

Kent Academic Repository

Full text document (pdf)

Citation for published version

Chu, Dominique and Ho, W. (2007) Computational Realizations of Living Systems. *Artificial Life*, 13 (4). pp. 369-381. ISSN 1064-5462.

DOI

<https://doi.org/10.1162/artl.2007.13.4.369>

Link to record in KAR

<https://kar.kent.ac.uk/14540/>

Document Version

UNSPECIFIED

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

Computational Realizations of Living Systems

Dominique Chu*
Computing Laboratory
University of Kent
Canterbury CT2 7NF
United Kingdom
D.F.Chu@kent.ac.uk

Weng Kin Ho
School of Computer Science
University of Birmingham
Birmingham B15 2TT
United Kingdom
W.K.Ho@cs.bham.ac.uk

Abstract Robert Rosen's central theorem states that organisms are fundamentally different to machines, mainly because they are "closed with respect to efficient causation." The proof for this theorem rests on two crucial assumptions. The first is that for a certain class of systems ("mechanisms") analytic modeling is the inverse of synthetic modeling. The second is that aspects of machines can be modeled using relational models and that these relational models are themselves refined by at least one analytic model. We show that both assumptions are unjustified. We conclude that these results cast serious doubts on the validity of Rosen's proof.

Keywords
Closure, (M,R) systems, analytic model, relational model, synthetic model

I Introduction

Computation in various forms has become a well-established part of the methodological spectrum in biological sciences. One aspect of this is the development of powerful database systems to store and efficiently retrieve vast amounts of experimental data. Computers are also increasingly used as modeling tools. Systems biologists [9, 10], for example, design detailed computational simulations of biochemical systems; at the moment these simulations are mostly confined to individual pathways, but the ultimate goal of the field is simulations of entire cells. Systems biology is not the first or only approach to modeling life in computers. Another field is artificial life (see for example [1, 19, 12]). Artificial life has its roots in computer science rather than biology. Accordingly its aim tends to be towards general underlying principles of organisms, with less emphasis on the detailed understanding of the chemical processes that go on within particular life forms.

Many aspects of organisms are very well modeled by and as computational processes [4, 3]. This does not necessarily mean that organisms as a whole are well described as or modeled by computational systems. In fact, the experience of research into artificial life seems to be (at least so far) that there is no useful abstraction of life that can be implemented as a computer program; the problem is not so much the lack of computational power as a lack of conceptual understanding of what precisely makes a system come alive [2, 8].

So far, these suggested limitations are only impressions, intuitions, or conjectures. Fact is that at the moment there is no convincing example of an implementation of a living system in silico; neither are there hard arguments to show why life cannot be modeled as a computational process. Yet one attempt to provide such arguments is a body of work by the late mathematical biologist

* Corresponding author.

Robert Rosen. In his 1991 book *Life itself* (LI) [21] Rosen presents an argument that living systems are fundamentally different to machines, in particular Turing machines; this result is often referred to as *Rosen's central theorem*. The proof is based on an analysis of various types of models and their power to describe classes of systems. An outline of the argument itself will be presented further below in Section 3.3; for the moment we will only introduce the most important concepts necessary to understand the central theorem.¹

- There are three types of models: *analytic*, *synthetic*, and *relational*. The last are essentially block diagrams summarizing high-level properties of systems. An example of a (simple) relational model is provided below (see Figure 2 in Section 3.2). Analytic and synthetic models will be defined in Section 2.
- Rosen defines a class of systems he calls *mechanisms*. We will not go into the details of the definition, but state two of the crucial properties of mechanisms: Firstly, in mechanisms analytic and synthetic modeling coincide. Rosen states that

if N is a mechanism, any mode of analysis whatsoever is equivalent to a process of anti-synthesis. Stated another way: in a mechanism, *analysis coincides with anti-synthesis*. [21, p. 212]

Rosen refers to this property as *fractionability*. Secondly every mechanism

has a unique largest model M^{\max} . [...] Epistemologically this model contains everything knowable about [the natural system] N , according to Natural Law. [21, p. 204]

- A particularly important class of mechanisms are machines. Turing machines are members of this class.
- Rosen then constructs a minimal relational model of an organism [sometimes called (M,R)-system]. He observes that this model exhibits a property that he called “closure with respect to efficient causation.” Loosely this means that all components of the system are created and repaired by other processes in the system.
- The final step is to show that systems that are closed with respect to efficient causation (i.e., in particular living systems) are not compatible with the idea of mechanism; they belong to a more general class of systems [16, 17]. He also shows that computational systems can at best be partial models of living systems. This is the central theorem.

Despite LI being first published in 1991, there is renewed interest in Rosen's work. Recent publications extending Rosen's work include Letelier and coworkers [14, 15], Wolkenhauer [25], and Casti [6, 5], to name but a few. Yet there are also a number of critical publications. Specifically one should mention Wells [23], who documents inconsistencies in Rosen's writings. Landauer and Bellman [11] point out mathematical errors in LI. Using a number of examples, they show that central parts of LI (in particular the minimal model of an organism) are mathematically questionable. While certainly illuminating, their approach is vulnerable to the objection that they simply chose the wrong examples to make their point. Chu and Ho [7] also looked at Rosen's minimal model of an organism. They argue that the entire concept of “relational modelling” that underlies Rosen's central

¹ The reader should note that the discussion in this article assumes some familiarity with Rosen's work. In addition to the original literature there exist a number of comprehensive reviews of Rosen's ideas [16, 17, 25, 22]; the reader who would like to familiarize herself with Rosen's work is encouraged to consult these.

theorem is inconsistently used by Rosen himself. The problem with Chu and Ho's article is that their treatment remains largely conceptual and lacks mathematical rigor.

Despite more or less clearly stated criticisms, all of the above authors seem to agree that Rosen's research program in theoretical biology is worthwhile to pursue. Yet, before his (or a similar) research program can be continued, it is necessary to understand whether or not Rosen's results (in particular the central theorem) hold. If the central theorem were indeed correct, this would have very wide implications on our understanding of living systems. However, in the present article we will show that Rosen's proof does not stand up to scrutiny.

This article is organized as follows. The next Section 2 defines analytic and synthetic models and derives a theorem describing the relation between these two models. The conclusion of that section will be that for finite systems analysis and anti-synthesis are never equivalent, thus contradicting a central tenet of Rosen's. This is followed by a discussion Section 3. Section 3.1 discusses the connection between analytic and synthetic models on the one hand and direct sums and products on the other. This section is not strictly required for the overall conclusion of this article; it was included anyway because direct products and sums feature prominently in Rosen's own discussion. Section 3.2 will show that even in the realm of mechanisms relational models cannot be extracted from supposedly largest analytic models. In Section 3.3 Rosen's proof of the central theorem will be briefly outlined. This outline will clearly show that the proof is based on false premises, namely that for mechanisms (i) analysis and anti-synthesis are equivalent and (ii) there is a largest analytic model that contains everything knowable about the system. Finally, Section 4 concludes this article.

2 Analytic and Synthetic Models

In this section we will define the notions of analytic and synthetic models; the objective is to do so as they were introduced by Rosen [21, 20]; furthermore, this section will provide some insight into the relationship between these two types of models, as they are relevant for the understanding of the central argument. This will culminate in the formulation of Theorem 1.

2.1 Analytic Models

We begin by defining analytic models (cf. [21, Chap. 6C]).

DEFINITION 1: *Let f be a mapping from a finite set S to some set U . An analytic model $M(S)$ of the system S is the set of equivalence classes on S induced by f . We say that two elements $x, y \in S$ are equivalent and write $x \sim_f y$ if $f(x) = f(y)$. The equivalence class of x on S induced by f is denoted by $[x]_f$ and is a subset of S . We denote by $\mathcal{M}^a(S)$ the set of all analytic models of S .*

Remark 1. The set S represents the system to be modeled and can be thought of as the set of states the system can take (for a more in depth discussion of this assumption see [7, Section 2.2]). If S is a natural system, then it is usually unknown; note that we assume here that S has a finite number of elements; this assumption will simplify the analysis to follow but, as it will turn out, will not affect our overall conclusions. The system can be probed via meters that indicate values of observables; in Definition 1, the observable is the function f , and the measurement results are elements of the set U .

We will henceforth indicate the fact that the model $M(S)$ is generated by the observable f by writing $M^f(S)$, or simply M^f . Given a specific system S , there typically exist a great number of different analytic models. Models can sometimes be compared with one another with respect to their degree of refinement, that is, the degree to which they differentiate between elements of S . Such models are called *compatible*. Two models may also describe unrelated properties of a system, in which case they cannot be compared and are *incompatible*.

DEFINITION 2: Given the analytic models M^f and M^g , we say M^f and M^g are compatible if either of the following conditions is true:

- $\forall x \in S [x]_f \subseteq [x]_g$,
- $\forall x \in S [x]_g \subseteq [x]_f$.

If there is no clear containment relation between the equivalence classes induced by one model and the other, then we say that these models are *incompatible*. Some equivalence classes of M^f might contain (or be contained in) equivalence classes of M^g , but if M^g and M^f are incompatible, then there must be at least one equivalence class of one model that has nonempty intersections with at least two equivalence classes of the other. This is expressed in the following definition.

DEFINITION 3: Given analytic models M^f and M^g , we say M^f is incompatible with M^g if $\exists x, y, z \in S$ s.t. $x \sim_f y, y \sim_g z, y \not\sim_f z$, and $x \not\sim_g y$.

It is easy to show that incompatibility is a symmetric relation, as it should be.

PROPOSITION 1: *Incompatibility is a symmetric relation.*

Proof. M^g is incompatible with M^f if $\exists x, y, z \in S$ s.t. $z \neq y, x, z \sim_g y, y \sim_f x, y \not\sim_g x$, and $z \not\sim_f y$. Through reordering of the variables in the equivalence relations one obtains the same set of conditions as in Definition 3: $y \sim_g z, x \sim_f y, x \not\sim_g y$, and $y \not\sim_f z$. This completes the proof.

Whenever two models are incompatible, they are not compatible, and vice versa. We will now show that this is the case.

PROPOSITION 2: *Incompatibility is the negation of compatibility.*

Let us first show that M^f, M^g compatible implies M^f, M^g not incompatible. If M^f, M^g are incompatible, then there exist $x, y, z \in S$ such that the following relations are true:

$$[x]_f = [y]_f, \quad [y]_g = [z]_g,$$

$$[y]_f \parallel [z]_f, \quad [x]_g \parallel [y]_g,$$

where $A \parallel B$ means that A and B are disjoint sets. This is Definition 3 written in a different form. If M^f, M^g are compatible, then we are entitled to assume that $[\omega]_f \subseteq [\omega]_g$, for all $\omega \in S$; this follows directly from the definition of compatibility. If we assume that M^f and M^g are both compatible and incompatible, then $[y]_f \subseteq [y]_g = [z]_g$. This means that the intersection of $[y]_f$ and $[z]_g$ contains at least the element y , thus contradicting the condition $[y]_f \parallel [z]_g$. Hence if two models are compatible, they are not incompatible.

Let us now show that the opposite direction is also true, that is, that not compatible implies incompatible, or equivalently that not incompatible implies compatible. If M^f, M^g are not compatible, then (without restricting generality) this means that there is at least one equivalence class Y_f of g that intersects at least two (disjoint) equivalence classes X_f, Z_f of f . Choose $x \in X_f \setminus (X_f \cap Y_g), z \in Y_g \setminus (X_f \cap Y_g)$, and $y \in X_f \cap Y_g$. It is straightforward to show that x, y, z fulfill the definition for incompatibility. This completes the proof.

There are various degrees of incompatibility. For the current purpose we will be specifically interested in models that measure completely different aspects of a system in the sense that knowing

about one does not tell us anything about the other, not even probabilistically; these models are then *totally incompatible*. Formally, total incompatibility means that each equivalence class of M^f intersects each equivalence class induced by M^g . Note that the models M^f and M^b in Figure 1 are totally incompatible.

DEFINITION 4: Given nontrivial analytic models M^f and M^g (i.e., both models induce at least two equivalence classes on S), we say M^f is totally incompatible with M^g if $\forall x \in S$ it is true that $\forall y \in S, S \supset [x]_f \cap [y]_g \neq \emptyset$.

Remark 2. Note that a pair of models is always compatible if one of the models is trivial, that is, if its function is constant and it induces only one equivalence class on S (namely S itself).

PROPOSITION 3: Total incompatibility implies incompatibility.

Proof. We want to show that if M^f, M^g are totally incompatible, then they are incompatible, that is, there exists a set of x, y, z that fulfills the equivalence relations in Definition 3. We start by choosing an arbitrary x and a set $U' = S \setminus [x]_g$. The set U is nonempty.

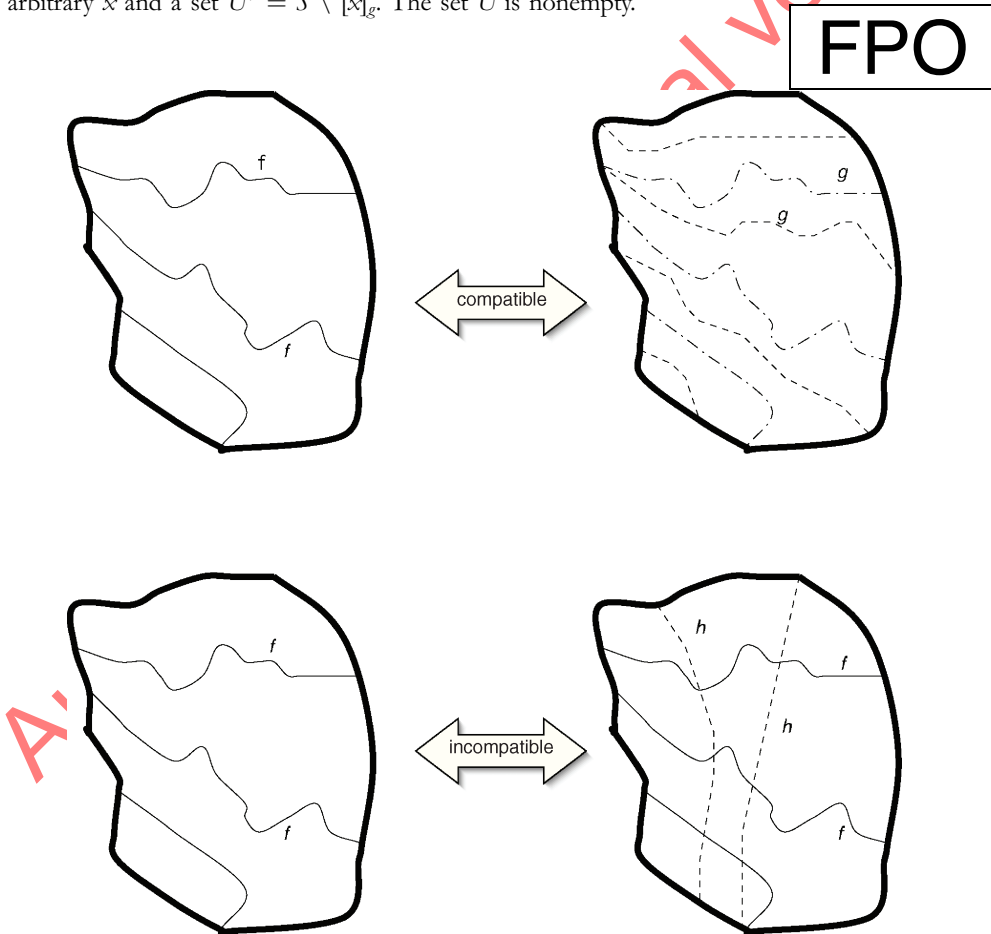


Figure 1. This illustrates the concept of compatibility and incompatibility. The system S is indicated by the thick closed line. The left hand side shows how the function f partitions S into equivalence classes (in this case into four classes). The right hand side (top) shows a function g that is compatible with f . Every equivalence class induced by g is contained in an equivalence class of f . The graph on the bottom right shows a function h that is incompatible with f (it is in fact *totally incompatible*).

Choose a point $\omega \in U'$. We know that the set $U = [x]_f \cap [\omega]_g$ is nonempty. Choose an element $y \in U$. It is now true that $x \sim_f y$ and $x \not\sim_g y$. This is the first half of the incompatibility conditions. To show the second half, choose an element $\omega' \in V'$ where $V' = S \setminus [x]_f$. Again it is clear that $V' \neq \emptyset$. We also know that the set $V = [\omega']_f \cap [y]_g$ is nonempty. Choose a $z \in V$. For any such z it will be true that $x \not\sim_f z$ and $y \sim_g z$, which completes the proof.

Given that two models are compatible, we can compare them in terms of how well they distinguish between states of S . Note that such a comparison is not well defined if the two models are not compatible. This leads us to the notion of refinement of models.

DEFINITION 5: *Given analytic models M^f and M^g , we say M^f is a refinement of M^g if every equivalence class corresponding to \sim_f is contained in one equivalence class induced by \sim_g . We use the notation $M^f \geq M^g$.*

Remark 3. There is a very close relation between refinement and compatibility. In fact, whenever two models are compatible, there is a refinement relation between them, and whenever there is a refinement relation between models, they are compatible. The distinction between those two concepts is twofold: Firstly, there is a subtle difference in meaning of refinement and compatibility. The former expresses the fact that of two models one is “better” than the other, whereas compatibility only means that it is meaningful to compare two models in the first place. Secondly, compatibility is a symmetric relation, whereas refinement is not.

Refinement relations between models introduce a partial order into $\mathcal{M}^a(S)$. Hence $\mathcal{M}^a(S)$ is a *partially ordered set*, or simply *poset*. In the remainder of this subsection we will explore the structure of this partial order.

PROPOSITION 4: *The set of all analytic models of S , $\mathcal{M}^a(S)$, is a partially ordered set with respect to the refinement relation \geq .*

Proof. Trivial, since such a refinement relation is obviously reflexive, antisymmetric, and transitive. This completes the proof.

Posets are a category in the sense of category theory (see for example [18, 13]). This is of some interest in the current context, because on posets there is a clear notion of direct products and direct sums of models (see for example the discussion by Chu and Ho [7]). Rosen himself seems to assign some significance to this observation, as evidenced by the prominence of the notion in the overall discussion in [21]. Yet, as will be discussed below (see Section 3.1), in the current context it seems that there is no fundamental significance to direct products and sums. This thread will therefore not be followed up any further.

Instead, we will now continue to ask how analytic models are related to one another. We start by defining the mutual refinement of two models. Even if two models are in no special order relation, one can combine the incompatible aspects of both to form a new model. This new model will be a refinement of both.

PROPOSITION 5: *Let M^f and M^g be incompatible analytic models of S with $f: S \mapsto U$ and $g: S \mapsto V$. Then the analytic model M^F of S induced by $F = (f, g): S \mapsto U \times V$ is the least mutual refinement of M^f and M^g . We will henceforth write $M^F = M^f \otimes M^g$.*

Proof. First let us show that M^F is indeed a refinement of M^f and M^g . In order to show that M^F is a refinement of M^f it is sufficient to show that every equivalence class $[x]_F \subseteq [x]_f$. This means that for all $x_1, x_2 \in S$ it is true that if $F(x_1) = F(x_2)$ then $f(x_1) = f(x_2)$. This is trivial from the definition of F . The analogous argument can be made for g .

Next we need to show that it is indeed the least mutual refinement. Assume M^G refines both f and g . This means that every equivalence class of G is contained in an equivalence class of f and of g , that is, if $G(x) = G(y)$ then also $f(x) = f(y)$ and $g(x) = g(y)$; this however means precisely that G refines $F = (f, g)$. Therefore F is the least mutual refinement of f and g . This completes the proof.

We know now that given two analytic models, we can always combine them into a refined model. The next proposition states that it is also true that, given an analytic model (that is not entirely trivial), it is always possible to find two incompatible models of which it is the mutual refinement.

PROPOSITION 6: *If $M^F \in \mathcal{M}^a(S)$ is an analytic model partitioning S into $n > 2$ equivalence classes $S = \cup_i X_i$, then there exist two models $M^f, M^g \in \mathcal{M}^a(S)$ that fulfill the following two conditions:*

1. M^f and M^g are incompatible.
2. M^F is the least mutual refinement of both M^f and M^g .

Proof. The model M^F can be represented by the set of equivalence relations it induces on S ; we assume that S is partitioned by F into several equivalence classes $X_k = [x_k]_F$. Two cases can be distinguished.

The first case is that F induces an uneven number of equivalence classes, that is, $S = X_1 \cup X_2 \cup \dots \cup X_{2m+1}$. We can define a new observable by its set of equivalence classes. Choose f, g to be given by the following:

$$f \doteq \{X_1 \cup X_{m+1}, X_2 \cup X_{m+2}, \dots, X_m \cup X_{2m}, X_{2m+1}\}$$

$$g \doteq \{X_1 \cup X_2 \cup \dots \cup X_m, X_{m+1} \cup X_{m+2} \cup \dots \cup X_{2m+1}\}$$

By this we mean that the new observable f is obtained from F by merging equivalence classes respectively. M^f and M^g are clearly incompatible. The proof for $n = 2m$ is analogous. This proves the first part of the proposition.

The second part can be seen as follows: M^f can be refined by successively de-merging unions of X_i ; this leads to a series of new models of which only the last one, where all X_i are de-merged, is also a refinement of M^g . Similarly, one can argue with successive refinements of M^g . In both cases, the only model that refines both M^f and M^g is M^F . Hence, M^F is the least mutual refinement of both M^f and M^g . This completes the proof.

Without providing a proof, we will also state that an analytic model can only under certain circumstances be considered the least mutual refinement of two other analytic models.

PROPOSITION 7: *Models M^f, M^g in Proposition 6 are totally incompatible if and only if the number of equivalence classes induced by F on S is even.*

The next proposition states that it is meaningful to say that a model is a minimal refinement of another model (as opposed to the minimal mutual refinement of a *pair* of models). In fact, every model is the minimal refinement of a whole set of other models in $\mathcal{M}^a(S)$. We will henceforth refer to those models as *minimally coarsened* models of M^F .

PROPOSITION 8: *If M^F is a model and $\exists x_1, x_2, \dots, x_n \in S$, $n > 2$, s.t. $\forall i, j \leq n$, $i \neq j$: $[x_i]_F \cap [x_j]_F = \emptyset$, then there exist $l = \sum_{i=1}^{n-1} i$ models of S such that for each of those models M^i it is true that $M^F > M^i$ and for all $M^G > M^i$ it is true that $M^G \geq M^F$. We will henceforth refer to models M^i as the minimally coarsened models of M^F .*

Proof. The proposition simply states that every observable that induces more than three equivalence classes on S can be coarsened in l different ways. This is easy to see: The only way to coarsen an observable without losing compatibility is to pairwise merge equivalence classes. There are exactly l ways to merge equivalence classes. This completes the proof.

Thus there are two modes of refinement: mutual refinement between incompatible models, and refinement of a single model. This latter mode of refinement corresponds to the splitting of equivalence classes. Correspondingly every model is associated with a number of minimally coarsened models.

These results allow us now to describe the structure of the poset $\mathcal{M}^a(S)$ of all analytic models of S . There exists a maximal model M^{\max} , that is, $\forall M \in \mathcal{M}^a(S)$ it is true that $M^{\max} \geq M$. This maximal model corresponds to an observable f_{\max} with the property $\forall x, y \in S$, $x \neq y$, $f_{\max}(x) \neq f_{\max}(y)$. Furthermore, there is a minimal model M^{\min} with the property $\forall x, y \in S$, $x \neq y$, $f_{\min}(x) = f_{\min}(y)$. This minimal model is trivial and refined by every model. Finally there is a set of simplest nontrivial models, of which M^{\min} is the minimal coarsening. Each of those models partitions S into exactly two equivalence classes; we will henceforth refer to those models as *bottom models*.

2.2 Synthetic Models

We will now turn our attention to synthetic models. A synthetic model is essentially the “union” of two analytic models of disjoint systems. So, for example, if one has two (separate) models of two (separate) particles, then one can formally merge those two models in one larger model. This new model would be a model of a larger system encompassing two particles. We will now formulate this idea more precisely:

DEFINITION 6: *Suppose S_1 and S_2 are (disjoint) systems, functions $f_1 : S_1 \mapsto U$, $f_2 : S_2 \mapsto V$ are not surjective, and M^{f_1} and M^{f_2} are analytic models of S_1 and S_2 respectively. Then $M^{\tilde{F}} = M^{f_1} \oplus M^{f_2}$ is a synthetic model of $S = S_1 \sqcup S_2$ (the disjoint union of S_1 and S_2), where $F = (\tilde{f}_1, \tilde{f}_2) : S \mapsto U \times V$ and $\tilde{f}_1 : S \mapsto U$, $\tilde{f}_2 : S \mapsto V$ are as follows:*

$$\tilde{f}_1(x) = \begin{cases} f_1(x) & \text{if } x \in S_1 \\ c_1 & \text{if } x \in S_2 \end{cases} \quad \tilde{f}_2(x) = \begin{cases} f_2(x) & \text{if } x \in S_2 \\ c_2 & \text{if } x \in S_1 \end{cases}$$

where

$$\forall x \in S_1 : c_1 \neq f_1(x)$$

$$\forall x \in S_2 : c_2 \neq f_2(x)$$

We denote by $\mathcal{M}^s(S)$ the set of all synthetic models of S .

Remark 4. In Definition 6 we require that f_1, f_2 be not surjective. This requirement does not limit the generality of the definition, but is necessary in order to ensure that there exist suitable constants c_1, c_2 . In case f_1 or f_2 are surjective, it is always possible to extend the sets U and V by one more element. This would be sufficient for the current definition.

This definition immediately leads to a first observation regarding the relation between the variables that are induced by the new system:

PROPOSITION 9: *If $M^F \in \mathcal{M}^s(S)$ with $M^{\bar{f}_1}, M^{\bar{f}_2} \in \mathcal{M}^a(S)$ and $M^F = M^{\bar{f}_1} \oplus M^{\bar{f}_2}$, and furthermore F, \bar{f}_1, \bar{f}_2 are as in Definition 6, then $M^{\bar{f}_1}, M^{\bar{f}_2}$ are not totally incompatible.*

Q2

Proof. Assume $S = S_1 \sqcup S_2$, $x_1 \in S_1$, and $x_2 \in S_2$. In order to be totally incompatible all equivalence classes of \bar{f}_1, \bar{f}_2 need to intersect. This is not the case. A counterexample is

$$\underbrace{[x_1]_{\bar{f}_1}}_{\subset S_1} \cap \underbrace{[x_2]_{\bar{f}_2}}_{\subset S_2} = \emptyset$$

because S_1 and S_2 are disjoint. This completes the proof.

Both analytic and synthetic models are associated with activities, namely the analysis and synthesis of models. Synthesis is simply the merging of two analytic models into a larger analytic model. By analysis of a model we essentially mean the process of decomposing a model M^F into pairs of models of which it is the mutual refinement, that is, finding models M^f, M^g such that $M^F = M^f \otimes M^g$. There are now two questions we want to ask. Firstly, under which circumstances are analytic models synthetic models, that is, when is an analytic model of the form $M^F = M^{\bar{f}_1} \oplus M^{\bar{f}_2}$? The answer to this, as it will turn out is that they always are. The second question we want to ask is: Under which circumstances are synthesis and analysis simply inverse processes?

DEFINITION 7: *We say that analysis and synthesis of S are inverse if for all analytic models $M^F \in \mathcal{M}^a(S)$ and all pairs of models $M^{\bar{f}_1}, M^{\bar{f}_2}$ with $M^F = M^{\bar{f}_1} \otimes M^{\bar{f}_2}$ there exists a set $S_1 \subset S$ s.t. $\forall x, y \in S_1$ we have $f_2(x) = f_2(y)$ and furthermore $\forall x, y \in S \setminus S_1$ we have $f_1(x) = f_1(y)$.*

Remark 5. Compare Definition 7 with Definition 6 of synthetic models. The essence of this definition is that we can only think of a pair of models $M^{\bar{f}_1}, M^{\bar{f}_2}$ as splitting S into two parts if f_1 and f_2 are constant on complementary parts of S .

We can now turn to the question under which circumstances analytic models are synthetic models and when analysis is anti-synthesis. The answer to this is summarized in the following theorem.

THEOREM 1: *Given a nontrivial analytic model $M^F \in \mathcal{M}^a(S)$ with the cardinality of S greater than 3, then M^F is always a synthetic model, but analysis and synthesis are not inverse.*

Proof. Let us first show the second statement is true. Assume it is false, that is, for some systems S analysis and synthesis are inverse. This statement is false, as Proposition 9 together with Proposition 7 provides counterexamples: If two models $M^f, M^g \in \mathcal{M}^a(S)$ are totally incompatible, then S cannot be split into complementary parts as required in Definition 7.

For the first part of the theorem, consider as an example a model M^F that partitions S into five equivalence classes $\{X_1, X_2, X_3, X_4, X_5\}$. In order to show that the theorem holds in this particular case, we need to find functions f_1, f_2 that partition S into two parts as required in Definition 6. There exist several such pairs of functions. One is as follows:

$$\bar{f}_1 \doteq \{X_1, X_2, X_3, \underbrace{X_4 \cup X_5}_{=S_2}\}$$

$$\bar{f}_2 \doteq \{\underbrace{X_1 \cup X_2 \cup X_3}_{=S_1}, X_4, X_5\}$$

From this it is clear that (i) S_1 and S_2 are disjoint but $S = S_1 \sqcup S_2$, (ii) M^F is the mutual refinement of M^{f_1} and M^{f_2} , and (iii) f_1, f_2 are of the form required for synthetic models (as in Definition 6). Hence in this particular case M^F is a synthetic model.

We will not prove the general case, but point out that the technique applied to this particular case can be analogously used for any model $M^{F'}$ as long as $M^{F'}$ partitions S into at least two equivalence classes.

3 Discussion

This theorem shows that the relation between analytic and synthetic models is in general not one of unconditional equivalence, that is, analysis is not anti-synthesis. For a given finite system S one will be able to find models M^f, M^g such that analysis and synthesis are inverse in the sense defined above, but there exist no finite systems such that this is always true. This result is not entirely surprising: Analysis and synthesis are formally very different. Whereas a mutual refinement of two incompatible models of S gives another model of the system S , the synthesis of two models of S gives a model of a system $S^U = S \sqcup S$, hence of an entirely different system.

Before continuing with the main discussion, it will be necessary to answer to a possible objection to what is to follow: The notion of mechanisms does not enter anywhere in Section 2. Theorem 1 states that analysis is never anti-synthesis. This might be right, one could object, but Rosen only claimed analysis and anti-synthesis to be equivalent for mechanisms, not in general. Hence there is no contradiction between Rosen and Theorem 1.

This is not so. The theorem only assumes that S has more than three but a finite number of elements. Apart from that, no specific assumptions are made. Mechanisms are certainly a subclass of finite systems (see Rosen [21]). Hence, the theorem applies in particular to mechanisms. In essence the theorem is a consequence of the fact that totally incompatible models do not split a system neatly into two disjoint parts; or inversely, if they do split a system into two parts, then they are not totally incompatible (see Proposition 9). This conclusion holds whether or not a system is a mechanism.

3.1 Direct Products and Sums

One leitmotif of LI is the equivalence of direct products and direct sums in the realm of mechanisms; this is itself just another aspect of the claimed equivalence of analysis and anti-synthesis. Given the current treatment, the significance of this claim is hard to see.

Analytic models can be associated with direct products in two different ways: Firstly, the operation of mutual refinement of models M^f, M^g is essentially the process of taking the Cartesian product of two observables f, g to build the new observable $F = (f, g)$. The resulting state space of the new model is the direct product of the state spaces of the two original models. Analytic models are also associated with direct products in a category theoretic sense: $\mathcal{M}^a(S)$ is a partially ordered set and thus a category (see [18, 13]). This means that the model M^F can be interpreted as the direct product of M^f and M^g ; see Chu and Ho [7] for an explanation. Similarly, the set $\mathcal{M}^s(S)$ is also partially ordered. This order can be defined in such a way that the synthesis of two models M^f and M^g is their direct sum (see [7]).

This association of analytic (synthetic) models with direct products (sums) is however not as clear as it seems at a first glance. Firstly, while it is true that the state space of $M^F = M^f \otimes M^g$ is the Cartesian product of the state spaces of M^f and M^g , it is similarly true that the state space of $M^F = M^{f_1} \oplus M^{f_2}$ is the Cartesian product of the state spaces of M^{f_1} and M^{f_2} . Secondly, whether or not M^F is the direct product or the direct sum of models M^f, M^g (in a categorical sense) entirely depends on the interpretation of the partial order relation in terms of categorical maps, that is, whether $M^F \leq M^G$ is interpreted as $M^F \mapsto M^G$ or $M^G \mapsto M^F$. The two choices lead to dual categories; what is a direct product in one category is a direct sum in the other. In this sense, whether or not analytic models are associated with the direct product or the direct sum is a question of choice, not a fundamental property of the system. The same is true for the poset of synthetic models.

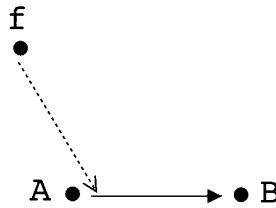


Figure 2. Rosen's ("relational") model of a Turing machine. The vertex f denotes the hardware that induces a software flow from input (A) to output (B). Based on Figure 9B.3 in LI.

Finally, even if one decided for one or the other reason to associate analytic and synthetic models with direct products and sums respectively, then their identity in the realm of mechanisms would still not hold (this is again a consequence of the above theorem). In summary, the significance of direct products and sums in the current context is hard to see and, most of all, not grounded in any mathematical fact.

3.2 Models of Machines and Mechanism

Rosen's argument crucially rests on the notion of analytic model. This notion can be objected to on several grounds. A detailed discussion and criticism, particularly of the philosophical underpinnings of analytic models, would go beyond the scope of the present contribution. We will thus only briefly indicate that the notion of analytic model is not consistently used by Rosen; this inconsistency will turn out to be relevant for the validity of Rosen's central theorem.

One of Rosen's basic assumptions is that (at least) every mechanism has a unique maximal (analytic) model. However, in what follows we will suggest that this is not the case, that is, analytic models in Rosen's sense encode nearly nothing about a concrete system. In particular, even in the realm of mechanisms, relational models of systems are not always recoverable from the largest analytic model of a system.

The difficulty for the analysis of relation models and how they relate to analytic models is that Rosen's own discussion leaves the notion of relation model somewhat unclear. In this particular case that turns out not to be a problem. This is best illustrated using one of Rosen's examples of a relational model of a Turing machine (TM) as reproduced in Figure 2. While the precise meanings of the nodes and arrows of this relational model is left unexplained in LI, f seems to represent the reading head of the TM. The nodes A and B are input and output software configurations; the transition from A to B is effectuated by the reading head f . A detailed justification and explanation of this relational model is both outside the scope of this article (and in fact, there are arguments that it is invalid as a model [23]) and unnecessary for what follows. For the sake of the argument, let us simply assume that Figure 2 is indeed a model of a TM as suggested by Rosen.

According to Rosen, any TM is a member of the class of mechanisms, and thus there exists a largest model M^f of the TM capturing everything knowable (see quotation above); in particular it contains all the information that is contained in the relational model. Hence, the relational model should be recoverable from the largest analytic model. In what follows we will show that this is not so, thus suggesting that either the notion of relational modeling is meaningless or the largest analytic model does not encode all information about the TM.

We will assume that the TM is implemented in a $n \times n$ cellular automaton (CA) c where each cell has k states (see for example Wolfram [24]). The CA c is certainly also a mechanisms; hence according to Rosen we are entitled to assume that there exists a (unique) largest model of c , namely,

$$F : c \mapsto \{1, 2, \dots, k\}^{n^2} \quad (1)$$

This model encodes everything there is to know about c , and hence specifically it should also encode the relational model of Figure 2.

Let us now assume that there is a second automaton c' that is identical to c in all aspects except in its initial state configuration. Assume, for example, that every cell of c' takes a random initial state.

In nearly all configurations c' will then not implement a TM. On the other hand, if the model in Equation 1 is a model of c , then it is also a model of c' . This must be the case because the only difference between c and c' is that they start from different initial conditions; these however do not enter into the relational model. At the same time c' does not implement a TM and is thus not modeled by a relational model as in Figure 2.

These observations allow for two possible conclusions: (i) If M^F is indeed a complete model of c and c' , then the relational model in Figure 2 is not a model of the TM, because it is not recoverable from M^F ; it cannot be recoverable, because M^F is also a complete model of c' that is not modeled by the relational model. (ii) If on the other hand Figure 1 is a model of the TM, then there are aspects of the TM that are not captured by M^F , that is, there exists no largest analytic model of the TM.

3.3 Rejecting Rosen's Central Argument

We are now in the position to reexamine the validity of Rosen's central proof that living systems are fundamentally different from mechanisms and machines. For the current purpose it will be sufficient to reproduce an outline of Rosen's proof; the reader who is not familiar with the material is encouraged to consult LI [21, Chap. 9]. The main structure of the proof is as follows:

1. *Assumption:* R is a minimal relational model capturing essential properties of a living system S (closure with respect to efficient causation). S is a mechanism.
2. *Premise:* If S is a mechanism (or machine), then it is fractionable, that is, analysis and anti-synthesis are equivalent.
3. *Premise:* If S is a mechanism (or machine), then there exists a largest model that contains all information about R .
4. Rosen shows that assumption 1 and premises 2 and 3 are not consistent, that is, if S is fractionable and R is a model of S , then there exists no largest model of S that captures R .
5. *Conclusion:* S is not a mechanism.

The reader who has followed so far will immediately see that premises 2 and 3 are invalid. Premise 2 is contradicted by Theorem 1; premise 3 is contradicted by the conclusion of Section 3.2 stating that at least some machines have no largest model that contains all the information about their relational models.² Thus showing that R does not have a largest analytic model does not tell us anything about the relation between organisms and mechanisms.

4 Conclusion and Outlook

Altogether this is a strong indication that the proof of Rosen's theorem needs to be rethought before it can be accepted as correct. Despite serious doubts about the correctness of the proof of the central theorem, we believe that Rosen's research program is a worthwhile one to pursue. Probably one of the more important aspects of Rosen's legacy is the notion of closure with respect to efficient causation. The idea is that organisms continuously build components that themselves are necessary to build other components. This chain of component-building components is closed in such a way that the structure as a whole can function autonomously (in this context see Letelier and coworkers [14]). There are arguments that such a closed system is difficult to achieve in computational systems [8].

Are mechanisms fundamentally different to living systems? Is computational artificial life possible? These questions are certainly still open. Rosen's central theorem might be correct, but as it stands now it is unproven. We think that the next step should be to find a novel mathematical ansatz for a modeling framework of both life and mechanisms to provide new insights into the organizational

² The alternative conclusion (i) of Section 3.2 is that Rosen himself failed to understand his own concept of relational model.

features of life, thus also clarifying the relation between living systems and mechanisms. In particular, we believe that attempting to salvage Rosen's proofs is perhaps not worthwhile, although his intentions and aims will certainly continue to inspire research in theoretical biology and artificial life.

Acknowledgments

We thank Dr. Boris Mitavskiy for comments on draft versions of this manuscript. D. C. thanks the Microsoft Research–University of Trento Centre for Computational and Systems Biology, where parts of this work were done for hospitality.

References

1. Adami, C. (1998). *Introduction to artificial life*. New York: Springer-Verlag.
2. Bedau, M., McCaskill, J., Packard, P., Rasmussen, S., Green, D., Ikegami, T., Kaneko, K., & Ray, T. (2000). Open problems in artificial life. *Artificial Life*, 6(4), 363–376.
3. Bhalla, U. (2003). Understanding complex signaling networks through models and metaphors. *Progress in Biophysics and Molecular Biology*, 81, 45–65.
4. Bray, D. (1995). Protein molecules as computational elements in living cells. *Nature*, 376, 307–313.
5. Casti, J. (1992). *Reality rules: II The frontier*. New York: John Wiley & Sons.
6. Casti, J. (2002). Biologizing control theory: How to make a control system come alive. *Complexity*, 7(1), 10–13.
7. Chu, D., & Ho, W. (2006). A category theoretical argument against the possibility of artificial life. *Artificial Life*, 12(4), 117–135.
8. Groß, D., & McMullin, B. (2003). The creation of novelty in artificial chemistries. In *Proceedings of the Eighth International Conference on Artificial Life* (pp. 400–409). Cambridge, MA: MIT Press.
9. Kitano, H. (2002a). Computational systems biology. *Nature*, 420, 206–210.
10. Kitano, H. (2002b). Systems biology: A brief overview. *Science*, 295, 1662–1664.
11. Landauer, C., & Bellman, K. (2003). Theoretical biology: Organisms and mechanisms. In D. Dubois (Ed.), *Computing anticipatory systems: CASYS 2001—Fifth International Conference* (pp. 59–70). Liege: Springer-Verlag.
12. Langton C. Ed. (1989) *Artificial life*. Reading, MA: Santa Fe Institute, Addison-Wesley.
13. Lawvere, F., & Schanuel, S. (1997). *Conceptual mathematics*. Cambridge, UK: Cambridge University Press.
14. Letelier, J., Marin, G., & Mpodozis, J. (2004). Autopoietic and (M,R) systems. *Journal of Theoretical Biology*, 222, 261–272.
15. Letelier, J., Soto-Andrade, J., Abarzua, F. G., Cornish-Bowden, A., & Cardenas, M. (2006). Organizational invariance and metabolic closure: Analysis in terms of (M,R) systems. *Journal of Theoretical Biology*, 238, 949–961.
16. Louie, A. (2005). Any material realization of the (M,R) systems must have noncomputable models. *Journal of Integrative Neuroscience*, 4(5), 423–436.
17. Louie, A. H. (2006). (M,R)-systems and their realizations. *Axiomathes*, 16(1–2), 35–64.
18. Pierce, B. (1991). *Basic category theory for computer scientists*. Cambridge, MA: MIT Press.
19. Ray, T. (1996). An approach to the syntheses of life (pp. 111–145). Oxford, UK: Oxford University Press.
20. Rosen, R. (1978). *Fundamentals of measurement and representation of natural systems*. New York: Elsevier North-Holland.
21. Rosen, R. (1991). *Life itself*. New York: Columbia University Press.
22. Rosen, R. (1999). *Essays on life itself*. New York: Columbia University Press.
23. Wells, A. (2006). In defense of mechanism. *Ecological Psychology*, 18(1), 39–65.
24. Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media Incorporated.
25. Wolkenhauer, O. (2002). Systems biology: The reincarnation of systems theory applied to biology? *Briefings in Bioinformatics*, 2(3), 258–270.