving genetic unit's fitness contribution. If P is low, system will act as if its parts were independent. If P is high, the system will act as an integrated whole. The same logic can be applied

---

[106] Sean Carroll (2006, pp. 80-81) defines purifying selection like this: Purifying selection is the elimination of harmful nonsynonymous mutations. Nonsynonymous mutations are genetic changes that alter the amino acid sequence of a protein. Synonymous mutations are the changes in a gene which do not affect the amino acid sequence produced by that gene. Purifying selection conserved 500 immortal genes since the differences in DNA sequences are more abundant than the proteins themselves. See also Futuyma, D. *Ibid.* p.453

[107] Kaufmann, S.: 1993, p. 53

to parts and subparts. Thus, we may find the optimal division (and integration) among constituents which will allow the system to have reachable and high optima.

There may be two types of search procedures in a fitness landscape. Only one of them was mentioned before: One step mutation strategy. The other strategy consists of jumping randomly across the space of possible genotypes and staying on one of them if it is fitter and jumping back if it is less fit. Let's remember an important point. The effects of a point mutation are usually considered to be small in fairly smooth fitness landscapes. If the lanscape is correlated (i.e. neighboring genotypes have neighboring fitness values), point mutations will carry the population to similar fitness values. However, every mutation might not have the same degree of phenotypic effect. Some mutations, especially those on genes governing early development or essential biochemical pathways, will have drastic phenotypic effects due to their higher generative entrenchment.[108]

The K values are not identical for every gene in complex multicellular animals like vertebrates. Genes affecting early development will have greater K values (i.e. more epistatic interactions) than those affecting later stages of development. This fact forces us to modify our simple models to include different K values for different genes. When K is greater, the landscape will be less correlated and the average time to find a fitter alternative decreases logarithmically with every trial. This may allow us to restate Von Baer's law. In its original formulation, Von Baer underlined the fact that early embryonic stages of different species show greater similarity and similarity decreases in the course of development. We can restate this law such that K values for genes acting on early development will be higher than the K values for other

---

[108] Wimsatt, W., Schank, J. C.: 1986

genes. Since a greater K value means that fitter variants will decrease exponentially with every trial, the evolution of those genes will resemble evolution on an uncorrelated landscape. Genes regulating early development will be preserved and evolution will only tinker with later stages of development.[109]

The orderliness of NK networks can be tuned either by tuning K (i.e. number of connections per unit) or P (i.e. the Boolean function's bias). By tuning these values, one can achieve a state where both stability and flexibility of behavior can be observed in a network. A state, in which K and P values are tuned such that the system shows both stability and flexibility, is said to be "on the edge of chaos".[110]

There are two broad regimes where an NK network might be. The first broad regime is the ordered regime in which the network is completely frozen. In such a regime, changes in the states of connected units will not affect the stasis of the system. This stasis is a result of having a few attractor states which have great basins of attraction. An attractor is a state of a complex system where many neighboring states flow into. The basin of attraction is the set of states that flow into an attractor. If a basin of attraction is great (i.e. the ratio of the number of states belonging to that set to the number of all possible states is great) state changes will be immediately reversed. Most of the possible states of such an ordered system will be a member of the basin of attraction, and any perturbed state will also be in one of those basins and flow into an attractor.

The second type of systems are said to be chaotic. They have great many attractors and little basins of attraction. Thus, any small state change will carry

---

[109] Kaufmann, S.: 1993, p75

[110] Kaufmann, S.: 1995, Chapter 4: Order for Free

the system to a different attractor and change the whole system. Some systems on the border of these two extremes have both frozen and maleable sites. A change in the maleable sites will possibly change the attractor of that part but it won't affect the overall organization. The changes will be buffered by the frozen sites. But the system will still be modifiable in the borders between changeable and frozen sites.[111]

The systems on the edge of chaos are produced and maintained by natural selection because of their behavioral flexibility. In a sense, these systems are products of two forces: Self-organization and natural selection. The emergence of biological norms in this kind of systems is driven by both at the historical level by adaptive evolution and at the synchronic level by means of the self-maintaining properties of those systems. Here again we witness the interplay of historical and synchronic causes that shape the organism. I call this interplay because the two types of factors may replace each other at different levels of organization. Self-maintaining properties of far from equilibrium systems may be results of adaptive evolution and adaptive evolution itself is constrained by the self-maintaining organization of the systems at hand.

Simulating the evolution of complex organisms by means of abstract mathematical models has many advantages. First of all, it enables us to locate the contributions of different forces of evolution in the huge space of possibilities. This allows us to compare the relative effects of each evolutionary force. The second advantage these models bring is that evolutionary changes that take billions of years in reality can be simulated in minutes. This allows us to understand both macro and micro aspects of evolutionary change. The third advantage is that these models allow us to avoid actual details of historical contingencies and focus on the generic properties of evolving populations.

---

[111] Kaufmann, S. *Ibid.*

Abstract mathematical models of evolution give us some hints about how normativity may emerge in evolution. For instance, modularity (or near decomposability) is one of the properties that allow us to analyze organisms functionally. The emergence of modularity is explained by the necessities of adaptive evolution in those models. These necessities may be physical necessities (self-organization) or products of history (natural selection). The models give us the opportunity to compare the relative weights of historical and synchronic forces that shape the organism into a viable collection of diverse functions. Hence, we can decide on which organismic functions are best analyzed as natural and generic dispositions of complex systems and which are best analyzed by means of idiomorphic historical conditions.

A lesson we can draw from those models is that biological norms may emerge from the interaction between generic properties of complex systems and their interaction with the environment of selection. Self-organization is a force that explains the emergence of life as a necessity (i.e. not historical coincidence). Since life is self-maintenance and self-maintenance is one source of biological norms, self-organization in complex systems is one source of normativity. However, organisms have diversified so much after the emergence of life, it can not be the sole source of biological norms. The other source is the competition among individual organisms in order to exploit environmental resources in order to reproduce better than their rivals. The enormous diversity of organisms along with their adaptedness to a wide variety of environments can not be explained by only self-organization. Thus, it is necessary to refer to the environments in which the organisms evolved, in order to explain diverse functional traits.

CHAPTER 6

Discussion and Conclusion

The thesis I defended so far can be summarized in this sentence: Biological normativity is an illusion created by natural selection and self-maintaining capacities of organisms and since it is an illusion, it is not an obstacle against reductionism. Let's begin from reductionism in biology.

Reductionism in biology is the idea that molecules and their interactions will ultimately explain any biological phenomena. There are two types of reduction in biology. The first is synchronic reduction by which we explain the current workings of a complex system in terms of its component parts and their interactions. The other type of reduction, which I dubbed diachronic reduction, explains complexity by means of a gradual design accumulation. The two senses of reduction refer to two research strategies in biology. Synchronic reduction is the strategy carried out by molecular biologists (and biochemists, neuroscientists, etc.) who merely try to analyze the complex machinery of life into the components, subcomponents and the causal pathways connecting them. Diachronic reduction is the strategy used by evolutionists to find out how the apparent complexity and adaptivity of organisms emerge by means of simple achievable steps. The two strategies complete each other in our search for a deeper understanding of adaptive complexity and enormous diversity of organisms.

Reductionism was once a totally negative label that people used in methodological discussions to show that their opponents had blinders that prevented them from realizing the intricacy of a subject matter. Nowadays the intellectual climate has dramatically changed in favor of reductionism. After the completion of the human genome project and other surprising successes of molecular biology have a definite role in this change of climate. In addition, given the successful examples from sciences of the nervous system (e.g. functional neuroanatomy, neurophysiology, pharmacology, etc.), reductionism can claim ownership of the domains of social sciences and philosophy. This seems to me good news both for sciences of nature and sciences of mental. However, there are still many obstacles against the triumph of reductionism. Ignoring moral aspects of the matter, intellectual difficulties in distinguishing genuine scientific knowledge from mere oversimplification of phenomena by misapplication of reductionist schemes in inappropriate domains is a real threat. The threat is real that we are afraid of losing the essence of those phenomena using too simplistic or too artificial models.

One of the essential properties of organisms that we fear to lose as a result of reduction is the apparent normativity of functions. We judge the performance of a trait and evaluate it in accordance with a norm. For instance, we evaluate the performance of a nicotine receptor by its capacity to bind nicotine. Functions and norms span the whole realm of biology, from macromolecules to ecosystems.

The debate about functions and functional explanations has revolved around two apparently inconsistent demands. One such demand is to place functional explanation into the general scheme of scientific explanation. Norms are not proper objects of scientific inquiry, unless they are naturalized. To naturalize norms is to show that either norms are result of certain non-normative processes and hence are resultant properties which can be reduced to non-normative phenomena or they are completely irreducible aspects of the universe that show “brute emergence”. [112]

The second demand is to preserve the essential properties of biological functions. Finding a place for functions and norms in the realm of science is driven by two such demands. One solution might be to find a unifying definition for functions. Another solution might be to accept a pluralist stance and postpone judgment unless the concrete phenomenon is before one’s eyes. Another solution would be to refuse any place to functions in scientific practice. The last possible position, which I have defended so far, is to accept that functions or norms are as real as any illusion. Even if the physical basis of an illusion (e.g. a mirage) is as real as any other physical phenomena, our beliefs about them are wrong.

Another tension in functions debate concerns the nature of functional explanations. So far, I have discussed the current positions in this debate and classified them into historical/etiological accounts, synchronic/causal accounts and combinatorial accounts. Proponents of historical/etiological accounts claim that in functional explanations, wee seek an explanation for the existence of a functional entity by means its causal history. Proponents of synchronic/causal role accounts claim that function ascriptions explain the contribution of a component part to a general capacity of a containing system. Proponents of combinatorial accounts give credit to both accounts and try to combine their virtues either by unifying them under one primordial concept (e.g. Kitcher’s concept of “design”) or by simply holding a pluralist position (e.g Amundson and Lauder).

---

[112] Bedau, .M., “Downward Causation and the Autonomy of Weak Emergence”, Unpublished draft for *Principia*

The most important discussion topic between the positions above concerns the normativity of functions. Functional categories (e.g hearts, lighters, etc) bring with them some criteria of proper performance. Proponents of historical/etiological accounts claim that, only a historical analysis provide those criteria, especially in biology, because, normal conditions of proper performance of a trait are the historical conditions in which the trait has evolved.[113] Proponents of causal role accounts either deny the reality of biological norms (e.g. Paul Sheldon Davies) or add further constraints (e.g. hierarchy) upon containing systems to escape from referring to history.

A new approach, discussed in the 3rd Chapter, is an example of the latter response to the criticism against a-historical functional accounts. According to the proponents of this approach (e.g Mossio et. al., Bickhard), some generic properties of living systems (e.g. self-maintenance, autocatalysis) can provide the criteria to judge the proper performance of biological traits without referring to history. This is an important contribution, even if Millikan's (1989a) criticism against a-historical concepts of functions still holds. The criticism holds because the property of being alive or self-maintaining can not be defined unconditionally. The conditions in which organisms are alive or maintain themselves should still be identified in order to attribute functions to traits that contribute to those capacities. The contribution of this new approach is important because being alive or maintaining one self is an unconditional end in a means-ends reasoning.[114] The normativity of biological functions may

---

[113] Millikan, R. G.: 1989

[114] MacLaughlin, P.: 2009 in Krohs, U. and Kroes, P. p.98

emerge from a means-ends reasoning.[115] Ends determine which traits are appropriate as means, hence, they act as the criteria for judging whether some organismic entity is working properly. However, Millikan's criticism still holds because these criteria include *ceteris paribus* clauses (e.g. given the temperature is such and such, pressure is such and such, etc) which should be replaced with real/historical conditions if they are to be used to judge any concrete organism's current properties.

The intuition behind the insistence on a-historical analyzes of biological functions comes from our tendency to see organisms as complex machines and analyze them into simpler subsystems. The intuition proved useful in many areas including molecular cell biology, molecular developmental biology, physiology, etc. However, normativity of functions, and the need to explain malfunction forces us to transgress the temporal and spatial boundaries of organisms. The history of organisms shows us two things at the same time. There are generalizable patterns (e.g. modularity), common properties (e.g. the genetic code) and reinvention of similar functional structures in highly differentiated lineages (i.e. convergent evolution) in this history. There are also idiosyncratic paths followed by different linages, which is the source of the diversity of organismic forms. In this diversity, we still find similarities if we look at the organization of developmental regulatory genes (e.g. Hox genes). Thus, we can build abstract models to simulate generic properties of evolving populations (and organisms) and add historical details afterwards.

Reduction unifies diverse fields. By the help of abstract models (e.g. NK type fitness landscapes, catalytic task spaces), we can unify historical and

---

[115] *Ibid.*

synchronic aspects of the emergence of biological norms. Biological norms were considered to be an obstacle against reduction. However, even history – the realm of idiosynracy – might be hiding generalizable patterns, expected outcomes under her veil of diversity. One such source of generalizable patterns comes from simulations. The other comes from evolutionary developmental biology.

Evolutionary developmental biology is the search for the genetic and developmental mechanisms that create animal forms and the evolutionary origins of those mechanisms. The apparent hierarchy of organisms is a source for our attributing them functions and analyzing them accordingly. In fact, Cummins type functional analysis is suitable for only those systems demonstrating some level of hierarchy. The generation of hierarchy in complex animals is achieved by means of the developmental process. The spatial patterning of the embryo is a result of differing gene expression patterns in different parts of the embryo. Nonsynonymous changes in genes regulating development mostly have dramatic phenotypic effects (e.g. *antennapedia* mutation). The diversity of animal functional parts are proximately explained by differences in structural genes, and perhaps more importantly, by changes in regulatory gene networks. Causal organization of regulatory gene networks is one synchronic source of biological normativity.

Molecular cell biology and biochemistry provide intuitions about how fuzzy may the border between function and malfunction be. In biochemistry, structure is essential in explaining function. Macromolecules, the molecules of life, perform their functions usually by means of weak interactions. Weak interactions also explain the stability and flexibility of macromolecules.[116]

---

[116] Garrett, R. H. and Grisham, C. M.*:* 1999, p.10

These weak interactions, their positioning on the macromolecule, give us another synchronic source of biological normativity. In some cases, the criteria for proper performance may be strict. For instance, you change one aminoacid in the active site of an enzyme and you lose all functionality. In other cases, loss of function is gradual. In a sense, weak interactions determine the boundaries of life and death.[117]

The first and most important conclusion of my research is that biological norms are statistical regularities in normative disguise. One of the essential properties of a norm is the physical freedom to disobey it (i.e. its arbitrariness in terms of physical conditions). If some states are physically constrained so that their change is not likely, we may not call being in that state a norm of action. In the case of biological norms, the probability distribution of genotypes or phenotypes is not even. For instance, if the fitness landscape is rugged and not correlated (i.e. chaotic), the initial genotype and the shape of that population constrains where a population can move. Since the population will be trapped in a local optimum and it is very unlikely to escape, we may not call that state a norm. There is very little freedom to disobey it.

The second conclusion of my research concerns the sources of biological norms. I have defended so far that there are two sources of biological norms. One is the self-maintenance of living systems. Living systems adapt to changes in their environment. They include regulatory causal networks that keep the internal state of the system unless it is disturbed beyond some boundary conditions. They act on their own behalf (i.e. they are autonomous). This

---

[117] *Ibid.*

property of organisms have been omitted in favor of natural selection. I beleive that self-maintenance is a genuine source of biological normativity.

The other source of biological norms is natural selection. Natural selection is a mechanism that explains the adaptive complexity of organisms. The adaptive properties of organisms reflect the demands of their environment. They show symptoms of a lengthy design process, from the intricate metabolic pathways to mating displays.

None of these factors can be considered as mutually exclusive. Self maintenence is an essential property of living systems and there is yet no satisfactory theory that explains the origin of life in terms only of natural selection. If self-maintenance is a direct result of passing a certain complexity threshold (as stated by Kaufmann), then it should be considered as a genuine factor in explaining the origin of biological norms. In other words, the precondition of life is also the precondition of biological norms, and if it is not a product of natural selection and it is a conserved generic property of life, it should appear in our analysis of biological functions.[118]

---

[118] If we decide to include every precondition and every generic property of life in functional analysis, then we will have to include generic properties of matter such as mass, gravitational forces, the etc. This would be a dead end for sure. However, I hereby assume the conditions that enabled the emergence of life from non-living matter.

BIBLIOGRAPHY

Amundson, R., Lauder, G. V.: (1998), "Function without a Purpose: The Uses of Causal Role Function in Evolutionary Biology", in Hull, D. L and Ruse, M. (eds.), *The Philosophy of Biology* (Oxford: Oxford University Press,

Bechtel, W. and Richardson, R. C.: 1993, *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton: Princeton University Press

Bedau, .M.: (2003) "Downward Causation and the Autonomy of Weak Emergence", *Principia* 6: 5-50.

Bickhard, M.: (1993), 'Representational Content in Humans and Machines', *Journal of Experimental and Theoretical Artificial Intelligence,* 5, pp. 285–333.

Brandon, R.: (1999), "The Units of Selection Revisited: The Modules of Selection"*, Biology and Philosophy 1*4: 167–180,

Brigandt, I. and Love*,* A.*:* (2009)*, "*Reductionism in Biology" in *Stanford Encyclopedia of Philosophy,* Spring 2009 Edition

Buller, D.: (1998), "Etiological Theories of Function: A Geographical Survey", B*iology and Philosophy*: 13 , pp. 505-527

Buller, D.(Ed.): (1999), *Function, selection, and design,* Albany, N.Y.: State University of New York Press

Carroll, S. B., Grenier, J. K., Weatherbee, S. D.: (2005), *From DNA to diversity : molecular genetics and the evolution of animal design,* Malden, MA : Blackwell Pub.

Carroll, S. B.: (2006), *The Making of the Fittest: DNA and the ultimate forensic record of evolution,* NY: Norton

Churchland, P.M.: (1988), *Matter and Consciousness ,* Cambridge Mass: MIT Press

Churchland, P.S.: (2005), "A neurophilosophical slant on consciousness research" P*rogress in Brain Research* 149, pp. 285-293

Cornish-Bowden, A.: (2004), *The pursuit of perfection : aspects of biochemical evolution,* Oxford : Oxford University Press,

Cummins, R.: (1975), "Functional Analysis", *The Journal of Philosophy*, Vol. 72, No. 20. (Nov. 20,), pp. 741-765.

Davies, P. S.: (2003), *Norms of Nature: naturalism and the nature of functions,* Cambridge, Mass. : MIT Press

Dawkins, R.: (1998), "Universal Darwinism", in David. L. Hull and Michael Ruse (eds.), *The Philosophy of Biology* (Oxford: Oxford University Press,)

Dennett, D. C.: (1996), *Darwin's Dangerous Idea,* New York : Simon & Schuster,

Fodor, J.: (1983), *The Modularity of Mind,* Cambridge, Mass: MIT Press,

Futuyma, D.: (2005), *Evolution*, Sunderland, MA: Sinauer

Garrett, R. H. and Grisham, C. M.: (1999), *Biochemistry, 2nd Edition,* Florida: Saunders

Godfrey-Smith,P.: (1991, July), "What Darwinism Explains", Paper presented at the Darwin Festival, Cambridge University

Gould, S. J.: (1984, April-May) "Challenges to Neo-Darwinism and Their Meaning for a Revised View of Human Consciousness", Paper presented at Tanner Lectures on Human Values. Clare Hall, Cambridge University, pp. 55-73

Gould, S. J., Vrba, E.: (1982), "Exaptation – A Missing Term in the Science of Form", *Paleobiology*, Vol. 8, No. 1, pp. 4-15

Hull, D. L. and Ruse, M. (eds.): (1998), *The Philosophy of Biology,* Oxford: Oxford University Press

Kaufmann, S.: (1993), *The Origins of order : self-organization and selection in evolution*, New York : Oxford University Press,

Kaufmann, S.: (1995), A*t Home in the Universe: the search for laws of self-organization and complexity,* New York : Oxford University Press,

Kaufmann, S.: (2000) *Investigations*, Oxford ; New York : Oxford University Press

Kitcher, P.: (1998), "Function and Design" in Hull, D. L. and Ruse, M., *The Philosophy of Biology*, Oxford: Oxford University Press

Krohs, U. and Kroes, P. (eds): (2009), *Functions in Biological and Artificial Worlds*, Cambridge, Mass: MIT Press,

Lewens, T.: (2004), *Organisms and artifacts: design in nature and elsewhere,* Cambridge, Mass.: MIT Press

Macdonald, G. and Papineau, D. (eds.): (2006), *Teleosemantics: New Philosophical Essays,* Oxford ; New York : Clarendon Press

MacLaughlin, P.: (2009), "Functions and Norms" in Krohs, U. and Kroes, P. (eds), *Functions in Biological and Artificial Worlds*, Cambridge, Mass: MIT Press,

Millikan, R. G.: (1989), "In Defense of Proper Functions". *Philosophy of Science* 56: pp. 294-295

Millikan, R. G.: (1995), "Propensities, Exaptations and the Brain" in *White Queen Psychology and Other Essays for Alice,* Cambridge, Mass. : MIT Press

Millikan, R. G.: (1999), "Wings, Spoons, Pills, and Quills: A Pluralist Theory of Function", *The Journal of Philosophy*, Vol. 96, No. 4 , pp. 191-206

Nelson, D. L., Cox, M. M.: (2000), *Lehninger Principles of Biochemistry, 3rd edition*, NY: Worth Publishers

Pinker, S.: (1997), *How the Mind Works*, New York : Norton

Preston, B.: (1998), "Why is a Wing Like a Spoon? A Pluralist Theory of Function", *The Journal of Philosophy*, Vol. 95, No. 5. , pp. 215-254.

Rosenberg, A.: (2001), "Reductionism in a Historical Science", *Philosophy of Science*, Vol. 68, No. 2 , pp. 135-163

Sober, E.: (1993), *The Nature of Selection,* Chicago: University of Chicago Pres,

Sober, E.: (2000), *Philosohy of Biology*, Colorado: Westview Press

Sterelny, K. and Griffiths, P.: (1999), *Sex and Death: An Introduction to Philosophy of Biology* Chicago: University of Chicago Press

Wimsatt, W.: (1986), "Forms of Aggregativity", In Donagan,Perovich Jr.,Wedin (eds), H*uman Nature and Natural Knowledge*, pp.259—291.

Wimsatt, W.: Schank, J. C.: (1986), "Generative Entrenchment and Evolution" PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Volume Two: Symposia and Invited Papers (1986), pp. 33-60

Wood, P. J.: (2006), Understanding Immunology, 2nd Edition, Dorchester: Pearson Education Ltd.,

Wright, L.: (1973), "Functions", Philosophical Review, 82, pp. 139-168