Integrating Autopoiesis and Behavior: An Exploration in Computational Chemo-ethology

Matthew D. Egbert and Ezequiel Di Paolo

Adaptive Behavior 2009 17: 387
DOI: 10.1177/1059712309343821

Adaptive **Behavior**

Integrating Autopoiesis and Behavior: An Exploration in Computational Chemo-ethology

Matthew D. Egbert, Ezequiel Di Paolo
CCNR, University of Sussex, Brighton

It has been argued that the difference between an autonomous entity and an agent is in the ability of the latter to perform behaviors supplemental to processes of self-maintenance (autopoiesis). Theories have been proposed concerning how such behaviors might relate to autopoiesis, but so far, computational models of autopoiesis have paid little attention to these relations. In this article we present a new model designed to explore the relationship between mechanisms of autopoiesis and behavior. We report on three clarifications of the theory provided by the model: (a) mechanisms of behavior can be related to mechanisms of autopoiesis while remaining operationally distinct, (b) the organization of an operationally closed system can change over time while remaining operationally closed, and (c) behavior modulation based upon autopoietic efficacy has limitations that can be avoided through the use of a partially decoupled behavioral system. Finally, we discuss questions that have surfaced during examination of the model.

Keywords autopoiesis · agency · behavior · operational closure · chemo-ethology · artificial chemistry

1 Introduction: Behavior

Fundamental to the concept of agency is the notion of the intrinsic origin of behavior. Agents *act*. Moreover, natural agents are self-produced entities, as in the case of living systems. Following this observation, it has been proposed that self-production implies a form of autonomy, which is a fundamental requirement for agency (Di Paolo, 2005; Ruiz-Mirazo & Moreno, 2004). Theory surrounding the idea of autopoiesis (Maturana & Varela, 1980), the process of self-production of a distinct entity, has furthered our understanding of autonomy, but "bare" autopoiesis fails to fully address behavior and explains only direct self-maintenance (Di Paolo, 2005, 2009). The relation between autopoiesis and behavior is a subtle one. For instance, it has been proposed that the mechanisms underlying the production of behavior (regulated coupling with the environment) have acquired complexity in the history of life through a succession of "decouplings" from underlying metabolic levels (Barandiaran & Moreno, 2008; Moreno & Etxeberria, 2005). These are complex theoretical proposals that can be difficult to grasp in the absence of a simple model of the ideas at work. The model presented in this article has been developed to serve as such a demonstration. We hope that by defining a specific system that has the organizational features described in the autopoietic literature, we shall facilitate the study of autopoiesis and its relationship with behavior and relevance to agency.

Most computational models of autopoiesis demonstrate only self-maintenance but no behavior. Generally the simulated autopoietic entities exist in an environment which requires no organism-scale action

Correspondence to: Matthew Egbert, CCNR, University of Sussex, Brighton, BN1 9QJ, UK. *E-mail:* mde@matthewegbert.com. *Tel.:* +44 (0)1273 872948

(2009), Vol 17(5): 387–401.
DOI: 10.1177/1059712309343821

to continue to exist (e.g., McMullin, 2004; Varela, Maturana, & Uribe, 1974). A few more recent models have demonstrated agents performing a slightly extended autopoiesis; extensions such as incorporating a simple behavior such as osmotic crisis avoidance (Ruiz-Mirazo & Mavelli, 2007) or chemotaxis (Suzuki & Ikegami, 2009). In these cases, the added behaviors are actually *extensions of the mechanisms of autopoiesis*—they are inseparable from the autopoiesis. To stop the mechanism of behavior is to stop the mechanism of autopoiesis. However, this is not the case for the majority of behaviors observed in nature that stop and start while autopoiesis continues.

This raises the question of how mechanisms of behavior can be related to, but somewhat independent of, the mechanism of autopoiesis. How can behavior be integrated with and yet "decoupled" (Barandiaran & Moreno, 2008; Moreno & Etxeberria, 2005) from mechanisms of autopoiesis?

In this article, we describe two tools of analysis that we use to study this question. The first is the notion of "operational closure" which we describe in detail in the next section. The second is the computational model of autopoiesis that we have developed to study the relationship between autopoiesis and behavior. This model does not demonstrate new behavior, rather it demonstrates a different organization of behavior relative to autopoiesis.

2 Operational Closure

Operational closure allows us to identify distinct sets of interdependent processes. Operationally closed sets are not completely independent of other processes, but rather each member process both depends upon and enables other processes within the set.[1] The aim of this definition is to provide a formal framework for the notion of a self-sustaining system. Other expressions for this concept can be found in Varela (1979), Thompson (2007), Di Paolo (2009), and elsewhere.

Given a collection of processes $\mathcal{C}$, a subset $\mathcal{P}$ of those processes is operationally closed if, for every constituent process $P \in \mathcal{P}$, the following holds:

1. Another process $P' \in \mathcal{P}$ is conditioned by process P
2. Process P is conditioned by another process $P'' \in \mathcal{P}$. (P' and P'' can be one and the same process.)

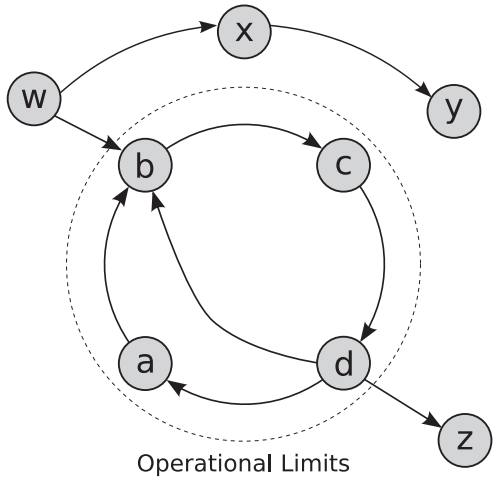

Figure 1 A schematic of an operationally closed system and its environment. Filled circles represent processes and arrows represent dependencies. The processes within the dashed circle fit the criteria for operational closure and are thus part of an operationally closed system.

In a weak sense, a process A is conditioned by another process B if alterations to B result in changes in A. In a stronger sense, the relation can be understood as one whereby B enables A: in this case, for A to sustain itself as a process, B must be present. In both situations the relation of conditioning captures the idea of a process influencing another in a weaker, but also less problematic sense, than that of causality.

With these criteria, given a set of processes and their relationships of condition, it is possible to identify operationally closed networks of processes. As an example, consider the hypothetical system depicted in Figure 1. Each of processes a, b, c, and d depend upon another process within the set *and also* enable another process within the set. There is also a second, nested operationally closed network, consisting of nodes b, c, and d.

So, given a collection of processes and their interdependencies, it is possible to identify operationally closed sets. However, the situation becomes less well defined when we recognize the dynamic nature of interdependencies. Process A might depend upon process B *now*, but not later (or vice versa). The point we wish to underline here is that system interdependencies can change, not only because the processes are dynamic but also because relationships of contingency change. It is already established that the realization of an oper-

ationally closed system can also change over time but, in addition to these structural[2] changes, the very relations between the constituent processes of an operationally closed system may also vary in important ways. As we shall show in this article, an operationally closed system can recruit behavioral processes, while conserving operational closure. Perhaps understanding this dynamic aspect of operational closure can help us understand the relationship between autopoiesis and behavior.

This concept of operational closure is related to, but should not be confused with, the concept of closure as the property of mathematical groups nor with the similar idea of closure of a set of molecular reactions (Dittrich & di Fenizio, 2007). The latter specifies that given a set of molecular species that react with each other, there will be no new product out of these reactions that does not already belong to this closed set. Operational closure in the current context relates to the organization of processes which may include physical, chemical, mechanical and/or behavioral aspects. What matters is that there is a relation of dependence between these processes such that they form a closed network and each process is precarious in the sense that it cannot be sustained on its own and requires the presence of the closed network.

3 Computational Chemo-ethology

Our model lies between two established approaches: artificial chemistries (Dittrich, Ziegler, & Banzhaf, 2001) and computational neuroethology (Beer & Chiel, 2008). We define a set of abstract artificial chemicals (reactants) and a set of rules of interaction and then observe the resulting dynamics.

Among the dynamics that take place within our model are a set of processes that maintain an autopoietic system. This is similar to previous models of autopoiesis in which an autopoietic unity emerges and maintains itself. However, previous models served only to demonstrate autopoiesis. These earlier models neither required nor performed any system-level behavior. In contrast, our model has been developed to explore the relationship between autopoiesis and behavior. We have therefore constructed the reactants and rules of interaction to produce an environment in which the autopoiesis of the agent depends upon a behavior that it performs.

The "computational neuroethology" approach, proposed by Dave Cliff (1991) advocated the study of cognition through the study of behavior of embedded, embodied agents. We adopt the dynamical analysis and behavioral analysis aspects of this approach, studying not only the low-level chemical reactions but also the agent-level behavior.

The computational neuroethology approach has been largely associated with the evolutionary robotics (ER) approach (Harvey, Di Paolo, Wood, Quinn, & Tuci, 2005), where parameters of agents are optimized by a genetic algorithm and the behavior and dynamics of the agent are analyzed. However, it is very difficult to study autopoiesis-based theories of autonomy and agency using typical ER methodology because the agents are not modeled as autopoietic; they neither degrade nor self-maintain. In contrast, in our model, the behavior generating mechanism and the autopoiesis are modeled in the same (artificial chemical) domain. This allows us to study not only the low-level and agent-level processes, but how behavioral and autopoietic processes interact.

4 Model

To explore the relationship between autopoiesis and behavior, we have developed a model inspired by biological cells. A primary motivation in the construction of this model was to produce a simulation that models behavior and autopoiesis in a single, unified manner. The autopoiesis and the chemotaxis of the agent are both the result of the same type of dynamics, namely the interactions between enzymes, a membrane, and a high-energy resource. This makes it possible for us to study the relationship between the two phenomena as there is no "gap" between them in the model.

The model consists of a set of particulate enzymes surrounded by a flexible membrane. These components are simulated in a two-dimensional arena and interact in such a way as to produce a mobile, autopoietic cell-like entity: the agent (see Figure 2). Before explaining the details of the model, we outline here how the agent is both an *autopoietic* and a *behaving* entity.

To show that the agent is autopoietic we must demonstrate that it is composed of a set of reactions that are operationally closed. Each process must be precarious in the sense that it depends upon some other process

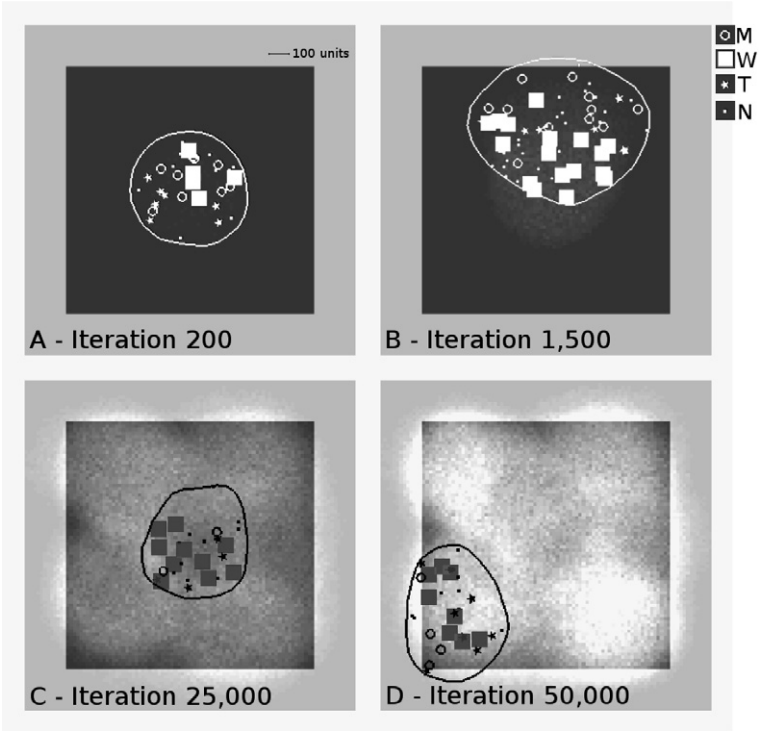

Figure 2 A time series of a healthy agent. To improve readability, only 10% of the enzymes are drawn. See text for a detailed description of the processes involved. (a) Soon after the start of a trial, stochastic processes have already produced some small, random motion in parts of the membrane. (b) The agent has begun moving upwards. Note the asymmetry in concentrations of M and W. Also note the increased size of the agent as a result of the high local levels of resource, R. (c) After having visited most parts of the arena, the agent returns to the center where the R is now high again relative to the agent's previous location. Note the agent has shrunk because of the decrease in availability of R. (d) The end of a typical run for a healthy agent.

in the network to continue. Each process must also make possible at least one other process in the network. This is most easily understood when we consider the two processes of *producing the membrane* and *the autocatalysis of the enzymes*.

First, let us consider the *precariousness* of these processes—their tendency to stop in the absence of each other. This precariousness is included in our model by causing both the membrane and the enzymes to degrade over time. The membrane continually shrinks in size and the enzymes have a significant chance of degrading into substrate particles at each iteration. These processes of degradation can be countered in a healthy agent by the production of newly catalyzed enzymes. But, if the membrane becomes too small, the random Brownian motion of the enzymes will cause them to leave the agent at a rate greater than their production and their population will fall to zero. Thus, the membrane must be kept above a certain size if new enzymes are to be produced. The agent has a problem, however, in that as we have already mentioned, the membrane steadily shrinks. Fortunately for the agent, certain enzymes (enzyme M and W) interact with the membrane in a way that causes it to grow. Here we have a cycle of dependence—the rate of enzyme production depends upon the membrane being sufficiently large and the maintenance of a large membrane depends upon the ongoing production of enzymes. This cyclical, reflexive dependence means that the agent is operationally closed. As a boundary maintaining, precarious, operationally closed system, we consider the agent to be autopoietic.

The agent performs the simple behavior of moving toward high concentrations of reactants (chemotaxis), thus modulating its coupling with the environment. The particular reactant that the agent seeks out is required

for its self-maintaining reactions. As such it can be conceptualized as a food or resource, upon which the agent depends. This reactant is named *resource*, R. In each trial, a finite amount of it is placed in an arena, and the behavior of a single agent is observed as it consumes the resource.

We shall now describe in more detail the various reactants and reactions within our model.

4.1 Resource

The agent exists in a 2-D, infinite environment initialized with a patch of resource (R). Reactions 1a and 2a in Table 1 show the use of R in the production of M, an enzyme that is fundamentally important in the maintenance of the membrane. Thus, the agent requires access to R if it is to be able to produce sufficient M to counteract its degradation. The finite quantity of R in the environment implies a finite maximum possible life span of the agent.

R does not diffuse. Its initial distribution is a square of concentration 1,000 R/unit^2, with dimensions 1200 × 1200 units (about three times the diameter of an agent) with a border of lower concentration (500) of width 200. This distribution of resource was selected to be large enough to observe the agents behavior in a healthy environment, but limited enough to determine how the agent responds as it consumes R and finds itself in an increasingly R impoverished environment. As the agent consumes R in the production of M, the distribution of R changes. This can be seen in the background values of Figure 2 which indicate higher concentrations of R in darker colors. No R is added after the initialization of the simulation.

To be clear, the resource (R) should not be confused with the substrate particles (S_1 and S_2) discussed below. Substrate particles are transformed into enzymes in a process that consumes R. R can be thought of as a high-energy molecule, similar to ATP, that contributes energy to what would otherwise be an endergonic reaction. Substrate particles are the molecules that are transformed into the product.

Table 1 Metabolic reactions. In the rate column, *r* represents the concentration of the resource (R). See main text for further details

Rate	Reactants		Products	
2×10^{-4} R	$M + S_1 + R$	$\Rightarrow$	$M + M$	(1a)
0.15	$M + M$	$\Rightarrow$	$M + S_1$	(1b)
2×10^{-4} R	$W + S_1 + R$	$\Rightarrow$	$W + M$	(2a)
0.15	$W + W$	$\Rightarrow$	$W + S_1$	(2b)
0.4	$T + S_2$	$\Rightarrow$	$T + T$	(3a)
0.15	$T + T$	$\Rightarrow$	$T + S_2$	(3b)
0.4	$N + S_2$	$\Rightarrow$	$N + N$	(4a)
0.15	$N + N$	$\Rightarrow$	$N + S_2$	(4b)
f(R, V)	$T + M$	$\Rightarrow$	$T + W$	(5)

4.2 Enzyme Properties and Chemical Reactions

There are four enzymes that are simulated as particles in Brownian motion: M, W, T, and N. The membrane is permeable to these enzymes. However, the enzymes degrade immediately if outside of the membrane. This is a simple way for us to capture a dependence of the reactions upon the membrane that they maintain, and it can be conceptualized as the membrane preventing an enzyme toxin from entering the agent.

The Brownian motion of these enzymes is simulated by modifying each enzyme's two spatial coordinates by values selected from a Gaussian distribution (Mean = 0, *SD* = 20). This occurs every iteration.

In addition to the aforementioned enzymes, the model includes substrate particles S_1 and S_2. These are not enzymes as they do not catalyze reactions, but rather are reactants that are transformed into products by autocatalytic reactions (reactions 1a, 2a, 3a, and 4a in Table 1). Note that we used two types of substrate. S_1 can be transformed into M and S_2 can be transformed into T and N. Our motivation for utilizing two different types of substrate was to simplify the analysis of the system. Specifically, to explore the relationships between behavior and autopoiesis we damaged parts of the behavioral mechanisms by removing T and N from the simulation (explained below in the section on lesion studies). We wanted to explore the behavioral ramifications of these lesions experiments without the additional complexity of indirect metabolic competition between the autocatalysis of the "behavioral" enzymes (T and N) and the "autopoietic" enzymes (M and W). Using different substrates made

Table 2 Summary of enzyme roles.

Enzyme	Role
M	Adds phospholipids to the membrane. Activates cilia to produce a local outward acceleration
W	Adds phospholipids to the membrane. Activates cilia to produce a local inward acceleration
N	Modifies V according to local concentration of R
V	Stores a representation of the mean quantity of resource experienced by the agent as a whole
T	Transforms M into W when the local concentration of R is lower than the average (represented by the concentration of V)
S_1	A substrate particle that can be transformed into M
S_2	A substrate particle that can be transformed into T or N

it possible for us to keep separate the behavioral and metabolic ramifications of the lesions.

Note that each of the autocatalytic reactions can run in reverse (reactions 1b, 2b, 3b, and 4b). The relative rates of reaction were selected to produce an equilibrium for each reaction pair such that typically a small number of substrate particles would be available for the autocatalytic reactions at any given location within the agent.

To simulate the enzyme–enzyme reactions described in Table 1, every iteration, the arena is divided up into a grid of "pockets," each containing the enzymes in that area. To avoid boundary effects, the offset of the grid is varied randomly each iteration. For each enzyme within each pocket a randomly selected reaction in which it takes part is selected. If the relevant reactants are present, there is a chance (indicated in the "Rate" column of Table 1) that the reaction will occur, that is, the reactants will be replaced by the products.

It should be noted that this method of simulating enzyme–enzyme reactions means that, as in real chemistry, the reactant concentration affects the chance of a reaction occurring during any iteration. For example, if an area contains 1,000 T (and sufficient S_2) the chance of autocatalysis of T (reaction 3a) occurring somewhere in that area is more likely than if it only contained 10 T. This aspect, in conjunction with the bidirectionality of the autocatalytic reactions, means that if one of the autocatalytic enzymes reaches a high concentration, the backward reaction (breakdown of the autocatalyst into substrate) becomes more likely than the forward reaction. In this way, the system does not transform all substrate to enzymes, but instead finds a balance consisting of a small population of both types of substrate.

There is one last enzyme which we must introduce. V is different from the other enzymes in that it is assumed to have a very high rate of diffusion (relative to its production and consumption). On this basis it is modeled as becoming instantaneously distributed (i.e., of equal concentration) across the cell within each time step. Computationally, this is most efficiently realized by representing V with a single, cell wide, continuously valued, concentration variable, rather than by multiple discrete particles with specific positions which have to be updated each iteration.

The concentration of V is increased by N when N is in a higher concentration of R than V. Similarly, N decreases the concentration of V when the concentration of R is lower than the concentration of V. Thus the concentration of V is updated by each N, each iteration, according to Equation 6 where c is the rate constant (1×10^{-3}) and G_n is the local concentration of R at the location of the nth N enzyme. We assume that this is the only reaction that affects the concentration of V.

$$\frac{d\mathrm{V}}{dt} = \sum_n c(G_n - \mathrm{V}) \tag{6}$$

In addition to the autocatalytic reactions in Table 1, there are two non-autocatalytic reactions (2a and 5). Both of these reactions involve W. This enzyme can

be thought of as an alternative form of M in that it participates in the same reactions as M and degrades to the same substrate. We shall see below that the only actual difference between these two enzymes is in how they interact with the membrane; M tends to cause a local motion of the cell away from the center and W tends to cause the opposite motion. As we shall see, these enzymes are fundamental to the chemotaxis behavior. W is not autocatalytic, but instead catalyzes the production of M. This is represented in reaction 2a. W is only produced by the transformation of M into W by T. The rate of this reaction is determined by the relative concentrations of R and V according to the function $f(\mathsf{R}, \mathsf{V}) = 0.9 \cdot H(\mathsf{V} - \mathsf{R})$ where $H()$ is the unit step function.

The rate of reaction 5 can be thought of as a comparison mechanism. Where the local concentration of R is lower than the concentration of V, there is a high probability of T transforming M into W. This mechanism is inspired by and similar to the type of mechanism hypothesized by Macnab and Koshland (1972) to occur in chemotactic *Salmonella typhimurium* (for a discussion of the chemotactic biochemical pathways found in bacteria see: Blair, 1995; Falke, Bass, Butler, Chervitz, & Danielson, 1997). The biochemical mechanism is slightly different from that in our simulation, in that it compares a current concentration to information about a previous concentration rather than comparing the concentration of two currently present chemicals. However, as can be seen in the following description given by Blair, the principal concept of a differential reaction based upon a comparison of one chemical concentration to another is clearly part of the biological system.

> The cell measures the concentration encountered during the past second and compares it with that encountered during the previous three or four seconds, basing decisions to run or tumble on the difference. Bacterial chemotaxis thus involves a simple, very short-term memory …[that] allows the cell to make the temporal comparisons that guide its choices to run or tumble. (Blair, 1995)

The reactions between V, R, T, M, W, and N produce a comparison mechanism. Speaking anthropomorphically, the agent keeps track of the average concentration in its local area of R by varying the concentration of V. It uses this value to transform M into W when they are in areas of low concentration of R relative to the average. This produces an asymmetrical distribution of M and W that induces chemotaxis.

All non-substrate enzymes (excluding V) degrade into substrate. Every iteration, there is a $p = 0.001$ chance that an enzyme will degrade into its substrate (M degrade into S_1 and the other non-substrate enzymes degrade into S_2). As mentioned above, if an enzyme leaves the confines of the membrane, it is immediately transformed into substrate (M or W into S_1 and the other non-substrate enzymes into S_2).

4.3 Membrane Dynamics

The flexible membrane is modeled as a circle of mass-points connected with linear and rotational springs.[3] Each mass-point with its associated linear and rotational spring represent one "membrane-section" (Figure 3). The rest-length of the linear springs and the mass of the membrane-section is proportional to the number of phospholipids (P) in the membrane-section. The linear springs apply a force $F = -kx$ to the mass-point, where k is the spring constant and x is the distance that the spring has been displaced from its rest-length. In this way, P influences the size and mechanical properties of the membrane.

The rotational spring applies torque $\tau = -\kappa\theta$ (where κ is the spring constant and θ is the rotational displacement of the spring) to the associated mass-point and its two neighbors. These forces can be thought of as a crude simulation of membrane rigidity. The rest-angle for each of the rotational springs is π (a straight line with the mass point in the center).

The simulated membrane consists of 32 such points connected in a circle. This number was selected because it is relatively low, reducing computational load, while high enough to produce a relatively smooth membrane shape.

Each membrane-section degrades over time. This degradation takes the form of a steady exponential decrease in the number of phospholipids at each membrane section over time. The particular rate of degradation was selected to produce a system that degrades rapidly in the absence of the system's self-production, but is stable when the agent is healthily self-producing. Left unchecked, the degradation of the membrane causes the agent to eventually shrink to a size too small to maintain the populations of enzymes. This is the equivalent of death within our model.

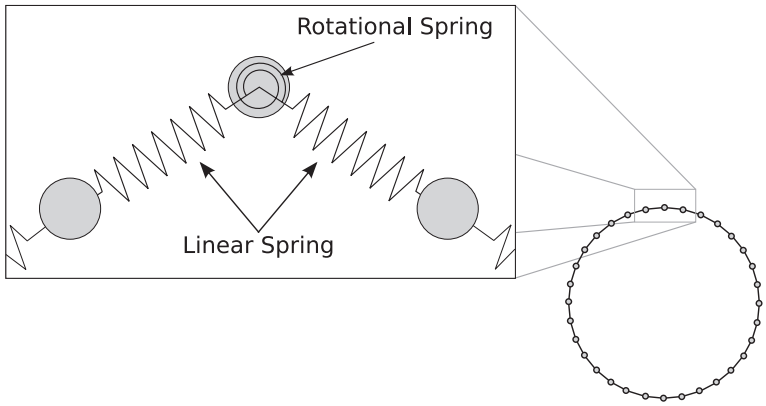

Figure 3 Schematic of a membrane-section.

It is interesting that quantifying this size is non-trivial. The viability (likelihood of survival) of the agent is not only a matter of the size of the membrane, but is the result of many interacting factors including the number and location of various enzymes, the amount of resource available, and so forth. We have begun investigating how the viability of autopoietic systems could be measured (Egbert, Di Paolo, & Barandiaran, in press).

Diffusion of P occurs between neighboring mass points according to Equation 7 where k is the rate of diffusion (0.01) and P_x represents the number of P at mass point x. This rate of diffusion was selected as it tends to maintain a roughly circular membrane.

$$\frac{dP_x}{dt} = -P_x^2 \times 10^{-6} + k((P_{x-1} - P_x) + (P_{x+1} - P_x)) \quad (7)$$

Membrane-sections have inertia and are subjected to a drag force proportional to the square of their velocity. This force represents the drag that would be present for a cell in a viscous medium.

$$F_{drag} = -0.5v^2 \quad (8)$$

4.4 Membrane–Enzyme Reactions

Membrane particles do not prevent the motion of enzymes. However, upon contact with the membrane (considered to occur if the Brownian motion of an enzyme causes it to cross a membrane section), M and W contribute 185 phospholipids (P) to the membrane. This number was found by experimentation to be sufficient to counter the degradation of the membrane, but not so high as to make the cell grow to an inappropriately large size. In addition to contributing phospholipids to the membrane, M particles briefly activate cilia in the local region where they collide with the membrane. This is modeled by applying a small force upon the membrane point perpendicular to the tangent of the surface of the membrane. M and W activate cilia in different manner such that for M, the resulting force is outwards (away from the center of the agent) and for W the force is toward the center of the agent. These interactions between enzymes and membrane cause the deformations of the membrane as well as the motion of the agent.

4.5 Initial Conditions

We initialized the model described above with a circular membrane of radius 300, placed in the center of a square arena, with 10 particles each of T, N, and M placed inside the membrane. Also, S_1 and S_2 particles were randomly distributed around the arena at a density of 0.001 particles per unit square for each type. As mentioned at the beginning of this section, the arena is initialized with a square of R, of concentration 1.0, and with dimensions 1200 × 1200 units. This square of R has a border of lower concentration (0.5) R of width 200. Outside of this border, the concentration of R is zero.

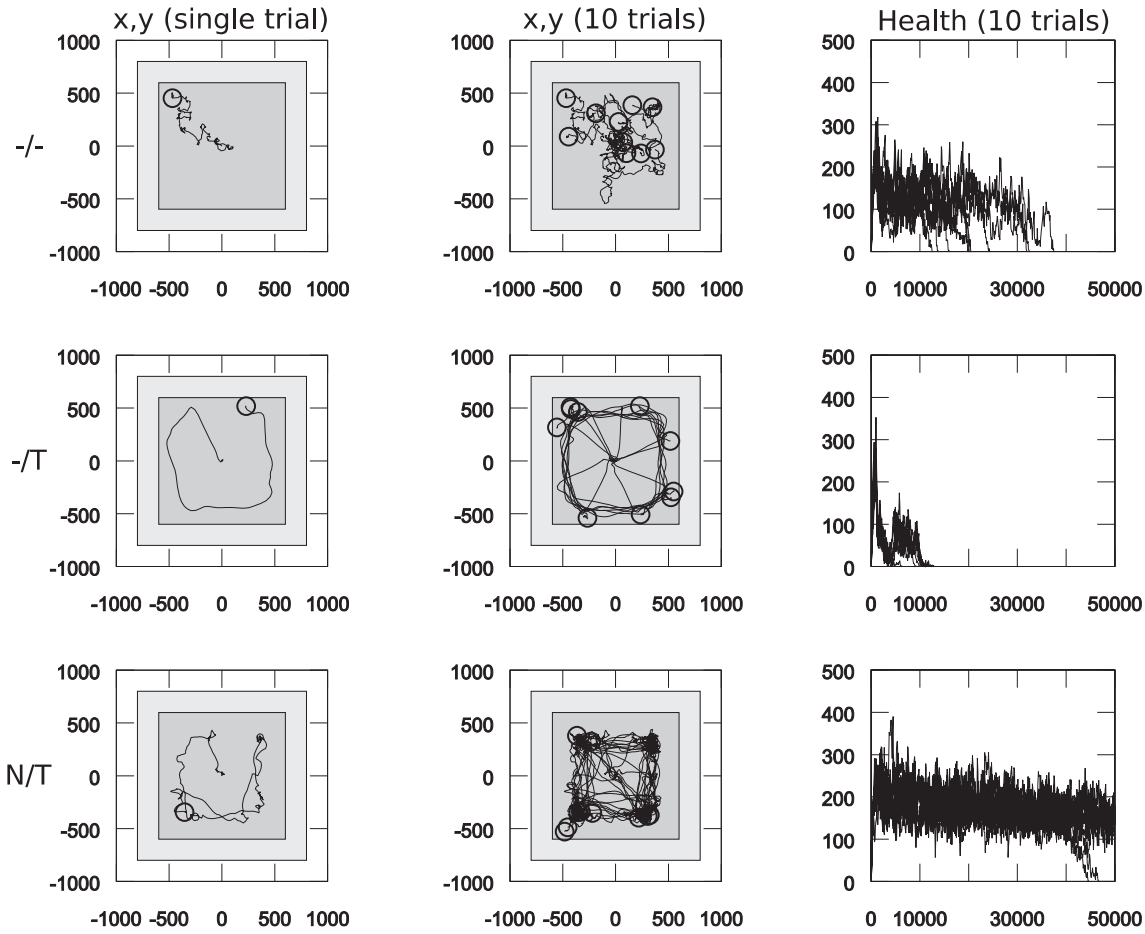

Figure 4 Paths and health of the healthy (N/T) and lesioned agents. Paths are indicated for a single trial (left column) and for 10 different trials (center column). The circles represent the location of the agent at the end of the trial and do not represent agent size. Note the doubling back in the healthy agent and the avoidance of previously visited areas by the –/T agent. The rightmost column indicates agent health (number of M particles) for the same 10 trials.

5 Analysis

Initially the agent remains momentarily stationary (see Figure 2a). Soon however, the autocatalysis of M begins to significantly decrease the local concentration of R. Brownian motion of the enzymes and their stochastic and therefore slightly asymmetric reactions with the membrane result in part of the agent extending into an area higher in concentration of R than those areas that have been depleted by the agent. N particles in this area increase the concentration of V. Then, T transforms M into W in the parts of the agent that are in the R-depleted area. The asymmetrical distribution of M and W particles, their Brownian motion, and their different effects upon collision with the membrane result in a motion of the agent away from the area low in concentration of R. Figure 2b shows an agent after it has begun to move in this way. The motion in this case is upwards but varies randomly as can be seen in Figure 4.

As the agent moves, it continues to reduce the local concentration of R. Areas that the agent has occupied for a longer period of time are more depleted of R. This causes areas toward the rear of the agent to be lower in R which causes further production of W which propagates the motion of the agent. In this way, the agent moves around the arena in a relatively directed manner. That is to say, once it starts moving in a particular direction, it tends to continue in the same direction.

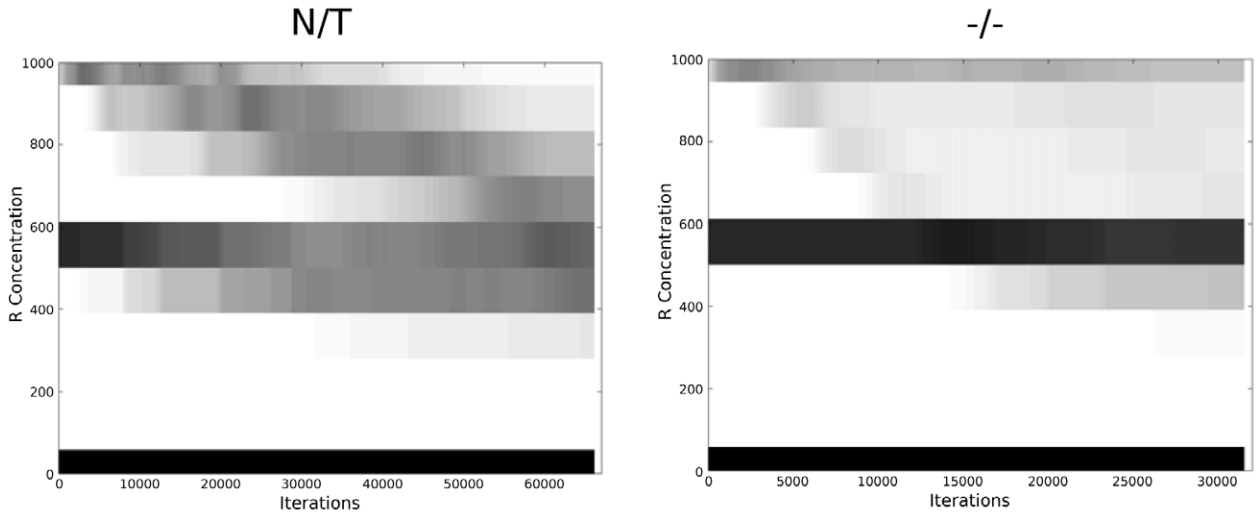

Figure 5 R density plotted against time for healthy and lesioned (–/–) agents. Brighter values indicate that a larger portion of the arena has the indicated R-concentration.

This directed motion is a good demonstration of how local molecular interactions can acquire a global coherence, creating a spatial asymmetry that is the enabling factor for the agent's motion. Although the underlying mechanisms are quite different, local to global bidirectional interactions have been found in other minimal systems capable of generating self-movement in a homogeneous spatial situation, for example, self-moving oil droplets studied by Hanczyc, Toyota, Ikegami, Packard, and Sugawara (2007). In their physical system and in our model, the conditions that cause movement are maintained by the movement itself.

The lower right plot in Figure 4 shows the motion of the center of a healthy agent during a typical run. It moves around the arena, turning when it approaches the end of the high-R concentration square, occasionally doubling back and often returning to areas it has been before. As can be seen in the left plot in Figure 5, a healthy agent consumes all of the high-R-concentration areas at least briefly at different times during a typical run.

6 Lesion Studies

To demonstrate the interactions between the mechanisms of behavior and of autopoiesis, we performed three lesion studies of our agent. By *mechanism*, we mean a set of processes. These studies consisted of the complete removal of either N or T or both. The removal was performed after an initial settling period of 1,000 iterations which was identified, by observing non-lesioned trials, as sufficient time for the system to "relax" from initial conditions. Both N and T are only produced by autocatalysis. Thus, when lesioned their population remains zero for the duration of the trial.

Lesion studies are typically used to analyze the dynamics of complex systems. It is no different with their use here. Though we designed the system, it is complex enough that this kind of analysis is useful for understanding how different aspects of the system are interacting.

The lesion experiments are identified by two-letter acronyms that indicate which of these particles were not transformed into substrate: N/T ≡ all enzymes intact; –/T ≡ T are intact, but N is transformed into S_2 at iteration 1000; –/– ≡ T and N are both transformed into S_2 at iteration 1000.

6.1 Behavior Determined Directly by Autopoiesis Efficacy is Limited

Let us first examine the lesioned agent –/–. In this agent, the main behavior-producing enzymes have been destroyed. The only reactions that remain active

in these agents are the autocatalysis of M, and the interactions between M and the membrane.

The rightmost column of Figure 4 shows the number of M particles in the simulation plotted against time, each plot showing 10 trials. We use this number as an approximation of health of the agent as M particles must be produced for the membrane to maintain sufficient size to enclose enzymes. It is clear when comparing –/– to N/T that the lesioned entities tend to have a shorter life-span. The longest surviving –/– agent dies shortly after iteration 35,000 whereas almost all of the N/T agents survive past the end of the trial (past 50,000 iterations). The mean survival-time of –/– agents was 22,787 iterations (*SD* 8,486.1). As can be seen in Figure 4 the –/– agent moves around the arena, but in a less directed, more random walk than the N/T agent.

The motion of the –/– agent is caused by the agent depleting the local R concentration to levels where further production of M becomes extremely unlikely. In areas where R has been so reduced, M levels fall, causing an asymmetrical distribution of M within the agent, causing an asymmetry in the activation of cilia which produces motion. This mechanism is reminiscent of one of the extensions of the original model of autopoiesis made by Suzuki and Ikegami (2009), in that lesioned agents only begin to move when part of them slows its contribution to autopoiesis. That is to say, the agents only move after they have come close to "death," that is, close to their viability boundary. When in this fragile state, the random variation in, for example, the particle trajectories can be sufficient to inhibit motion for long enough as to cause death.

Unlike the non-lesioned agents, the chemotaxis of –/– agents is modulated quite directly by the efficacy of the autopoietic system. This direct relationship between behavior and autopoietic viability is a fragile organization in two ways. First, in a system such as –/–, response to a dangerous phenomena requires a change (decrease) in efficacy of the autopoietic system. This is, by definition, a decrease in viability which correlates with an increase in fragility of the agent. This increase in fragility makes the agent less capable of surviving various scenarios, for example, turning corners or enduring unlucky stochastic dynamics such as a brief unlucky motion *down* the R-gradient. Second, direct response to change in the efficacy of the autopoietic system can trap the agent in a local maximum. This local maximum is surrounded by more damaging situations, so the direct response would keep the agent in the local maximum. However, under certain circumstances, just on the other side of the more dangerous situations there might be more advantageous situations. An agent whose behavior is determined directly by the efficacy of the autopoietic system would never find these superior conditions, but an agent whose behavioral mechanisms are more decoupled from the autopoiesis might be capable of forging out into a less safe environment in search of a superior environment. It is interesting to consider this dynamic while examining the paths taken by the agents (Figure 4). The N/T agents (with their partially decoupled behavioral mechanism) clearly make better use of resource, tending to feed on the higher resource areas before consuming areas down to dangerous levels (see Figure 5).

What we wish to highlight here is that a decoupled behavioral system, a system that is not as directly determined by the efficacy of the autopoietic system, is not subject to the same limitations. A behavioral mechanism could, for example, continually move the agent from one precarious situation to the next, and in such a motion, produce an environment for the autopoiesis that is actually less precarious than any of the situations are on their own. This is a capability of a decoupled behavioral mechanism that as far as we can see is absent in bare autopoietic systems.

6.2 Mechanisms That Are Related But Distinct

Let us now examine the –/T agents. As a result of lesioning the N enzymes, the concentration of V becomes fixed. Speaking anthropomorphically, these lesioned agents are incapable of changing their standards for what is good and what is bad. Without this ability, they move around the arena once, consuming some R from each area and then lie still. At this stage inappropriately high value of V (compared to a healthy agent) will cause T to change *all* M to W preventing any asymmetry.

In the long term, these agents have a lower average health and life span compared with the non-lesioned agents (see right column in Figure 4), but in the short term the autopoiesis continues. The same can be said for the –/– agents. Furthermore, it is easy to imagine a small change in the environment (e.g., if the R diffused through the environment) that makes the motion

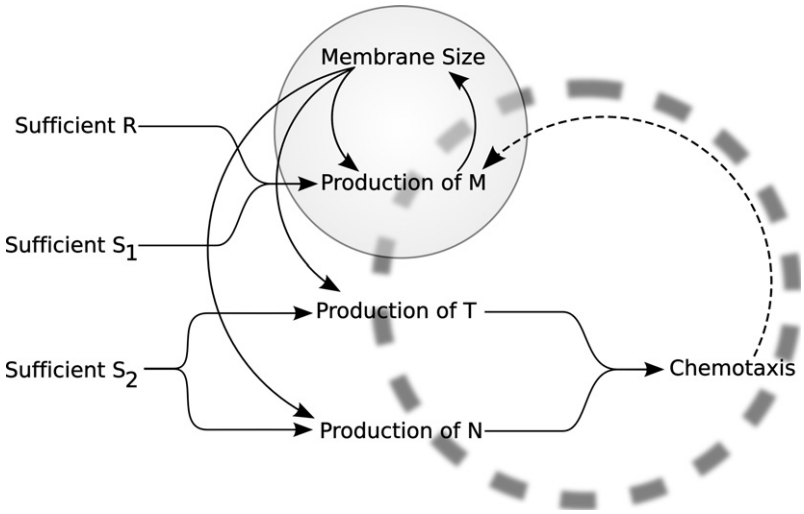

Figure 6 Interdependencies within our model. Arrows represent interdependencies as in Figure 1.

of the agent unnecessary for the autopoiesis to continue. What we are driving at is that the behavioral mechanism in all of these agents is related to autopoiesis *but also distinct*. We have removed a component of the behavioral mechanism without removing a component of the autopoietic mechanism. The behavior affects the autopoiesis indirectly rather than being directly part of the autopoietic process. Because the behavioral mechanisms are somewhat decoupled from the mechanisms of autopoiesis, it is possible for the entity to continue to exist *without* performing the behavior. This is, as far as we know, unlike previous models of autopoiesis (Suzuki & Ikegami, 2009) which if they were self-maintaining were always performing their behavior (e.g., chemotaxis).

This relationship between the mechanisms of autopoiesis and behavior is asymmetrical. Without autopoiesis there is no agent to perform the behavior but the autopoiesis is not dependent (in the short term) upon the behavior. In the case of our model, the chemotactic mechanism depends upon the autopoietic mechanism but the autopoietic mechanism can operate in the absence of the chemotactic mechanism. This is not to say that the behavior is irrelevant to the autopoiesis. The health plots in Figure 4 clearly show the behavior tends to have an effect on the duration of autopoiesis, but the relationship is more indirect and more subject to change than the dependence of the behavior upon autopoiesis.

To understand the relationships between these mechanisms it is helpful to view the system from a perspective of operational closure. This can be most easily realized through study of Figure 6 which depicts the interdependencies present within our model.[4] The clearest example of operational closure in our model is the autopoietic system, which is highlighted by the gray filled circle. For the M particles to be produced, the membrane must be large enough and for the membrane to be large enough, M particles must be made. A membrane (physical boundary) is maintained by these processes and therefore this collection of processes is not only operationally closed but also autopoietic.

Just as the autocatalysis of M relies upon the membrane encircling sufficient substrate particles, so does the autocatalysis of T and N. These relationships are indicated by the arrows from "Membrane Size" to "Production of T" and "Production of N." The T and N are integral to the reactions that cause chemotaxis. But what is less clear is the relevance of chemotaxis to the production of M. This depends upon the environment in which the agent exists. For example, we can imagine that in an environment in which the consumption of R is insignificant compared to the replenishment of R, then no motion is necessary for the agent. It can remain more or less stationary, existing indefinitely. If, however, the agent depletes R in its local environment, if it is to survive it must move to another area to find more R. Because this dependence changes depending upon the environment in which the agent exists, we have depicted this dependence as a dashed arrow.[5] We wish to draw attention to these organiza-

tional relationships in order to highlight the non-intuitive notion that the operationally closed system can change organization over time. In this case, depending upon the environmental conditions, the operationally closed system can include only the autopoietic cycle (the filled circle) or it can include both the autopoietic cycle and the behavioral mechanisms (dashed circle).

Finally, it is worth briefly considering a hypothetical alternative system organization. What if the production of T and N depended upon the chemotaxis? This would produce two new operationally closed loops (one consisting of "Production of T" and "Chemotaxis," the other consisting of "Production of N" and "Chemotaxis"). These loops have their own optimal operating conditions and it would be possible for such behavioral loops to have norms different or even in conflict with the norms of the primary autopoietic system (some discussion of this can be found in DiPaolo, 2009). There is not space to discuss this further here, but we intend to return to this idea of behavior in conflict with autopoiesis in future explorations of this model.

7 Conclusions

This article reports on a new computational model that enables us to explore the relationships between mechanisms of behavior and autopoiesis. The model, which incorporates aspects of computational artificial chemistry and neuroethology has helped us examine the idea of decoupling in the context of mechanisms of autopoiesis and behavior. Specifically, three clarifications have been made through analysis of the model.

First, behavior modulation based upon autopoietic efficacy has limitations that can be avoided through the use of a partially decoupled behavioral system. An agent whose behavioral mechanism is based on its autopoietic efficacy must approach its "viability boundary," becoming more fragile before it can respond. Furthermore, direct sensitivity to autopoietic rates can trap an agent in local viability maxima that are not optimal solutions in the long term. These limitations can be avoided by having a behavioral mechanism that is not as directly modulated by autopoietic processes.

Second, mechanisms of behavior can be related to mechanisms of autopoiesis while remaining operationally distinct. We have been able to make more explicit the idea of decoupling as a form of changing dependence and independence within operationally closed networks. This relationship of *related but independent* is built upon the idea that operationally closed systems change over time.

Third, the organization of an operationally closed system can change over time while remaining operationally closed. In this context of dynamic operational closure, behavioral mechanisms are those mechanisms that are not always in operation and can, depending upon environmental conditions, affect viability. In contrast, autopoietic mechanisms are those mechanisms that are always required, extenuating circumstances aside, to occur if the agent is to continue to exist.

The simplicity afforded by our model (compared with biological systems) has enabled us to see more clearly how behavioral and autopoietic processes can operate over the same physical components and have relations of dependence that can change depending upon the situation.

This relation between behavior and self-production reveals an interesting possibility for operationally closed systems, one that has rarely been made explicit in theoretical terms. Namely, an identity can be maintained while the network of interdependencies changes over time. While the metabolic loop remains always active and unaltered in its closure, the dependence of some of its processes and inner relations on the efficacy of behavior is an organizational, not merely a structural, change to the system. We suspect that this possibility of entering temporarily different modes of system organization can play an essential role in developing the theory of autopoiesis toward an account of biological transformations of organization (e.g., from embryo to adult, or the appearance of novelty across generations), which is at the moment one of its blindspots.

To summarize, this model has allowed us to see the relationship between self-production and behavior-production in a more concrete manner. We intend to continue to further explore how simultaneous mechanisms, operating over the same physical space can interact to produce the cognitive abilities that we associate with the notion of agency. Future study may include increasing the variety of environmental conditions within the model. For example, one area of the arena might contain non-diffusing R, while another area contains diffusing R. In these different environments, the dependence of the autopoiesis upon the

chemotactic behavior depends upon the location of the agent. This would allow us to study the changing dependency and operational closure in one continuous system. In the more distant future it would be interesting to implement an evolutionary process in this environment to see what the long term trends would be in terms of the relationship between behavioral and autopoietic mechanisms.

Notes

1 Developing the formal definitions of *process*, *dependence* and *enabling* is outside of the scope of this article. For the present discussion, it suffices to resort to intuitive understanding of the terms. This is an area in need of attention (see Virgo, Egbert, & Froese, in press).

2 In the autopoietic literature *organization* and *structure* have significantly different connotations. The organization of a system is the way the different parts of the system relate to each other and the structure is what the different parts of the system are made out of. The structure of a membrane (the molecules out of which it is made) changes over time, but the organization (e.g., the membranes ability to maintain high concentrations of reactants within the cell) can not change without the system failing.

3 The model we have used for the membrane is more complicated than is strictly necessary. We developed this mass–spring model with future extensions in mind.

4 It is worth mentioning that this Figure could be drawn at different scales. That is to say we can describe the relationships of dependence at different levels of detail (e.g., we chose not to include the intermediate process of the creation of W by T as it is not important to our argument. It suffices to describe chemotaxis as being dependent upon the production of T).

5 It is also possible to imagine contrived scenarios in which the other interdependencies cease to exist. In this sense the dependence of continued enzyme production upon chemotaxis is not unique. However, we have chosen to highlight the changeability of this dependence as an example of how certain natural organizationally closed systems can be formed and cease to exist, as different natural environmental conditions are encountered.

Acknowledgments

Thanks to Xabier Barandiaran, Tom Froese, Nathaniel Virgo, Lucas Wilkins, Marieke Rohde, and the three anonymous reviewers for their helpful comments. Also, thank you to the participants and organizers of the Kyoto Agency Workshop for the interesting conversation that inspired this work.

References

Barandiaran, X., & Moreno, A. (2008). Adaptivity: From metabolism to behavior. *Adaptive Behavior*, *16*, 325–344.

Beer, R. D., & Chiel, H. J. (2008). Computational neuroethology. *Scholarpedia*, *3*(3), 5307.

Blair, D. (1995). How bacteria sense and swim. *Annual Review of Microbiology*, *49*, 489–522.

Cliff, D. T. (1991). Computational neuroethology: A provisional manifesto. In J. A. Meyer & S. W. Wilson (Eds.), *From Animals to Animats: Proceedings of the 1st International Conference on Simulation of Adaptive Behavior*. Cambridge, MA: MIT Press.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, *4*(4), 429–452.

Di Paolo, E. A. (2009). Extended life. *Topoi*, *28*, 9–21.

Dittrich, P., & di Fenizio, P. S. (2007). Chemical organization theory: towards a theory of constructive dynamical systems. *Bulletin of Mathematical Biology*, *69*, 1199.

Dittrich, P., Ziegler, J., & Banzhaf, W. (2001). Artificial chemistries – a review. *Artificial Life*, *7*(3), 225–275.

Egbert, M. D., Di Paolo, E. A., & Barandiaran, X. E. (2009). Adaptive behavior of a motile protocell: Sensor-less sensitivity to emergent viability conditions. In *Advances in Artificial Life, Proceedings of the 10th European Conference on Artificial Life, ECAL*. Berlin: Springer Verlag.

Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A., & Danielson, M. A. (1997). The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annual Review of Cell and Developmental Biology*, *13*, 457–512.

Hanczyc, M., Toyota, T., Ikegami, T., Packard, N., & Sugawara, T. (2007). Chemistry at the oil-water interface: Self-propelled oil droplets. *Journal of the American Chemical Society*, *129*(30), 9386–9391.

Harvey, I., Di Paolo, E., Wood, R., Quinn, M., & Tuci, E. (2005). Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life*, *11*(1–2), 79–98.

Macnab, R. M., & Koshland, D. E. (1972). The gradient-sensing mechanism in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, *69*, 2509–2512.

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition : the realization of the living*. Dordrecht, Holland: D. Reidel Pub. Co.

McMullin, B. (2004). Thirty years of computational autopoiesis: A review. In *Artificial Life 10* (pp. 277–295). Cambridge, MA: MIT Press.

Moreno, A., & Etxeberria, A. (2005). Agency in natural and artificial systems. *Artificial Life*, *11*(1–2), 161–176.

Ruiz-Mirazo, K., & Mavelli, F. (2007). Simulation model for functionalized vesicles: Lipid-peptide integration in minimal protocells. In F. A. e Costa, L. M. Rocha, E. Costa, I. Harvey, & A. Coutinho (Eds.), *ECAL* (Vol. 4648, pp. 32–41). Berlin: Springer.

Ruiz-Mirazo, K., & Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life*, *10*(3), 235–259.

Suzuki, K., & Ikegami, T. (2009). Shapes and self-movement in protocell systems. *Artificial Life*, *15*(1), 59–70.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind* (1st ed.). Belknap Press.

Varela, F. J. (1979). *Principles of biological autonomy*. North-Holland: New York.

Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organisation of living systems, its characterization and a model. In *Biosystems* (Vol. 5, pp. 187–196).

Virgo, N., Egbert, M. D., & Froese, T. (2009). The role of the spatial boundary in autopoiesis. In *Advances in Artificial Life, Proceedings of the 10th European Conference on Artificial Life, ECAL*. Berlin: Springer.

About the Authors

Matthew Egbert is a Ph.D. student, studying theoretical biology at the Centre for Computational Neuroscience and Robotics (CCNR) at the University of Sussex. His research interests include metabolism, the origin of life, self-production, self-organization, and embodied cognition. Matthew completed a B.Sc. in computer science from the University of St. Andrews in 2001 and an M.Sc. with distinction in evolutionary and adaptive systems from the University of Sussex in 2007.

Ezequiel Di Paolo is a reader in evolutionary and adaptive systems at the University of Sussex and co-director of the Evolutionary and Adaptive Systems (EASy) M.Sc. program. He is a member of the Centre for Computational Neuroscience and Robotics (CCNR) and the Centre for Research in Cognitive Science (COGS). He is the author of over 100 peer-reviewed publications and his interests include: adaptive behavior in natural and artificial systems, biological modeling, evolutionary robotics, embodied cognition, philosophy of mind, and philosophy of biology.