That's Life?-Yes, No, Maybe

John L. Casti

SFI WORKING PAPER: 1992-10-050

September 1992

That's Life?—Yes, No, Maybe

JOHN L. CASTI

Institute for Econometrics, OR and System Theory
Technical University of Vienna
A–1040 Vienna, Austria

AND

Santa Fe Institute
Santa Fe, NM 87501, USA

Abstract

Contrary to the cherished beliefs of many physicists and other unreconstructed materialists, there's a lot more to life than mere matter. What distinguishes life from non-life is how the matter is organized. This paper argues that by retreating from a resolutely Newtonian view of life and returning to the Aristotelian notion of causal categories, we open up new avenues of approach to understanding the interconnections among earthly life forms, artificial life in machines and possible types of extraterrestrial life.

 This paper will appear in the proceedings of the "Frontiers of Life" workshop, Blois, France, October 1991

1. The Why of Life

According to Caesar, all Gaul was divided into three parts. And so it is too when it comes to the study of life, where we find that the deepest questions seem to be parceled out among three seemingly distinct categories. The first, of course, is just the wet, squishy, carbon-based type of life we're familiar with here on Earth. Into this Category I fall all the usual origin-of-life (ORI) questions involving the way life got its start on this planet. Category II is of much more recent vintage, centering less upon how life got its start and more upon the even deeper question made famous by the title of Erwin Schrödinger's book, *What is Life?*. In this category we find the current work on artificial life (AL), which seeks to answer Schrödinger by constructing life based on patterns of information in silicon and electricity instead of in carbon and water. Finally, we come to close encounters with life of the third kind: extraterrestrial life forms (ET). Here in Category III reside all the puzzles involving the possible nature of life forms evolving under environmental conditions radically different (or, perhaps, indistinguishably similar) to those found here on Earth.

Superficially, these categories appear to be pretty much disconnected. Or at least that's the impression one gets from a reading of the scientific research literature. In the spirit of the "Frontiers of Life" meeting, I'll argue in this paper that not only are these categories inextricably intertwined, but that the adoption of a more functionally-oriented view of life opens up the possibility for a productive interchange of ideas among research groups pursuing work in the different categories.

Just about the first thing every journalist learns is that a good story answers the following questions: Who? What? When? Where? Why? How? Applying this tried and true principle the story of life, it's not too hard to see that these interrogatives can be apportioned out among our Categories I–III as follows:

- *Origin of Life: How?*—studies of the origin of life on Earth tend to focus almost exclusively on questions beginning with "How." Occasionally, as with various extraterrestrial theories of the origin of life, "Where" enters the picture too. But, by and large, ORI studies center on the construction of various scenarios for **how** life got going on Earth four billion years or so ago.

- *Artificial Life: What?*—the key issue around which most AL studies revolve is Schrödinger's famous question: What is life? Put more specifically, we ask if a suitably-organized pattern of information in a machine can ever constitute a genuine living organism? So the operative word here is "What".

- *Extraterrestrial Life: Who? Where? When?*—"Where are they?" This was Enrico Fermi's famous retort to the claim that the universe must be filled with ET's. Studies over the intervening half century or so have centered mostly on where to look, what to look for and when will we find something (someone?).

The most striking aspect of this list of journalistic interrogatives is the singular absence of "Why". So if Schrödinger's question governed studies of life in the 20th century, I'd like to predict that 21st-century studies of life will be driven by the question, "Why is life?" Let me take a page or two to explain why.

2. *From Newton to Aristotle*

The basic goal of modeling, mathematical or otherwise, is to answer the question 'Why?' According to Newton (a physicist), the corresponding 'Because' is given in terms of local interactions involving material particles and unexplainable forces. Aristotle (a biologist) had a quite different way of saying 'Because.'

In Aristotle's view, the 'Why' of things can be described in terms of three basic entities: (i) the material *substance* comprising physical objects, (ii) the abstract or geometric *forms* that objects can assume, and (iii) the processes of *change* by which either the substance or the form may be transformed. Thus, Aristotle's 'Because' results in four disjoint and inequivalent *causal categories* which, taken together, provide a complete answer to 'Why' the world is as it is. These causal categories are:

- *Material* cause—things are as they are because of the **matter** of which they are composed;

- *Efficient* cause—things are as they are because of the **energy** that went into making them as they are;

- *Formal* cause—things are as they are because of the **plan** according to which they were built;

- *Final* cause—things are as they are because of the **desire** or **will** of someone to have things take their current state.

Note that in the above scheme of things, material cause corresponds to substance, with efficient cause relating to processes for changing the substance. Similarly, formal cause explains the abstract or geometric form of an entity, with final cause describing how one changes the form.

This scheme explains why there are four basic causes in the Aristotelian view of the world, and not three or five or 3,469.

In the epistemology of Aristotle, all things can be explained by invoking the four basic causes, each cause illuminating a different fundamental aspect of the system at hand. We can also interpret these inequivalent causal categories by thinking of each category as being concerned with the manipulation of "something" as indicated in Table 1.

CAUSE	PROPERTY MANIPULATED
Material	Physical Matter
Efficient	Energy
Formal	Information
Final	Desire; Will

Table 1. Aristotelian Causal Categories and Manipulations

Interestingly enough, both the Newtonian and Aristotelian explanatory schemes talk about the same thing: a material substance and the process by which this substance can change. However, in the Aristotelian picture substance is not enough; one also needs the idea of form and some kind of dynamic by which one form can be transformed into another. This latter idea is totally absent from the Newtonian picture. In partial compensation, the Newtonian setup offers a mathematical apparatus by which we can encompass both the particles (material cause) and the forces (efficient cause) that constitute the heart of the Newtonian modeling paradigm. The Aristotelian picture provides no mathematical machinery, only a verbal description of the causes. It's instructive to examine this dichotomy in a bit more detail in order to get a feel for what must be done to extend the Newtonian formalism to accommodate the additional Aristotelian causes.

Newton's Second Law is usually written as the differential equation

$$\ddot{x}(t) = F(t), \qquad x(0) = x_0,$$

where $x(t)$ is the state of the system of particles at time t, the quantity F represents the unexplained external forces, and x_0 is the initial state of the system. For our purposes, it's more convenient to write this relationship in integrated form as

$$x(t) = x_0 + \int_0^t \phi(s)\, ds,$$

where

$$\phi(s) = \int_0^s F(r)\, dr.$$

Now we can ask the question: Why is the system in the state $x(t)$ at time t? Newton can give only two answers:

1) The system is in the state $x(t)$ at time t because it was in the state x_0 at time $t = 0$ (material cause);

2) The system is in the state $x(t)$ at time t because of the operator $\int_0^t(\cdots)$ which transformed the initial state to the state at time t (efficient cause).

Thus, the Newtonian framework has neither the need nor the room to accommodate the additional Aristotelian categories of formal and final causation. Some would argue, myself included, that this fact more than any other accounts for the banishment of formal and especially final cause from polite scientific conversation for the better part of three centuries. There is just no room to fit them in to the classical Newtonian framework.

In actuality, even the most die-hard Newtonian ultimately came to recognize, albeit implicitly, that the missing causal categories would somehow have to be grafted on to the classical setup. In particle systems, the role of formal cause is usually assigned to various parameters specifying important constants in the situation. So things like as particle masses, gravitational constants, electric charges and so on serve to characterize the "plan" of the system. It's through the specification of such parameters, and their incorporation into the mathematical framework, that formal cause enters by the backdoor into the Newtonian scheme of things. But what about final cause? How does Newton deal with the idea of desire or will? The simple answer is that he doesn't.

When reading Aristotle's account of the causal categories, one is struck by the great significance he attaches to the notion of final cause. In fact, for Aristotle it seems that final cause was just a little more equal than any of the other categories, and he reserved his greatest respect and kindest words for what would today be termed (by Newtonians) "teleology." For the kinds of problems that concerned Newton, it appears reasonable to omit final causation from consideration since it's difficult to imagine non-living, material particles having any particular kind of will, volition or consciousness. Thus, Newton and his successors had no need to invoke any of the ideas associated with final cause, notions like goals, plans, will, or even self-reference, in their analyses of physical processes.

From this point of view, it's rather easy to understand why the mathematical machinery they employed seemed perfectly adequate to the task at hand, even though it contained no natural way to account for final cause and dealt with even formal cause in a rather *ad hoc* manner. Unfortunately for the biologist, economist and psychologist, Newton's prescriptions were too successful in answering questions in physics, chemistry and engineering, leading to a gradual emergence of the view that it's a breach of scientific etiquette, if not downright unscientific, to allow anything even smacking faintly of final causation to enter into polite scientific discourse.

In other words, if you can't use the methods that work in physics, then you're not doing science. And using the methods of physics means working within the Newtonian paradigm.

Newton's world naturally sees organisms as being just special kinds of material objects; in short, biology $\subset$ physics. Aristotle's *Weltanschauung,* on the other hand, argues for just the opposite view, implying that the physics of particles and forces is just a special kind of biology. In other words, the Aristotelian world is one in which the Newtonian framework is just a special case of a broader-based paradigm suitable for characterizing life. Aristotle's problem was that he had no mathematical formalism at his disposal with which to describe his causal theory. But mathematical modeling has developed considerably in the two thousand years since Aristotle. So let me now sketch the outline of just such a formal framework. Then I'll show how this framework can help in putting several questions from ORI, AL and ET onto a common footing.

3. Formalizing a "Causal" Biology

As the role model of a system displaying the sort of features that concern us here, let me consider that most quintessential of earth-life objects, the living cell. From a functional standpoint, a living cell engages in three distinct kinds of activities: (1) *metabolic* activity by which the cell carries on its primary function of transforming various types of chemical compounds into other types needed for its existence; (2) *repair* activity by which the cell attempts to counteract disturbances in its operating environment and/or metabolic machinery; (3) *reproductive* activity where the cell acts to preserve its functional activities by building copies of itself.

It's important to note here that these functional activities are pretty much independent of the particular physical substrate in which they are carried out. Thus, while a biological cell may act as a small chemical plant transforming one sort of chemical into another, an economic system may perform exactly the same functional activities making use of no chemicals at all. Rather, the economic system transforms money or labor or some other sort of raw "material" from one form to another. It's in this sense that I would argue that the "systemness" of the situation consists in emphasizing formal and final causation over material and efficient.

An abstract formal mathematical structure suitable for capturing the essence of the three distinct types of cellular activity was proposed some years ago by the theoretical biologist Robert Rosen under the rubric "metabolism-repair" systems[1-3]. This formalism encompasses all of the requirements noted above but without sacrificing flexibility in formulation, thereby allowing considerable leeway in the details by which the processes of metabolism, repair and replication make their appearance in the overall structure. But rather than speak in such vague generalities, let me briefly sketch the main steps in the development of these (M, R) systems, leaving the interested reader to consult the references for the fine-grained technical details.

We begin with the almost self-evident observation that every living cell has evolved to perform some kind of metabolic function. What this means is that the business of a cell is to transform some collection of chemical inputs into a design metabolic output. So if we let Ω represent the set of possible environments that the cell may face, while letting Γ denote the set of possible outputs that the cell is capable of producing, then for any given cell there is a design environment ω^* and a design output $\gamma^* \in \Gamma$ such that the cell's metabolic machinery f^* maps ω^* to γ^*. In general, of course, there may be a whole range of acceptable environments ω^* and correspondingly acceptable outputs γ^*. But for sake of simplicity and ease of exposition, let's assume that each cell has just a single design environment $\omega^* \in \Omega$, which is transformed by the cellular machinery into a single design output $\gamma^* \in \Gamma$. So when all is working according to plan, the cell accepts the design environment ω^* and processes it according to the basal metabolism f^*,

f^*, producing the design output γ^*. For future reference, note that here the metabolic map f^* belongs to the set $H(\Omega, \Gamma)$, consisting of all physically possible cellular metabolic mechanisms. This set is determined by various constraints of a physical and chemical nature on just what the cell can do, and may differ widely from one cell type to another.

It's possible to compactly summarize the cellular basal metabolism by saying that it consists of a map

$$f^* : \Omega \to \Gamma,$$

$$\omega^* \mapsto \gamma^*$$

In passing, let me point out that this is exactly the setup from Newtonian dynamics when those dynamics are expressed in the integrated "input/output" form given earlier. In other words, the usual Newtonian picture as expressed by the differential equation

$$ma = \ddot{x} = F(t), \qquad x(0) = x_0,$$

is precisely a system of the above metabolic type when we take $\omega^* \equiv F(t)$, and let the system output $\gamma^* \equiv x(t)$. But this is where the Newtonian framework quits. To go beyond mere physics, we have to be able to make the processes of repair and replication emerge "naturally" out of the metabolic data Ω, Γ, and $H(\Omega, \Gamma)$.

Functionally speaking, the role of the repair operation is to stabilize the cellular metabolic activity in the face of disturbances to the ambient environment ω^* and/or perturbations in the internal metabolic processing operation f^*. For our purposes, we'll suppose that the way this is accomplished is for the cell to "siphon off" some of its metabolic output, and then processes this part of the output by a repair mechanism that produces a new metabolic map f. This "machine" is then used to process the next environmental input ω. So, we postulate the existence of a repair map

$$P_{f^*}: \Gamma \to H(\Omega, \Gamma)$$

$$\gamma \mapsto f$$

Here we include the subscript f^* to indicate explicitly that the role of the repair map is to stabilize the cell's metabolic operation. This fact leads to the obvioius boundary condition on P_{f^*} : "If it ain't broken, don't fix it," which can be translated into the mathematical statement

$$P_{f^*}(\gamma^*) = f^*. \tag{$*$}$$

This condition says that when the cell is operating as it should (i.e., producing the design output γ^*), the repair mechanism should generate the basal metabolism f^* itself as the machinery to be used to process the next input. We relegate to the references a consideration of how to construct the repair map P_{f^*} in a mathematically "natural" manner directly from the given design metabolism. Let it suffice to say that this can indeed be done—and rather simply at that. Now what about replication?

Biologically, the replication operation is a complicated processing of the cellular output in a special way to create a copy of the genetic information contained in the cellular DNA. In our abstract setting, the internal model the cell has of its own structure is embodied in the condition ($*$), which serves as a self-referential description of what the cell is supposed to be doing. Somehow we have to find a replication operator that accepts some subset of the cellular inputs, outputs, metabolic and repair operations, producing a new genetic component. In our framework, we will represent this component by the repair map P_{f^*}. There are probably several different ways to do this mathematically, but it turns out to be very convenient and natural to assume that the replication component accepts some part of the metabolic machinery f, and then processes this into the new repair map P_{f^*}. Thus, we postulate a replication map

$$\beta_{f^*} : H(\Omega, \Gamma) \to H(\Gamma, H(\Omega, \Gamma))$$

$$f \mapsto P_f$$

Again we have the obvious boundary condition

$$\beta_{f^*}(f^*) = P_{f^*}, \tag{$\dagger$}$$

which follows from the same line of reasoning employed for the repair map.

The following diagram summarizes the development thus far:

$$\Omega \underbrace{\xrightarrow{f}}_{\text{metabolism}} \Gamma \underbrace{\xrightarrow{P_f}}_{\text{repair}} H(\Omega, \Gamma) \underbrace{\xrightarrow{\beta_f}}_{\text{replication}} H(\Gamma, H(\Omega, \Gamma))$$

This diagram of sets and maps, together with the boundary conditions ($*$) and ($\dagger$), forms an abstract representation of a (M, R)-system. As noted earlier, if we neglect the repair and replication operations we obtain the standard Newtonian setup. So if all we're interested in studying is cellular metabolism, it's not necessary to go beyond the modeling framework of classical physics. It's for this reason that I refer to the usual Newtonian systems of physics and engineering as "metabolism-only" processes. But the distilled essence of life is not metabolism alone; what distinguishes living processes are the activities of repair and replication—exactly the aspects of the (M, R)-diagram that are not accounted for by Newton.

To make this (M, R)-formalism work, it's necessary to drape some mathematical meat upon the bones of this abstract collection of sets and maps. I refer the reader to the references[4–6] for these technical details. Now let's see how the combination of Aristotelian causation and the (M, R)-setup help illuminate some of the key questions in ORI, AL and ET.

4. The Hen or the Egg?

Almost all competing theories of the origin of life involve a scenario in which either the first living thing was a metabolizer that later acquired the ability to replicate, or is a replicator that then assumed metabolic functions. Thus the question: Which came first, the hen or the egg?

While the current flavor-of-the-month in the ORI business seems to be a replicator-first theory involving an evolution of life from self-catalytic RNA, the (M, R)-framework strongly suggests otherwise. On purely logical grounds, quite independent of physical substrates like nucleotide bases, amino acid chains and all the other paraphernalia of earthly life, it's very difficult to see how to start with an abstract replicator like RNA and make metabolic activity emerge from replication. There does not appear to be any straightforward way to begin with the process of replication and use "natural" mathematical operations on sets and maps to create a metabolizer. Part of the difficulty comes from the fact that the repair component of the cell acts as an intermediary between replication and metabolism, and it's just far simpler to see how to make this operation come about from *something* rather than having to emerge from nothing. And since replication involves no actual production of something new, but only the copying of something old, there's just not enough raw material to work with to get the job done. So if you believe in the (M, R) framework, then you'll also believe that the time is ripe for metabolism-first theories of ORI to begin making a comeback.

Turning now to Aristotle and his causes, the very nature of the ORI problem seems to call for an explanation of life on Earth by *material* causation. And, in fact, the distinguishing characteristic separating the competing theories—self-rep RNA, hypercycles, clay, iron—is the kind of material substance they postulate that the first organism was made of. But the (M, R)-framework is one grounded in formal and final cause, not material and efficient. In particular, by thinking in these alternative causal categories we are able to pin down the elusive notion of final cause: it is simply the cell's purpose or, in biological terms, its design metabolism. From what was said in the last section, we can express this mathematically by the boundary condition on the repair map: $P_{f^*}(\gamma^*) = f^*$. Now let me shift the venue from real life to artificial life.

5. *Life* in Silico

In summer 1991 Thomas Ray, a plant biologist in Delaware, created a computer program that reproduces, undergoes spontaneous genetic changes, passes them on to its offspring and evolves new species whose interactions mimic those of real biological evolution and ecology—all without human guidance or intervention. The significance of Ray's program was that it was the first logical demonstration of the validity of the Darwinian theory of evolution. This is exactly the goal of the artificial lifers: to show that the logical structure of life is completely independent of material cause. True or not, just as with work on ORI, in AL too there is something to be gained by thinking in terms of other causal categories.

For instance, Darwinian theory implies that the environment in which an organism develops will strongly influence the physical form and function of the organism. This is a material-cause explanation for why the vast majority of living things on Earth are shaped like cylinders with appendages. But suppose we have something like Ray's electronic organisms that evolve within the cozy confines of a computer's memory banks. Thinking in terms of material cause, we can ask: What is the shape of such an electronic "bug"? And what would such an organism see when it looks at its "environment"? And in what way does the bug have to adapt to such an environment? And so on and so forth. The point is that we are led to ask a quite different set of questions than those that naturally come to mind in the original setting of formal cause.

Before leaving this point, let me add that by taking a collection of abstract cells of the type in the (M, R)-framework and connecting them into a network, it's possible to construct an electronic ecology similar to that produced in Ray's computer. Some suggestions about how this might be done are given in the papers [1,6]. One result that comes out of such an exercise is the purely logical fact that immortality is flatly impossible. In other words, not all genetic lines can perpetuate themselves indefinitely. This is explanation by formal cause. It's reasonable to suppose that additional light will be shed on this phenomenon by looking at it from the vantage point of material causation. But this is a matter for future research.

6. Who Goes There?

From a physical life point of view, certainly the most central question surrounding the existence of ET is: What do "they" look like? In other words, what kind of physico-chemical structure is an ET likely to have? This, of course, is basically a problem of material and efficient cause. And for the spectrum of possible answers, it's hard to do better than to consult the "hard" science fiction literature.

The creature depicted in Figure 1 below shows an artist's depiction of a Cygnan, a race of alien beings that threatens the Earth in Donald Moffitt's classic science-fiction novel *The Jupiter Theft.* The Cygnan is from a race of creatures the evolved on the satellite of a gas giant planet orbiting a binary star system. He (she? it?) is about $1\frac{1}{2}$ meters tall, with six limbs that can be used as either arms or legs, and a long, three-petaled tail that folds to conceal the sexual organs. The slender, tubular body is built on a cartilaginous skeleton, with the brain located between the upper pair of limbs at the top of the spinal cord. The three eyes are placed on stalks in an equilateral triangle around a broad, flexible mouth. The Cygnan has a harsh, rasping plate in the mouth, and a spiked, tubular tongue. It has a well-integrated nervous system, with much faster synaptic reflexes than those of a human being. Cygnan speech is musical, consisting of chords produced by multiple larynxes, and depends on absolute pitch. The language is incredibly rich and varied; it has more than a million phonemes, and each word is made up of several phonemes.

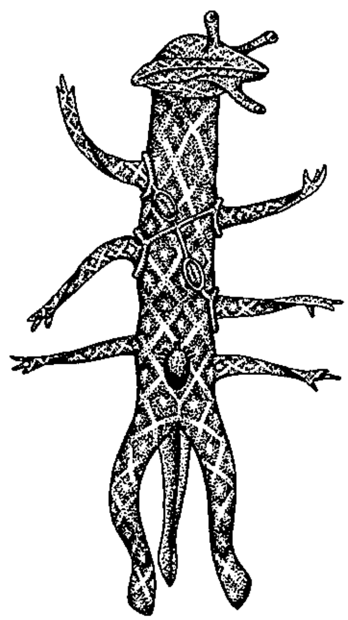

Figure 1. A Cygnan from *The Jupiter Theft*

So here we have a well worked-out example of an ET that differs radically from earthly organisms in almost every way that counts (unfortunately for the humans in Moffitt's story). But if we consider formal cause, the science-fiction literature offers up even stranger possibilities.

Figure 2 shows the Cryer, a creature from Joseph Green's book *Conscience Interplanetary.* The Cryer is an independently functioning unit of a planet-wide silicon-based plant intelligence inhabiting the planet Crystal, which has an atmosphere of 18 percent oxygen, the rest being nitrogen and hydrogen. Life on Crystal is based on silicon, with a high percentage of metallic elements.

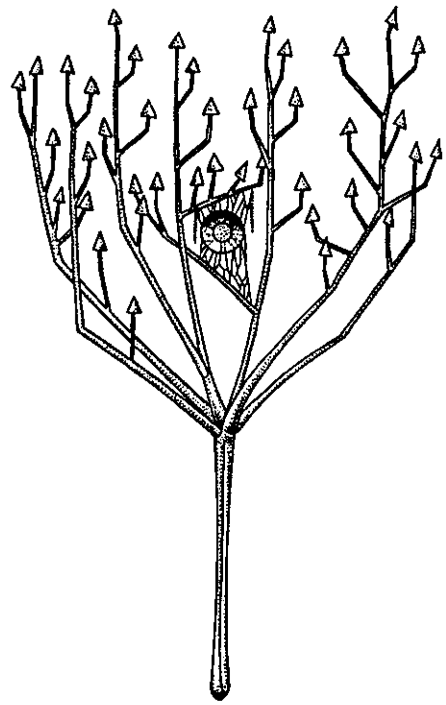

Figure 2. The Cryer from *Conscience Interplanetary*

The Cryer resembles a two-meter-high bush with a crystal-and-metal trunk and branches, with small, sharp glass leaves. The trunk contains silicon memory units, powered by a low-voltage solar storage battery and connected by fine silver wires. About six feet up the Cryer's trunk is an organic air-vibration membrane that enables the Cryer to speak with human beings. It is a broad, saucer-shaped leaf held in place by stretched wires to provide a vibrating diaphragm. A magnetic field generated in silver wire coils hanging on either side of the speaker causes it to vibrate and produce sound.

The planet-wide intelligence consists of thousands of smaller units like the Cryer, connected by an underground nervous system of fine silver wire. Each unit has a specialized function, some storing electricity generated by sunlight, some extracting silver for constructing the nervous system, some providing memory storage, and some acting as sensor units. The overall intelligence is able to perceive temperature, motion, position, electrical potential and vibrations through its member units.

Cygnans and Cryers are, of course, just thought experiments. Many more are reported in the imaginative volume[7]. By considersg the almost infinite range of possible life forms, and by examining the ET problem from all four of the Aristotelian causal perspectives, we can hope to get some feel for what may turn up on our radio telescopes or in our backyards one of these days. And, in fact, as the paper[1] points out, the metabolism-repair systems form what in mathematics is termed a "category" of objects. Therefore, each and every member of this category corresponds to a different biology, opening up the possibility of studying exobiologies by examining the mathematical structures and constraints imposed by these different objects in the category.

7. That's Life!

This paper has argued that to understand the difference between the living and the dead, it's not sufficient to concentrate just on the material aspects of living things. By adopting a less Newtonian and a more Aristotelian view of the problem, one is led to give equal time to explanations of life in terms of other causal categories besides the material. In particular, we have shown that there is much to be gained (and nothing to be lost but prejudices) by considering the other forms of life beyond those commonly-encountered here on Planet Three. The overall conclusion that emerges is that there are significant synergies possible between earth life, artificial life in machines and extraterrestrial life 'out there,' and that researchers in each area can benefit from paying careful attention to work by those in the other categories. To paraphrase a famous remark about war and generals, life is just too important to be left to the biologists.

References

1. Rosen, R., "Some Relational Cell Models: The Metabolism-Repair Systems," in *Foundations of Mathematical Biology,* Volume 2, Academic Press, New York, 1972.

2. Rosen, R., *Anticipatory Systems,* Pergamon, Oxford, 1985.

3. Rosen, R., *Life Itself,* Columbia University Press, New York, 1991.

4. Casti, J., "Linear Metabolism-Repair Systems," *Int'l. J. Gen. Sys.,* 14 (1988), 143–167.

5. Casti, J., "The Theory of Metabolism-Repair Systems," *Applied Math. & Comp.,* 28 (1988), 113–154.

6. Casti, J., "Newton, Aristotle, and the Modeling of Living Systems," in *Newton to Aristotle,* J. Casti and A. Karlqvist, eds., Birkhäuser, New York, 1989, pp. 47–89.

7. Jonas, D. and D. Jonas, *Other Senses, Other Worlds,* Stein and Day, New York, 1976.