EVOLUTIONARY BIOLOGY AND THE AXIOMATIC METHOD REVISITED

Andreea EŞANU*

Abstract. In evolutionary biology there has never been reached a consensus regarding the manner of presenting its theory. According to the skeptical view, evolutionary biology cannot even be presented as a proper scientific theory because of its lack of uncontroversial laws. The skepticism, however, is not shared by the logicians of science, who claim that biological explanations can be molded into proper theories by means of logical reconstruction. In this paper I will address three developments of the axiomatic method in evolutionary biology: the hypothetical-deductive, the semantic and, perhaps the most recent one, the natural deduction method. I will point out that the key concept in figuring out the logical structure of evolutionary theory concerns deductive consequence. Then I will argue that a syntactic approach to deduction in the sense of the third method might be the most promising option. Eventually, I will claim this approach does half justice to the skeptical view that there are no uncontroversial laws in evolutionary biology.

Keywords: evolutionary biology, axiomatic method, syntactic approach.

In evolutionary biology there has never been reached a consensus regarding the manner of presenting its theory. According to the skeptical view, evolutionary biology cannot even be presented as a proper scientific theory because of its lack of uncontroversial laws. The skepticism, however, is not shared by the logicians of science, who claim that biological explanations can be molded into proper theories by means of logical reconstruction, *i.e.* by using the methods of formal logic. The logical properties of biological language might tell the important difference between empirical observations and theoretical laws in a way that would not require anything else but an adequate understanding of the biological language itself. This approach is a classical formal approach, and it draws from the works of Alfred Tarski's and Rudolf Carnap's on the axiomatization of natural science.

^{*} Andreea EŞANU, Postdoctoral Fellow at University of Bucharest, the Department of Theoretical Philosophy. Email: aesanu2@gmail.com.



In the following, I will address three developments of the axiomatic method in evolutionary biology: the hypothetical-deductive, the semantic and, perhaps the most recent one, the natural deduction method. I will point out that the key concept in figuring out the logical structure of evolutionary theory concerns deductive consequence. Then I will argue that a syntactic approach to deduction in the sense of the third method might be the most promising option. Eventually, I will claim this approach does half justice to the skeptical view that there are no uncontroversial laws in evolutionary biology.

I. THE PROBLEM OF DEDUCTION IN EVOLUTIONARY BIOLOGY

Let us suppose that we start to investigate the logical properties of natural language. By doing this, we easily conclude that certain combinations of words express tautologies, like "It rains or it doesn't rain" - even though such statements look similar to all the factual statements that we could express in natural language. Take, for instance, "It rains but it doesn't pour". What the logical analysis reveals is that, in spite of the superficial resemblance, the two sentences in our example have different logical structures and different logical properties. In the same manner, it is thought that if we focus on the biological language in which biological explanations are cast, we might also come to the realization that some statements have different logical properties than usual factual statements, even though they would not be tautologies. In the tradition of approaching scientific theories from a logical point of view, it is often believed that, once discovered, the logical properties of the biological language could tell the important difference between factual observations and so-called theoretical laws, in such a way that isolating the laws of biological science could become entirely a logician's business.

But, in spite of its straightforwardness, the logical approach to biology is faced with a load of difficulties. I shall discuss here only the case of evolutionary biology, even though similar considerations would probably hold for other fields in natural science. It seems that no logical method of reconstructing evolutionary biology, although sided with pertinent arguments from different philosophical perspectives, can resolve something that might be called the "problem of deduction". In formal evolutionary theory, it seems that the concept of deductive consequence works in an either ambiguous or scarce manner, and therefore the distinction between factual observations and theoretical laws turns out to be equally ambiguous or scarce. In this section, I will consider two classical approaches to formalizing evolutionary biology: the hypothetical-deductive approach and the semantic approach. My point will be that the hypothetical-deductive approach rests on an ambiguous concept of deductive consequence, while the semantic approach relies on a scarce one.

In classical logic, a sentence like: (**P**) "It rains or it doesn't rain" could be deduced from a finite set of axioms, unlike "It rains but it doesn't poor"

which could not. The deduction from the set of axioms \mathbf{S} would validate the formula: (**P'**) "p V ~p", whose instance is **P**, so that the confirmation of **P** would rely exclusively on a logical procedure. This aspect alone tells the difference between the logical properties of "It rains or it doesn't rain" and "It rains but it doesn't pour". In the standard hypothetical-deductive representation of scientific theories we find something almost similar. A biological theory, for instance, has its own propositions P'. These are called theoretical hypotheses. One such example is Fisher's law: (\mathbf{P}') "Given any species with sexual reproduction, the sex ratio inside that species is approximately 1:1". The validation of **P'** is very problematic since it is impossible to check for each individual of each and every species populating the biosphere; and there are cases in which the ratio is, in fact, not Fisher's ratio (*i.e.* there are anomalous distributions across populations). So, validating P' is not a question of blunt empirical observation, just as it is not a question of empirical observation to validate "p V ~p". This indicates that theoretical hypotheses in biology play a different role and possess different logical properties (e.g., validation properties) than empirical statements, even though they belong to the same biological language. In the hypothetical-deductive representation of scientific theories, the P' sentences are defined as *deductive steps* in a logical system equivalent to the theory in question (for Fisher's law, that would be a theory of natural selection). As with "p V ~p", Fisher's law can be consistently deduced from a finite set of axioms, and consistently entail factual statements like: (P) "In the population of Hawaiian finches, the sex ratio is approximately 1:1".

To put it more formally: given **T** a biological theory and given Δ a set of propositions in **T**, then all the propositions in Δ are ordered by the *entailment relation* '=>'. In this sense, a biological theory is a logical system.

In the validation process of **T**, two important steps occur:

- (i) concerns how the subset P' in △ (the subset of theoretical hypotheses) is constructed *i.e.* the axioms can be chosen freely and the set of axioms and their consequences needs to be *consistent*;
- (ii) concerns the manner in which the subset **P** in Δ (the subset of factual consequences) are confirmed *i.e.* for the reference class of **T** to be non-empty, the factual consequences **P** in Δ must be *equivalent* to correct empirical observations.

But if this is the hypothetical-deductive frame of evolutionary theory, then the entailment relation is at least ambiguous. In evolutionary biology, incompatible theoretical hypotheses can be consistently generated from the same set of axioms (*e.g.*, the standard Williams set¹). Moreover, incompatible theoretical hypotheses can to do justice to the same known facts from a given



¹ See Williams 1970.

reference class. For example, it might be the case that both a Lamarckian and a Darwinian hypothesis regarding natural of selection² could be formulated on the standard Williams axiomatization of evolutionary biology, in order to explain the observed sex-ratio or the beak size in a population of Hawaiian finches. Another bizarre situation arising from this ambiguity (iwhich is, in the end, the "right" evolutionary theory?) concerns Fisher's law - the one law I introduced earlier in full confidence. The interesting thing is that Fisher's law does follow from evolutionary theory axioms, but on the assumption that natural selection is Darwinian (i.e. frequency-dependent). This assumption, however, hangs in the air and the consequence for the logical approach to evolutionary biology is bleak. Even if we may develop successful logical tools in order to detach theoretical laws from factual observations in the language of a biological theory T, we are still facing the following difficulty: how do we know that what the logical system reveals by the logical ordering of T are proper laws of nature and not mere theoretical possibilities that the biological language simply allows for? It seems that we cannot decide whether Fisher's law is a law, unless we figure what natural selection is really about.

A considerable number of logicians of science (Suppes 1967; Beatty 1980; van Fraassen 1987; Thompson 1988; Lloyd 1994) argued that the most pertinent way to avoid the problem of deductive ambiguity is to eliminate entailment altogether from the logical systems that reconstruct biological theories. Thus, it was suggested to replace the entailment relation '=>' with the *satisfiability relation* ' |=', from standard Model Theory. In this sense, a biological theory is to be built not as a deductive system, but as a class of abstract models in the mathematical sense of model theory, *i.e.* a class of models satisfying a given mathematical structure (or a Suppes structure³).

Formally: Given **T** a biological theory, and *S* a Suppes structure in set theory, it is the case that $S \models T$. In this sense, a biological theory is said to be a set-theoretic entity.

This general view is called the semantic view of scientific theories. According to the semantic view, biological theories are not stating claims about real biological entities (*e.g.*, species, populations, organisms etc.), but only about models, although models are used to represent real biological entities (Horan, 1988). Also, the deductive structure is replaced with something *prima facie* non-inferential, *i.e.* representation. In this manner the controversial concept of "biological law" is simply eliminated, along with the distinc-

² By the Lamarckian hypothesis I mean the idea that selection is directional (Reece *et. al.* 2002) or environment-dependent. By the Darwinian hypothesis I mean the idea that selection is frequency-dependent (which means that the environment has only an indirect effect on how individuals from a given population get selected), and therefore non-directional.

³ See Suppes 1967.

tion between "theoretical" and "empirical" sentences, given that the theory ceases to be a linguistic entity altogether.

For instance, Beatty (1981) and Thompson (1989) show that the mathematical structure of evolutionary theory (ET) is captured by the Suppes structure: $\beta = \langle P, A, f, g \rangle$, where β defines an abstract Mendelian breeding system.

In the set-theoretic approach, the evolutionary theory (ET) is validated if:

- (i) β is *isomorphic* to real biological systems (like the population of finches in the Hawaiian Islands);
- (ii) The predictions obtained in the abstract system *match* empirical distributions in the real populations.

In the set-theoretic formulation of ET, $\beta = \langle P, A, f, g \rangle$ could be freely interpreted as having a Darwinian or a non-Darwinian selection function⁴ in a consistent manner with the axioms of ET (the Williams set), and preserving the isomorphism required in (i). In the hypothetical-deductive framework the resulting equivalence was a problem, but in the semantic view, equivalence coupled with a general antirealist understanding of natural science is something intrinsic. In the semantic view's terms, ET is not a deductive system at all, but a class of models; it is not committed to anything other than set-theoretic entities, *i.e.* abstract sets of equivalents.

Nevertheless, if we are not already full-blooded antirealists with respect to natural science, it is hard to see how there could be a theory of evolution with no concept of evolution whatsoever. For our general understanding of the process of biological evolution on Earth, there is a huge difference between, let's say, a Darwinian and a non-Darwinian model of ET. For instance, non-Darwinian theories (e.g. neo-Lamarckist theories) hold that evolution might be intrinsically cumulative⁵, *although* phenotypic variation is not directly inheritable, but it is inherited through a particular mechanism called cellular development (Goodwin 2001; Kirschner & Gerhart 2006); Darwinians (e.g. the Modern-Synthesis tradition), instead, claim that evolution is not intrinsically cumulative, given that phenotypic variation is not directly inheritable, and ignore cellular development theories. The two positions cannot be both correct, so it is difficult to concede that evolutionary biologists could work with ET using only isomorphism and equivalence classes and, also, do it in a meaningful manner. Maybe this is where lies the main reason why partisans of the semantic approach to evolutionary biology insist on stating that the abstract set-theoretic entities behind their reconstruction have "implicit" or "intended" interpretations.

⁵ Lamarck believed that, since selection is directional, evolution exhibits a trend towards more complicated and fine-tuned adaptations in the living world. His belief was reformulated in more recent terms as "the arrow of complexity" hypothesis. (Bedau 2009).



⁴ In other words, a frequency-dependent or a non- frequency-dependent function.

Andreea Eşanu

Formally: It is said that $\beta \models ET$ if and only if β and the axioms of ET *implicitly* characterize Darwinian evolution.

But this is at least bizarre. In what sense do the set-theoretic structure β and the axioms of **ET** *implicitly* characterize a Darwinian process? This appeal to implicit features of biological theories is very problematic, although it probably draws from the tradition of using implicit definitions in the axiomatization of science. The highly abstract models of **ET**, *i.e.* the Mendelian breeding systems, are considered to "relate to phenomena through a complex hierarchy of theories" (Thompson, 1989) that constrain the interpretation of **ET** as Darwinian evolution. One of the most important links in this hierarchy of theories is the theory of natural selection (**NT**).

Formally: $\beta \models ET$ if and only if $\beta \models NT$, so that β is a set-theoretic model of **NT**.

In plain words, what the above sentence says is that the evolutionary theory (ET) is true for Mendelian breeding systems (β) if only if the theory of natural selection (NT) is also true for Mendelian breeding systems. Assuming that a population of Hawaiian finches is isomorphic to a Mendelian breeding system, then the Hawaiian finches' beak mutations, are *lawfully* correlated, through NT, to the selective advantages provided by certain finch genotypes over other genotypes, and called "evolutionary adaptations". The same holds for the sex ratio estimation: the Hawaiian finches' 1:1 sex ratio is lawfully correlated to the lack of selective advantages provided by the male genotypes over the female genotypes (or vice versa) in the reproductive strategy of the species. But why does the insertion of NT make ET implicitly Darwinian? The answer is straightforward: Mendelian systems are defined by means of genetic inheritance only and, given that natural selection operates on organisms isomorphic to Mendelian breeding systems, only genetic inheritance is correlated with evolution. This is exactly what Modern-Synthesis is about.

The existence of implicit interpretations of **ET** seems to deal with the problem of theoretical ambiguity in evolutionary biology elegantly, but it also opens the door to another major problem. As philosopher Mario Bunge points out (2005, 350), if theories are definitional equivalents of abstract systems, then all of them, irrespective of their position in the hierarchy, are abstract and so none of them can have any relation whatsoever to real biological or systems other than by stipulation. If we look again, we may easily see that the manner in which the intended interpretation of evolutionary theory (**ET**) is chosen is arbitrary, *i.e.* it simply happens that we all know that Lamarckism is false and we rely on the other convenient interpretation. But, actually, nothing in the axioms of **ET** forces us to do so: the isomorphism (*e.g.*, between Mendelian breeding systems and the population of finches in



Hawaii) still obtains with a different mechanism of natural selection (*e.g.*, the cellular development mechanism, which does not contradict the fundamental claim that living organisms could be represented as abstract Mendelian breeding systems; it only shows that something else should be considered). So, in this case, evolutionary theory is no longer ambiguous, but it is obviously scarce. It does not capture all that might be relevant in its representation of evolution.

The canonical solution (Thomson 1989, 82) to the scarcity problem is to say that the intended interpretation of **ET** is not, in fact, arbitrary because it captures all that is needed to be captured in order to represent evolution. It is claimed that the abstract Mendelian breeding systems are causally similar to the real biological systems. This means that, if a certain selection mechanism (*i.e.* Darwinian selection) is causally responsible for the sexual distribution, or for the distribution of beak phenotypes, in the population of Hawaiian finches, the same selection mechanism is caught as such in the abstract model as well. Possible alternative interpretations of **ET** are eliminated on the basis of actual causal constraints that affect the abstract Mendelian systems themselves.

But this solution looks very confusing at least for two reasons: first of all, it appears to be circular; secondly, it seems to rely on a category mistake. I will insist a little bit on the second issue because it throws a significant light on the scarcity problem. Let us think at the following situation, often found in forensic psychological profiling. A psychological profile is considered something like a model of someone's behavior: it is based on previous behavioral observations, some general personality traits etc., and it is used to predict future behavior. But a psychological profile could be done in at least two distinct ways: a) it could be designed to predict future behavior using only correlations between past behavior and some general personality traits; or b) it could be designed to predict future behavior by emphasizing, beside what we find in (a), certain mental "dispositions" that causally influence action (like the disposition to avoid pain). The second model relies on mental causation to provide a richer explanation for someone's habitual behavior than the explanation provided in the first model. But mental causation assumes something very problematic. It assumes that even if mental dispositions are not physical objects, when it comes, for instance, to move my hand from a hot spot in the room both my disposition to avoid pain and a physical chain-reaction chemically controlled by my brain could play the same role, *i.e.* they both could be the cause of my hand's displacement and, in this sense, dispositions and physical objects are causally similar. But this way of thinking relies on an obvious category mistake. It explains physical effects by invoking non-physical causes and the explanation is simply incoherent. The second model of behavior displays this kind of incoherence, even though it looks richer than the first model.



In a similar manner, Thomson's solution to the scarcity problem in **ET** relies on a category mistake that makes the theory look richer, but at the same time incoherent. To hold that abstract Mendelian breeding systems have causal properties and embed selection mechanisms into set-theoretic entities is oddly confusing, given that no mathematical structure could ever display the same causal properties as physical systems.

II. ZSYNTAX

In spite of the wide spreading of semantic antirealism in the philosophy of science during the last decades, we can still hold that the key concept in order to understand the logical structure of evolutionary biology is the concept of deductive consequence. On one hand, deductive structure is fundamental in natural science, because it provides a simple formal instrument to organize scientific knowledge in such a way as to ease scientific progress. In "Biology and the Axiomatic Method" (1962), J.H. Woodger was making the following appraisal:

"the aim (...) is to provide an exact and perfectly *controllable* language by means of which biological knowledge may be *ordered*. In a former book⁶ (...) it was shown how largely these disputes were traceable either to failure to eliminate metaphysical elements from biological topics, or to difficulties created by the shortcomings of current biological language. Both of these sources of confusion can be avoided by paying attention to the requirements of an ideal scientific language. Because if we have a language we need not dispute, we need only calculate and experiment." (Woodger 1962, 11).

On the other hand, the semantic view is faced with deep conceptual inconsistencies, coming especially from the attempt to conflate the deductive structure of evolutionary biology into a causal chain of bizarre set-theoretic entities. Such difficulties may not be overcome unless the focus is changed again from this rich ontology to the modest address concerning the epistemology of evolutionary biology: *i.e.* the purpose of a logical system is not to represent (as in instantiate) a scientific theory, but just to order it.

For the rest of the paper, I will suggest a returning approach to the concept of deductive consequence. I hope that in the long run this suggestion will offer a fresh perspective on the axiomatization of evolutionary biology and, also, an insight into how to understand theoretical unification in this controversial field. I will argue that a "natural" concept of deductive consequence is a sound option in order to formalize evolutionary biology and, still, follow in the footsteps of empiricism. My argument stems from a recent development of natural deduction systems in molecular biology, *i.e.* the

⁶ In The Axiomatic Method in Biology (1937).

Zsyntax. But, before going into that, I will make a brief point about theoretical unification in evolutionary biology. As J.H. Woodger pointed out, the concept of deductive consequence has an important bearing on the matter of unification, even though this might not come straightforward.

For instance, the standard Mary Williams' (1970) axiomatization of evolutionary biology was found lacking by many theoreticians because it did not rely on a properly formalized concept of heredity. Hence the deductive fuzziness of her logical system: what does the natural selection *really* apply to? how mutations spread in the Darwinian sub-clans? what really are the Darwinian sub-clans? etc. The rival axiomatization belonging to Zanardo and Rizzoti (1986a, 1986b) – in fact, an improved version of Woodger's (1937) in the genetics of populations – managed to account for heredity in a formally consistent and clear manner. But it did it reductively, so Zanardo and Rizzoti's system was also found lacking. Due to the restrictive definition of heredity (*i.e.* as genetic heredity), this axiomatization holds only for a limited number of evolutionary processes, and it concerns only eukaryotic organisms. Therefore, it fails to provide theoretical unification. This issue makes, in fact, the come-back point for Mary Williams. As she reveals, ambiguity might actually serve unification in evolutionary biology:

"The existence of several different, non-trivial, interpretations of the axiom system greatly increases the power of the theory; it enables the transformation of every creative insight gained from the study of one level of selection into formally analogous insights on the other levels. These insights, though formally analogous, will not usually be obviously analogous." (Williams 1970, p. 349)

Zanardo and Rizzoti's axiomatization was reductionist because it focused exclusively on the frequencies of genes and genotypes as if they were the self-existing, ultimate entities of evolutionary biology (Bunge, 2005, p. 344). By contrast, the Williams axiomatization was receptive to the import of extensive new biological knowledge coming from modern developmental biology or from ecology. The Williams axiomatization is compatible with both Darwinist (the Modern Synthesis) and non-Darwinist variations of ET (neo-Lamarckist theories). This is the case precisely because Marry Williams' focus is on the concept of natural selection leaving heredity aside. This move allows for a lucrative approach to issues like multiple levels of selection, favoring unification in evolutionary biology. The only major problem is that the resulting logical system ends up in deductive inconsistency, as I have pointed out already (section 1). Nothing in the Williams axiomatization tells us that the non-trivial interpretations should be seen as analogous, and not as concomitant. The "multiple levels of selection" trick is just a supposition external to the set of axioms itself.

A way to address the fuzziness of Mary Williams' axiomatization would be to make the supposition concerning "multiple levels of selection" internal



to **ET**, while keeping the deductive structure intact. To this end, *Zsyntax* from molecular biology may provide an useful insight.

Zsyntax is a non-standard proof calculus that elaborates in Rudolf Carnap's concept of "languages as calculi" (Carnap 2001, p. 4). We could imagine that the language of evolutionary biology is something like the language of chess. The language of chess is systematic because it involves a calculus: *i.e.*, any finite series of symbols/positions is called an expression of that calculus/game if and only if it is the result of a given set of rules for that calculus/game. There are two types or rules in an abstract calculus:

- (a) the formation rules: *i.e.* the conditions under which an expression/position can be said to belong to a certain category of expressions/positions in the calculus (Carnap, 2001, p. 4);
- (b) the transformation rules: *i.e.* the conditions under which the transformation of one or more expressions/positions into another or others may be allowed in the calculus (Carnap 2001, p. 4).

The first kind of rules are also called "syntactical rules", while the second kind of rules are also called "logical laws of deduction". She second kind of rules need not be axioms; they could simply be inference rules – *i.e.* inferring a future position in chess is based on applying transformation rules to the current position. In logic and proof theory, *natural deduction* is a kind of proof calculus in which deduction is expressed by such inference rules. These rules are closely related to the "natural" way of reasoning which works rather with inference rules than with abstract sets of axioms.

The universal operation in natural deduction systems is the *reasoning from assumptions*. For example, consider the following derivation:

(1) A&(B&C)
(2) B&C 1,&E
(3) B 2,&E

This derivation does not establish that *B* is the case as such; rather, it establishes the following fact: if A & (B & C) is the case then *B* is also the case. This is a *hypothetical derivation*, which could be written formally as follows:

 $A\&(B\&C) \mid -_{Nded} B$

This is the chess-game kind of reasoning A future position in the game n is derived from the well-formed initial position, given a finite set of transformations (or "moves"), so that the future position k is the case given that the initial position i is the case.

The general form of a hypothetical derivation would be:



(i) $P_{1'} P_{2'} \dots P_n$

(k) Q

In short, each hypothetical derivation has a collection of antecedent derivations (the P_{y}) and a succedent (Q).

Now, the notion of hypothetical derivation could be *internalized* in the transformation rules of the language, and what we get is a very interesting situation in the calculus. Let's take a look at the rule for the introduction of logical implication:

(i) P a(k) Q (m) P -> Q, i,k ->I

The antecedent *a* named "assumption" is discharged in the conclusion. In fact, discharging assumptions is a mechanism for delimiting the *scope* of the hypothesis: the sole reason for assuming *P* is to establish *Q*; *P* cannot be used in any other way and it cannot be used below the introduction rule.

Let' say that we want to prove the following first-order validity: "A -> (B V ~B) is the case". The proof would be:

(1) A	а
(2) B V~B	the excluded-middle
(3) A -> (B V ~B)	1,2 ->I

Let's move now to *Zsyntax*. In molecular biology biological processes are defined in terms of their participant molecules⁷. The language of molecular biology is made of molecules and chemical reactions, just as the language of chess is made of positions and moves.

The formation rules for *Zsyntax* are:

(i) If two types of molecules, *A* and *B*, are able to interact in some way, the outcome is denoted as the complex molecule *A***B*, where *"**" indicates the operation called *Z*-interaction.

⁷ "These formulae are assembled into chains, in accordance with rigorous logical rules in order to represent chains of biological reactions, so that the latter are treated as logical deductions." (Boniolo *et al.* 2010, p. 2).



- (ii) If *n* molecules of types *A1*,..., *An*, interact the aggregate is denoted by introducing the operator called the *Z*-conjunction: *A1* & ... & *An*.
- (iii)Let us consider an initial aggregate of molecules A & C and a final aggregate B. In this case, it is said that all aggregates of type C are also of type A -> B. This means there is a transition from A to B if there is C allowing it. ,,->" is the Z-conditional.

The transformation rules of *Zsyntax* are:

- (i) Elimination of the *Z*-conditional (->E): If *A* -> *B* can be derived from *C* and *A* can be derived from *D*, then *B* can be derived from *C* & *D*.
- (ii) Introduction of the *Z*-conditional (->I): If *B* can be derived from *C* & *A*, then $A \rightarrow B$ can be derived from *C* alone. In the logical language one says that the assumption *A* is discharged, "which in our context means that the availability of an aggregate of type *A* is incorporated as a (sufficient) condition in the antecedent of $A \rightarrow B$. Hence, the derivability of (an aggregate of type) $A \rightarrow B$ no longer depends on the availability of (an aggregate of type) A."⁸
- (iii) The elimination and introduction rules for the *Z*-conjunction (I will not detail them here).

In order to understand how *Zsyntax* works, I shall give a brief example of logical validity in *Zsyntax*. Similar to the hypothetical derivation discussed earlier, in *Zsyntax* it is claimed that "*A* -> *B* is the case" or $|-_{z} A -> B$ if the empty aggregate (namely the aggregate consisting of zero molecules), denoted by Ø, allows the path from *A* to *B*; that is, if Ø & *A* -> *B*. Consider the case of the tumor suppressor TP53⁹. The protein TP53 can bind the MDM2 gene to activate its transcription, so that the MDM2 protein is produced in the cell. In this case there is an intermediate product, that is, MDM2*TP53. By definition of the *Z*-conditional, the "empty aggregate", Ø, must be of type MDM2&TP53->MDM2, since from any aggregate of type Ø&MDM2&TP53 we can arrive at some aggregate of type MDM2. Under these circumstances we can say that "MDM2&TP53 -> MDM2 is the case".

((1) MDM2&TP53 -> MDM2	Ø
((2) MDM2&TP53	а
((3) ((MDM2&TP53 -> MDM2) & (MDM2&TP53)) ->MDM2	1,2 (f)Z-conditional
((4) MDM2&TP53 -> MDM2	3, I->

That is: Ø & (MDM2&TP53) $|_{z}$ MDM2, which says that the validity of MDM2&TP53 -> MDM2 depends on the "null" assumption or on no assumption at all. In other words: $|_{z}$ MDM2&TP53-> MDM2.

⁹ *Ibid.,* p. 3.



⁸ Boniolo *et al.* 2010, p. 4.

Thus, the form of the deduction theorem for molecular biology could be written as follows:

IF $\emptyset \& P$ entail in *Z*-calculus *Q*, then $|_{-_Z}P \rightarrow Q$.

The "laws" of molecular biology are generated by applying the above formula and detaching the "null" assumption from chains of molecular reactions.

Before closing this illustration, it is important to stress something problematic about *Zsyntax*. At the first glance, Boniolo *et al.*(2010) claim that logical deductions in the *Z*-calculus are analogous to the biochemical processes they describe, in the sense that the mechanisms behind chemical processes are equivalent in the calculus to "lawful" hypothetical derivations. Of course, biochemical mechanisms are not deductions, since biochemical processes are concrete entities, while deductions are logical entities. In this respect, *Zsyntax* is as heuristic as the empiricist systems of Rudolf Carnap's or J.H. Woodger's: it allows to structure prediction problems pretty much like in the standard hypothetical-deductive frame, by setting a demarcation line between theoretical hypotheses (mechanisms) and empirical observations (individual processes). However, there is a stronger claim made by the authors, according to which *Zsyntax* delivers nothing less than the "constitutive logic" of molecular biology itself:

> "Logic is not just an auxiliary tool for analyzing biological models based on some external formalism, but becomes the core of a research program in which biological processes are the intended semantic interpretation of a non-classical logical system" (Boniolo *et al.* 2010, p. 2).

Even if there are no axioms in Zsyntax, the isomorphism issue enters the picture again, in a confusing manner. It is hardly coherent to hold that the interpretation of the deduction theorem in molecular biology is the causal molecular mechanism *per se* – or, in other words, that logical relations fully capture causal molecular agencies. A question in the epistemology of science is forcedly translated, again, into a question of ontology, and this fact alone generates incoherence.

I chose to mention it, however, because I believe that such confusions should be avoided in the logical reconstruction of evolutionary biology.

III. THE EVOLUTIONARY THEORY AS CALCULUS

In the following, I will introduce **ET** as a calculus in the sense of section **2**. Due to the reduced length of this paper, I will give only a sketch of the calculus, leaving other details to future presentations. Thinking about **ET** as a calculus involves writing down a deduction theorem. In this section I will



show what the deduction theorem in **ET** might look like if we take **ET** to be a logical calculus. Then I will argue that this formalization of deduction opens the possibility to express important parts of evolutionary biology, parts that previous axiomatizations failed to express.

The computational approach to evolutionary biology (Bedau 2009; Hunemam 2012) will be very insightful in this respect. I will not argue here why this is the case. My main intention fort this section is to illustrate how the overall idea of reasoning from assumptions might work in evolutionary biology (since it works in molecular biology), if we take for granted that evolutionary theory could be written down as a computational theory. The discussion of why evolutionary biology could be reconstructed as a computational theory goes beyond the scope of this paper and, probably, will be carried on somewhere else. However, the basic consequences of writing down a scientific theory as a computational theory are very significant (Humphreys 2004; Winsberg 2009) if we are interested in articulating the logical structure of evolutionary theory. This will be the main topic of the present section.

The **ET** could be expressed as chains of CA-states (cellular automata states) in the so-called "Game of Life" (**GL**). Thus, the language of **ET** would consist of CA-states and transition rules in **GL**. The Game of Life is a computational device that generates evolutionary patterns in artificial populations of cells. In this logic, the demarcation in **ET** between theoretical hypotheses and factual observations will be drawn as a distinction between evolutionary patterns and singular histories in **GL**.

The formation rules for **GL**:

A cellular automaton such as **GL** consists of a grid of *cells*, each in one of the two states *on* and *off*. The grid is usually two-dimensional. For each cell, a set of cells called its *neighborhood* (including the cell itself) is defined relative to the specified cell. An initial state (time *t*=0) is selected by assigning a state for each cell. A new *generation* is created (advancing *t* by 1), according to some fixed *rule* (usually a mathematical function, but not necessarily) that determines the new state of each cell in terms of the current state of the cell and the states of the cells in its neighborhood. The rule to update the state of cells is the same for each cell and does not change over time, and is applied to the whole grid simultaneously. One interacts with **GL** by creating the initial configuration of cells (*i.e.* units of selection) and observing how the configuration evolves in time, once the transformation rules are applied.

The transformation rules for GL:

- i) Any live cell with fewer than two live-neighbors dies, as if caused by under-population.
- ii) Any live cell with two or three live-neighbors lives on to the next generation.
- iii) Any live cell with more than three live-neighbors dies, as if by overcrowding.



iv) Any dead cell with exactly three live-neighbors becomes a live cell, as if by reproduction.

These rules are designed as to embed a process of natural selection. This differentiates **GL** from other kinds of cellular automata.





What is most interesting about the computational approach to evolutionary biology is the flexibility that **GL** allows in interpreting the fundamental concept of *unit of selection*. For instance, the **GL** cells could be built as universal von Neumann constructors (von Neumann, 1951) and endowed with genetic make-up so that they transfer genes to future generations. In this case the main source of variation in the CA-setting is, as in the standard picture of the Modern Synthesis, genetic variance. But, as in the traditional Darwinian picture, the unit of selection is the organism (*i.e.* a cell or a functional group of cells), since selection is understood simply in terms of cells' survival (there is no phenotypic characterization of fitness).

In Fig.1 something particularly interesting happens: the local rules of evolution in **GL** generate multi-cellular groups of cells (*i.e.* the glider-guns), changing the level of selection in the game.

Let us look now at Fig.2. What we see in Fig.2 might be simply called digital parasitism: the glider reproduces as an organism, given the fulfillment of certain conditions in its evolutionary niche. Its survival depends upon the existence of another "species", *i.e.*, the backrakes. In coevolution, the level of selection changes again, from groups of cells to groups of organisms (or even species).





(2) Reproduction of the Mitchell parasite in the presence of two backrakes.

Source: http://pentadecathlon.com/lifeNews/2011/01/sprouts_and_parasites. html

What is fundamental about cellular automata like **GL** is the fact that such evolutionary patters show *robustness*. A robust **GL** simulation is, thus, a proof. Therefore, we can claim that **GL** is a calculus. For instance, the CA-game in Fig.2 displays evolutionary emergence (*i.e.* the parasitism phenomenon), which is actually robust in the simulation.

If we go back to section **2**, we see that the standard formal approach to **ET** deals with the problem of different levels of selection by treating the concept of "unit of selection" ambiguously and saying that the corresponding evolutionary processes are analogous. The CA approach says something different. The evolutionary processes involving different levels of organization and selection are not analogous, but in fact emergent. So, there should be nothing ambiguous about **ET** as long as the concept of "evolutionary emergence" is properly formalized.

The CA reconstruction of **ET** allows for a consistent formalization of evolutionary emergence as computational incompressibility, so that no external supposition to the calculus itself needs to be made.

Let us take an incompressible instance of **GL** (like the one in Fig.1). The idea of computational incompressibility in the glider gun simulation is that, while there is a deterministic rule so that one can go from a set of cells C<(i – j, i + j), n) to state of cell C<(i),n+1>, there is no shortcut from **C**<(1,**M**),**n**> to **C**<(1,**M**),**n**+1> (M being the number of cells)¹⁰, but we get there *only* by running the game (*i.e.* there is no other mathematical function, equation or rule).

¹⁰ Huneman, 2010, pp. 197-198.

1. (n) C(i-j, i+j) (n+1) C(i) (n) and one of the rules (i)-(iv).

But:

2.		
(n)	C(1,M)	(the glider in step n)
(n+1)C (1,M)	(the glider in step n+1)	(n) and the GL incompressible
		simulation

The glider transition displays computational incompressibility. The same thing holds for the coevolution of backrakes and Mitchell parasites in Fig.2.

But, although incompressible, this kind of CA-process could be described in terms of *counterfactual dependencies* between sets of CA-states at various CA-steps:

> "A glider can be characterized as a set of states of cells; gliders have typical behaviors, so that if at step n the glider is at one point, analysis of the cellular automaton may entail that it will be somewhere else a while after (after k steps, let's say). One can then claim that, on the background of the rules of the automaton, the glider would not be where it is at step n + k if it had not been where it was at step n. This is a counterfactual dependency, and it illustrates the fact that even in cellular automata showing incompressibility, regularities of this type can be produced." (Huneman 2010, 198)

Thus:

(n) C(1,M) (n+k) C (1,M) (the glider in step **n**) (the glider in step **n+k**) (n) and the **GL** counterfactual dependence rule

So that: $|-_{GL} C(1,M)_n \rightarrow C(1,M)_{n+1}$

Any lawful correlations in the CA-game could be expressed as counterfactual dependencies between CA-states, when emergent levels of selection are concerned.

The incompressibility problem sheds a significant contrast on Mary Williams' standard axiomatization of **ET**. In the Williams axiomatization we say that cell selection, organism selection, species selection etc. are all models of the axiomatic system that constitutes **ET**, so **ET** is deductively ambiguous, but non-reductionist. By contrast, what the computational incompressibility of **GL** shows is that we might have different deductive structures in **ET**, depending on how emergence between the levels of selection is considered. These several deductive structures of **ET** are reflected in the form of the deduction theorem.



If we go back to the previous section, we see that $|-_{GL} C(1,M)_n \rightarrow C(1,M)_{n+1}$ is, in fact, equivalent to: $\emptyset \& C(1,M)_n |-_{GL} C(1,M)_{n+1}$.

If we write down the deduction theorem of **ET** for computational incompressibility, we obtain: (i) IF $(\emptyset \& T) \mid -P$, then $\mid -T \rightarrow P$

But not all the kinds of evolutionary emergence can be defined as computational incompressibility *simpliciter*. (i) holds for *P* only when *P* refers to the class of complex adaptations called ", bounded", for which there is at least an incompressible finite CA-description.

There are at least two more classes of evolutionary complex processes, for which the expression in the **ET** calculus would need different deduction theorems.

(ii) IF ($\Phi \& T$) |-P then Φ |- (T->P)

where *T* is the evolutionary theory, Φ is a general assumption that natural selection is directional, and *P* is an empirical statement about complex adaptations for which there is *no* finite CA-description, unless assuming Φ .

For instance, most explanations of evolutionary novelty fall in this class.

(iii) IF $\Psi \rightarrow \Phi$ and ($\Phi \& T$) |- P, then Ψ |- (T->P)

where *T* is the evolutionary theory, Φ is an implicit assumption behind directional selection (such as the cellular development hypothesis), and *P* is an empirical statement about complex adaptations for which there is *no* finite CA-description, unless assuming Φ .

For instance, (iii) challenges the standard insight of the Modern Synthesis tradition (i) that the sole basis of evolution is genetic variance. The implicit assumption says that genetic variance is in fact biased by phenotypic constraints, at least for some kinds of adaptations. Such constraints channel genetic variance into potentially useful phenotypic directions (Klingenberg 2005; Kirschner and Gerhard 2006), so that natural selection appears to be directional. Most explanations concerning niche formation fall in this class.

Once we establish the forms of the deduction theorem for **ET**, we can easily see how **ET** is, in fact, a family of evolutionary calculi that discharge assumptions unevenly. This aspect might tell us something significant about unification and laws in evolutionary biology from a formal point of view. It seems that not all the evolutionary processes that populate the "reference class" of evolutionary theory fall in the first, "no assumptions" class. This means that important regularities in evolutionary biology cannot be treated as lawful correlations, unless under specific constraints on the form of the theory. As pointed out already, large-scale evolutionary processes that involve niche construction most often fall in the third class (*e.g.*, explanations of why the Hawaiian finches have different beak phenotypes from the continental finches), while most frequent models of evolutionary novelty still



fall in the second class (*e.g.*, explanations of why finches have beaks in the first place and not, for instance, muzzles). But the most common processes of evolution by natural selection, with bounded accumulation of organization and function, could be placed – with precaution – in the "no assumptions class", given there is at least a CA proof for the "|-T - P", where *P* is our specific process.

Before closing this section, I will make only another brief observation. As we have seen, the calculus approach to evolutionary biology avoids both gene realism and gene-reductionism, and also the deductive ambiguity of the standard non-reductionist axiomatization, by absorbing external suppositions into internal assumptions. But there is something else. Unlike *Zsyntax* in molecular biology, the calculus approach to evolutionary biology rejects the idea of causal isomorphism. *Zsyntax* mixes the "causal structure" of molecular processes with the "deductive structure" of the logical system, pretty much like the semantic view in evolutionary biology does with the causal interpretation of the Mendelian breeding systems. The calculus approach to evolutionary biology stresses firmly on the *syntactic* construction of **ET**. Even though it does not embrace the linguistic representation of biological theories, as in the standard empiricist views of J.H. Woodger (1937) and Rudolf Carnap (1939), it neither conflates the epistemology of evolutionary biology into a strange kind of ontology, as *Zsyntax* eventually does.

"Consider Holland's chess analogy. Rules or laws have no causal efficacy; they do not in fact "generate anything". They serve merely to describe regularities and consistent relationships in nature. These patterns may be very illuminating and important, but the underlying *causal agencies* must be separately specified (though often they are not)" (Corning 2002, p. 26).

In this sense, the syntactic approach to **ET** is as heuristic as the standard empiricist approach. It only correlates causal agencies in nature with counterfactual dependencies in the calculus, in order to create a logical *understanding* of lawfulness in evolutionary biology. However, counterfactual dependence is not *what makes* the evolutionary processes lawful. In order to grasp "the laws of evolution", counterfactual dependence is obviously insufficient (Corning, *Ibidem*).

IV. FINAL REMARKS

In this paper I suggested a picture on the axiomatization of evolutionary biology. I identified three main approaches: two "classical" approaches (the hypothetical-deductive and semantic), and a "non-classical" one (the natural deduction), and I argued that the "non-classical" approach might be promising for the logicians of science. For instance, it could tackle with interesting results the issue of theoretical unification in evolutionary biology, given that Mary Williams' standard axiomatization is inconclusive, because of



its ambiguous treatment of deductive consequence, and the genetic axiomatization is highly reductionist. My analysis also emphasized that ET comes in several forms, depending on how the deductive structure and the "reference class" of the theory are defined. In the understanding of ET as a family of calculi, an important role is given to the concept of evolutionary emergence – a concept so far ignored in the literature on the axiomatization of evolutionary biology.

Last but not least, I am inclined to point out that the calculus approach to **ET** does partial justice to the skeptical claim that there are no universally agreed upon laws in evolutionary biology (Gould 1989). There cannot be any such laws if **ET** is a family of calculi with different assumptions. Nevertheless, the skepticism might not be entirely justified since the logical systems behind **ET** heuristically reveal, once deductive ambiguity is properly dealt with, some lawful relations in evolutionary biology as formalized counterfactual dependencies.

REFERENCES

- Beatty, J. 1980. What's Wrong with the Received View of Evolutionary Theory? In *PSA*: *Proceedings of the Biennial Meeting of the Philosophy of Science Association*: 397 - 426.
- Bedau, M. A. 2009. The Evolution of Complexity. In Barberousse A. et al. (eds.) Mapping the Future of Biology. Evolving Concepts and Theories, Springer, 111-132.
- Bedau M. A. 1998. Philosophical Content and Method of Artificial Life. In Bynam, T.W. and Moor, J.H. (eds.) *The Digital Phoenix: How Computers are Changing Philosophy*. Oxford: Basil Blackwell, 135-152.
- Van Balen, G. 1988. A Critique of the Williams/Rosenberg Argument. *The British Journal for the Philosophy of Science* 39(4): 441-448.
- Boniolo, G. *et al.* 2010. Zsyntax: "A Formal Language for Molecular Biology with Projected Applications in Text Mining and Biological Prediction". *PLoS ONE* 5(3): e9511. doi:10.1371/journal.pone.0009511.
- Bunge, M., Mahner, M. (1997) 2005. *Foundations of Biophilosophy*. Berlin, Heidelberg, New York: Springer Verlag.
- Carnap, R. 1958. Introduction to Symbolic Logic. New York: Dover.
- Carnap, R. (1937) 2001. The Logical Syntax of Language. London: Routledge.
- Corning, P. 2002. Re-emergence of Emergence: a Venerable Concept in Search of a Theory. *Complexity* 7(6): 18-30.
- Emeche, C. 1994. The Computational Notion of Life. *Theoria Segunda Epoca* 2(9): 1-30.



- Ereshefsky, M. 1991. The Semantic Approach to Evolutionary Theory. *Biology and Philosophy* 6: 59-80.
- van Fraassen, B.C. 1987. The Semantic Approach to Scientific Theories. In Nersessian, N.J., (ed.) *The Process of Science: Contemporary Approaches to Understanding Scientific Practice*. Dordrecht: Kluwer, 106-124.
- Goodwin, B. 2001. *How the Leopard Changed its Spots*. Princeton: Princeton University Press.
- Gould, S. 1989. Wonderful Life. New York: Norton.
- Holland, J.H. 1995. *Hidden Order: How Adaptation Builds Complexity*. New York: Basic Books.
- Horan, B. 1988. Theoretical Models, Biological Complexity and the Semantic View of Theories. In *PSA*: *Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2: 265-277.
- Humphreys, P. 2004. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press.
- Huneman, P. 2012. Computer Science Meets Evolutionary Biology: Pure Possible Processes and the Issue of Gradualism. In Huneman P., (ed.) *Special Sciences and the Unity of Science*, Springer, 137-162. doi:10.1007/97 8-94-007-2030-5_9.
- Huneman, P. 2010. Determinism, predictability and open-ended evolution: lessons from computational emergence. *Synthese*, 185 (2): 195-214. doi:10.1007/s11229-010-9721-7.
- Kirschner, M.W., Gerhard, G.C. 2006. *The Plausibility of Life: Resolving Darwin's Dilemma*. Yale: Yale University Press.
- Klingenberg, C.P. 2005. Developmental constraints, modules, and evolvability. In Hallgrimson B. and Hall B. K., (eds.) *Variation: A Central Concept in Biology*. Elsevier Academic Press, 219-248.
- Langton, C. et al. 1990. Transition Phenomena in Cellular Automata Rule Space. *Physica* D 45(1-3): 77-94.
- Langton, C. 1986. Studying Artificial Life with Cellular Automata. *Physica* D 22: 120-149.
- Lewontin, R. 1974. *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.
- Lloyd, E. 1994. *The Structure and Confirmation of Evolutionary Theory*. Princeton: Princeton University Press.



- von Neumann, J. 1951. The General and Logical Theory of Automata. In Jeffress, L.A. (ed.) *Cerebral Mechanism and Behavior*. New York: John Wiley and Sons, 1-31.
- Reece, J.B. et al. 2002. Biology. San Fracisco: Benjamin Cummings.
- Rizzotti, M., Zanardo, A. 1986a. Axiomatization of Genetics. 1. Biological meaning. *J. Theoret. Biol.* 118: 61–71.
- Rizzotti, M., Zanardo, A. 1986b. Axiomatization of Genetics. 2. Formal development. J. Theoret. Biol. 118: 145-152.
- Rosenberg, A. 1985. *The Structure of Biological Science*. Cambridge: Cambridge University Press.
- Ruse, M. 1997. Is Biology Different from Physics? In Colodny, R.G. (ed.) Logic, Laws, and Life: Some Philosophical Complications. Pitsburgh: Pitsburgh University Press, 89-128.
- Suppes, P. 1967. What Is a Scientific Theory? In Morgenbesser, S. (ed.), *Philosophy of Science Today*. New York: Basic Books, 55–67.
- Tarski, A. (1945) 1995. Introduction to Logic and the Methodology of Deductive Sciences. New York: Dover.
- Tennant, N. 1990. Natural logic. Edinburgh: Edinburgh University Press.
- Thompson, P. 1989. The Structure of Biological Theories. Albany: Suny Press.
- Thompson, P. 1988. Explanation in the Semantic Conception of Theory Structure. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*: 286 – 296.
- Thompson, P. 1986. The Interaction of Theories and the Semantic Conception of Evolutionary Theory. *Philosophica* 37 (1): 73-86.
- Williams, M.B. 1986. The Logical Skeleton of Darwin's Historical Methodology. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*: 514 521.
- Williams, M.B. 1985. Species Are Individuals: Theoretical Foundations for the Claim. *Philosophy of Science* 52 (4):578-590.
- Williams, M.B. 1973. Falsifiable Predictions of Evolutionary Theory. *Philosophy* of Science 40 (4):518-537.
- Williams, M. B. 1970. Deducing the Consequences of Evolution: A Mathematical Model. J. Theoret. Biol. 29: 343-385.
- Winsberg, E. 2009. Computer Simulation and the Philosophy of Science. *Philosophy Compass*, 4 (5): 835-845. doi: 10.1111/j.1747-9991.2009.00236.x.



- Wissner-Gross, A.D. 2008. Pattern Formation without Favored Local Interactions. *Journal of Cellular Automata* 4: 27-36.
- Woodger, J.H. 1962. Biology and the Axiomatic Method. *Annals of the New York Academy of Sciences*, 96: 1093–1116. doi: 10.1111/j.1749-6632.1962. tb54121.x.
- Woodger, J.H. 1937. *The Axiomatic Method in Biology*. Cambridge: Cambridge University Press.