Towards a theory of biological organization
A Programmatic statement

Walter Fontana
Theoretical Chemistry
University of Vienna Währingerstraße 17
A-1090 Wien, Austria
and
IIASA
A-2361 Laxenburg, Austria

Leo W. Buss
Department of Biology
and
Department of Geology and Geophysics
Yale University
New Haven, CT 06520-8104 USA

In this document we broadly outline our research intentions "towards a theory of biological organization". We first identify a "distant shore", but we cannot plan in advance the exact path that will take us there. We then try to motivate and justify the tools with which we intend to undertake the trip. With that in the background, we proceed to sketch briefly an approximate plan of action. We conclude with a few thoughts about what may lie ahead once that distant shore has been reached.

1 The distant shore

Biology has claim to two theories unto itself: Darwin's natural selection and Mendel's transmission rules. Both are correct, their joint operation (known as the "Modern Synthesis") can be nicely formalized, and together they are insufficient to account for the history of life as we know it.

What are we missing? We lack a theory of the genesis of biological organizations (that is, organisms or phenotypes), their modification, and their combination into superstructures.

Why are we missing it? It is lacking because the major theoretical tool employed so far – the theory of dynamical systems – is inadequate.

The limits of dynamical systems

When reasoning about some aspect of the world with the tools of dynamical systems theory, we associate with each pertinent entity a variable that holds the value of a quantifiable property (such as, position, momentum, concentration, wealth, energy). The dynamical systems approach views the net effect of the interactions among entities as causing their associated quantities to change. The real world interactions are formalized by fixing the functional form of the dependencies among the variables of the dynamical system. This abstraction yields a set of coupled differential equations describing the space and time evolution of quantitities related to entities that interact (possibly with floating coupling coefficients) in a predefined and fixed way. The collective phenomena, or units of aggregation, captured by this methodology are self-sustaining spatio-temporal patterns of *quantitative* change. The theory provides a classification of such patterns as attractors in phase space (e.g., fixed points, limit cycles, chaos) and techniques to assess their stability.
The price for this framework is keeping the entities outside of it. As a point in fact,that which changes endogenously are quantities associated with entities, not the entities themselves. The entities that constitute a system cannot change endogenously, for their internal structure is either irrelevant in the formal setup of a dynamical system, or it is disconnected from any action that *causes* structural change in other entities.

The dynamical systems view formally separates the entities and the interactions in which they engage. To go beyond this framework requires removing that separation.

The goal of this project is to develop a theory of biological organization by adding a *constructive component* to dynamical systems. This raises two issues. (i) What entity class supporting the constructive component is the appropriate one to start with? (ii) What is the appropriate formalism to capture it?

A theory of biological organization must be grounded in a representation of that which organisms are composed. The theory must be grounded in chemistry, the real-world entities of interest here are molecules. The characteristic feature of chemical action is the generation of specific new molecules. As it stands, the reaction "equation"

$$CH_3OH + CH_3COOH \longrightarrow CH_3COOCH_3 + H_2O$$

does not express the change of any numerically quantifiable property. Of course, thermodynamics and reaction kinetics register the change of several quantities. The point, however, is that a chemical reaction goes beyond that; it expresses a relation of production between structures. The arrow "$\longrightarrow$" has the flavor of a *logical consequence,* where premisses – CH_3OH and CH_3COOH – imply specific conclusions – CH_3COOCH_3 and H_2O. Since the conclusions can be premisses for further inferences, the composition of an initial set of molecular entities changes from within. This is a quite different kind of "flow" than a dynamical system. As a consequence one obtains a different kind of aggregation unit. A specific ensemble of entities can be tied together into a unit by, for example, sustaining invariant relationships of mutual production. In traditional dynamical systems an aggregation unit is defined exlusively in terms of *dynamical* dependencies. Here we aim at an extended framework that includes *constructive* (or, equivalently, logical or functional) dependencies.

A central theme of this project is our emphasis on the action of "agents" as consisting in a mapping from "agents" to "agents". The foundational distinction between this project and others sharing similar emphasis is that none of the others have the feature of being interpretable as mathematical structures. To bootstrap a theory of biological organization, requires a formalization of chemistry that is useful for biology. Ultimately this means defining "interaction" and "structure of an entity" mutually. In our case "structure of an entity" will specifically come to mean a formula, expression, or term within a syntactical system resulting from a *theory* about the possible interactions characterizing that class of entities. To be illuminating, a linkage between structure and interaction must connect to a *mathematical* theory.

2 Tools

An interaction in which entities generate other entities seems at first a rather special and contrived one. While the chemical motivation is obvious, this mode of interaction is far more general than chemistry. Here we attempt to convey the philosophy of the mathematical "materials" to be deployed. All of these are well known (but they need to be interpreted).

2.1 The macro: combinatory algebra

An interaction in which two elements a and b of a set $\mathcal{C}$ yield another element $a \cdot b = c$ in the same set is but a "multiplication" that turns $\mathcal{C}$ into some ordinary algebraic structure. An intriguing perspective arises when we wish to view any element of $\mathcal{C}$ as being also a *function* from $\mathcal{C}$ to $\mathcal{C}$; think of fixing a while varying b over $\mathcal{C}$ in $a \cdot b$, or, chemically speaking, as having CH_3OH react with some molecule b that varies over the set of all molecular species. A mathematical picture of a minimal chemistry requires *an algebraic universe $\mathcal{C}$ in which every element of $\mathcal{C}$ can be rigorously interpreted as an element of some set of functions $\mathcal{C} \to \mathcal{C}$ and vice versa.* As a matter of fact a characterization exists for a universe in which this shuttling back and forth between entities as both elements-of-the-universe and functions-on-it is possible in a *consistent* way. Such a universe is known as a *combinatory algebra* [3, 4, 8].

Combinatory algebras

In algebra the nature of elements (henceforth "objects") is irrelevant. Nevertheless, it helps to have some notion of "object" in mind to get a taste of combinatory algebra. One such notion considers an "object" as being any *set of properties* expressed in some (first order) language. If that lacks imagery, then simply consider as "object" anything that can be *"object of thought"* (cf. [4]). What can be said about a universe $\mathcal{C}$ based on such an abstract notion of "object" and some (as yet entirely unspecified) notion of "multiplication" within $\mathcal{C}$? A number of things, as it turns out.
To make an example, take three particular objects in $\mathcal{C}$, say a, b, c, and combine them in a certain way, say, $(a \cdot (b \cdot a)) \cdot c$. By following the "multiplication" table of $\mathcal{C}$ this combination will result in, say, z. Now, any three objects could be combined in this way, and "to combine arbitrary three objects, x_1, x_2, x_3 into $(x_1 \cdot (x_2 \cdot x_1)) \cdot x_3$" is itself a (first order) property. As such it specifies some object that must, by definition, exist in $\mathcal{C}$. Call that object P. The defining property of P is precisely that for its "multiplication" with any three objects x_1, x_2, x_3 the relationship $((((P \cdot x_1) \cdot x_2) \cdot x_3) = (x_1 \cdot (x_2 \cdot x_1)) \cdot x_3$ holds in $\mathcal{C}$. By this reasoning we conclude that for any arbitrary combination of an arbitrary number of objects in $\mathcal{C}$ there must be another object that expresses the possibility of combining things in that way, and there will be a corresponding relationship (an equality) stating that fact. This remarkable property is known as "combinatory abstraction". A series of other properties follow from it, but we won't pursue this matter further. Universes in which it holds are "combinatory complete", and are known as "combinatory algebras".

We can think of a combinatory algebra as a phenomenological characterization of a universe that has embedded within it an infinity of algebraic structures of all kinds. Algebraic structures are specifications of interrelations among components, and as such they constitute the abstract core of "systems". They are abstract in the sense of not specifying how the components must be internally structured to satisfy these relations (they may not even exist).

Why is this important? In the physical world relational structures are being formed and destroyed at any time. Certain structures have become robust beyond the time scales set by the nature of their constituent processes; examples include metabolisms, organisms, ecosystems, money, states, cultures. Stable relational structures are what conventional dynamical systems take as a given. If we aim at a description of their genesis, we must extend our description of the world to include a sound and sufficiently general *dynamics of relationships* capable of generating stable relational structures as stationary outcomes. What enables any such dynamics from within, we submit, is at the very minimum a notion of entities that lead this subtle double-life: entities that simultaneously act as relationships between entities (by being interpretable as functions on them). Of course, it would make no physical sense to think of such a universe as existing ready-made in its totality, equipped with a random-access mechanism capable of picking any one of the possible relational structures on demand. It is the new dynamics that must construct them. Its concrete implementation forces us to think about the internal structure of "objects" and their binary "multiplication". A dynamics of relationships needs a causal mechanism, a "physics". This mechanics, however, must imply a combinatory algebra, or else the interpretation of our base entities as both elements and functions could not be logically sustained. The desired mechanics exists.

2.2 The micro: λ-calculus

The micro-foundation of combinatory algebra is known as λ-calculus.

λ-CALCULUS

λ-calculus [2, 9, 1] is a *syntactical* (as opposed to an algebraic) system intended as a convenient means for writing expressions which denote functions.
The "objects" are symbolic strings (i.e., expressions) constructed according to certain rules of grammar. The "multiplication" or binary operation among expressions is *application*. It's operation is defined in terms of a scheme for rewriting the structure of expressions. That scheme derives from an analysis of the concept of "substitution" (i.e., the process of replacing a variable in an expression by an argument expression). The ingredients of λ-calculus are syntactical concepts, such as "variable", the associated notion of "scope", and "substitution". The rewrite scheme of λ-calculus generates an equality relation between different combinations of expressions, much like the rules of arithmetic allow to deduce the equality of different symbol sequences, such as $2 + 2$ and $3 + 1$.
There exists a translation from λ-expressions to terms of a combinatory algebra such that all equations which hold in the λ-world are exactly those that hold in combinatory algebra.

One can roughly think of combinatory algebra as a purely behavioral characterization of the λ-calculus engine in which every reference to the syntactical machinery has been removed. This is why we refer to λ-calculus as the "micro", or the "physics", and to combinatory algebra as the "macro", or the space of possible behaviors.

A number of indicators single out λ-calculus as the unique formal "physics" for the purpose of a *first attempt* at a theory of chemical organization. For one, substitution is a key mechanism in chemical reactions. Here the similarity is obvious and unique to λ-calculus. Then there is our stance that algebra is a necessary mathematical instrument for framing the intuitive concept of "organization". Here λ-calculus is unique in being the formal "mechanics" that generates the combinatory algebra (the universe of "all imaginable" organizations). Last but not least, chemical objects share with certain mathematical objects, such as programs or processes, a tight linkage between structure and applicative behavior. The relevance of λ-calculus for our agenda resides precisely in formalizing the distinction between behavior and that which behaves, thereby making syntax a vehicle for reasoning about action.

To summarize, the need to formalize a concept of biological organization has brought up the need to formalize chemistry at a different level of analysis. A first guess at a useful abstraction of chemistry has led us to consider chemistry as a universe of elements that are functions on those same elements. The mathematical analogue of such a universe is a special algebra known as combinatory algebra. Combinatory algebra is a world of possible relational structures about which one could reason on its own grounds, i.e., irrespective of a "micro-foundation". As such it epitomizes the *concept* "chemistry" (that's precisely the blind spot of quantum mechanics). Underneath there is the causal "micro"-picture of chemistry as the science of molecules whose internal structure determines their functional action. The mathematical analogue to start with is the syntactical machinery of λ-calculus. Combinatory algebra obtains as a behavioral characterization of λ-calculus where the specific representational machinery has been removed[1].

3 Routes for navigation

In past work we have amalgamated λ-calculus with a conventional dynamical system inspired by chemical mass action kinetics [5, 6]. Upon imposition of biologically motivated boundary conditions, we obtained a process that generated algebraic structures (instantiated as ensembles of λ-calculus expressions) that were self-maintaining. This is an example for how the design of a combination `quantitative dynamics (kinetics) + relational dynamics (λ-calculus) + boundary conditions` can be used to access relational structures that possess a desired feature.

Two issues arise from this. Can such a self-organizing process be designed for *any* feature? (Or, for which features can one be designed?) This topic may be labelled tentatively as **organizational learning**. The other issue arises from taking a feature as a given, and asking what are the possible ways of realizing it at an algebraic level, and how are the instantiations distributed at the micro-level of λ-calculus? This topic may be labelled **charting organization space**.

The above threads build on our existing theoretical and computational platform. Their potential impact on a theory of biological evolution should not be underestimated. Said impact, however, is obviously limited by the grotesque although heuristically quite useful abstraction of chemistry by means of λ-calculus. The

[1] Think, by analogy, of an algebraic structure like a group vis à vis one of its presentations, say, the symmetry operations in 2-space represented as matrices (plus the rules of matrix calculation).

thread labelled **chemistry and logic** aims at a better formal framing of chemistry to ease this discomfort.

Parallel to improving the chemical formalism, we need a systematic rethinking of the axiomatization of the evolutionary process in the light of this agenda and the mathematical tools it suggests. This is simply labelled **biology**.

In the following we briefly summarize these issues.

- **Charting organization space**

In past work on this project we implemented a model and a simulation platform that produced a variety of self-maintaining organizations. A basic question still remains: in what ways can a set of functions be self-maintaining? The goal is a classification scheme aimed at identifying basic organizational modes of self-maintenance and the possible transitions among them. Answers to these questions are sought at first strictly within the familiar λ-calculus and combinatory algebra framework for which a suite of tools, both theoretical and computational, is available.

- **Organizational learning**

Can we systematically construct ensembles of λ-expressions together with interaction laws (cast again in terms of λ-expressions) such that they instantiate predefined algebraic structures? In other words, can we formulate a learning procedure that does for our relational dynamics, say, what back-propagation does for neural networks? Aside from the biological motivation, such a learning, or training, process would be of considerable interest for computer science where algebraic structures are used as specifications of abstract data types. Such a learning scheme would automatically construct *implementations* of abstract data types. Our past model can be seen as a procedure to single out algebraic structures with self-maintenance as a desired feature. Going beyond, includes as "desired feature" the satisfaction of prespecified equations.

- **Chemistry and logic**

A discussion of the limits of λ-calculus as an abstraction of chemistry is given in [7] together with a detailed agenda for improvement. Here it suffices to say that interaction symmetry, mass conservation, and concepts such as "shape" and "rate constants" can be accomodated within existing formal systems (logics and proof-theories), all of which have a natural connection with λ-calculus. However, a major conceptual problem towards a respectable chemistry remains. We must account for reactions that have two or more products, such as splitting a molecule in two pieces. In the world of traditional algebra this is not the case, as $a \cdot b$ yields *one* element at a time, while we really need something like $a \cdot b = c, d$ where the comma indicates the simultaneous occurrence of two elements of the set in which the "multiplication" occurs. We need a logic where one inference can yield more than one conclusion – not in the sense of a choice among several alternatives, but rather in the sense of concurrent conclusions.

• **Biology**

Present evolutionary theorizing is based on an axiomatization of Darwinism and an axiomatization of genetics. Our goal is to add an axiomatization of "organization" on the basis of a high-level formalization of chemistry. We believe that the formal areas briefly introduced previously offer techniques that are sufficient for constructing a perturbation theory of organizations. By this we mean an understanding of the linkage between the (algebraic and kinetic) properties of an organization and the ways in which it can be innovated – extended or merged – including an assessment of how such innovations constrain its further evolvability. All this is specific to the particular representation we have chosen, but we argue that the tools underlying it – combinatory algebra, λ-calculus, and reduction systems – are germane to biology.

Restating by way of analogy, think of how conservation principles (read "constraints") shape our understanding of physical motion (read "possible phylogenies") without necessarily being sufficient in pinning down actual trajectories. In classical physics such basic constraints arise from the symmetries of space and time, the media in which motion plays out. Reverse the analogy, and expect the imputed connection between a logic of concurrency and chemistry to yield basic organizational constraints deriving from the symmetries underlying logic.

Connections with the empirical data base of biology are necessary for grounding a theory. Yet *both* our approach and biology may be immature for this interplay at the present moment in time. The strength of our agenda is twofold. It goes to the conceptual core of major unsolved problems central to biology, and it does so with formal instruments at hand. The interplay between experiment and proto-theory need not be sought at any price before the connection between the conceptual and the formal has been properly made. While a number of little excercises in applied math could be ventured right away (and published in top journals), they would retain the descriptive character of *ad hoc* models, rather than being a statement about what one can predict in biology and why. Suppose, for example, that only a handful of major bauplans for self-maintaining organizations can show up in our framework. This would make the actual architectures generated by our approach relevant, not just its abstract concept of organization. At this point a meaningful interplay with biological knowledge can be sustained. To get there, we must first progress in charting organization space, in sharpening the connection between chemistry and logic, and in understanding how exogenous constraints can shape organizational structure.

4 The continent behind the shore

Once we go beyond the traditional subject matter of physics, radically new ways of being organized become apparent. What remains elusive, however, is what exactly the "newness" consists in. The search for distinguishing basic principles that make different modes of organization tick is what we understand as the "study of complexity". Chemistry has a special status in this, for it marks the transition to a notion of *functional* organization that is conspicuously absent in physics (as a discipline). Framing a view of chemistry with appropriate formal concepts and demonstrating how it gives rise to a particular notion of functional organization was our contribution to this theme.

The "chemical perspective" can play a number of roles. While the concept of organization conveyed by chemistry is at best of metaphorical use in the social sciences and economics, it seems nevertheless more "to the point" than the undiscriminated use of spin glass physics and "landscape" themes. Furthermore, because a concept of organization can actually be formalized at the chemical level, its limits with respect to the social sciences and economics are rendered visible. The further role of the chemical perspective, then, is to serve as a stepping stone to step away from. What are, from an abstract point of view, the differences between social action and chemical action? Which mechanisms are capable of progressively weakening the tight chemical linkage between syntax (physical agent structure) and action? Which mechanisms are capable of progressively feeding back the global into the local (think of memory, cognition)?

On the way to the organizational themes of social science and economics (on the way to a "science of organization"?) we need to understand how action is liberated from physical syntax.

References

[1] H. G. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*. Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, second revised edition, 1984.

[2] A. Church. A set of postulates for the foundation of logic. *Annals of Math. (2)*, 33:346–366, 1932.

[3] H. B. Curry. Grundlagen der kombinatorischen Logik. *Amer. J. Math.*, 52:509–536; 789–834, 1930.

[4] E. Engeler. *The Combinatory Programme*. Birkhäuser, Berlin, 1995.

[5] W. Fontana and L. W. Buss. 'The arrival of the fittest': Toward a theory of biological organization. *Bull. Math. Biol.*, 56:1–64, 1994.

[6] W. Fontana and L. W. Buss. What would be conserved 'if the tape were played twice'. *Proc. Natl. Acad. Sci. USA*, 91:757–761, 1994.

[7] W. Fontana and Leo W. Buss. The barrier of objects: From dynamical systems to bounded organizations. In J. Casti and A. Karlqvist, editors, *Boundaries and Barriers*, pages 56–116. Addison-Wesley, Redwood City, 1996.
The tutorial appendices on λ-calculus (Appendix A), type theory (Appendix B) and proof-theory (Appendix C) can be obtained with the main text as IIASA WP-96-27, as paper 41 from the Yale Center for Computational Ecology, or from the World Wide Web at http://peaplant.biology.yale.edu:8001/alchemy.html.

[8] A. R. Meyer. What is a model of the lambda-calculus? *Information and Control*, 52:87–122, 1982.

[9] M. Schönfinkel. Über die Bausteine der mathematischen Logik. *Math. Annalen*, 92:305–316, 1924.