

Machine-Likeness and Explanation by Decomposition

Arnon Levy
The Van Leer Jerusalem Institute

© 2014 Arnon Levy
*This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivatives 3.0 License.
<www.philosophersimprint.org/014006/>*

"The entire cell can be viewed as a factory that contains an elaborate network of interlocking assembly lines, each of which is composed of a set of large protein machines." (Alberts, 1998).

"The ribosome is the universally conserved, RNA-based molecular machine that uses an mRNA template to direct the synthesis of protein." (Tinoco & Gonzalez, 2011).

"Now that the innards of the GTP-binding machine are laid open, we can begin to understand its action." (Bourne, 1986)

1. Introduction

Analogies to machines are commonplace in the life sciences, especially in cellular and molecular biology and closely related areas. In these disciplines, the image of a machine often plays an organizing role — shaping conceptions of the phenomena and expectations about how they are to be explained. But while such analogies are common and intuitive, their content is rarely made explicit: it is difficult to find a detailed statement of what makes a system machine-like or an indication of the explanatory contexts where we can expect machine analogies to be fruitful. In this paper I offer a framework for thinking about these questions.

My discussion is guided by the connection between machines and what is sometimes called *decompositional explanation*, *i.e.* an explanation that teases apart underlying components and attends to their structural features and interrelations. When they can be had, decompositional accounts are powerful vehicles of understanding, and they are readily linked to methods of discovery and confirmation. I believe that machine analogies crop up in contexts in which it is believed that a system is amenable to decompositional analysis. And so my aim is to spell out the features that make machine-like systems appropriate targets for explanation by decomposition, so as to better understand what is implied by drawing these analogies and where the limits of decompositional understanding lie. The discussion is oriented toward biology (especially cell biology), because it seems to me that the issues are especially pertinent there. But my central claims can be exported

with minor changes to other relevant contexts, and I have written the paper so that no biological background is presupposed.

The paper begins by looking at a pair of examples (section 2) — one of these illustrates a highly machine-like system, while the other a highly non-machine-like system. From there I work towards a general account, which centers on the idea that machines are systems that exhibit underlying causal order. I elaborate on the relevant notion of order and its connection to decompositional explanation in section 3. In section 4 I discuss the relation between machines, thus understood, and the concept of a mechanism, which has received much attention in recent philosophy of science. Having laid out an account of what machines *are*, I will discuss modeling (sections 5 and 6). A model can depict its target as more or less machine-like, and one key way of doing so involves abstracting from parts and internal relations. I extend this discussion by looking at one significant line of work within recent cell biology, which involves such abstraction and a concomitant move away from machine analogies.

Before I delve in, a remark concerning terminology: sometimes ‘machine’ and ‘mechanism’ are used interchangeably. This is not how I understand the terms. The class I shall focus on is a subset of mechanisms, consisting of orderly mechanisms. In other words, as I use the terms some mechanisms, the orderly ones, are machine-like while others are not. This will be given more precise sense in section 4, and at that point terminology will matter less. But in order to forestall confusion, I will mostly avoid the terms ‘mechanism’, ‘mechanistic explanation’ and their cognates.

2. Motivating examples

The examples I begin with lie at two ends of a spectrum: one is a highly machine-like phenomenon, the other rather unmachine-like. The first example (and another one which I will discuss later) comes from work on the phenomenon of bacterial chemotaxis. Speaking generally, chemotaxis is the self-propelled movement of bacteria and other microorganisms toward or away from specific chemicals in their environment.

This allows them to forage, avoid toxins and adjust their population density. In many bacteria, propulsion is due to a specialized system called a *flagellum*. Consider the following description (and the accompanying Figure 1), drawn from a paper by Howard Berg, one of the foremost experts on the topic:

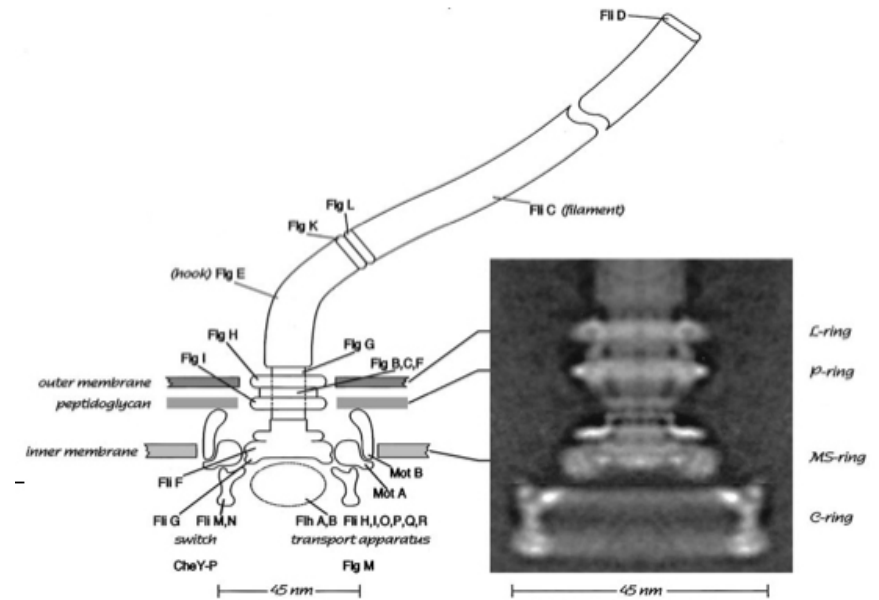


Figure 1. The bacterial flagellar motor. Left: a scaled drawing, noting various subparts. Right: image derived from electron microscopy. [Source: Berg, 2003].

“The bacterial flagellar motor is a nanotechnological marvel, no more than 50 nanometers in diameter, built from about 20 different kinds of parts. It spins clockwise or counterclockwise at speeds on the order of 100 Hz¹, driving long thin helical filaments that enable cells to swim... A [bacterial] cell is propelled by a set of four helical flagellar

1. Hz = cycles per second

filaments... Each filament is driven at its base by a rotary motor embedded in the cell envelope. A cell swims steadily in a direction roughly parallel to its long axis for about a second — it is said to “run” — and then moves erratically in place for a small fraction of a second — it is said to “tumble” — and then swims steadily again in a new direction. When a cell runs at top speed, all of its flagellar filaments spin counterclockwise, the filaments form a bundle that pushes the cell steadily forward. When a cell tumbles, one or more filaments spin clockwise; these filaments leave the bundle, and the cell changes course.” (2003, 19–20)²

Berg goes on to describe the various parts of the flagellum, their specific composition, shape and location; their motions and how these give rise to rotation of the filaments, etc. This is a paradigmatic instance of the machine analogy and decompositional explanation. It attends to parts in detail, specifying their structure and the way they are situated in relation to one another. It aims to show how the concerted action of diverse parts results in the overall phenomenon of flagellar propulsion.

The second example also pertains to the behavior of a large set of molecules. It concerns a very fundamental and widespread phenomenon, namely diffusion. I will discuss a simple case, the first law of diffusion, because it is highly illustrative and recounting it here can be done concisely. But phenomena that essentially involve diffusion, or closely related processes, occur in many other contexts, including in cellular biology (Berg, 1993; Murray, 1992; Nelson, 2007). The first law of diffusion (a.k.a. Fick’s first law) states that a diffusing substance moves from higher to lower concentration - the flux is proportional to (minus) the concentration gradient. To see why this holds, consider the behavior of a single diffusing particle. Suppose the particle moves in small increments of size δ every small unit of time τ . And suppose that it is equally probable to move either to the right or to the left. Let $N(x)$ denote the number of particles at location x . How many particles will move across x in one direction, say, to the right? Well, at each

2. I’ve made minor changes to this paragraph, replacing a number of (presently unnecessary) acronyms.

time-step, half of the particles that are within δ of x will move across it from left to right, and half will move across it from right to left, *i. e.*:

$$-1/2 [N(x+\delta) - N(x)]$$

Let us define the flux, J , to be the number of particles that move across a given area A , per unit time:

$$J_x = -1/2 [N(x+\delta) - N(x)] / A \tau$$

Some simple algebra leads to:

$$J_x = -D \frac{1}{\delta} [C(x+\delta) - C(x)]$$

Where D is the diffusion constant, and the C ’s are the concentrations (*i. e.* number of particles per unit volume) at x and $x+$, respectively. For small enough δ ’s, we can apply the definition of the derivative and obtain³:

$$J_x = -D \frac{\partial C}{\partial x}$$

This is the first law, which states that diffusion is more intense the steeper the concentration gradient of the diffusing substance.

The explanation of the first law pays minimal attention to the structure of components or to how they are laid out in space and time. Instead it begins with a coarse-grained description of a typical particle and derives the first law by aggregating the behavior of many such particles. Correspondingly, there seems to be little temptation to regard a diffusive flux as analogous to a machine.

3. Machine-likeness as underlying order

The examples discussed both concern the dependence of a system-level feature on interactions among constituents. One way to think about

3. I assume a familiarity with the derivative. If it’s been a long time since you’ve thought in these terms, note that the basic message is already contained in the next-to-final step: as the difference in concentration across the interval x to $x+\delta$ grows, the flux grows too, but with the opposite sign. In other words, particles will move against the direction of a concentration difference.

the contrast between them is in terms of underlying order. Flagellum-powered propulsion is a very orderly phenomenon: every component must be in the right place, at the right time, playing the right role. Diffusion, on the other hand, is a disorderly phenomenon: the role of individual particles is insignificant and so is their layout. All that matters are the on-average, in-aggregate properties. In this section I flesh out this informal notion of order and explain its connection to machines and decomposition. Let me begin with a summary statement, in the form of a definition. After providing some general remarks, I will discuss each clause in turn.

Suppose we have a system S , exhibiting a behavior B . S is orderly to the extent that:

- (a) Distinct components of S play different roles in bringing about B .
- (b) Components play their roles in virtue of local relations to other components.

Speaking somewhat metaphorically, we can describe the key idea as follows. An orderly system exhibits an internal division of labor, analogous to that present in many manmade machines: each part does something distinct and recognizable, but there is also interdependence among parts, so that the system's overall behavior is an integrated product of their activities.

Note that under the suggested characterization, whether a system is orderly or not depends on the behavior under consideration. The flagellum is highly orderly with respect to cellular locomotion. But it dissipates heat in a much less orderly manner. So while it is natural to talk in terms of orderliness (and/or machine-likeness) simpliciter, and while I shall do so below, strictly speaking these notions are relativized to an overall effect. (Where this relativization matters, I draw the reader's attention to it.) Moreover, orderliness is a matter of degree. Each of the conditions can be regarded as a separate dimension

of order, and different systems may depart from paradigm cases to a greater or lesser extent along one or both of these dimensions.

I will speak below of decomposition and justify this definition in terms of the relation between order and decomposition. Let me highlight that I have in mind, primarily, a certain mode of explanation and understanding: one that pays close attention to parts, their structure and activity and the way they are situated relative to one another. I suggest that this mode of explanation is closely tied to the notion of order and motivates machine analogies. 'Decomposition' often also refers to empirical methods that literally break down a system into parts. There is, of course, a connection between these two ideas, but it is explanation that is my focus here.

Finally, implicit in the definition is a notion of difference-making: I take it that to say, for instance, that components contribute in virtue of the relations among them, is to say that such relations *make a difference* to the system's overall behavior. The notion of difference-making is closely linked with causal explanation and has received much attention in the literature on that subject. One influential recent view is Woodward's (2003), which treats difference making as a matter of manipulability. To a first approximation, this means that A makes a difference to B just in case B can be manipulated via A , where a manipulation is understood as an ideal controlled experiment. Put differently: A makes a difference to B if by performing a perfectly controlled experiment on A one can effect a change in B . Woodward offers this as an account of causal relations. It can be adapted to situations where A is a component and B is a system-level behavior (Craver, 2007). The manipulability view has distinct advantages and it is quite widely accepted. But there are alternative accounts (*e.g.* Hitchcock, 1993; Strevens, 2008). In principle, I think the present discussion can proceed ecumenically, without committing to an account of difference making. But it seems that clarity is better served by filling in some details, and I will do so by appealing to Woodward's ideas. That said, let me highlight that the big picture does not depend on

this choice. With these points in mind, let us move to a more detailed discussion of conditions (a) and (b).

3.1. *Differentiation of parts.*

Condition (a) says that an orderly system is one in which distinct components play different roles. The underlying idea is that in a machine-like system, an overall job is allocated to distinct sub-elements, and therefore understanding the system involves identifying these elements and their contribution to the system's overall behavior.

Talk of distinct components brings to mind spatially defined parts. Components are often spatially separate, but need not be. The important point concerns functional distinctness, *i.e.* distinctness with respect to the difference made by components. This can be fleshed out in terms of modularity (Woodward, 2003). A component's contribution is modular, *sensu* Woodward, if the difference it makes is independent of the difference made by other components.⁴ That is: component C's contribution is modular if it is possible to disrupt the activity of other components without affecting the contribution of C. In the flagellum, for instance, it is possible to shorten or deform the filament, thereby disrupting its activity, without affecting the motor's rotation, and vice versa (indeed such experiments are routinely performed). These, therefore, are distinct components of flagellum-driven locomotion.

Components may be distinct, yet contribute in the same way to the system's overall behavior. Continuing with the same example: each bacterium is equipped with several flagella. During a "run", they form a bundle that rotates counterclockwise in unison. Each flagellum can still be seen as making a distinct contribution to the bundle, but these contributions are qualitatively similar. By contrast, within each flagellum the motor and the filament play qualitatively different roles. The motor generates torque, while the filament exerts directional force on the surrounding medium. Thus, the idea that components make

4. Woodward defines modularity in terms of the representation of casual relations - rather than in terms of the disruptability of causes "in the world" (2003, 7.4). But this doesn't matter for present purposes.

differential contributions is a special case of modularity — it requires distinct *and* qualitatively different contributions.⁵

The rationale for the requirement of differential contributions follows directly from the idea that orderly systems are good targets for decompositional explanation. Decomposition is useful where it allows one to discern explanatorily relevant aspects of the phenomenon — when attending to parts shows how its features arise. This holds when the contributions of parts are distinct and different. In the starkly contrasting case of diffusion, contributions are qualitatively identical. So one does not get traction on the system by tracking its internal workings in a part-by-part manner. Of course, whether parts contribute differentially is a matter of degree: all else equal, the more distinctions among parts matter, the more orderly and machine-like the system.

I said above that components may, but need not, be spatially distinct. For some people, spatial distinctness, and geometrical properties more generally, have a special status. It is natural to view systems in which components are spatially distinct as especially machine-like. Although I share this inclination, I will not take a definite stand on this point. From the perspective of the general analysis I am offering here, it is possible to treat systems in which parts are localized and spatial layout is important as especially machine-like, or as one among several ways of having distinct components, and hence of being machine-like. (The same holds for spatial relations, which I discuss next.)

3.2. *Importance of local relations*

Condition (b) concerns the importance of local relations among components. Again this should be understood in terms of difference making — the idea is that in machine-like, orderly systems, local relational properties of components make a difference to the system's overall behavior.

Here too the spatial connotation is intentional: geometrical match and physical proximity among components is often important in

5. I do not think the notion of a *qualitative* difference can be given a productive general explanation.

orderly systems, as in the case of the flagellum. But ‘local’ is intended to encompass other kinds of relations. In the general sense, local relations are ones a component has with a designated, typically relatively small, subset of the system’s other components — its causal neighborhood, so to speak. Thus, we have order to the extent that a system’s behavior depends on interactions among small subsets of its components.

The requirement that local relations make a difference is closely related to the idea that an orderly system exhibits internal integration. It is possible for a system to fulfill condition (a), *i. e.* for its parts to contribute differentially, without these contributions being integrated. One kind of case that illustrates this involves properties that arise out of a simple mixture of elements — *e. g.* the acidity of a solution, as measured by its pH — and depend primarily on concentrations and densities. Here different components (solvent molecules, an acidic solute) play different roles. But they do so by freely mixing, so that the specific layout of elements does not matter much. Such a phenomenon appears less orderly than the flagellum — and indeed the value of decomposition is diminished in these contexts.

The importance of local relations can be approached from a different direction. William Wimsatt (1986) has described a notion of aggregativity — phenomena in which the whole is, sometimes literally, no more than a sum of its parts. The concentration of a substance, *i. e.* the number of particles divided by its volume, is a simple case in point. Wimsatt’s analysis of aggregativity proceeds by spelling out counterfactual invariance conditions: an aggregate is a system whose focal properties are indifferent to certain manipulations — especially to various ways of reshuffling components and substituting one for another. If you swap two molecules of the same substance its concentration will not change, because concentration is an aggregative property. In contrast, if you swap parts in the flagellum, it is unlikely to work normally, if at all. This, in a sense, is an operational counterpart to my condition (b). Where there is order, local relations matter, and so moving around parts is disruptive. Conversely, if the system is invariant under Wimsatt-style manipulations, then it is unlikely to be orderly.

3.3. *Machine-likeness and design.*

Typically, manmade machines are the products of deliberate design (and, derivatively, can be assigned proper functions). It is tempting to view machine analogies as embodying a commitment to the designed nature of the analogues. I will not enter into a discussion of ascriptions of design and function, in general or in biology (Buller, 1999; Lewens, 2004). Nor do I need to: I have deliberately resisted the temptation to tie the concepts of machine-likeness and design (or proper function). The notion of order I have outlined makes no reference to the processes through which the system in question has come about — it has to do only with its synchronic causal structure.

My key reason for not wanting to bring in design is that I think that if machine analogies are understood in relation to decomposition (and I believe this to be a central motivation for such analogies in science), then designedness and machine-likeness can and often do come apart. There are products of natural selection or artificial design that are not orderly and do not invite a machine analogy — because they have been designed (or selected) to do what they do in a disorderly fashion. Vice versa too: physical systems that have no apparent design can be helpfully analogized to machines (*e. g.* the solar system and other multi-part mechanical systems). So as far as I can tell, the connection between machine-likeness and decomposition is orthogonal to the question of design.

That said, there can be a connection between the function of a system — in the sense of what it was designed to do — and the degree or character of underlying order. For one thing, recall that whether a system is orderly is defined relative to some focal behavior. In many cases, at least in biology, the focal behavior will be the system’s designed function, because functional behaviors are of great interest to biologists. So the system’s design often enters indirectly into judgments about order and machine-likeness, by fixing the focal behavior. There may also be empirical connections between design and order. It seems plausible, for instance, that designed systems, whether manmade or natural, tend to exhibit more underlying order. This may in part be due

to the need for such systems to have a substantial degree of modifiability over time (Calcott, 2014).

3.4. *Comparison with Woodward.*

In a recent paper James Woodward (2013) has discussed machine analogies, outlining a view similar to the one presented here. Let me offer a brief comparison. Woodward's account also treats machine-like systems as a subset of causal systems.⁶ And like the present account, he too accounts for machine-likeness in a non-design-related fashion.

Woodward specifies three features as key to machine-likeness: modularity, fine-tunedness and stability of intermediate links. I have already discussed modularity, and I have done so by relying on Woodward's earlier work on this notion — here our accounts are identical as far as I can tell. Woodward's fine-tunedness amounts to the idea that in machine-like systems, specific relations among components and particular differences among their intrinsic causal properties matter. I have expressed a similar view but I have broken down this requirement into a condition on parts and a condition on relations. With respect to relations I believe the two accounts to be similar. But with respect to parts, my account requires qualitative differences among components, whereas Woodward makes no such requirement. This is due to the centrality of decomposition to my view, which is closely associated with differences among components. I also tend to think that breaking the fine-tunedness requirement into two sub-conditions, as I have done, affords a better perspective on the range of possible systems and their relation to paradigms of machine-likeness, and allows one to offer a better characterization of newer explanatory trends in biology (as I discuss in section 5, below). Finally, Woodward holds that in machine-like systems internal causal links exhibit greater stability than the system they are part of. This means that interactions among

6. Woodward's terminology is slightly different from the present one: he uses 'mechanistic' and 'machine-like' as synonyms. This, of course, does not affect the substance of the account. But I think it somewhat obscures the relationship between Woodward's ideas and those of the "mechanistic school" in philosophy of biology. I discuss these issues in my own terms in the next section.

components are invariant over a broader set of background conditions than the system's overall behavior. While I believe that stability is a significant feature of causal systems, I am less confident that it contributes specifically to machine-likeness. In part, this is because stability appears to be orthogonal to the appropriateness of decomposition.⁷

3.5. *Analogies and disanalogies.*

To round off this section, let me remark on the cognitive role of analogies and disanalogies between manmade and natural machines.

It is fairly apparent (though worth highlighting) why thinking of a natural system as akin to a manmade machine may be helpful: it enables one to visualize the system more easily, primes one for qualitative causal reasoning and suggests ways of testing hypotheses and amending them in the light of evidence. But the cognitive utility of machine analogies isn't merely a matter of the subjective thinking habits of scientists. I have suggested that there is a genuine, objective similarity between some natural systems and manmade machines. Machine analogies work well wherever decompositional explanation does, and for similar reasons: both are appropriate where fine-grained features of components and their interrelations matter, *i. e.* where the conditions for orderliness are met. To the extent that such properties are absent, we should expect machine analogies to be less compelling, and decompositional analysis to be less explanatory.

That said, let me be clear that my analysis stands independent of whether one wishes to draw an analogy between orderly systems and manmade machines. One may dislike machine analogies or find them unhelpful. But that would not be a reason to reject the idea that, say, the flagellum is more orderly than a diffusive flux. Furthermore, even if (like myself and many biologists) you find the machine analogy helpful, that does not mean that orderly natural systems (*i. e.*

7. A further difference between the present discussion and that of Woodward is that the latter does not highlight the contrast between machine-like systems and aggregative systems (like the diffusive flux). But I think this is largely a matter of exposition and emphasis.

natural machines) resemble manmade machines in every respect. Indeed some important differences may exist. For instance, Daniel Nicholson (2013) points to a number of disanalogies between biological systems and manmade machines. Chief among these is the tendency of organisms to be self-sufficient in certain respects, *e.g.* in that they have a capacity to repair malfunctioning parts or the ability to procure their own energy. To these one may add other differences, such as the degree of noise in the system and, at least in some cases, the level of plasticity, which is generally greater in biological cases.⁸ Now, a person may take such differences to tell against the machine analogy. To some extent they surely do. But I have argued that the analogy plays a significant role in facilitating decompositional explanation, and I regard this as sufficient reason to retain it. I would add that oftentimes, good analogies are useful not only because they direct our attention to shared aspects, but also because they highlight disanalogies.

4. Machines and mechanisms

In recent philosophy of science, considerable attention has been given to mechanisms. The foregoing discussion allows us to make clearer the relation between machines and mechanisms. This will serve as a further elaboration of the ideas presented so far, and will also clarify what seems to me a less-than-perspicuous aspect of the literature on mechanisms.

Consider three well-known philosophical characterizations of the notion of mechanism:

A mechanism underlying a behavior is a complex system which produces that behavior by the interaction of a number of parts according to direct causal laws. (Glennan, 1996, 52)

8. Some of these features (*e.g.* self-sufficiency) will matter more when we are thinking of whole organisms, while others (*e.g.* noise) will be most significant at the cellular or molecular level.

Machine-Likeness and Explanation by Decomposition

Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions. (Machamer, Darden & Craver, 2000, 3).

A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena. (Bechtel, 2006, 26).

These characterizations do not distinguish cases in which underlying interactions are orderly and machine-like from those that aren't. They apply equally well to systems in which parts have differential roles and local relations matter and to those in which such features are absent. The flagellar motor is a mechanism according to these characterizations, but so is a diffusive flux. This is a deliberate choice; in several places these writers state explicitly that they intend their characterizations to encompass more than machine-like systems (*e.g.* Glennan, 1996, §2; Machamer, Darden & Craver, 2000, §5; Craver, 2007, Ch. 1.). One reason is that they aim for good agreement with scientific usage, according to which 'mechanism' usually functions inclusively to denote any underlying causal structure.

Given this inclusive construal, it is reasonable to think that a wide variety of explanations, in a wide range of disciplines, are mechanistic. If one counts all manner of underlying causal structures as mechanisms, then any explanation that appeals to underlying causes is mechanistic. And if one accepts that explanation is (in the main) causal, as most scientists and philosophers of science do, then any phenomenon that involves a relationship between a system's overall behavior and its underlying parts will call for a mechanistic explanation. I regard this as the official stance of advocates of mechanistic explanation. According to it, mechanisms are ubiquitous and so is mechanistic explanation.

Machines are a subset of mechanisms, and explanations that appeal to machine-like structures are a sub-species of mechanistic explanation.

However, some features of the discussion of mechanisms are in tension with this official picture. For instance, mechanists often treat as exemplar systems that have a highly machine-like character, such as a mousetrap (Craver & Bechtel, 2007), a vending machine (Glennan, 1996) or blood-pumping by the heart (Craver, 2001). More importantly, discussions of mechanistic explanation focus, almost exclusively, on decompositional accounts and related empirical methods (*e.g.* Bechtel & Richardson, 1993; Darden, 2006). This gives the impression that mechanistic explanation invariably involves detailed attention to parts and their local relations. But while this is true for orderly, machine-like mechanisms it does not hold in general: aggregative phenomena like diffusion are mechanistic but cannot be effectively accounted for in decompositional terms. And indeed, some mechanists explicitly deny that aggregative phenomena are mechanistic (Bechtel & Richardson, 1993, Ch. 2; Craver, 2001, 2007, Ch. 4). This situation has resulted in some confusion over the scope of the mechanistic program, including developments and criticisms of it which, in one way or another, presuppose an identification of mechanisms with orderly, decomposable, machine-like systems (*e.g.* Matthewson & Calcott, 2011; Milstein & Skipper, 2005; Nicholson, 2012; Woodward, 2011).

I think we would do well to distinguish two pictures, one narrow and the other broad. On the narrow picture, some systems in nature are machine-like and call for decompositional explanation. Such systems, I have suggested, involve differentiation among parts and sensitivity to local relations. To account for the behavior of these kinds of systems, one must attend to parts and relations in a fine-grained manner. There is of course room for disagreement over the details of this account. But I think it is highly implausible to regard it as a general treatment of explanation in science, or in cellular and molecular biology. In contrast, the broad picture concerns mechanisms, *i.e.* any system with multiple interacting parts. Here too the details may be contested. But the picture naturally aspires for generality, since so

much of science, including a lot of biology, is concerned with figuring out how various phenomena arise out of interactions among constituents. My focus here is on the former, narrower category. This can be seen as a contribution to the literature on mechanisms, or at least to one strand within it.

5. Modeling orderly phenomena

The discussion so far has been directed at the question: what makes a system machine-like? This is a question about the world, so to speak. To get an overall picture of the role of machine analogies, we must also attend to questions about how machine-likeness is represented, or modeled.⁹ I will focus on one key kind of operation: abstraction. In this section, I will argue for the general claim that, all else equal, the more abstract a model the less it depicts its target as machine-like. In the next section I will look at an example and comment on some recent developments within cell biology that may point to a shift away from machine analogies.

In speaking of abstraction, I have in mind a simple idea: to abstract is to remain non-specific, to give a coarse-grained description.¹⁰ Abstraction occurs in almost any descriptive context. In thinking about science, it is helpful to distinguish abstraction from idealization (Jones, 2005; Godfrey-Smith, 2009). While abstraction is the omission of information, idealization involves simplifying misrepresentation. A model that says that a cell has a finite volume, without specifying a particular value, is abstract; a model that treats the cell as a perfect sphere is idealized. Thus, as I use the term, abstraction is a matter of saying less than one could (in principle) but it doesn't involve saying something incorrect. That said, while an abstract model is by definition not fully informative (since more could be said about the phenomenon it

9. I do not presuppose any specific account of scientific representation, nor of modeling.
10. One may distinguish the process of abstracting, *i.e.* the leaving out (or at any rate the non-inclusion) of detail, from the product, an abstraction, *i.e.* a detail-poor representation. But this will not matter much here.

targets), it may still be the best answer to some question. In particular, abstracting often enhances explanatory power, because the abstract model omits irrelevant details and highlights difference makers (Strevens, 2008).

Thinking back to the examples I began with, we can see intuitively that there is a connection between abstraction and machine-likeness. The first example, the flagellar motor, contained substantial details regarding the flagellum's parts, their structure, geometry, how they are situated relative to one another and how they move and change throughout the process of locomotion. Diffusion, on the other hand, was given a very skeletal explanation: we derived the first law by considering the average particle, taking very few attributes of it into account. There was no mention of the particles' structure, the relationships among them or any other local aspect of the goings-on inside the flux.

This observation generalizes, and the reasons are apparent given our characterization of order. As one provides less detail, one tends to ignore or smooth out distinctions among parts as well as relations among them. Therefore, an abstract description will portray its target as less sensitive to these features, *i. e.* as less machine-like. Molecular models tend to focus on the structural specifics of parts and on delicate aspects of their organization. Descriptions of such processes as diffusion, on the other hand, go in the opposite direction. They treat the system's constituents (individual particles) as similar, largely independent entities, each behaving in a relatively simple way. Such treatments are common in statistical mechanics and related areas of physics, but they occur in population biology as well, and in other parts of the life sciences that study animal collectives. But note that the connection between abstract description and order has little to do with size. Diffusion is a disorderly process regardless of whether the diffusing elements are ions or organisms. On the other hand, a large structure, such as an organ or an entire organism, can be described in a very concrete way that emphasizes structural and relational properties, and such a system will tend to look more orderly and machine-like.

6. The middle of the spectrum

The derivation of Fick's law stands at or near the end of a spectrum with respect to abstraction. It depicts the phenomenon as all but devoid of internal order and hence as very un-machine-like. Macromolecular explanations, such as Berg's description of the flagellar motor, lie at the other end of this spectrum. They attend to very concrete structural properties of parts and to their organization. But the inverse relation between abstraction and machine-likeness isn't just a matter of the ends of the abstraction spectrum; it applies to points along the way. Let us look at one middle-of-the-spectrum example, which will also enable us to glimpse a potential decline in the prominence of machine analogies in cell biology.

6.1. *The Barkai-Leibler model of chemotactic adaptation.*

This example also comes from the study of bacterial chemotaxis. As noted before, the flagellum serves bacteria to move toward food and away from toxins. A bacterium will "run" in a sustained direction only if the stimulus is graded, *i. e.* only if it detects an increase in the level of food (or, in the case of a toxin, a decrease). This holds irrespective of the absolute level of stimulus. Thus, suppose a population of *E. coli* is floating in a solution, and, all of a sudden, the level of some food item changes sharply, to a higher but uniform level. Initially, the bacteria will respond, beginning to run. But they soon "realize" that there is no gradient and settle down. This is an instance of the phenomenon of *sensory adaptation*, analogous to how our eyes adjust when we move from a dark room to a sunlit outdoor environment.

Bacterial sensory adaptation is simpler than its more familiar counterparts, but it is an impressive feat nonetheless. Barkai & Leibler (1997) offered the presently accepted model of how it works. As we will see, the model focuses on only a handful of the system's elements. It will be instructive to look first at a more inclusive description, which contains some features that Barkai and Leibler ended up leaving out.

Figure 2 summarizes the overall structure of the bacterium's sensory-motor apparatus (some of the details appear only in the caption):

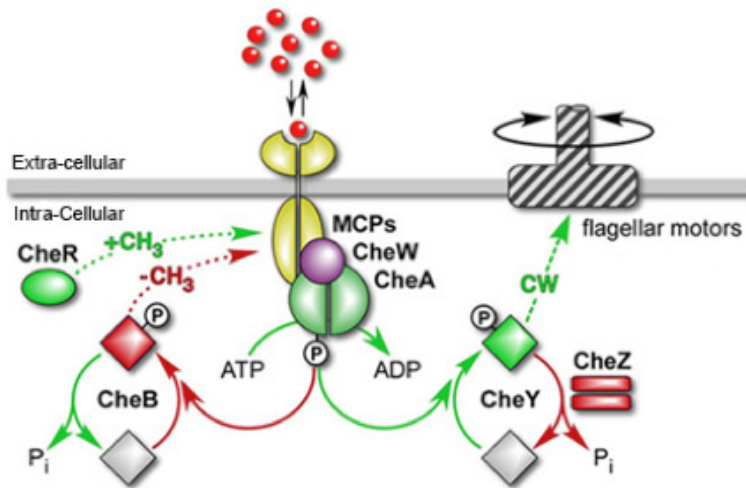


Figure 2. The bacterial chemotactic sensory-motor apparatus. Incoming stimulus (red circles) binds to a receptor complex (MCP-CheW-CheA), thereby activating two secondary complexes: CheY, which modulates flagellar rotation, and CheB, which modulates the sensitivity of the receptor complex to new stimuli. P_i and CH_3 indicate chemical modifications (phosphorylation and methylation, respectively) which facilitate the interactions among the key components. (source: http://chemotaxis.biology.utah.edu/Parkinson_Lab/projects/ecolichemotaxis/ecolichemotaxis.html)

An external stimulus such as a food item (red circles) binds to receptors on the outside of the cell. Via a number of intra-cellular reactions, this results in the dispatch of two secondary elements: (1) CheY (green diamond on the right), which directly modulates the flagellum's rotation. (2) CheB (red diamond on the left), which together with another enzyme, CheR, modulates the sensitivity of the cell to stimuli by chemically modifying the receptor complex.

As noted, this figure is a summary of the processes occurring during sensory modulation of chemotaxis. It contains much less detail than one would find in a standard molecular biology text. Still, it gives some indication of the molecules involved, their structure and the interactions among them. But in working out an explanation for sensory adaptation, Barkai and Leibler moved to a substantially less detailed description. It is given by figure 3. Here the receptor complex is collapsed to an entity that transits from an inactive state E to an active state E_m . This transition is mediated by CheR (abbreviated R). The active form E_m upregulates the "output" (i.e. the flagellum's rotation) and, via CheB (abbreviated B), reverts back to an inactive form. Thus, only five elements are described, their structures play little role, and the interactions among them (activation, inhibition etc.) are treated generically.

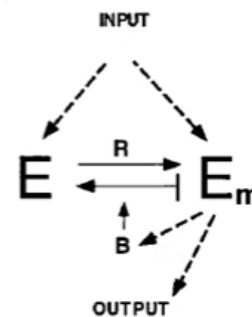


Figure 3. The Barkai-Leibler model. The receptor complex transits between two states: inactive (E) and active (E_m). The active state upregulates the activity of the flagellar motor ("output") while also modulates the $E_m \rightarrow E$ transition. (adapted from Barkai and Leibler, 1997)

Figure 3 embodies a key assumption made in the model: B acts only on the active form E_m . The upshot is that the stimulus generates an “output”, *i. e.* flagellar activity, but does so via a pathway that feeds back fully on itself (because B acts *only* on E_m), performing what is known as integral feedback (Yi et al., 2000). Under these assumptions Barkai and Leibler used standard biochemical principles¹¹ to derive equations for the rate of change of E_m , which I shall not reproduce here. They show that at a steady state, the system always returns to the same level of activity, irrespective of the absolute level of stimulus; unless there’s a gradient, the bacterium returns to the same level of “running”. It is noteworthy that on the Barkai-Leibler model, chemotactic adaptation is robust, in the sense that it does not depend on the concentrations or activity rates of the various components. Insofar as returning to the steady state is concerned, all that matters is that the network exhibits the feedback structure discussed above. In other words, the Barkai-Leibler model depicts adaptation as arising from the way the chemotaxis network is set up — from the system’s connectivity.

The Barkai-Leibler model is less concrete than the description of flagellar propulsion, with which we started. On the other hand, it isn’t quite as abstract as the derivation of the first law of diffusion. I have suggested that there is a correspondence, in descriptions and models, between abstractness and apparent machine-likeness. This holds in the Barkai-Leibler case as well: bacterial sensory adaptation is treated as requiring a certain measure of differentiation among parts, as well as a basic relational setup. Internal structure is depicted sparsely, but not wholly ignored. So the system is treated as less machine-like than a flagellum, but more machine-like than a diffusive flux.

I think these observations generalize: connectivity models — *i. e.* models depicting a pattern of connections without much additional information concerning the structure and behavior of parts — represent their targets as less orderly and less machine-like. They are inherently

so because they aim to distill a pattern of connections, and hence must place less emphasis on the concrete structural and organizational aspects that matter in paradigmatic machines. We might put the point by saying that connectivity models represent a *kind* of machine, a blueprint from which many different concrete machines can be built. Such representations, therefore, do not correspond to any specific machine, and hence the cognitive tool kit that goes along with machines — visualizations, qualitative causal reasoning, etc. — has less of a role to play in this context.

6.2. *A shift in cell biology?*

The Barkai-Leibler model is representative of a growing genre of research in cell biology. Until fairly recently, explanation of intra-organismal phenomena, including the cellular level, focused almost exclusively on relatively machine-like, orderly explananda. Such accounts are couched in highly concrete descriptions of cellular and molecular structures. Knowledge in cell biology was primarily centered on the structure and detailed operation of “large protein machines” (as Alberts, quoted above, puts it). This situation might now be changing, and the Barkai-Leibler model is an illustration of this. The change I have in mind involves both the kinds of phenomena cell biologists are studying and the level of abstraction of their models. Many more instances have emerged over the last decade or so. Such work targets the dynamical properties of cell-level properties and operates via models that treat the cellular architecture in a fairly coarse-grained, often network-based way, and apply associated forms of mathematical modeling.

Much of this work goes under “systems biology”. One example is the study of so-called network motifs, simple circuits that recur in gene regulation networks and are thought to have special roles in virtue of their dynamics. (Alon, 2007a; see also Levy & Bechtel, 2013). Broadly similar work, focusing on basic patterns of connectivity and their dynamical signature, is also conducted under other headings, such as “modular biology” (Hartwell et al., 1999), “network biology” (Barbasi &

11. Specifically, they apply Michaelis-Menten kinetics.

Oltvai, 2004) and the study of “design principles” (Alon, 2007b; Eldar & Elowitz, 2010;). Let me label this, somewhat tendentiously, as the “connectivity turn” in cell biology.

The reasons for the connectivity turn are not entirely obvious at present, nor is its eventual impact. It is probably true that such a turn could not have occurred until fairly recently, because constructing and, especially, testing plausible connectivity models requires a lot of detailed cell biological information — which was not available, say, fifty years ago. The explosion of efficient, cheap, high-throughput experimental techniques in biology, especially since the Human Genome Project, has enabled such fine-grained, often quantitative, delineation of explananda as well as testing. Another, complimentary set of reasons for the connectivity turn is sociological: there has been a significant influx of physicists, computer scientists and other mathematically inclined researchers into cellular and molecular biology — and a concomitant burgeoning of methods, thinking habits and evaluation standards that tend to emphasize abstractness (and generality, which often comes along with abstract description).

Needless to say, all this does not spell the end of structural-molecular work. Surely that is still a central pillar of current cell biology and will continue to be in the future. But the connectivity turn I am describing does signal, I think, a change in the importance of the machine image. As developed in the 20th century, cell biology drew heavily on machine analogies, seeking a detailed understanding of cellular constituents and their structural properties. That program has been enormously successful, but its limitations have become increasingly apparent since around the turn of the century. Many biologists, including some who have led the effort to view the biological world in machine-related terms, are emphasizing the limitations of this way of thinking, in one way or another (Alberts, 1998; Moore, 2012). It has been argued that low-level “machine-ist” work has limited predictive power, that it lacks sufficient generality and that there are inherent difficulties in extrapolating from the structural-molecular level to higher levels of organization (Kirschner et

al., 2000; Bornholt, 2005). Not all of these problems stem from the machine analogy per se, but they are connected to work that seeks concreteness and detailed structural knowledge of parts. These are the features that a machine analogy guides one toward, and that connectivity work directs one away from.

7. Summary

I have outlined a notion of causal order, according to which orderly phenomena are those in which multiple parts make distinct contributions and local relations are important difference makers. Such phenomena, I have argued, have a machine-like character — the more a system meets the conditions for order, the more machine-like it is. No natural phenomenon is orderly to the fullest imaginable extent, but something like flagellum-powered locomotion comes close. The flagellum has many distinct parts, each making a different difference to locomotion; the layout of parts and the specific interactions among them are of utmost importance. On the modeling side, I suggested that abstraction is key in making a model depict the machine-like aspects of a system, and I commented on the growth of relatively abstract, and less machine-like, connectivity models in recent cell biology.

When we make explicit the commitments of drawing the machine analogy, various issues come into sharper relief. One of these is the relationship between machines and mechanisms, which I discussed in section 4. We are also better able to pose big-picture empirical questions, such as: Where in nature are machine-like systems most common, and why? Are there features of the biological world — for instance, certain modes of evolution by natural selection — that tend to give rise to machine-like phenomena? Are there systematic differences between orderly and disorderly processes, such as their degree of robustness or modifiability? My hope is that the present analysis can help organize and orient our thinking about this interesting and under-explored cluster of topics.

Acknowledgements

For comments and discussion I am indebted to Dov Barak, Brett Calcott, Adrian Currie, Susie Fisher, Peter Godfrey-Smith, Eva Jablonka, Ehud Lamm, Sandra Mitchell, Daniel Nicholson, Maureen O'Malley, William Wimsatt and James Woodward.

References

- Alberts, Bruce. (1998), The Cell as a Collection of Protein Machines: Preparing the Next Generation of Molecular Biologists, *Cell* 92: 291–294.
- Alon, Uri. (2007a), Network Motifs: Theory and Experimental Approaches, *Nature Reviews Genetics*, 8: 450–461.
- Alon, U. (2007b), An introduction to systems biology: Design principles of biological circuits. Boca Raton, FL: Chapman & Hall/CRC.
- Barabási, Albert-László & Oltavi, Zoltán N., Network biology: understanding the cell's functional organization, *Nature Reviews Genetics*, 5: 101–113.
- Barkai, Naama & Stanislas Leibler. (1997), Robustness in Simple Biochemical Networks, *Nature*, 387: 914–917.
- Bechtel, William. (2006), *Discovering Cell Mechanisms: The Creation of Modern Cell Biology*, Cambridge: Cambridge University Press.
- Bechtel, William & Robert C. Richardson. (1993), *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton, NJ: Princeton University Press.
- Berg, Howard. (1993), *Random Walks in Biology*, Princeton, NJ: Princeton University Press.
- Berg, Howard. (2003), The Rotary Motor of Bacterial Flagella, *Annual Reviews in Biochemistry* 72: 19–54.
- Buller, David J. (Ed., 1999), *Function, Selection and Design*, New York: SUNY Press.
- Calcott, Brett (2014), Engineering and Evolvability, *Biology & Philosophy*, DOI: 10.1007/s10539-014-9425-3.
- Craver, Carl F. (2001), Role Functions, Mechanisms, and Hierarchy, *Philosophy of Science*, 68: 53–74.
- Craver, Carl F. (2007), *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, NY: Oxford University Press.
- Craver, Carl F. & William Bechtel, (2006) Mechanism. In S. Sarkar and J. Pfeifer (eds.), *Philosophy of Science: an Encyclopedia*. New York: Routledge..
- Darden, Lindley. (2009), *Reasoning in Biological Discoveries: Essays on Mechanisms, Interfield Relations, and Anomaly Resolution*, Cambridge: Cambridge University Press.
- Eldar, Avigdor & Elowitz, Michael B. (2010), Functional roles for noise in genetic circuits, *Nature*, 467(9): 167–173.
- Glennan, Stuart. (1996), Mechanisms and the Nature of Causation, *Erkenntnis*, 44: 49–71.
- Godfrey-Smith, Peter. (2009), Abstractions, idealizations, and evolutionary biology. In A. Barberousse, M. Morange & T. Pradeu (Eds.), *Mapping the Future of Biology: Evolving Concepts and Theories*, Boston: Springer.
- Hitchcock, Christopher. (1993), A Generalized Probabilistic Theory of Causal Relevance. *Synthese*, 97: 335–364.
- Jones, Martin. R. (2005), Idealization and abstraction: A framework. In Martin R.-Jones & N. Cartwright (Eds.), *Idealization XII: Correcting the Model*. Amsterdam: Rodopi.
- Kirschner, Marc, Gerhart, John & Tim Mitchison. (2000), Molecular “Vitalism”, *Cell* 100(1): 79–88.
- Levy, Arnon & William Bechtel. (2013), Abstraction and the Organizations of Mechanisms, *Philosophy of Science*, 80(2): 241–261.
- Lewens, Tim. (2004), *Organisms and Artifacts: Design in Nature and Elsewhere*, Cambridge MA: MIT Press.
- Machamer, Peter, Darden, Lindley & Carl F. Craver. (2000), Thinking About Mechanisms, *Philosophy of Science*, 67: 1–25.
- Matthewson, John and Brett Calcott. (2011), Mechanistic Models of Population-Level Phenomena, *Biology and Philosophy*, 26: 737–756.

- Moore, Peter B. (2012), How Should We Think About the Ribosome? *Annual Reviews of Biophysics*, 41: 1–19.
- Nelson, Philip. (2007), *Biological Physics: Energy, Information, Life*, New York: W.H Freeman.
- Nicholson Daniel, J. (2012), The Concept of Mechanism in Biology. *Studies in History and Philosophy of Science Part C*, 43(1): 152–163.
- Nicholson, Daniel J. (2013), Organisms ≠ Machines, *Studies in History and Philosophy of Biological and Biomedical Sciences* 44 (4): 669–678.
- Skipper, Robert & Roberta Milstein. (2005), Thinking About Evolutionary Mechanisms: Natural Selection, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36: 327–347.
- Strevens, Michael. (2008), *Depth: An Account of Scientific Explanation*, Cambridge, MA: Harvard University Press.
- Wimsatt, William. (1986), Forms of Aggregativity, in A. Donagan, A. Perovich & M. Wedin (Eds.), *Human Nature and Natural Knowledge*, Dordrecht, Holland: Reidel.
- Woodward, James. (2003), *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.
- Woodward, James. (2011), Mechanisms Revisited, *Synthese*, 183(3): 409–427.
- Woodward, James. (2013), Mechanistic Explanation: Its Scope and Limits, *Aristotelian Society Supplementary Volume*, 87(1): 39–65.
- Yoshida, Masasuke, Muneyuki, Eiro & Toru Hisabori. (2001), ATP Synthase – a Marvelous Rotary Engine of the Cell, *Nature Reviews Molecular Cell Biology*, 2: 669–677.