A.J. Wells

In defense of mechanism

Article (Accepted version)
(Refereed)

Original citation:
Wells, Andrew J. (2006) In defense of mechanism. *Ecological psychology*, 18 (1). pp. 39-65
DOI: 10.1207/s15326969eco1801_2

In Defence of Mechanism

A.J. Wells

Institute of Social Psychology

The London School of Economics and Political Science

London, United Kingdom

Abstract

In *Life Itself* and in *Essays on Life Itself*, Robert Rosen (1991, 2000) argued that machines were, in principle, incapable of modelling the defining feature of living systems which he claimed to be the existence of closed causal loops. Rosen's argument has been used to support critiques of computational models in ecological psychology. This paper shows that Rosen's attack on mechanism is fundamentally misconceived. It is, in fact, of the essence of a mechanical system that it contains closed causal loops. Moreover, Rosen's epistemology is based on a strong form of indirect realism and his arguments, if correct, would call into question some of the fundamental principles of ecological psychology.

Introduction

The pregnant question ‘What is Life?’ has received many different types of answer derived from a variety of different research strategies. One strategy is to examine the denizens of the world and to look for a common element or condition possessed by all living things and by nothing else. That common element, once identified, can then be proclaimed as the answer to the question. There are many difficulties with this strategy. One is that it is circular. It presupposes that living things can be identified in order that the essence of life can be extracted from them. A second difficulty is the essentialist assumption that there is a common thread binding all living things together. The forms of life may simply be too varied and disparate for this to be true.

Another strategy is to proceed by a process of elimination. By becoming clear about what life is not, one hopes to become clearer about what it is. The bio-physicist Robert Rosen adopted essentially this strategy in his book *Life Itself*, Rosen (1991). Rosen’s principal target was the idea that living things were biological machines. In his note to the reader at the start of the book he said

> It has turned out that, in order to be in a position to say what life *is*, we must spend a great deal of time in understanding what life is *not*. Thus, I will be spending a great deal of time with mechanisms and machines, ultimately to reject them, and replace them with something else.
>
> Rosen (1991, p.xv)

Rosen’s central argument reached the conclusion that ‘*there can be no closed path of efficient causation in a mechanism’* (Rosen 1991, p.241). By contrast, in his view, organisms are characterised precisely by the presence of such paths: ‘*a material system is an organism if, and only if, it is closed to efficient causation*.’ (Rosen 1991, p.244) The centrality of the argument based on causal loops can be gauged from the fact that it is repeatedly reasserted in *Essays on Life Itself* and is also a key part of Rosen’s description of *Life Itself* in a short paper called *Autobiographical Reminiscences* which can be found on the Internet at ‘http://www.rosen-enterprises.com/RobertRosen/RRosenautobio.html’. The statement there is clear and trenchant:

> I argue that the external, public, material world is full of closed causal loops, just as the internal, mathematical world is full of closed inferential ones (impredicativities). The “world” of the mechanism, or machine (or, as I call it, the simple systems), and which I believe is an artificial human limitation on reality, does not allow such loops. Accordingly, as a class, these simple systems are extremely poor, or limited, in entailment and hence extremely nongeneric.

Rosen described the idea that science is essentially mechanistic as a prejudice and argued that it had ‘disastrous consequences’ (Rosen 1991, p.xvi). In particular he claimed that the assumption of mechanism made it impossible to answer the question ‘What is life?’ This was not just dramatization for emphasis: that the charge was intended seriously is clear from the context.

> [T]he initial presupposition that we are dealing with mechanism already excludes most of what we need to arrive an answer. No amount of refinement or subtlety within the world of mechanism can avail; once we are in that world, what we need is already gone.
>
> (Rosen 1991, p.xvi)

It is not immediately clear why the assumption of mechanism should be so damaging to the attempt to understand the essential characteristics of organisms. Theories are always wrong in some respects but rarely make it impossible to answer core questions. The explanation for the strength of Rosen's anti-mechanism is to be found in his conception of science and his view of the relation between minds and the external world.

Rosen's critique of mechanistic ideas in biology has been endorsed by leading ecological psychologists including Michael Turvey and Robert Shaw (cf. Turvey and Shaw 1995; Shaw 2003; Turvey 2004). If the critique is correct it applies, among other things, to computational models of psychological processes because they are machine models. It thus becomes apparent that the critique is very wide ranging in its scope. It would show, if it were true, that the mainstream approach to cognitive science must be mistaken in its foundational assumption that the mind can be understood in computational terms. It would also show that proposals to deploy computational thinking in ecological psychology, for example the modelling of affordances as the configurations of Turing machines by Wells (2002), must be flawed or, at best, radically incomplete. Fortunately, for those who value machine models in psychology, Rosen's critique is not correct. It is, in fact, fundamentally misconceived because it is of the essence of a machine that it contains precisely the closed causal loops that Rosen claims cannot be realised in machines.

This paper defends machine models and the mechanist philosophy against Rosen's attack. The simplest, and most obvious, way to rebut Rosen's critique is by displaying machines which exhibit closed loops. Several such machines are described and discussed. These machines provide conclusive counterexamples to Rosen's central argument. A rebuttal along these lines alone, however, does not address the question of how Rosen came to reach his conclusions nor does it tackle the details of his arguments. These are valuable supporting exercises. The paper begins, therefore, with an overview of *Life Itself* (Rosen 1991). It then summarizes and offers a critique of Rosen's central argument. This is followed by discussion of some subsidiary arguments which appear principally in *Essays on Life Itself*. Turing machines are then introduced and are readily seen to exhibit closed causal loops. In the final section of the paper some aspects of Rosen's philosophy and methodology are compared with some of the principles and methods of ecological psychology.

Overview

One of the curious features of *Life Itself*, which Rosen described as a book about biology, is that there is rather little in it which most people would recognize as biology and a great deal which looks like mathematics. Category theory, set theory and recursive function theory all feature in the development of the argument. The reason for this is that Rosen understood science in terms of a specific conception of Natural Law and in terms of an epistemology which laid fundamental emphasis on modelling relations between natural systems and the accounts that we give of them in both natural and formal languages. The first task for a student of Rosen should, therefore, be to understand his epistemology and philosophy of science.

Science, Rosen said, is built on dualities and he suggested that ‘the most fundamental dualism’ is that between the self and everything else. Section 3B of *Life Itself* describes this dualism in Cartesian terms. According to Rosen the self encompasses or contains

> ‘our perceptions, our thoughts, our ideas, our imaginings, our will, and the actions that spring from them. This is the *inner world*. Everything else is *outside*.’
>
> (Rosen 1991, pp. 40-1)

That which is other than the self Rosen calls the ‘ambience’ and he thinks of science as a kind of relation between the self and the ambience. Science he says

> ‘requires an external, objective world of phenomena, and the internal, subjective world of the self, which perceives, organizes, acts, and understands. Indeed, science itself is a way (perhaps not the only way) of bringing the ambience *inside*, in an important sense, a way of importing the external world of phenomena into the internal, subjective world that we apprehend so directly.’
>
> (Rosen 1991, p.41)

It is, I think, of great importance for ecological psychologists to understand, as this quotation shows, that Rosen's epistemology is based on a strong form of indirect realism. It is a key part of the context needed to understand his attack on mechanism and it suggests, rather obviously, that Rosen's philosophy is much less compatible with ecological psychology than some of his apologists may like to think.

A 'second basic dualism' as Rosen describes it concerns how we take the ambience to be partitioned into systems and their environments. The way that Rosen characterizes this second dualism reinforces the point just made about his indirect realism. Talking about how we manage our perceptions of the external world he says,

> 'At this level, we have no universal principles to guide us, nothing *given* to us, like the distinction between the inner world of the self and the outer world, what we called the ambience. It rests rather on a consensus *imputed* to the ambience, rather than on some objective and directly perceptible property of the ambience.'
>
> (Rosen 1991, p.41)

The idea that there is a consensus that imputes or ascribes properties and meanings to the external world suggests the view that everyday reality is, in part, a social construction. Notice the striking contrast here with Gibson's view of the external world as the locus of ecological reality:

> The world of physical reality does not consist of meaningful things. The world of ecological reality, as I have been trying to describe it, does. If what we

> perceived were the entities of physics and mathematics, meanings would have to be imposed on them. But if what we perceive are the entities of environmental science, their meanings can be *discovered*.
>
> (Gibson 1986, p. 33)

The significance of the partition of the ambience into systems and their environments lies, says Rosen, in the fact that it leads us to treat the components in fundamentally different ways. Systems are described in terms of states and environments are described in terms of their effects on systems. This, Rosen says, is a source of fundamental trouble but it is of less concern for the present paper than the first basic dualism between the self and the ambience. Rosen takes perception of the external world to be mediated by the inner self. In section 3D of *Life Itself* he says that language plays ‘an essential role as an intermediary between the self and its ambience’ (p.43) but he does not say clearly in that section what the essential role of language is. Instead he describes the differences between natural and formal languages and between syntax and semantics and introduces what will become a key idea, namely the fact that the syntactic production rules of languages, both natural and formal, are vehicles for inferential entailment. Further discussion of inferential entailment in section 3E includes a brief discussion of Aristotle’s four categories of causation, a topic which is treated further in due course. A key point to note about inferential entailment is that it can be understood in terms of causal relations between syntactic elements. For example, given the rules of elementary arithmetic, the string of symbols 2 + 2 = ? entails, or can be said to cause, the symbol 4, whereas the string

2 – 2 = ? entails the symbol 0. In section 3G Rosen discusses entailment in the external world. The key question he asks is the following:

> ‘[I]s there, in this external world, any kind of *entailment*, analogous to the inferential entailment we have seen between propositions in a language or formalism?’
>
> (Rosen 1991, p.55)

The discussion which follows further confirms his view of perception as indirect. Rosen says, with explicit reference to Kant, that things in themselves are ‘inherently unknowable’ except through the perceptions they elicit in us and acknowledges that this is a potent source for scepticism. He concludes, nonetheless, that it is reasonable to posit relations of entailment between phenomena in the external world because natural language ‘imputes hordes of entailments to the ambience’ (p.56) without leading us seriously astray. These entailments are taken to be causal. Notice the second use of the term ‘impute’.

To summarize the discussion so far, Rosen has, before the end of chapter 3 of *Life Itself* made the following points: there is a fundamental distinction between the internal world of the self which, in Cartesian fashion, is known directly and indubitably and the external world which is known only indirectly. The internal world contains a variety of languages which instantiate systems of inferential entailment. Semantic projections from these inferential systems outwards onto the external world are generally reliable and support the assumption that the external world contains entities which are systematically related to one another by causal entailments.

Building on this foundation, Rosen discusses natural law and modelling relations in section 3H. The point of this section is to discuss whether, and if so how, inferential entailment in the inner world can be related to causal entailment in the external world. Rosen argues that they can be related by the establishment of a modelling relation between them. According to his view modelling relations are possible by virtue of Natural Law which provides the explicit underpinning for science. Natural Law, in Rosen's formulation of it, rests on two fundamental assertions: first, that events in the external world are ordered and exhibit systematic causal relations; second that these orderly causal relations can be grasped by the human mind.

> '[T]he causal relations manifested by a natural system provide the orderliness required of the ambience. Inferential entailment in a formal system is a way of providing the orderliness required of the self. The art of bringing the two into correspondence, through the establishment of a definite modeling relation between them, is the articulation of the former within the latter; it is in effect science itself.'
>
> (Rosen 1991, p.59)

A modeling relation is thus a form of congruence between entailment structures. Causal entailment in the external world is mapped onto inferential entailment in the inner world. If the mapping captures the essential features of the causal system, it will be possible to use the inferential structure of the formal system to make predictions about events in the external world. The discovery of the planet Pluto is a well known example. Prior to its discovery, astronomers had modelled the orbits of the known planets in a mathematical model of the solar system. The mathematical model was a system of inferential entailment in Rosen's terms whereas the actual orbits of the planets constitute a system of causal entailment. Calculations based on the model allowed the astronomers to predict the future positions of the planets. Observations showed that the model was faulty because the planets did not behave as the model predicted. Further mathematical reasoning, which can be described as the tracing of possible paths of entailment in a revised model, showed that the positional discrepancies of the known planets could be accounted for by the postulation of another planet beyond Uranus. The revised model was used to predict the position of this planet and telescopic search subsequently led to the discovery of Pluto.

Two points of particular importance should be noted about the modeling relation. First, models are essential because they provide the means whereby we obtain knowledge about the external world and its future behaviour. Second, there are two-way relations between natural systems of causal entailment (*N*) and formal systems of inferential entailment (*F*). Rosen describes the situation thus:

> '[T]he modeling process compares causal entailment in *N* with inferential entailment in *F*; if we are successful in establishing such a relation, then *F* is the model; *N* is a *realization* of that model. But it is essential to note that the roles of *N* and *F* can be interchanged. That is, instead of starting with a natural system *N*, and looking in effect for a formalism *F* that models it, we could start with a formal system *F* and ask for a natural system *N* whose causal entailment provides a model for inferential entailment in *F*.'

(Rosen 1991, p.61)

Here, I think, we have reached the foundation of Rosen's objections to machine models in biology and the basis for his belief that such models make it impossible to discover the causal principles governing organisms. The core problem, in his view, is that mechanistic philosophy starts with a formal machine model in the inner world and projects this onto the outer world. It assumes, that is, that organisms are machines and will be found to exhibit causal entailments that are congruent to the inferential entailments found in machines. This is a double mistake in Rosen's view: it is wrong because, as a matter of fact, organisms exhibit patterns of causal entailment that are not, and cannot be, captured by the patterns of inferential entailment available in machines, and it is wrong because the strategy prevents us from understanding the patterns of causal entailment that do exist in organisms. Formal models are our windows on the external world and the machine model is a window that makes it impossible to understand how organisms work. It is this, in Rosen's view, that makes mechanism such a pernicious philosophy of biology.

From this standpoint there are two major issues for Rosen to tackle. First, machine models must be characterized and shown to be inadequate; second, an analysis of the kind of modeling strategy needed to provide the theoretical foundation for biology has to be given. Notice that it is not enough for Rosen simply to reject mechanism. On his view, models are an essential part of the process of theory construction and development. A critique of mechanism leaves a gap that has to be filled by formal models of some kind. Chapters 4 to 10 of *Life Itself* are concerned with these issues and the final chapter outlines Rosen's proposal that biology should be based on relational models and compares the proposed relational biology with existing theories and models.

The anti-mechanist argument of *Life Itself*

Rosen’s anti-mechanist argument is complex and multi-faceted. It is also obscure in places. Even his most ardent supporters would hardly claim that Rosen’s exposition is always clear. I think, however, that the following list of propositions is a fair summary of the principal points of the argument and, roughly, of its sequence. I have indicated in parentheses the primary places in *Life Itself* in which the points of the argument are presented and/or elaborated. It should also be noted that some points depend on earlier chapters of the book. Points 5-7, for example, draw on the material of chapter 6, and the argument in chapters 9 and 10 draws on the ideas about relational modeling developed in chapter 5.

1. A system, natural or formal, is a machine if it simulates, or can simulate, something else. Simulation is what machines do (7B, 7D, 7E).
2. Simulation involves a fundamental distinction between hardware and software (7D, 9A, 9B, 9D).
3. A natural system, *N*, is a mechanism if and only if all of its models are simulable (8B).
4. A natural system, *N*, is a machine if and only if it is a mechanism such that at least one of its models is already a mathematical machine (8B).
5. If *N* is a mechanism it has a unique largest model (8C).
6. If *N* is a mechanism it has a finite set of minimal models (8D).

7. The largest model of a mechanism is the direct sum of its minimal models and is therefore synthetic (8E).
8. In mathematical machines efficient causality and material causality are segregated into disjoint structures; hardware is the embodiment of efficient cause while material cause is embodied in software (9A, 9B).
9. Hardware entails the flow of software states from input to output, but software cannot entail hardware (9B, 9D).
10. A hardware component can only be entailed by another hardware component (9E).
11. A system S_1 in which a previously unentailed hardware component is entailed must have more states than S_0 the system in which the component was unentailed (9E).
12. The larger system S_1 will itself have at least one unentailed hardware component (9F).
13. Only in the limit as n approaches infinity will every component in a system S_n be entailed (9F).
14. Such a system cannot be a mechanism because a mechanism has a finite largest model (9F).
15. Hardware cannot be entailed by adding further constraints to existing components because this involves splitting states into direct summands and leads to an infinite regress of fractionation. Such a regress contradicts the fact that a machine has a finite set of minimal models (9F).
16. The argument of points 1 – 15 shows that there can be no closed path of efficient causation in a mechanism (9G).
17. Organisms exhibit closed paths of efficient causation (10A, 10C).

18. Organisms are not machines (10B).

The argument is vulnerable to a range of criticisms and fails to establish the central claim made in point 16. I shall discuss its principal weaknesses but it will be useful, before doing this, to consider briefly the distinction between efficient and material causes that is mentioned in points 8 and 16.

The distinction stems from Aristotle's classification of 'causes' into four types, material, efficient, formal and final. There is a consensus among contemporary scholars that the Greek word 'αιτιον' which has traditionally been translated as 'cause' should, in fact, be translated as 'because' (Hocutt, 1974; Moravscik, 1974). The four-fold distinction made by Aristotle is, therefore, best thought of as a classification of different types of explanation not different types of causes. The only type that would be considered causal in modern terminology is the category of efficient cause. Rosen uses Aristotle's terminology primarily to distinguish causal relations among the hardware components of a system which he describes as instances of efficient causation, from causal relations in software which he describes as instances of material causation (cf. *Life Itself*, sections 5H, 9D). Part of the reason why Rosen uses Aristotle's terminology stems, I think, from the fact that he wants to compare entailment in material systems with entailment in formal systems and is able to do this by using Aristotle's analysis to distinguish the different types of cause in both cases. The discussion in section 3G of *Life Itself* (p.57) seems to me to support this understanding. However, given the fact that material 'causes' in the Aristotelian scheme of things are not causes as we would understand them the distinction obscures more than it clarifies. It is, in any event, not a significant aspect of the argument. What is significant is Rosen's claim that the causal relations between hardware components of a system are distinct from the causal relations embodied in software. It should, of course, be remembered that in any real machine to which the hardware/software distinction applies, software components are just as real and material as hardware and, therefore, just as much subject to causal influence. The distinction which Rosen marks by the use of the terms ‘efficient’ and ‘material’ is, therefore, a distinction of causal role, not causal type. Provided, then, it is understood that the central claim in point 16 of Rosen’s argument refers to causal relations among the hardware components of machines, and provided also that the causal roles of hardware and software are distinguished where necessary, nothing is lost by ignoring or eliding the distinction between so-called efficient and material causes.

Turning now to the status of the argument itself, point 1, which takes simulation to be the key criterion for distinguishing machines from other systems, is simply wrong. Most machines do not simulate. Think of the functions of the everyday machines which surround us: refrigerators cool things; cars transport us from place to place; scissors cut things; pianos produce music; and so forth. None of these machines is a simulator. Machines which do simulate, of which computers are the most obvious examples, are a special, highly organized class. Rosen’s definition of a machine is, therefore, far too narrow. However, even if it were correct, the definition does not establish what he wishes to establish. Part of his reason for focusing on simulation as the key activity of machines is his belief that there is a fundamental divide between simulation and modelling. He claims that whereas a model lays bare the entailment structure of the system it models, a simulation hides it.

> ‘In causal terms, simulation involves the conversion of efficient cause, the hardware of that being simulated, into material cause in the simulator. In

> essence, this means that one can learn nothing about entailment by looking at a simulation.’
>
> (Rosen 1991, p.193)

This is an extraordinary claim which, once again, is simply wrong. If it were true there would be absolutely no point in the many computer simulations, for example of airflows over aircraft wings and the development of weather patterns, which are used to study causal relations in complex systems. Far from hiding the details of causal interactions in the systems which are modelled, simulations enable them to be studied in great detail and at a variety of time scales. A simulation is rather like a high speed film of the impact of a bullet on a particular material in which time can be slowed down on playback precisely to enable the investigator to understand better the extremely rapid succession of causal interactions between the bullet and the material on which it impacts. A simulation should, if anything, be described as a kind of ‘supermodel’.

It is worth pursuing this point a little further because Rosen’s claim about simulation is couched in terms of the unhelpful distinction between efficient and material causes which is a source of confusion rather than clarification. Suppose we were to construct a simulation of the impact of a bullet on a sheet of glass. To do this we would need, *inter alia*, to construct a simulation of the bullet, a simulation of the sheet of glass and a simulation of the dynamic relations between them. One obvious way to do this would be by constructing symbolic representations of the terms of the simulation in a computer program. Thus the bullet might be represented as an array of quantities which collectively simulate its shape, mass, density, velocity and so forth and the glass might be represented as a lattice or similar regular structure. It is this process of translating into software the causally active components of the hardware system represented in the simulation that Rosen refers to as the conversion of efficient cause into material cause. It should immediately be clear that whether or not one can learn anything about entailment in such a simulation depends not on the process of conversion but on whether one has chosen the correct aspects of the system to represent. A simulation which included only the name of the maker in the representation of the bullet and the reflection of the external environment in the representation of the glass would not be informative about the impact of the bullet on the glass but a simulation which included detailed representations of the salient characteristics of the bullet and the glass and which modelled the changing dynamics of the relationship between them at a microsecond time scale might be extremely informative. The quality of a simulation depends on the quality of the modelling on which it is based. It is simply incorrect to claim that the conversion of terms which a simulation requires hides the entailment structures of the system simulated.

Point 2 of the argument is correct. There is a fundamental distinction between hardware and software, at least in computers that simulate. However, the distinction does not support the argument in the way that Rosen wishes to develop it in points 8 and 9.

Points 3 and 4 of the argument inherit the deficiencies of point 1. They are also rather strangely worded. One does not normally think of calling a system a machine if and only if its models are of a particular kind. Rosen explains this curious usage in terms of Natural Law:

> ‘[T]his peculiarity stems only from my expression of these concepts in terms of the models of *N*, rather than try to talk directly about *N* itself. This is all that Natural Law entitles us to do.’
>
> (Rosen 1991, p.203)

The point is interesting, not so much for what it tells us about the argument but for the fact that it is a further indication of the indirect realism which provides the foundation for Rosen’s epistemology. In his view we have to discuss natural systems in terms of their models because these are all that is directly available to us.

Points 5, 6, and 7 are lemmas to the main argument. They are used to support points 14 and 15. There are both general and specific comments to be made. The important general comment is that the arguments of points 5-7 rest on the discussion of models in chapter 6. I do not propose to go into the material of chapter 6 in great detail but I think it is clear that it does not establish the fundamental claims about models that Rosen wants to make. The distinction between analytic and synthetic models that Rosen develops in Chapter 6 is based on equivalence relations over sets. Analytic models are constructed in terms of Cartesian products and synthetic models in terms of direct sums of disjoint subsets. Rosen takes the analytic/synthetic distinction to be applicable to models of any kind, but it is not clear that the concept of an equivalence relation is a suitable tool for developing a generally applicable typology of models or for exhibiting the relationship between a model and the system, natural or formal, of which it is a model. In order for an equivalence relation to be specified the elements of the set on which it is defined have to be known and it is precisely this that is not generally known when a model of a natural system is being developed. Rosen himself acknowledges this obvious disanalogy in his discussion of the equivalence relations over a set *S* that he calls ‘observables’:

> I cannot emphasize too strongly that, in the formal world, *S* is already a determinate entity (in this case, a set), so that in general, looking at *S* as imaged in the spectrum *f(S)* of an observable inevitably “*loses* information” about *S*. In the case of a natural system, on the other hand, the counterpart of *S* is initially *unknown*, veiled completely in its noumenal and phenomenal shrouds. The whole purpose of measurement in this case is to *provide information* about it.
>
> (Rosen 1991, p.157)

It is hard to understand, given the above, why Rosen persists in using equivalence relations over sets as the basis for his definition of types of models of natural systems. He suggests on the page following the quotation just given that set theory provides ‘a formidable battery of inferential structure’ (p.158) to study a given set *S* but the various operations on a set, which include the generation of further sets, tell us about the structure of the set not about the system which the set represents. Consider a simple example to make the point clear. The Cartesian product of a set with itself, a structure about which Rosen says a great deal, is the set of ordered pairs of its elements. Suppose, as a result of observation, we have developed a simple model of a car in terms of the set *CAR* = {body, wheels, engine, transmission}. This is a very simple model but it is a start. The Cartesian product of the model, *CAR* × *CAR* is the set of ordered pairs {(body, wheels), (body, engine), (body, transmission), (wheels, engine), (wheels, transmission), (engine, transmission)}. Does the Cartesian product tell us anything we didn't already know about the car? Rather clearly, the answer is no. The Cartesian product tells us something about the structure of the set but not about the system of which the set is a model. The difficulty with Rosen's view is that he assumes that the structures of sets like Cartesian products are somehow necessarily informative about the models represented by the original sets. In general that is not the case. In order to improve the model we need to make further observations of the car, not engage in an analysis of the set that constitutes the original model.

The more specific points to be made concern the status of the argument that purports to establish point 5, the claim that every machine has a unique largest model. The argument in section 8C of *Life Itself* is faulty. It takes the form of a *reductio ad absurdum*. Starting from the assumption that *N* is a mechanism but that the category of all its models *C(N)* contains no largest model, Rosen claims that we can find an infinite sequence of increasingly refined models. Because *N* is a mechanism, all its models must be simulable and each of them must have a program of finite length. We can then form the intersection of all the models and, by hypothesis, this is also a model. Unless the sequence of models terminates after a finite number of iterations, the model formed from the intersection is larger than any of the other models. Because it is simulable it must also have a program. The conclusion of the argument, according to Rosen, is the following:

> We thus end up with a countable family of distinct programs, each of which is a distinct word of finite length on a finite alphabet. This is clearly impossible.
>
> (Rosen 1991, p.205)

In fact it is perfectly possible. Thus even if the premises of the argument were correct, which is questionable, the conclusion would not follow. The decimal representations of the integers, for example, are the elements of a countably infinite set of distinct words of finite length on a finite alphabet and integers are commonly used to represent programs. It is also always possible to extend a program without altering its functionality by adding new instructions which do nothing (cf. Rogers 1967, p.22, Theorem III). This possibility contradicts the assertion that a machine must have a unique largest model. Given the failure of the argument for a largest model, the argument for a finite family of smallest models is of less interest because Rosen's argument requires both points 5 and 6. It is, however, pertinent to ask what is meant by a 'minimal' model. Various possibilities exist but it is hard to form a clear idea of what Rosen intended.

Points 8, 9, and 10 constitute the heart of the argument because they purport to show that the distinction between hardware and software in machines entails the segregation of causal factors which, ultimately, is the basis for the claim that machines cannot contain closed loops of 'efficient' causation. There are two principal flaws in the argument. The first is the fact that, contrary to Rosen's view, machines can and do have causal entailments flowing from software to hardware. There is no incompatibility between this fact and the fact that hardware and software are distinct parts of a machine. In consequence, there can be closed loops of entailment in machines involving both hardware and software. Rosen is unable to see this because he thinks of the partition between hardware and software as 'absolute' (cf section 9D, p.228). The second flaw in Rosen's argument is that it simply fails to recognize the possibility that there can be direct causal links between the hardware components of a machine, as well as links via software, and there is no reason why these links cannot form loops. Indeed, as I show in the discussion of Turing machines, it is of the essence of a Turing machine that there are such closed loops. I think the explanation for Rosen's failure to see this is reasonably straightforward. In the argument presented in chapter 9 of *Life Itself*, Rosen treats hardware as whatever it is that executes a mapping from inputs to outputs. Given a simple function $f : A \rightarrow B$ which Rosen takes to represent a component of a machine, f constitutes the hardware which, as Rosen puts it, induces the software flow from the input set A to the output set B. He is at pains to point out, in section 9B (p.222) that the hardware and the flows it induces are different things and that the essence of hardware is to generate flows (p.224). As a result of this way of looking at things, when he comes to consider a system with more than one component, (cf. Figure 9C.2, p.224) he assumes that all components are of the same type and that the only function of the hardware is to induce flows on software. He simply misses the possibility that there can be direct links between the different hardware components representing other types of causal interaction than the generation of software flows. The kind of causal interaction needed can be described as a 'component-component' interaction. There is no argument in *Life Itself* to say that such interactions cannot exist in machines and in fact they both can and do exist. The combined effect of the two flaws is to make Rosen's picture of what can be done with machines irretrievably narrow and limited.

Points 11 to 18 of the argument depend entirely on the earlier points. They contribute nothing extra. Since the earlier points do not in fact support the claims made, the argument as a whole fails. The failure of the argument shows that machine models of organisms are perfectly possible but it also shows that Rosen's criterion, which claims that a material system is an organism if and only if it is closed to efficient causation, is invalid. Machines can be closed to efficient causation in just the same way as organisms but are not organisms simply by virtue of that fact.

It remains possible that Rosen's positive proposal, based on the concept of an (M,R)-system, provides a different criterion for distinguishing organisms from machines. However, the description of an (M,R)-system in *Life Itself* is muddled and inconsistent and has been criticized by Landauer and Bellman (2002). In any case, the crucial concept of self-replication was shown to be consistent with mechanistic hypotheses by John von Neumann as early as 1948. The entailment of replication by the functions of metabolism and repair is, therefore, unlikely to distinguish organisms from machines. Rosen claims (p.234) that von Neumann's construction rests on an equivocation between software and hardware but that is not correct. Von Neumann's discussion of the logic of self-reproduction makes clear and consistent use of the distinction between hardware and software. He shows that a machine can, in principle, reproduce itself if it has access to a store of elementary parts (hardware) and contains a set of instructions (software) for its own construction which can be copied and passed to its clones.

Rosen's contradictory accounts of open systems

The argument of *Life Itself* is the major, but not the only source of Rosen's anti-mechanism. In *Essays on Life Itself*, a collection of papers written mainly after *Life Itself* was published and brought to press by Rosen's daughter after his death, other related arguments can be found. One such type of argument concerns a distinction between simple and complex systems which Rosen used to argue for the fundamental separation of organisms from machines. The argument is different in form from the central argument of *Life Itself* and is worth separate comment. It centres on the concept of an open system and tackles the question of how open systems can be made amenable to analytical study. Different versions of the argument appear in two chapters of *Essays on Life Itself.* They come to contradictory conclusions.

In Chapter 1 of *Essays on Life Itself*, Rosen discusses open systems in the context of Schrödinger's famous essay *What is Life?* In the section of Chapter 1 entitled *The Forcing of Open Systems* Rosen defines an open system as follows:

> A system that is open in *any* sense is one whose behaviors depend on something outside the system itself, whereas in a closed system, there *is* no outside.
>
> Rosen (2000, p.21)

Rosen says that physics has typically had trouble in modelling open systems and that one way to deal with them is to try to internalise the external influences so as to get a bigger system which is closed and to deal with that. However, he says, this strategy does not generally work:

> Indeed, what we end up with in this fashion is generally a bigger open system, which is in some sense even more open than the one we started with…[W]hat one typically ends up with after carrying out such a strategy is the entire universe, which is not very helpful.
>
> Rosen (2000, p.21-2)

After some further discussion of other possibilities for augmenting open systems Rosen concludes that the resulting models are still not generally stable and that further expansions of the model to include more of the external forces acting on the system are needed. At this point, he says, ‘we have a glimpse of an incipient infinite regress establishing itself.’ Rosen (2000, p.24). The regress can be avoided, Rosen says, if the forces internalised at stage N of an expansion process have already arisen at earlier stages.

> A source for such an N^{th}-stage internalised forcer is *a mechanism for its replication*, expressed in terms of the preceding N-1 stages, and not requiring a new $N+1$ stage. Thus replication is not just a formal means of breaking off a devastating infinite regress, but it serves to stabilize the open system we arrived at in the N^{th}stage.
>
> Rosen (2000, p.24)

At this point the crux of the argument has been reached. Rosen claims that the price to be paid for escaping the infinite regress is that the systems thus arrived at are complex, non-computable and contain closed loops.

> Breaking off such an infinite regress does not come for free. For it to happen, the graphs to which we have drawn attention, and which arise in successively more complicated forms at each step of the process, must fold back on each other in unprecedented ways. In the process, we create (among other things)

> closed loops of efficient causation. Systems of this type cannot be simulated by finite-state machines (e.g., Turing machines); hence they themselves are not machines or mechanisms. In formal terms, they manifest impredicative loops. I call these systems *complex*.
>
> Rosen (2000, p.24)

The particular point to notice is the claim that it is the process of breaking off the infinite regress of system expansions which generates complex systems.

The argument in chapter 1 contradicts and is contradicted by another of Rosen's arguments in chapter 20 of *Essays on Life Itself*. In chapter 20 Rosen discusses the differences between therapeutic interventions in medicine and control engineering. His target is the idea that an organism is a biological machine. If this were so, he says, we would expect therapies to have few or no side effects whereas in practice side effects are the rule. To explain why this is the case Rosen constructs an argument which uses temperature control as an example. A room without temperature control is an instance of an open system. If the room is large the temperature may appear to be constant, but it will eventually change 'because the room is open to ambient influences we do not see directly.' Rosen (2000, p.299). A first level of control over the room temperature can be achieved by installing a thermostat. However, in order to close the room to the effects of changes in the ambient temperature in this way, the system as a whole has had to be enlarged.

> The thermostat itself is new material structure, which we have had to bring into the system to control the effects of unpredictable temperature fluctuations.

> The thermostat closes the room off to ambient temperature, but it itself is now open to other interactions—i.e., to new sources of noise. For instance, parts of it may corrode because of humidity and oxygen in the air in our room…[W]e may indeed end up with more noise than we had originally.
>
> Rosen (2000, pp.300-301)

Notice the strong parallel here with the first example. In each case, the effect of internalising a source of noise or openness is to produce, potentially, a more open or noisier system. Rosen again explores the possibility of further expansions of the system to bring the new sources of noise under control and again the possibility of an infinite regress is noted:

> Even in this simple example, we see an incipient and deadly infinite regress yawning before us…The real question arising here is whether, and if so, when, this potential infinite regress can be broken off.
>
> Rosen (2000, p.301)

As with the first example, Rosen says that the regress can be broken off if the system can be turned back on itself.

> [I]t is conceivable that such a potential infinite regress actually breaks off. It will do so if, and only if, we can arrange matters so that the noise arising at the *N*th step of this sequence produces consequences that are subject to controls

> instituted at earlier stages. In such a case, the sequence breaks off at the *N*th stage.
>
> Rosen (2000, p.303)

Rosen notes that the example of the thermostatically controlled room is characterised by a single state variable but says that the analysis can be generalised to deal with any finite number of state variables.

> The analysis, of course, grows increasingly complicated, but, in effect, we now have a much larger family of cascading control loops, each of which creates the potentiality for infinite regress.
>
> Rosen (2000, p.303)

The crux of the argument has again been reached. In the earlier example Rosen claimed that the price to be paid for breaking off the infinite regress was that the resulting systems were complex. In the latter example, however, he reaches the opposite conclusion:

> [I]f every such cascade breaks off after a finite number of steps, then the system itself, and its environment, *must both be simple*. Conversely, if a system (or its environment) is not simple, then there must be at least one cascade of simple controls that does not break off…A system that is not simple in this sense (i.e., is not a mechanism) I call *complex*.

Rosen (2000, p.304)

To drive home the point, a little later in the chapter Rosen argues that '[t]here is a sense in which complex systems are infinitely open' Rosen (2000, p.307) and it is for this reason, he claims, that side effects are the norm rather than the exception in medical interventions.

The contradictory nature of the two arguments cited is perfectly clear. In one case Rosen argues that the existence of a break point that prevents an infinite regress of system openings leads to complex systems, in the other he argues that the break point leads to simple systems. In one case he argues that complexity is the result of system closure, in the other that complex systems are 'infinitely open'. The core problem is that the distinctions he wants to insist on, between organisms and machines, and between simple and complex systems are not to be had on his terms. Both organisms and machines are systems with closed causal loops and the distinction between complex and simple systems, if there is one, does not demarcate organisms from other machines.

Rosen's subordinate arguments

Life Itself and *Essays on Life Itself* also contain subordinate arguments which deserve mention. An argument which is largely implicit in *Life Itself* but which is given greater prominence in *Essays on Life Itself* claims that machines cannot contain closed causal loops because such loops are 'impredicative' and are forbidden in formal systems such as Turing’s. One instance of this argument, from Rosen’s autobiographical sketch, was mentioned in the introduction. Here is another:

> *Impredicativity*…was identified as the culprit in the paradoxes springing up in Set Theory. Something was impredicative…if it could be defined only in terms of a totality to which it itself had to belong…Formalizations are simple systems (in my sense) and, in particular, cannot manifest impredicativities or self-references or “vicious circles”. This is precisely why such a simple world seemed to provide a mathematical Eden, inherently free from paradox and inconsistency. Alas, as Gödel showed, it was also free of most of mathematics. We cannot dispense with impredicativity without simultaneously losing most of what we want to preserve.
>
> (Rosen 2000, pp.293-4)

There are several flaws in this argument. One is the implicit claim that all closed loops are impredicative. They are not: the closed loops in the finite state control automata of Turing machines are not necessarily defined impredicatively although they may be. Thus, closed causal loops could be found in machines even if impredicative loops were forbidden. Another is the claim that formalizations cannot manifest self-references. They both can and do. The recursive definitions that Turing used in the construction of his universal machine are frequently self-referential. Third, it is misleading to suggest that impredicativity was identified as *the* culprit in the paradoxes identified in set theory. It is true that Bertrand Russell thought it to be the source of the problem, at least with respect to the paradox that he devised, but it has been clear since the analysis of Gödel (1944) that impredicative properties are troublesome only under very special circumstances. The idea, therefore, that abstract machines such as Turing machines cannot exhibit closed causal loops because these are always impredicative and hence forbidden is mistaken.

A second argument, related to the first, is the claim that machines are, as Rosen puts it, ‘non-generic’, exceptionally rare, and feeble in their entailment structures. The point is hinted at in the quote above with the idea that formalization loses most of what is needed from mathematics and that formalizations cannot ‘manifest’ impredicative loops. In another passage from *Essays on Life Itself* Rosen sums up his view of the attempts at formalization which he attributes to Hilbert’s program:

> The status of all these formalizations is informative. They turn out to be infinitely feeble compared with the original mathematical systems they attempted to objectivize. Indeed, these attempts to secure mathematics from paradox by invoking contructibility, or formalizability, end up by losing most of it. This is one of the upshots of Gödel’s celebrated Incompleteness Theorem (Gödel 1931), which showed precisely that “self-referential” statements (e.g., “this proposition is unprovable in a given formalization”), which are perfectly acceptable in the context of ordinary Number Theory, fall outside that formalization.
>
> (Rosen 2000, p.92)

One claim that Rosen seems to be making in this quotation is that self-referential statements cannot be constructed, and hence are not expressible, in formal systems. That appears to be the force of the suggestion that they ‘fall outside’ a formalization. This shows a misunderstanding of what Gödel achieved. The point that Gödel made was not that you can’t express or construct a self-referential statement in a formal system. It was precisely the opposite. Gödel showed that you *could* construct a self-referential statement as a well-formed formula of any formal system powerful enough to contain arithmetic. In the sketch of his proof at the start of the famous 1931 paper, having explained how formulas of Russell and Whitehead’s system *Principia Mathematica* could be used to express metamathematical notions, Gödel went on to explain that the proof rested on the construction of a specific proposition using the formalism of *PM*:

> We now construct an undecidable proposition of the system *PM*, that is, a proposition *A* for which neither *A* nor *not-A* is provable.
>
> (Gödel 1931, p.147)

The crucial point was that the undecidable proposition was both syntactically well-formed and asserted its own unprovability. From this Gödel showed that if the formal system containing the proposition was consistent the proposition had to be true and hence unreachable by a finite sequence of inferences from the axioms. Thus the formal system was incomplete. The proof rests on a fundamental distinction between constructibility and provability which Rosen seems not to have fully grasped. A short but telling passage from *Life Itself* supports this suggestion:

> [G]iven any finite set of axioms for Number Theory, there are always propositions that are in some sense theorems but are unprovable from those axioms (unless, of course, the axioms are inconsistent to begin with—in which case everything is a theorem).
>
> (Rosen 1991, p.35)

The crucial point, which Rosen explicitly fudges, is that unprovable propositions are not theorems in any sense. They are syntactically well-formed formulas but they cannot be reached by finite sets of inferences from the axioms. It is precisely this point that distinguishes constructible, but unprovable, propositions from theorems of the system. To say, therefore, that unprovable propositions are 'in some sense' theorems clouds exactly the distinction that needs to be kept clear. The distinction between constructibility and provability undermines the claim that closed or impredicative loops cannot be expressed in formal systems.

A second point which Rosen makes is that formalizations are in some sense 'infinitely feeble'. Rosen repeatedly asserts that mechanisms are a vanishingly small proportion of mathematical systems. The assertion rests, I think, on the distinction between countable and uncountable infinities. This distinction was originally made by the mathematician Georg Cantor in the late nineteenth century. Cantor showed, using the technique of diagonalization, that, in a strictly definable sense, there are infinitely many more real numbers than natural numbers even though there are infinitely many of these. The natural numbers are countably infinite but the real numbers are uncountably infinite. Turing machines can be paired one for one with the natural numbers and this shows that there are countably, infinitely many of them. In comparison with the uncountably infinite number of real numbers the Turing machines can be said to be a vanishingly small proportion of mathematical systems

It is very hard, however, to know what to make of the distinction between countable and uncountable infinities because the sizes, even of countably infinite collections, violate our everyday intuitions about collections of things. If one thinks, for example, about natural numbers, it seems to common sense that the totality of them can be divided into two halves, the odd numbers and the even ones. For any finite totality which is divisible into two halves, each half clearly has half as many members as the totality. If I have five green apples and five red ones I have ten apples in total with the green apples forming half the totality and the red ones the other half. By the same token, because there are *n* odd numbers and *n* even numbers, there should be 2*n* natural numbers in total. But it isn't so. Infinite totalities don't work like that. There are exactly as many odd numbers as there are natural numbers and exactly as many even ones. Countably infinite totalities are all of the same size even though they appear to common sense to have different numbers of elements. Given this, it is very hard to know what follows from the fact that countable infinities are infinitely small compared with uncountable ones even though there is a clear mathematical sense in which this is so. What Rosen is suggesting seems to rely on combining this point with the supposed exclusion of closed loops from the world of machines. Look, he says, the collection of machines is infinitely small by comparison with the collection of real numbers. Moreover, there are no closed loops in the collection of machines. Thus this collection must be infinitely feeble in the properties it can express. If it really were the case that there were no closed loops in machines the argument might have some force. As it is, we have infinitely many machines with as many closed loops as we care to define. That is a perfectly satisfactory foundation on which to construct a mathematical account of biological, and for that matter psychological, systems.

Causal loops in Turing machines

The principal line of defence to Rosen's anti-mechanist claim does not depend on the fact that Rosen's arguments about open systems are contradictory or that the central argument of *Life Itself* is unsound. The principal defence of mechanism rests on the demonstration that Turing machines, which underpin mathematical and computational thinking about mechanisms, contain closed causal loops. This demonstration falsifies Rosen's fundamental claim. A Turing machine is an abstract entity but one that could perfectly well be built. In this section I provide a very brief introduction to Turing machines focused on the issue of closed loops. A full account of Turing's work and its place in psychology can be found in Wells (2006). Readers are also encouraged to study Turing (1936), the seminal paper in which Turing set out his theory.

A Turing machine is a model of a human agent engaged in a paper and pencil calculation. This simple fact is not often mentioned in psychological discussions of computational models but it is of great significance and bears consideration by ecological psychologists. I shall take a few moments to comment on it in the light of Rosen's epistemology which constitutes the framework for his objections to machine models. Rosen says, as reported in the overview, that the primary feature of natural law is to bring systems of causal entailment in the external world into correspondence with systems of inferential entailment in the inner world of the self. He also made the important point that one could start from either side. Starting from the external world the primary entity is a natural system whose causal entailments one attempts to discover and model in formal terms. Starting from the internal side, the primary entity is a formal system for whose inferential entailments one attempts to find a matching natural system. One of Rosen's principal objections to the machine metaphor, particularly in the form in which it was inherited from Descartes, was precisely that it started from a formal model and attempted to force the phenomena of life to fit that model. Even worse, the Cartesian machine metaphor was incompletely specified. Near the beginning of *Life Itself* in a discussion of Descartes, Rosen says of him:

> What he had observed was simply that automata, under appropriate conditions, can sometimes appear lifelike. What he concluded was, rather, that *life itself was automaton-like*. Thus was born the machine metaphor, perhaps the major conceptual force in biology, even today. Descartes took this fateful step with only the haziest notion of what a mechanism or automaton was (Newton was still a generation away), and an even dimmer notion of what an organism was.
>
> (Rosen 1991, p.20)

It appears from what Rosen says later in the book that, at the time of writing it, he felt there was still no canonical machine model on which to base his assessment. At the start of chapter 7, for example, he refers to 'the vague concept of machine' (p.182) and although he refers to Turing in that chapter he does so in a way that suggests he was unaware of the derivation of Turing's machine model from the example of a human calculating with paper and pencil. Had he been aware of this derivation he would hardly have claimed as he did (p.185) that Turing machines are the formal counterparts of clockwork, i.e. the machinery that drives the hands of a clock, and he would also hardly have claimed that his quasi-Newtonian definition of an algorithm in section 7C of *Life Itself* 'is essentially a Turing machine' (p.189). Neither of these claims is accurate. It is a great pity that Rosen was unaware of the origins of Turing's modeling enterprise because Turing machines are exactly the kinds of models that Rosen proposes are needed to embody natural law. Turing began by observing the natural system of causal relations that is involved in a paper and pencil computation. From this he abstracted what he felt were the essential components and used these to construct his formal model. He was then able to use the model in the way Rosen says that models should be used, to reflect on and reason about causal relations in the natural system modelled. One important result was the construction of the universal machine, which is a simulator *par excellence*. What the universal machine shows is not, as Rosen asserts, that simulation hides the details of the modelled system from the observer. Quite the contrary: a simulation makes available in an explicit symbolic format the details of the entailment structures that the model expresses. Turing's universal machine also demonstrates why symbolic notations are so important for almost all systematic human activities. Among other important characteristics they make available, in forms that do not have to be remembered, structures of inferential entailment that model causal relations in the external world that we find useful or pleasing or both.

The Turing machine has two primary components, a finite control automaton, which is a model of the mind of the human, and a one-dimensional tape which is a model of the paper on which a human writes out a calculation. For present purposes, the structure of the finite control automaton is of particular importance. Closed causal loops are most evident there although they are also found in the dynamic relationship between the automaton and the tape. Turing machines were originally devised to explore the issue of effectiveness in calculation and the numbers they can compute, which are called the computable numbers, are a subset of the real numbers and include irrational numbers like π which have infinite decimal expansions. The opening sentence of Turing's paper makes a point which is of fundamental importance in the present context. He said:

> The "computable" numbers may be described briefly as the real numbers whose expressions as a decimal are calculable by finite means.
>
> (Turing 1936, p.230)

The point to note is that computable numbers like π which have indefinitely long decimal expansions are nevertheless said to be calculable by 'finite means'. What this means is that a finite structure or set of resources, the control automaton, which is defined in advance of the calculation is sufficient to produce the endless sequence of digits representing a number like π. The way this is done, and has to be done, is to build the finite control structure with components connected in loops whose processing can be iterated as many times as necessary. It is thus of the essence of a Turing machine that its fixed processing resources are structured as one or more closed loops. It is precisely because this is so that infinite sequences can be produced by finite means. Let us consider three examples briefly. Much more extensive discussion can be found in Wells (2006).

The first example is one of the simplest Turing machines imaginable. It was defined by Turing in his 1936 paper and outputs the endless sequence 010101…I shall call the machine TM1. The finite control automaton of TM1 has four internal states, which we can call q1, q2, q3 and q4. These, collectively, can be described as a set Q = {q1, q2, q3, q4}. Internal states describe the causal relations among the parts of a Turing machine in an abstract way. The concept of state, as it is defined and used in Turing machine theory, is very different from the concept of state used by Rosen in *Life Itself*. In Turing machine theory states are relational entities, the sorts of things that Rosen calls 'components' in Chapter 5 of *Life Itself*. The operations of TM1 are described by two functions, each of which takes the 'configurations' of TM1 as its arguments. A configuration is an ordered pair consisting of an internal state and an input. One function determines the next state of TM1, the other its output. The next state function is more important for present purposes. This function is a map from Q to Q and is implemented as a closed loop of entailments. The starting state of the machine is q1: q1 entails q2, q2 entails q3, q3 entails q4 and q4 entails q1. Once started, the machine cycles endlessly through this processing loop. It is precisely because the four internal states of its control automaton are structured as a closed loop that the fixed, finite machine TM1 can output the infinitely long sequence 010101… TM1 provides a definitive and conclusive rebuttal of Rosen's lengthy and complex argument in *Life Itself*. TM1 is a machine and it has a closed causal loop of the kind that Rosen says machines cannot have. The causal relations between the states which implement the loop are of the type that I described as component-component interactions in the discussion of points 8 – 10 of Rosen's argument from *Life Itself*.

TM1, by itself, is quite sufficient to rebut Rosen's argument but it is worth considering briefly two more examples. The first is a machine used to illustrate the argument made by Wells (2002) that the configurations of Turing machines provide natural and informative models of Gibsonian affordances. The machine is called HP and is described in detail in Wells (2002, pp.163-7). Like TM1, HP has four internal states, but whereas TM1 has a single closed loop connecting its states, HP has six closed loops. In addition, it also makes use of closed loops connecting its internal states with symbol structures on its tape, which is its environment. This is not a feature of TM1. Most Turing machines are like HP in this respect, however, and it is partly because these machines have causal loops connecting their internal states to their environments that they are potentially of such interest and importance to ecological psychology.

Finally, it is worth remarking that universal Turing machines, the abstract ancestors of contemporary digital computers, exhibit complex patterns of closed loops to implement their processing. It is not possible to construct a universal machine which does not have such loops and they involve both hardware and software. Thus it is of the essence of Turing machines, in general, that they contain closed causal loops.

Rosen and ecological psychology

Rosen’s principal concern in *Life Itself* and in *Essays on Life Itself* was fundamental theory in biology and he says relatively little which is specific to psychology. The general thrust of his theorizing, however, appears to be antithetical to the interests and motivations of ecological psychology. His epistemology, as the overview shows, assumes a strong form of indirect realism. He also used the (untenable) distinction between simple and complex systems to rule out at least some of the methods that have been discussed favourably in the ecological psychology literature and his emphasis on the importance of closed loops of causation *within* the organism downplays the significance of the environment.

Organisms and environments

Rosen's epistemology is concerned with the sources of answers to 'why' questions about systems of various kinds. He perceives a spectrum of possibilities. At one extreme are systems with components about which 'why' questions can be answered only by reference to the environment within which the system is embedded. According to his view, mechanisms are systems of this kind:

> Most of the "why?" questions we can ask about such a system are unanswerable within the system, and therefore, must be referred to its environment. Put another way: most elements of an abstract block diagram arising from a mechanism are *unentailed*.
>
> (Rosen 1991, pp. 248-9)

Rosen's view is that organisms lie at the other end of the spectrum. It is generally, he claims, possible to answer 'why' questions about organisms from within the organism without the need to make reference to the environment:

> My claim is that organisms lie at the other extreme as far as entailment is concerned. Their abstract block diagrams manifest *maximal entailment;* in particular, if f denotes a component of such a system, the question "why f ?" has an answer, in terms of efficient causation, *within the system*.

(Rosen 1991, p. 249)

Setting aside for the sake of the discussion the fact that the distinction between mechanisms and organisms doesn't hold for the reasons discussed earlier, it is curious, given Rosen's view of organisms as essentially self-contained systems, that ecological psychologists should view his work favourably. As the quotations show, Rosen's view implies that organisms can be understood as largely independent of their environments whereas it is a fundamental principle of ecological psychology that organisms and their environments exist in relationships of mutuality and reciprocity and that each has to be understood in terms of the other. This has always been a clear feature of Gibson's work:

> The fact is worth remembering because it is so often neglected that the words *animal* and *environment* make an inseparable pair. Each term implies the other. No animal could exist without an environment surrounding it. Equally, although not so obvious, an environment implies an animal (or at least an organism) to be surrounded.

(Gibson 1979/1986, p. 8)

The reciprocity of animals and environments implies that perception involves causal loops connecting the animal to the environment. Gibson was very clear about this when he developed the concept of a perceptual system:

> Instead of looking to the brain alone for an explanation of constant perception, it should be sought in the neural loops of an active perceptual system that includes the adjustments of the perceptual organ. Instead of supposing that the brain constructs or computes the objective information from a kaleidoscopic inflow of sensations, we may suppose that the orienting of the organs of perception is governed by the brain so that the whole system of input and output resonates to the external information.
>
> (Gibson 1966, p.5)

Rosen, by contrast, viewed the relation between organism and environment in the classical representationalist sense that ecological psychologists reject. He also, as reported earlier, specifically endorses the traditional view of perception as indirect:

> As philosophers have pointed out for millennia, all we perceive directly are our selves, together with sensations and impressions that we normally interpret as coming from "outside" (i.e., from the ambience), and that we merely *impute*, as properties and predicates, to things in that ambience. The things in themselves, the *noumena,* as Kant calls them, are inherently unknowable except through the perceptions they elicit in us; what we observe are *phenomena,* which are to an equally unknowable extent corrupted by our perceptual apparatus itself (which of course also sits partly in the ambience).
>
> Rosen (1991, p.56)

One further corollary of Rosen’s emphasis on the organism as an essentially self-explanatory system, all of whose ‘why’ questions have intra-systemic answers, is a denial of the value of evolutionary explanation.

> To me, it is easy to conceive of life, and hence biology, without evolution. But not of evolution without life. Thus, evolution is a corollary of the living, the consequence of specialized somatic activities, and not the other way around…biology is wrapped up with soma and how it operates; thus we cannot invoke evolution as an explanatory or causal principle for these purposes.
>
> (Rosen 1991, p. 255)

Once again the contrast with Gibson’s views is striking. The facts of evolution came to play a central part in the development of his ecological theory of visual perception.

> [T]he fact of information in the light falling upon an organism, is the situation to which animals have adapted in the evolution of ocular systems. The visual organs of the spider, the bee, the octopus, the rabbit, and man are so different from one another that it is a question whether they should all be called *eyes*, but they share in common the ability to perceive certain features of the surrounding world when it is illuminated. The realization that eyes have evolved to permit perception, not to induce sensations, is the clue to a new understanding of human vision itself.

(Gibson 1966, p.155)

Methodological issues.

One of the consequences of Rosen's mistaken supposition that machines cannot exhibit closed causal loops was the distinction between what he called *simple* and *complex* systems. All mechanisms and all systems that could be simulated on computers were said to be simple. A system was complex if, and only if, it had non-computable models. This claim, alone, has a consequence that ought to make ecological psychologists wary of Rosen's work. The consequence is that dynamical systems theory is an inadequate theoretical basis for the study of living systems. The reason for this is that it is based on computable equations. Rosen was quite clear about its insufficiency:

> [T]o assert that organisms, or human systems, are thus complex, is a radical thing to do. For one thing, it says that differential equations, and systems of differential equations (i.e. dynamical systems), which are inherently simulable, miss most of the reality of a complex system…just as any attempt to formalize, for example, Number Theory misses most of its theorems.

(Rosen 2000, p.325)

Many ecological psychologists believe that dynamical systems theory offers a particularly useful set of tools for studying the complex interplay of organismic and environmental variables that characterizes human situated activity. If Rosen were correct, this belief would be erroneous. I should perhaps add that the claim made by some proponents of dynamical systems theory, for example Port and van Gelder (1995), that dynamical systems approaches are an alternative to computational ones is also a mistake. Rosen was right about that.

Conclusion

The anti-mechanist arguments of Robert Rosen have been used by some ecological psychologists to support an attack on computational methods and mechanist thinking generally. Close examination of Rosen's views shows that his epistemology assumes a strong form of indirect realism and his arguments, if valid, would constitute a denial of some of the fundamental principles on which ecological psychology is based. However, his arguments are not valid and do not show that organisms have properties which cannot be captured by machine models. For me this is important because I believe that anti-mechanist thinking in contemporary ecological psychology is based on a restricted view of the possible types of computational model and is acting as a barrier to the development of a genuinely ecological form of computational psychology. Computational thinking in psychology need not, and should not, be tied to models derived from stored program, serial, digital computers. A genuine alternative, which is distinctively ecological, can be built on the foundations laid by Turing in his groundbreaking paper of 1936. In recognition of the striking parallels between Turing's work and Gibson's ecological approach I have suggested that the alternative approach should be called 'ecological functionalism' (Wells, 2006). Had Rosen been aware of the possibilities which open up when one combines Turing's account of inferential entailment in the world of machines with Gibson's account of causal entailment in the ecological world I think he might have reconsidered his anti-mechanism.

Acknowledgements

I am grateful to Peter Kugler and Bob Shaw for their comments on the first version of this paper. Among other points they persuaded me that a discussion of Rosen's philosophy was needed rather than just a rebuttal of his arguments. I am also grateful to Bill Mace for his editorial sagacity.

References

Gibson, J.J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.

Gibson, J.J. (1986). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. (Original work published 1979).

Gödel, K. (1931). On formally undecidable propositions of *Principia Mathematica* and related systems I. In *Kurt Gödel. Collected Works. Volume 1. Publications 1929-1936*. Oxford, UK: Oxford University Press, 144-195.

Gödel, K. (1944). Russell's mathematical logic. In *Kurt Gödel. Collected Works. Volume I1. Publications 1938-1974*. Oxford, UK: Oxford University Press, 119-141.

Hocutt, M., (1974). Aristotle's Four Becauses. *Philosophy,* 49, 385-99.

Landauer, C. & Bellman, K.L. (2002). Theoretical Biology: Organisms and Mechanisms. CP627, *Computing Anticipatory Systems: CASYS 2001—Fifth International Conference,* 59-70.

Moravscik, J.M.E., (1974). Aristotle on Adequate Explanations, *Synthese*, 28, 3-17.

Port, R.F. & van Gelder, T. (1995). *Mind as Motion. Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.

Rogers, H. Jr. (1967) *Theory of Recursive Functions and Effective Computability*. Cambridge, MA: MIT Press.

Rosen, R. (1991). *Life Itself. A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life.* New York, NY: Columbia University Press.

Rosen, R. (2000). *Essays on Life Itself*. New York, NY: Columbia University Press.

Shaw, R.E. (2003). The Agent-Environment Interface: Simon’s Indirect or Gibson’s Direct Coupling? *Ecological Psychology,* 15(1), 37-106.

Turing, A.M. (1936). On Computable Numbers with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, Series 2, 42, 230-265.

Turvey, M.T. (2004). Impredicativity, dynamics, and the perception-action divide. In V.K. Jirsa & J.A.S. Kelso (eds), *Coordination Dynamics: Issues and Trends. Vol.1 Applied Complex Systems* (pp. 1-20). New York: Springer Verlag.

Turvey, M.T. & Shaw, R.E. (1995). Toward an Ecological Physics and a Physical Psychology. In R.L. Solso & D.W. Massaro (eds) *The Science of the Mind: 2001 and Beyond*. New York, NY: Oxford University Press.

Wells, A.J. (2002). Gibson’s Affordances and Turing’s Theory of Computation. *Ecological Psychology*, 14(3), 141-180

Wells, A.J. (2006). *Rethinking Cognitive Computation. Turing and the Science of the Mind*. Basingstoke, UK: Palgrave Macmillan.