University of Sussex

A University of Sussex DPhil thesis

Available online via Sussex Research Online:

http://sro.sussex.ac.uk/

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

US

University of Sussex

Adaptation From Interactions Between Metabolism and Behaviour
Self-Sensitive Behaviour in Protocells

Matthew Deems Egbert

Submitted for the degree of Doctor of Philosophy
University of Sussex
December, 2010

ACKNOWLEDGEMENTS

From the non-academic world, I'd like to first thank my brother Noah for the first piece of inspiration that got me on this track. That simulation of predators and prey you made back in the early '90s stuck in my head for the following 20 years until I decided to go back to University!

Of course, a massive thank you to my parents for all of your support over the past 32 years. You are amazing parents – I couldn't ask for better. Thank you!

Love to my 9/11 Lade Brae's friends – to you all I dedicate this equation:

$$H \cdot \phi = h \sum^{\infty} e$$

Finally, thank you Alison for being my best friend, being supportive, for the great times we've had, great times we have planning for the future, and great times we'll have in the future. We sure know how to have a nice time together. I love you.

Matthew Egbert
Brighton, UK, 2010

PREFACE

All of the work presented in this thesis was undertaken during my doctoral studies. The vast majority of the analysis presented is my own, with the remaining minority the result of collaboration with the colleagues with whom I coauthored the following peer-reviewed publications.

[A] Egbert, M. D. and Di Paolo, E. A. (2009) Integrating behavior and autopoiesis: An exploration in computational chemo-ethology. *Adaptive Behavior* Vol. 17, No. 5: pps. 387-401.

[B] Egbert, M., Di Paolo, E. A. and Barandiaran, X. (2009) Chemo-ethology of an Adaptive Protocell: Sensorless sensitivity to implicit viability conditions in Proceedings of the Tenth European Conference on Artificial Life, ECAL09, Budapest, September 13-16, 2009, Springer Verlag.

[C] Virgo, N., Egbert, M. D., and Froese, T. (2009) The Role of the Spatial Boundary in Autopoiesis in Proceedings of the Tenth European Conference on Artificial Life, ECAL09, Budapest, September 13-16, 2009, Springer Verlag.

[D] Egbert M. D., Barandiaran X. E., and Di Paolo E. A., (2010) A Minimal Model of Metabolism-Based Chemotaxis. *PLoS Comput Biol* 6(12): e1001004. `doi:10.1371/journal.pcbi.1001004`

[E] Egbert, M. D., Barandiaran, X., and E. A. Di Paolo (2010) Behavioral Metabolution: Metabolism based behavior enables new forms of adaptation and evolution in Artificial Life XII: Proceedings of the Twelfth International Conference on the Simulation and Synthesis of Living Systems. Edited by H. Fellermann, M. Dörr, M. Hanczyc, L. Ladegaard Laursen, S. Maurer, D. Merkle, P. Monnard, K. Stoy and S. Rasmussen. pps 213-220.

Chapter 4 is a development of work first appearing in publications A, B, and C. Much of Chapter 5 first originally appeared in publication B. Chapter 6 was published as D, and a first version of the material presented in Chapter 7 appeared as publication E. Chapter 8 is an extension of publication A. I developed all of the computational models using tools developed by the open-source community such as emacs, g++, and octave.

Adaptation From Interactions Between Metabolism and Behaviour

Self-Sensitive Behaviour in Protocells

Matthew Deems Egbert

SUMMARY

This thesis considers the relationship between adaptive behaviour and metabolism, using theoretical arguments supported by computational models to demonstrate mechanisms of adaptation that are uniquely available to systems based upon the metabolic organisation of self-production.

It is argued how, by being sensitive to their metabolic viability, an organism can respond to the quality of its environment with respect to its metabolic well-being. This makes possible simple but powerful 'self-sensitive' adaptive behaviours such as *"If I am healthy now, keep doing the same as I have been doing – otherwise do something else."* This strategy provides several adaptive benefits, including the ability to respond appropriately to phenomena never previously experienced by the organism nor by any of its ancestors; the ability to integrate different environmental influences to produce an appropriate response; and sensitivity to the organism's present context and history of experience. Computational models are used to demonstrate these capabilities, as well as the possibility that self-sensitive adaptive behaviour can facilitate the adaptive evolution of populations of self-sensitive organisms through (i) processes similar to the Baldwin effect, (ii) increasing the likelihood of speciation events, and (iii) automatic behavioural adaptation to changes in the organism itself (such as genetic changes).

In addition to these theoretical contributions, a computational model of self-sensitive behaviour is presented that recreates chemotaxis patterns observed in bacteria such as *Azospirillum brasilense* and *Campylobacter jejuni*. The models also suggest new explanations for previously unexplained asymmetric distributions of bacteria performing aerotaxis.

More broadly, the work advocates further research into the relationship between behaviour and the metabolic organisation of self-production, an organisational property shared by all life. It also acts as an example of how abstract models that target theoretical concepts rather than natural phenomena can play a valuable role in the scientific endeavour.

CONTENTS

LIST OF FIGURES

1

INTRODUCTION

1.1 OVERVIEW

Metabolism, the transformation of the environment into the very same system that is transforming the environment, has been identified by a number of scientists and philosophers as a definitive characteristic of life (Gánti, 1975; Maturana and Varela, 1980; Rosen, 1991; Kauffman, 2000; Ruiz-Mirazo et al., 2004). All life metabolises, but what are the implications of this property? How is it relevant to evolution and the life-time adaptation observed in biological organisms? This thesis presents work that addresses these questions using a new methodology entitled "computational chemo-ethology" in which abstract, artificial, spatially embedded chemistries that support the formation of metabolic, protocell-like systems are simulated in computational models. In these modelled systems, we study metabolism-based mechanisms of behaviour; i.e. metabolism-based mechanisms through which the protocell-like systems modulate their interaction with their environment. Analysis of these models exposes adaptive behavioural capabilities that are uniquely available to systems that create and maintain themselves from their environment. We demonstrate forms of 'metabolism-based behaviour' in which the behaviour of organisms is in response not directly to their environment, but to the state of their metabolic processes of self-construction. This 'self-sensitivity' provides organisms with an indirect sensitivity to their environment and a direct sensitivity to their own metabolic health, enabling simple but powerful behavioural strategies like *'If I am healthy now, continue as before, otherwise do something else'*. We present various, substantial advantages to these kinds of metabolism-based behavioural strategies, including (i) the ability to respond to phenomena never previously experienced by the organism nor by

any of its ancestors, (ii) the ability to adapt to changes in the organisation of the organism itself, (iii) the ability to integrate various environmental influences into an appropriate behavioural response, and (iv) context sensitive adaptive behaviour.

In addition to these mechanisms of adaptation that operate at the level of the individual organism, we demonstrate how metabolism-based behaviour can substantially accelerate adaptation at the level of the population. Specifically, plasticity provided by metabolism-based behaviour can facilitate Darwinian evolution through Baldwin-effect like processes; metabolism-based behaviour can increase the likelihood of diversity-producing speciation events; and adaptation to changes in the organisation of the organism itself can make certain mutations, that would otherwise be neutral or deleterious, beneficial.

Supporting these theoretical contributions, several computational models are presented. One of which is a minimal model of metabolism-based chemotaxis. It recreates chemotaxis patterns observed in bacteria such as *Azospirillum brasilense* and *Campylobacter jejuni*. It also provides an explanation for the previously unexplained asymmetric distribution of bacteria performing aerotaxis.

More broadly, the work advocates further research into the relationship between behaviour and the metabolic organisation of self-production, an organisational property shared by all life. It also acts as an example of how abstract models that target theoretical concepts rather than natural phenomena can play a valuable role in the scientific endeavour.

1.2 ORGANISATION OF THE THESIS

The main line of argument is as follows. First, background material is provided, including a description of context into which this research fits (Chapter 2), methodology (Chapter 3), and theoretical framework (Chapter 4). Then, 'self-sensitive' metabolism-based behaviour is defined and demonstrated in a computational model (Chapters 4 – 5). This leads to a set of demonstrations of benefits of metabolism-based behaviour at the level of the individual (Chapters 5 – 6). These benefits are discussed and the broader concepts supported a model that recreates phenomena observed in bacterial chemotaxis (Chapter 6). We then demonstrate and discuss the facilitation of population-level adaptation (evolution) by self-sensitive behaviour (Chapter 7). Finally, we discuss some of the limitations of metabolism-based behaviour and how it could interface with metabolism-independent mechanisms of behaviour (Chapter 8) before summarising the main conclusions of the thesis (Chapter 9).

1.3 SUMMARY OF PRIMARY CONTRIBUTIONS

1. Most approaches in the study of adaptive behaviour, including the contemporary dynamical systems approach advocated by Beer, "take an agent's autopoiesis for

granted in behavioural and cognitive analyses" (Beer, 2004, p. 320), trivialising metabolic processes or omitting them completely from their analysis. This thesis argues, with support from computational models, that the following three phenomena are lost in this omission:

a) The potential for rich interplay between mechanisms of metabolism (autopoiesis) and behaviour.

b) Specifically, a mechanism through which an organism can evaluate the success of its behaviour or the quality of its environment with respect to its health. This self-sensitivity can enable simple but powerful mechanisms of adaptation (see primary contribution 2, below).

c) More philosophically, it becomes difficult, perhaps impossible to objectively (i. e. in a principled manner) distinguish between agent and environment without referring to 'macro-scale' (population level) or 'micro-scale' (biochemistry) dynamics.

This forms the basis of an argument advocating the inclusion of aspects of metabolic organisation (such as autonomy as defined by operational closure) in the future study of certain aspects of cognition and adaptive behaviour.

2. A new, general mechanism of adaptive behaviour that is based upon the interaction between behaviour and metabolism is proposed (Chapter 4) and demonstrated in computational models (Chapters 5 – 8). This "self-sensitive" or "viability-sensitive" behaviour stems from a measurable notion of viability that emerges when the metabolic organisation of organisms is considered. It is a simple, yet powerful mechanism of adaptation that allows individual organisms to (i) respond to phenomena never previously experienced by the organism nor by any of its ancestors, (ii) adapt to changes in the their own organisation, (iii) integrate various environmental influences into an appropriate behavioural response, and (iv) to adapt and behave in a context and history-sensitive manner.

3. We demonstrate how self-sensitive behaviour can facilitate adaptive evolution. This facilitation could have had a substantial impact on how life has evolved on Earth (Chapter 7). It can occur through processes similar to the Baldwin effect, through increasing the chance of speciation events, or through the ability provided by self-sensitive behaviour whereby an organism can adapt to changes in its own organisation, thereby transforming otherwise neutral or deleterious mutations into beneficial mutations. It is conceivable that this population-level adaptability provided by self-sensitivity may even have played a role in bootstrapping genetic evolution.

4. We present the first minimal model of metabolism based chemotaxis, a type of self-sensitive behaviour (Chapter 6). This model is the first of its kind and it

recreates chemotaxis patterns observed in bacteria such as *Azospirillum brasilense* and *Campylobacter jejuni*. It also proposes a new explanation for previously unexplained asymmetrical distributions of bacteria in aerotaxis.

2

CONTEXT

Whoso would describe and know aught that's alive
Seeks first the spirit forth to drive;
The parts he then hath in his hand,
But lacks, alas! the spirit-band.
Encheiresin naturae *chemists call it now,*
Mock at themselves and know not how.
(Goethe, translated by W. H. van der Smissen,
quoted in Oparin, 1961)

CHAPTER SUMMARY

The research presented in this thesis studies the relationship between metabolism and behaviour. We consider how metabolic processes could influence behaviour and conversely, how behaviour could influence metabolism. The interface between these phenomena is not well studied. Typically, research considering behaviour ignores metabolism and vice versa. This is not surprising as the two phenomena are typically considered in two very different ways. Metabolism is seen as a subset of chemistry in which the objects of study are low-level descriptions of molecules or populations of molecules, enzymes and chemical reactions, etc. In contrast, the study of behaviour tends to involve much "higher level" descriptions of phenomena (such as classical or operant-conditioning, emotional states, neural-dynamics, etc.).

There are however important relationships between behaviour and metabolism. Organisms must behave to acquire metabolic resources, and metabolism must occur to build the body and provide the energy necessary to accomplish behaviour. The two phenomena are interdependent and this interdependence is relatively self-evident, but there also are other, more subtle interactions. The later chapters of this thesis describe some of these interactions and present computational simulations of them. But, for the full implications of the studies to be communicated, it is important to first provide the context into which the research fits. That is the goal of this chapter.

We start by reviewing recent approaches taken in the study of behaviour. This initial review focuses upon the relatively contemporary movements of cybernetics and embodiment, as these developments set the stage for the subsequent reviews of research pertaining to self-maintaining systems and the research presented in later chapters. We highlight a pervading problem in the recent history of cognitive science: the absence of theory that aids in the identification of *the subject*. Whether one embraces the computationalist or the dynamical, embodied approach to cognition, in both frameworks, there is an absence of theory concerning how to identify an agent that is behaving with goals, limits, desires, emotions, and / or intentions *of its own*. In Section 2.2, we consider the argument that "ultra-radical embodiment" can help us address this void.

Ultra-radical embodiment refers to the idea that by including metabolism (actually, the closely related concept of self-production) in the description of an agent, it becomes possible to identify an individual organism and behaviour that is intrinsic to that individual. We review research that has focused upon these self-maintaining systems which, as mentioned above, has predominantly involved the analysis of low-level systems of chemicals and has not, for the most part, included higher-level phenomena such as behaviour. Finally, we review the predominantly philosophical work that connects 'life' and 'mind', i.e. the biological organisation of self-maintenance with the autonomy of living systems. This last review forms the basis of our working definition of behaviour and the metabolic organisation which are presented in Section 2.4.

2.1 APPROACHES TO BEHAVIOUR

What is behaviour? What is different from an animal foraging for food and a rock rolling down a hill? These questions have been approached in a variety of different ways throughout history. Vitalism, popular until the end of the 19th century, proposes that in parallel to the laws of physics that govern everyday physical matter, there exist a special and distinct set of "vital forces" that govern living systems. This is related to the idea of Cartesian dualism, that the mind is of some alternative, non-physical form of matter. In contrast to vitalism, the mechanistic view sees organisms and their minds as subject to the same laws of nature as non-living material, and the invocation of special non-physical vital forces as both unnecessary and unscientific. With the discovery of Darwinian evolution, and the abiotic synthesis of organic compounds, the apparent need

for a special vitalistic force disappeared, and the reductionistic, mechanistic approach gained favour setting the scene for the establishment of modern biology (Mayr, 1982).

This 'modern biology' has been highly successful; but many fundamental open questions remain. While it is generally accepted that the behaviour of life is driven by the same laws of chemistry and physics, we still have little idea about how organisms manage to behave in such an amazingly adaptive manner, how they generate their own goals and adapt them to novel situations. We struggle to build robots that can walk up stairs, while in nature, organisms navigate complex environments, stalk prey, build nests, write poetry, etc. Organisms appear to act with intent, to have their own desires, motivations, and goals. Something is different between the behaviour of human artifacts such as computer programs and robots. It is not some mystical soul or non-physical vitalistic "energy", so what is it? One possible answer is based upon the idea of *circular organisation of influence* in which a system influences itself. The recent history of the study of this idea begins with the first cybernetics movement of the mid-20th century.

2.1.1 *Cybernetics, feedback, and teleology*

Cybernetics can be conceived of as the study of systems with circular causality. Perhaps the simplest instance of circular causality is a basic feedback loop in which the action (or output) of a mechanism influences the mechanism itself. As an example, consider the *negative feedback* mechanism of a centrifugal governor, depicted in Figure 2.1. When the shaft (D) spins more quickly, the centrifugal force drives the weights (E) outwards, causing the fuel valve (A) to close, thereby reducing the power of the engine spinning the shaft. This results in the maintenance of a near-constant speed even under a variety of different operating conditions. This is called *negative* feedback, as the mechanism inhibits itself rather than exciting itself.

In the 1940s, Wiener and Bigelow employed negative feedback in their attempts to automate control of anti-aircraft fire (Bechtel, 2007, p. 271). Struck by similarities between their automated control systems and biological systems, Wiener and Bigelow (with physiologist Rosenblueth) "argued that negative feedback provided a means of resuscitating notions such as purpose and teleology, enabling these concepts to be applied to both biological and engineered systems without invoking vitalism" (Bechtel, 2007, p. 272). Their idea was simple: if a system accomplishes an action through an error-correcting negative feedback mechanism, then this system can be considered purposeful and the purpose of the system is whatever it is that is maintained by the negative feedback (e. g. the maintenance of the aircraft in the sight of the anti-aircraft gun). "Teleological behavior thus becomes synonymous with behavior controlled by negative feed-back." (Rosenblueth et al., 1943, p. 23).

Rosenblueth, Wiener and Bigelow (1943) present this negative-feedback explanation of teleology as part of a hierarchy of types of behaviour, distinguishing between active, passive, purposeful, teleological, non-teleological, and systems with different orders
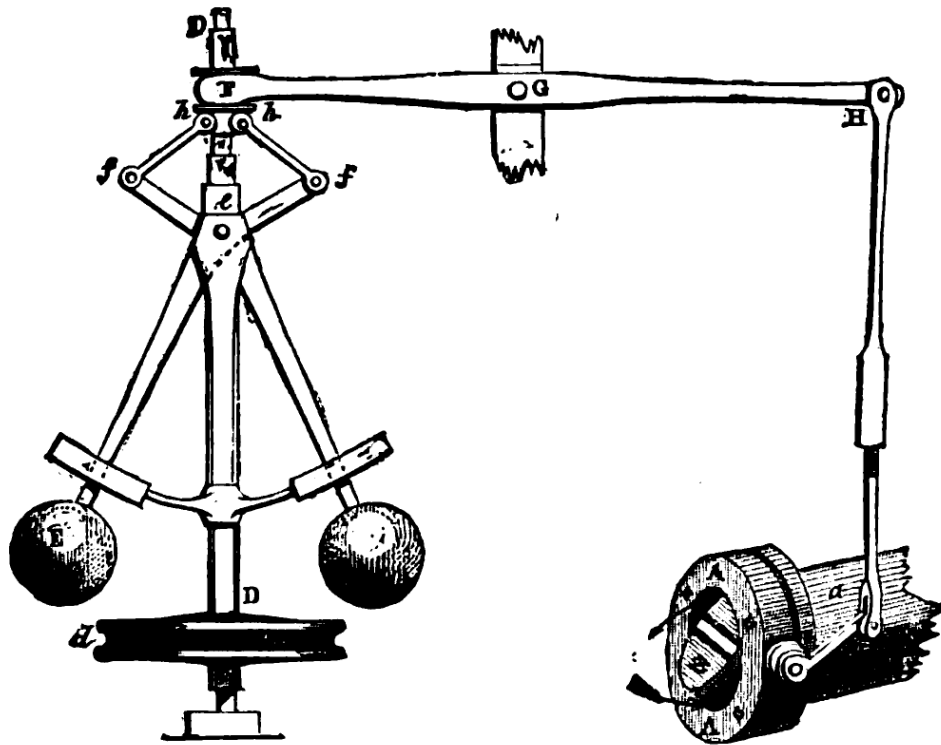
FIG. 4.—*Governor and Throttle-Valve.*

Figure 2.1: The centrifugal governor, a mechanical mechanism for negative feedback. When the shaft (D) spins more quickly, the fuel valve is closed to reduce the speed of the engine spinning the shaft.

of predictiveness. Active behaviour is seen as behaviour for which "the source of the output energy involved in a given specific reaction [...] the input does not energize the output directly" (Rosenblueth et al., 1943, p. 18) and in contrast, passive behaviour is that in which "all the energy in the output can be traced to the immediate input [...] or else the object may control energy which remains external to it throughout the reaction (e.g., the soaring flight of a bird)." (ibid). They also described purposeful behaviour "may be interpreted as directed to the attainment of a goal" (ibid) — a poorly defined concept in their paper as criticised by Taylor (1950) and they equated teleology with negative feedback, distinguishing between non-predictive and predictive (or extrapolative) systems, including the idea of higher-order prediction which involves the integrative prediction of multiple processes.

The drawing of attention to feedback and circular causality by Rosenblueth et al. was undoubtedly a great step forward for engineering and the scientific study of behaviour, leading to subsequent developments of embodiment, sensorimotor-loops, and the second wave of cybernetics (all discussed below), but arguably Rosenblueth et al. were over-enthusiastic in equating negative feedback with purposeful behaviour. Soon after Rosenblueth et al.'s influential paper, Taylor argued that their definitions were too inclusive; that they "(a) utilize criteria which render purposeful behavior a ubiquitous

phenomenon, and (b) thereby endow the word "purpose" with a meaning having no similarity to any meaning it has customarily been taken to possess." (Taylor, 1950, p. 317). Taylor goes on to criticise the poor specification of the concept of goal, pointing out that coincidental situations such as a rock falling off a cliff and hitting someone on the head could be interpreted under Rosenblueth et al.'s definition as purposeful.

There is also the question of the origin of the behaviour. Rosenblueth et al. describe behaviour as being "*directed* to the attainment of a goal" (Rosenblueth et al., 1943, p. 12, emphasis added), but as observed by Taylor, there is a difference between systems that *direct themselves* and those that are "directed by some purposeful being *other* than the behaving object itself" (Taylor, 1950, p. 312). This concept of autonomy, a system somehow directing itself is quite relevant to the work presented in this thesis, and we shall return to discuss it further below.

Is negative feedback a sufficient criterion to identify teleological behaviour? Undoubtedly, the provision of a description of teleological behaviour that does not require a mystical *vital* component was a major leap forward. But, negative feedback cannot be the defining requirement for teleological behaviour as it is too encompassing. The "behaviour" of (for instance) an automated anti-aircraft device is, at the very least, of a different *kind* than that of a predator chasing prey or a bird building a nest. The behaviour of the organisms is *intrinsic*; it is driven by the needs of the organism itself, not by its creator or some other external motivation or force. This is unlike the anti-aircraft device that serves the needs of its designers and users. To understand how these ideas were later extended and adapted to better capture the idea of behaviour, it is necessary to understand the context in which the advances arose: the computationalist boom of the second half of the 20th century.

2.1.2 *Computationalism*

With the meteoric advance in computing from the 1940s to present, cybernetics and notions of feedback, remained relevant to control theory and engineering applications, but fell out of the focus of scientific study of cognition and behaviour. In its place grew the new field of artificial intelligence (AI) with its strongly computationalist view. In this approach, cognition is seen as computation: the manipulation of abstract symbols that represent phenomena. Advances in this area continue to the modern day. As examples, consider the development of increasingly proficient artificial intelligences capable of a variety of tasks such as playing chess (Deep Blue), or mining data (the Google search engine).

Despite this progress, it is important to realise the short-comings of functionalist computationalist approaches to teleology, behaviour, adaptivity and cognition. Fundamental conceptual issues associated with this symbol-manipulation approach, pointed out early in its history (Dreyfus, 1965, 1979; Searle, 1980), remain problematic to this day. While artificial intelligences have managed to succeed in certain tasks such as chess and data

mining, they have failed completely at others such as natural language comprehension, demonstrating creativity, generating emotional responses, etc. Similar to to the absence of intrinsic motivation in the cybernetics grounding of teleology in negative feedback, in AI there is little theory that addresses how one could identify an intentional agent, or a subject that is thinking, experiencing, or feeling anything.

Various efforts have been made to overcome these limitations of Symbolic-AI, including connectionism, the focus on cognition as the manipulation of *sub*-symbols through the development of complex neural networks, a dynamical-systems study of cognition, and the embodiment and situatedness movements. The connectionist movement has little to do with the work presented in this thesis, so we do not address it further here (if interested, see Feldman and Ballard, 1982). The dynamical-systems approach and the embodiment and situatedness movements, however, are more closely related and comparable with the research described in this thesis, so we shall now discuss each of these areas.

2.1.3 *Embodiment*

Instead of focusing on disembodied, abstract programs that operate on abstract strings of symbols, Brooks (1991), Varela (1997), Pfeifer and Bongard (2006) and others have drawn attention towards the *embodiment* and *situatedness* of natural agents. In this new framework, artificial intelligences were not unmoving computers, but were given bodies that operated in a physical, spatial world. The inclusion of the body in the study of cognition opened up new approaches, that improved the abilities of robots to solve certain tasks such as navigation or obstacle avoidance. "The world is its own best model" argued Brooks (1991, p. 15), meaning that some of the effort needed to accomplish cognitive tasks can be off-loaded into the environment.

Embodiment makes it easier to solve some tasks, but is it fundamental to cognition? This remains a topic of debate. Clark differentiates between 'simple embodiment' where the interactions between the neural system, body and physical world act only as "*constraints upon a theory of inner organization and processing*" (Clark, 1999, p. 348. emphasis his) and 'radical embodiment' in which these same phenomena are seen as "*profoundly altering the subject matter and theoretical framework of cognitive science*" (ibid). Below, we elaborate upon the idea of ultra-radical embodiment which adheres to the second, more radical interpretation of embodiment, but also focuses on different aspects of 'the body' than are typical of the embodiment paradigm: the metabolic organisation.

Ziemke observes that "many discussions/notions of embodied cognition actually pay relatively little attention to the nature and the role of the body involved" (Ziemke, 2003, p. 1305). Typically, embodiment in AI involves the use of robotic 'bodies' such as the Khepera robot (Figure 2.2) or simulations of similar systems. The benefits afforded by this form of embodiment are primarily due to the spatial 'situatedness' of the robot, i. e. its coupling to a physical environment. It is, of course, important to recognise this role of
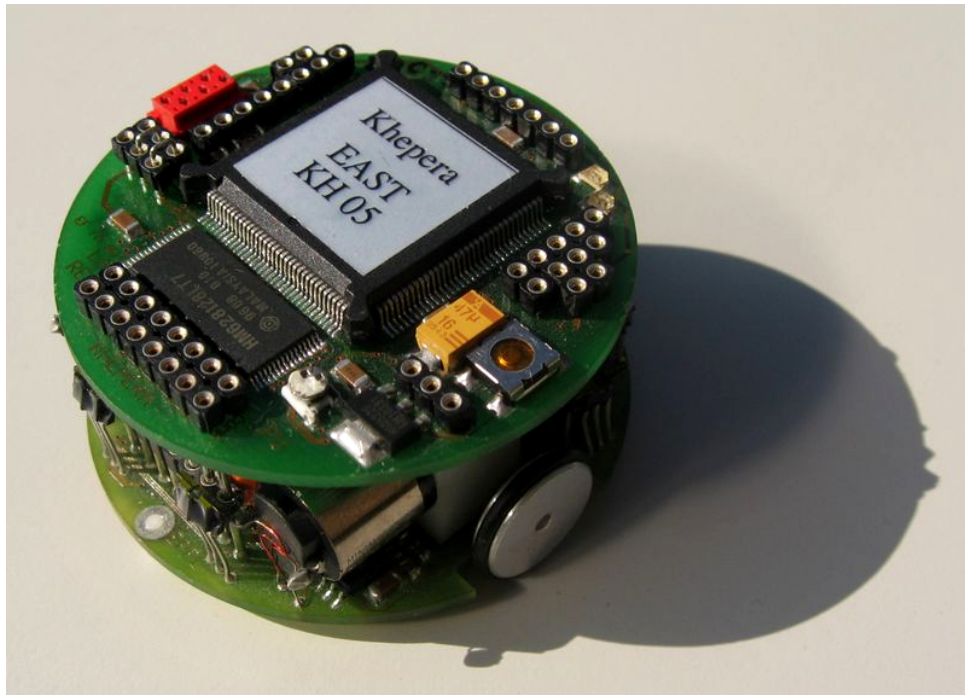
Figure 2.2: A 'Khepera' mobile robot. This "body" shares little in common with living bodies.

the body, but what are the other aspects of embodiment that are important in the context of artificial intelligence? One such aspect is 'organismic embodiment' (Ziemke, 2003), the idea that "there are crucial differences between living organisms, which are autonomous and autopoietic, and man-made machines, which are heteronomous and allopoietic" (Ziemke, 2003). Autopoiesis (Varela et al., 1974; Maturana and Varela, 1980), coined by Maturana and Varela in the 1970s means self-making, and it describes a property of living systems that is not shared with allopoietic ("made by other") robots. That property is the "metabolic organisation" of perpetual degradation countered by self-(re)-construction in which living systems depend upon their own influence for their continued existence. "[B]eing embedded does not necessarily necessitate being (truly) embodied [...] a body is not a mere lump of matter, but the physical aspect of a living system, created and maintained as a functional unity by an autonomous metabolism" (Boden, 1999, p. 239).

In this vein Di Paolo (2003) argues (and we agree) that "something is still missing" from the embodied AI of Pfeifer, Brooks and others. An "organism is unlike any machine in that the very nature of the flux [of matter and energy through them] is used to fabricate and maintain its own components. [...] If metabolism stops, the organism ceases to be" (Di Paolo, 2010, p. 137). And, therefore, the "biologically inspired" robots of Pfeifer, Brooks, Beer, Harvey and others are not "organismically inspired" in that they do not have the metabolic organisation. "[R]obot design may be getting important inspiration from the properties of biological neuronal mechanisms and from the dynamics of animal bodies, but it is getting little or no inspiration from the fact that the organisms that serve

as models are *living systems*" (Di Paolo, 2010, p. 132). Following this idea of 'organismic embodiment', this thesis considers 'ultra-radically embodied' systems, i. e. systems that are organismically embodied in that they have the metabolic-organisation, and for which their behaviour is fundamentally and inextricably linked to this organisation.

2.1.4 *The dynamical approach*

Beer (1995, 1997, 2003), studies 'minimal cognition' and adaptive behaviour by analysing artificially evolved abstract dynamical systems coupled to an environment. In this dynamical-system-based approach (and in the spirit of the embodiment movement), the body, the mind and the environment become, in a sense, equals – they are all described by a set of differential equations that are roughly of the same form and importance in determining how or why a certain behaviour is performed. This is part of the contemporary movement away from thinking of cognition as computation, in that the it focuses upon *temporal* coupling between agent and environment, and includes continuous variables rather than discrete "atemporal"[1] interactions associated with symbolic manipulation of computation.

This framework has has proven helpful in elucidating how minimal forms of cognitive behaviour can be accomplished without computation in abstract simulated conditions (e. g., Beer, 2003; Williams et al., 2008) and in models more closely connected to biological target organisms (Izquierdo and Lockery, 2010). It has also provided an alternative language for describing cognitive phenomena that is not based in the computational metaphor. However, the problem remains that there is no way of distinguishing between agent and environment, or between agent and non-agent. Typically, in this framework, one set of differential equations is arbitrarily chosen to represent the motions or actions of the agent, and another set is chosen to describe the environment in which the agent exists. The only difference between the agent and the environment is an arbitrary description made by the researcher. The behaviour of the predefined agent is in no objective sense *by* or *for* "the agent". In fact, from the ultra-radical embodiment perspective taken in this thesis, there is no objectively identifiable agent in these systems! Even if we were to accept the predefinition of the agent in these models, the behaviour of these agents is not grounded in (i. e., it does not stem from) intrinsic needs, desires, or goals of the agents. Instead, it is the result of the desires of the researcher to study a relatively simple system that accomplishes the target behaviour. Once again, the subject (in this case the cognitive subject) is, in a sense, missing from the modelled system.

We argue in this thesis in support of the idea that the metabolic-organisation of perpetual self-degradation countered by processes of self-maintenance can provide the

---

1 The dynamics of computers are, of course, temporal, but in a very constrained and peculiar way. A clock synchronises all of the interacting components in such a way to limit computers to perform precise, specific kinds of "computational" tasks. Dynamical analysis, such as that undertaken by Beer explores systems that do not have this constraint, and is thus an exploration of a wider variety of temporal dynamics.

basis for a definition of an agent that allows for (i) the distinction of agent from non-agent, and agent from environment, and (ii) the formulation of behaviour that is intrinsic to the agents. Beer is aware of this work and the philosophy that inspired it (reviewed below) and of the absence of such metabolic-organisations within his models. He was, in fact, inspired to his current framework in part by the theory of *autopoiesis* (Beer, 2004, p. 322), Maturana and Varela's theory concerning self-maintaining systems that we describe in detail below. He recognises that his dynamical, embodied approach does not address autopoiesis and that his models do not include autopoietic organisations, and he rightly observes that this shortcoming of DS analysis is not necessarily problematic, depending upon the questions that are being asked, and which aspects of the model and of cognition or adaptive behaviour are being evaluated.

> "Since it is meaningful to study an agent's behaviour only so long as that agent actually exists, we largely take an agent's autopoiesis for granted in behavioural and cognitive analyses." (Beer, 2004, p. 320)

Inherent in this statement is the assumption that the autopoiesis of an agent can be ignored when studying its behaviour – that autopoiesis is only serving to produce the body that behaves, and that the dynamics of self-production do not have a ongoing effect upon the agent's behaviour. That assumption may be appropriate in many circumstances. There may be many situations in which autopoiesis / metabolism has no effect upon behaviour. But, in this thesis we consider what is lost when that assumption is made. What is lost when autopoiesis is assumed out of a model, simulation, or conception of living systems?

At least three things are potentially lost. First, by ignoring the process that behaviour subserves (self-production), evaluation of success of behaviour, or the quality of an environment with regards to that self-production is made more difficult or impossible (discussed in Chapter 5). Second, the potential for a rich interplay between processes of self-production and behaviour is lost. We demonstrate in Chapters 5 – 7 that these interactions can produce simple yet highly adaptive mechanisms that adapt behaviour to suit the organism. Third, when behaviour is modelled without metabolism, it becomes difficult, perhaps impossible to objectively (i.e. in a principled manner) distinguish between agent and environment; the only difference between the agent and the environment is an arbitrary description made by the researcher. This ability to distinguish between agent and non-agent allows us to also distinguish between behaviour (of e.g. someone walking down a hill), and the passive suffering of events (e.g. a rock rolling down a hill). We review some research that addresses this aspect in later sections of this chapter, but do not address this issue directly in our computational models. Thus, this thesis is an exploration in which we include autopoiesis in the study of behaviour so that we may study the phenomena that are lost in the usual behavioural science frameworks that omit the metabolic-organisation.

In this section, we elaborate upon the idea of the metabolic organisation as a self-maintaining system. First, in Section 2.2.1, we introduce the central role of metabolism in biology, describing the ongoing debate as to whether metabolic systems or template replication systems came first in the origin of life. We highlight this debate as it is relevant to our discussion in Chapter 7 of how metabolism-based behavior can facilitate adaptive evolution. In Section 2.2.2 we review more abstract and theoretical research that considers abstract self-maintaining systems such as models of protocells and artificial chemistries. This leads to our brief review and discussion in Section 2.2.3 of research that considers metabolic organisations to be the minimal form of life.

2.2.1   *Metabolism*

Metabolism is the transformation of energy and material in the environment of the organism into the organism itself. Metabolism can take a variety of forms, involving different enzymes, substrates and products, but the underlying *metabolic organisation* – the organisation of a system that is inherently unstable, but counters its degradation through processes of self-production or self-repair – is universal to all known life (Varela et al., 1974; Gánti, 2003b; Morowitz and Smith, 2007). In a recursive cycle, energy is used by organisms to transform materials into the structures of the organism that harness the energy to perform further material transformations (Gánti, 2003b; Kauffman, 2000; Morowitz and Smith, 2007; Rosen, 1991).

A "primary metabolism" is shared by all autotrophs (organisms capable of producing all of their constituents from abiotic resources) and consists of reactions that "transduce energy; synthesize necessary monomers, cofactors, and other small molecules; and assemble the molecules of cellular functions; proteins, nucleic acids, carbohydrates, and lipids." (Morowitz, 1999, p. 46). This primary metabolism "cuts across all taxa and is clearly close to or at the evolutionary origin." (Morowitz, 1999, p. 46). In this primary metabolism, universal metabolic pathways have been identified – pathways shared by all known life with very little variation. These include glycolysis (Romano and Conway, 1996) and the reductive citric acid cycle (Morowitz, 1999). The latter or something very similar, has been proposed as part of the earliest biological chemistry as it only involves 11 molecule types, is autocatalytic, and it "lies at the core of biosynthesis and contains the precursors of all the major biosynthetic pathways" (Morowitz, 1999, p. 47).

A great deal of research has involved the analysis of various, specific metabolic pathways. This research has proven useful in developing antibiotics, understanding and fighting disease, etc. But the methods employed in such research and the results gathered are of a different category of research than that presented in this thesis, so we shall not discuss the analysis of specific metabolic pathways further.

Metabolism and the origin of life

The universality of metabolic pathways seems to reach so far back into the history of life that there is speculation as to which came first, metabolism or replication. The metabolism-first theory (Oparin, 1961) proposes that before template replicator molecules like DNA or RNA, natural selection operated over metabolic systems with limited heredity. These might have been 'composomes', self-assembling groups of chemicals (Segre et al., 2001), or compartmentalised autocatalytic sets (Fernando and Rowe, 2008). The primary difficulty associated with metabolism-first theory lies in the limited heredity of such organisations (Maynard Smith and Szathmáry, 1998). Unlike metabolisms, template replicator systems are capable of inheriting so many different states as to be considered unlimited in their heredity potential. For instance, a template with 4 nucleic bases and a length of only 100 bases, could take $\approx 10^{60}$ different forms. This high number of stable, heritable states facilitates evolution, but is unusual for chemicals or chemical networks and autocatalytic sets do not in general have anywhere near as large a number of stable states. In certain conditions, they have been shown to increase in complexity when under a selective pressure to maximise biomass (Fernando and Rowe, 2008), but the widely accepted view is that the small number of heritable states of metabolic networks suggest that the natural selection of metabolic networks undergoing random change is extremely unlikely, on its own, to have produced the template replication systems shared by all modern organisms.

Given the apparent limited heredity of metabolisms, it has been argued that the template replicators must have come first. Unfortunately this explanation of the history of life also seems to be unsatisfactory. Template replicators have been generated abiotically in laboratories, as of yet, there is no known mechanism for their synthesis in prebiotic conditions (Shapiro, 2006) although work continues in this area (see e. g., Ciciriello et al., 2008 and other work by Di Mauro's group). The conditions necessary for the production of template like monomers seem constraining to the point that even some proponents of replicator-first theory are starting to concede that the prerequisite "relatively pure, complex organic molecules might have been made available in large amounts via a self-organizing, autocatalytic cycles" (Fernando and Rowe, 2008). Alternatives to the metabolism-first or replicator first theories have also been proposed. For example, Cairns-Smith (1990) suggests that clay mineral crystals might have played an intermediate role of the template replicator, eventually replaced by nucleic acid template macromolecules.

Part of the appeal of the replicator-first theory was the power of neo-Darwinian evolution (i. e. evolution of template molecules such as RNA). Proponents imagined that once up and running replicator-theory suffices to explain the emergence of the complexity and richness of life found on earth. However, Kacian et al. (1972) demonstrated *in vitro* that template replication is insufficient to drive evolution towards anything other than the smallest template capable of replication (the shorter the molecule, the faster it can replicate). Moreno and Ruiz-Mirazo (2009) point out that, as Speigelman's experiments

show, template-replication on its own is insufficient for the evolution of complex life. That for natural selection to produce complexity, there must first exist an organisation of the evolving systems that makes possible functionally diverse mutations. So, even if the abiotic synthesis hurdles are someday overcome, it cannot be presumed that neo-Darwinian evolution on its own suffices to produce life as we know it.

In Chapter 7, we discuss how the behaviour of organisms can be in response to "metabolic well-being" (i. e. how well its processes of self-maintenance are faring against processes of degradation). We describe how this 'self-sensitive behaviour' can facilitate the evolution of metabolisms through a variety of mechanisms, and how it could have played a role helping to bootstrap the early evolution of life.

2.2.2 *A mathematical framework for the study of self-maintenance*

We now turn our attention to the study of mathematical or computational abstractions of metabolisms. The metabolisms of modern organisms are incredibly complex. To avoid this complexity, a number of researchers have decided to focus on mathematical abstractions of metabolisms. There have been many attempts to create such abstractions, but there remains no established mathematical framework for the study of self-maintaining systems (Fontana and Buss, 1996, p. 3). The traditional scientific tool of dynamical systems analysis is not well equipped for dealing with constructive (Fontana and Buss, 1996, p. 3) or metadynamical processes (Beer, 2004, p. 319) (i. e., systems in which the dynamics of the system influence how the rules that describe how the system changes in time).

Fontana and Buss (1996, p. 5) explain that dynamical system descriptions "never deal with the objects [of study] themselves [...] objects disappear into arrays of structureless variables confined to holding numerical values that quantify properties of an object class, such as the frequency of a gene, the concentration of a chemical, the density of an electromagnetic field, the position and velocity of an aircraft, the pressure of a gas, the earnings of a firm. The moon, for example, is never represented as an object in the equations that express its orbit; the 'moon' is defined as a time-dependent vector of numbers specifying position and momentum." They go on to point out that although dynamical analysis has been very successful in physics, its achievement in biological, cognitive and social sciences has been "mixed at best" with its failures being "nearly universally attributed to some combination of high dimensionality and nonlinearity [...] referred to as 'complexity'" (Fontana and Buss, 1996, p. 4), concluding that blaming the lack of progress in these areas on the fact that they are too complex is perhaps too simple, and that really what is needed is a new mathematical framework for analysing such systems.

Fontana and Buss are not alone in arguing for the development of a new framework for the study of self-maintaining systems. Kampis (1991) advocates the study of 'component systems', systems in which "there is a set of simple units from which higher structures of increasing intricacy can be built, by recursively putting larger and larger units together"

(Kampis, 1991, p. 200). In component systems (Kampis, 1991), the components, their associated dynamic variables, and the processes that determine the systems dynamics all can change as a result of the state of the system.

Also emphasising the limitations of orthodox dynamical analysis when applied to living systems, Rosen (1991) argues that the Newtonian approach to physics does not apply to biological systems! The former, he says, can be simplified to basic units of particles with location, velocity, acceleration and external force, whereas the the latter is *fundamentally* recursively self-influential. The arguments of Rosen and Kampis are opaque to say the least; it is not obvious how to apply their conjectures to real or simulated organisms.

So, while the proposals of Rosen and perhaps Kampis are extreme, the criticism of dynamical systems remains valid. Dynamical analysis has limitations and there is undoubtedly room for explorations of new frameworks better suited to the analysis of systems in which new objects are created and the properties of objects and the way they interact can evolve dynamically. To make this point more concrete, let us consider an example of such a limitation of dynamical systems. Typical dynamical models of chemistry require pre-specification of the relevant chemicals, allowing only for the analysis of the change in concentration of these predetermined chemicals; "classical systems theory does not, conveniently, take novelty into account." (Dittrich and di Fenizio, 2007, p. 1200). It is therefore impossible in such scenarios to study the likelihood of, for instance, the likelihood of the emergence of autocatalytic networks, or the chance of a particular chemistry producing new reactants capable of profoundly changing the existing dynamics of the system, (such as the introduction of a new catalyst, or the production of amphiphilic molecules that can self-organise into vesicles, thereby separating reactants and constraining their concentration). "Conventional dynamical systems [...] are well-suited to treat changes in the magnitudes of quantitative properties of fixed object species, but ill-suited to address interactions that change the objects themselves" (Fontana and Buss, 1996, p. 6).

To address this shortcoming of the dynamical framework, Fontana and Buss (1996) studied and modified λ-calculus in an effort to capture the dynamics of combinatoric construction of new entities that influence how the system changes over time.

The details of λ-calculus are outside the scope of this chapter, but the main idea is that instead of keeping mathematical functions and the target of those functions in separate conceptual classes, in λ-calculus, the target of a function can also be a function and vice versa. In the mathematics we learn at school, you can not ask "what is multiplication times multiplication?", but λ-calculus is a framework in which precisely these kinds of operations-applied-to-operations can be studied. Operations are applied to operations to produce new operations. It is not difficult to find analogies between λ-calculus and chemistry where molecules interact with molecules to produce new molecules, and this is, more or less, what Fontana and Buss did with λ-calculus.

Rather than pre-specifying the possible molecules (as is necessary in the typical dynamical models of chemistry), an initial set of λ-expressions (think molecules) are specified and some or all of their interactions are simulated. Which λ-expressions are 'present' determines what λ-applications (think reactions) are possible and therefore which λ-expressions will exist in the future. Thus, in λ-calculus, there is a cycle of influence between (i) the state of the system and (ii) how the state changes. This is an example of a constructive dynamical system, in which the object of study is motion through "object space", a potentially infinite-dimensional space of the possible objects, rather than the typically finite and predefined "phase space" of system states. Fontana and Buss extend λ-calculus to include other properties of chemistry. For example they introduce the possibility of some functions being present (available to be applied to other functions), while other functions are not. They make further extensions to include more complex notions related to chemistry such as thermodynamics and conservation of mass and energy.

Self-maintaining organisations in artificial chemistries

There has been a variety of work studying such self-maintaining systems in "artificial chemistries" (Dittrich et al., 2001). Fontana and Buss identify self-maintaining 'organisations' within their modified λ-calculus (Fontana and Buss, 1996, p. 16) commenting that an "infinity of such organizations are possible. Developing a taxonomy of their structure and properties remains a long-term goal of our program." (Fontana and Buss, 1996, p. 17). Similarly, Dittrich and di Fenizio providing formal descriptions of a variety of categories of self-maintenance. These include 'closed-sets' in which "all molecules that can be generated by reactions inside a set are already inside the set" (Dittrich and di Fenizio, 2007, p. 1203); 'semi-self-maintaining sets' in which "every molecule that is consumed within a set, is produced within that set" (Dittrich and di Fenizio, 2007, p. 1204); 'semi-organisations' (a combination of semi-self-maintenance and closure), and 'organisations' (a semi-organisation but capable of growth) (Dittrich and di Fenizio, 2007, p. 1199). Kauffman (1993) studied the likelihood of the emergence of self-maintaining autocatalytic sets, and Piantadosi and Crutchfield (2010), studied the influence of the spatial embedding of chemicals, showing that some autocatalytic networks can exist in spatially distributed systems that could not exist in well-mixed reactors.

Artificial chemistries make it possible to observe and study the dynamics in which self-maintaining systems form, persist and / or 'die' (cease to exist). The parts that form the self-maintaining individual and the parts that form the environment of the individual are of the same category – they are both chemicals. Yet in some scenarios, a formally specifiable individual emerges, such as the 'organisations' discussed by Fontana, Buss, and Dittrich, or the autocatalytic networks discussed by Kauffman and others. This 'individual' does not immediately appeal to the everyday, intuitive conception of an individual. Rather than being identified by its spatial limits, the "individuals" that exist in these systems are defined instead by the interdependence between their components.

These individuals can only continue to exist because of their own influence on themselves and on their environment. As we shall see in Chapter 4, this non-spatial, relational definition of an individual is actually a stronger defining concept for the biological individual; more robust to criticism than the widespread view of the individual as a spatial structure.

In the beginning of this section, we described some limitations of orthodox mathematical frameworks for the study of self-maintaining systems. The alternative frameworks explored by Fontana, Buss, Dittrich and others have helped to establish the theory of self-maintaining organisations that underlies the research presented later in this thesis and is elaborated further below.

It should be pointed out that most of the work presented herein does not focus on the same low-level self-constructing dynamics studied by these authors. Instead, we employ a higher level of abstraction in which we predefine a self-maintaining system, its constituent parts and processes, and study what kind of behaviour such a system can accomplish that a non-self-maintaining system could not. This abstraction allows us to study certain properties of self-maintaining systems using a combination of artificial chemistry and dynamical systems analysis. The criticisms raised above apply primarily to the study of *the emergence of* self-maintaining organizations. They do not as big a problem in the study of the dynamics of an established self-maintaining chemical organization, as the object of study is not newly emerging processes or objects and is therefore more amenable to typical dynamical analysis.

2.2.3   *A minimal form of life*

Metabolism is common to all life and thus has been considered to be the minimal definitive property of living systems. In an effort to describe the basic unit of life, Varela and Maturana introduced the word 'autopoiesis' to connote the perpetual processes undergone by living systems through which they produce and maintain themselves (Varela et al., 1974; Maturana and Varela, 1980, 1992). They argue that unlike "allopoietic" systems (those that are produced by things other than themselves), living systems are autopoietic; i.e. they are precarious networks of self-producing interdependent processes.

> "An autopoietic system is organized (defined as a unity) as a network of processes of production (synthesis and destruction) of components such that these components:
>
> (i) continuously regenerate and realize the network that produces them, and
>
> (ii) constitute the system as a distinguishable unity in the domain in which they exist." (Varela, 1997)

Similar to the research presented above, Varela and Maturana were focusing on the recursive nature of life. They viewed a living system as a network of interdependent processes,

defined as a relational organisation (i. e., a set of relationships between processes) rather than a structure of physical objects (Varela, 1979).

Unfortunately, Varela and Maturana's publications on autopoiesis (Varela, 1979; Maturana and Varela, 1980), are not always clear or consistent (Luisi, 2003), often suffering from "obscure and idiosyncratic language" (Beer, 2004, p. 310), and thus autopoiesis has been interpreted in a variety of ways. Some of these interpretations have been less than scientific, giving autopoiesis an unfortunate "new-agey" reputation (Luisi, 2003, p. 50). But recently, with the rise of synthetic biology, and with subsequent publications that have refined and clarified concepts introduced by Varela and Maturana (many of these are reviewed below) the idea of autopoiesis as the minimal organisation of life has gained traction in the biology and philosophy of biology communities. Biochemists have created autopoietic systems in real chemistry (Luisi and Varela, 1989) to study minimal forms of life, the challenges that such systems would have faced, and how they could have evolved (see e. g., Luisi et al., 2006; Rasmussen, 2009).

Also trying to capture the most basic unit of life, Gánti (1975) developed the concept of a 'chemoton': an integrated set of three self-producing subsystems: the cytoplasm (metabolic), template (genetic), and membrane. Each of these three subsystems is autocatalytic (a form of self-production), with a stoichiometric coupling between the systems such that the chemoton as a whole is a single self-producing system. Gánti proposed the unification of these three systems, "a programmed, self-reproductive, fluid chemical machinery capable of spatial division and thus, of real proliferation" (Gánti, 2003a) that "grows by metabolism, reproduces by fission [...] and [...] has rudimentary hereditary variation" (Maynard Smith and Szathmáry, 1998, p. 20) as the minimal unit of evolution and of life (Gánti, 1975).

The chemoton focuses on a self-maintaining minimal unit of life and therefore is closely related to autopoiesis. However, chemoton theory is intended to form the basis of an enquiry into the origin of life, and its subsequent evolution. In contrast, autopoiesis, does not focus on evolution or even reproduction, seeing these as "not intrinsic to the minimal logic of the living [but rather] an added complexification" (Varela, 1997, p. 76). Rather than studying evolutionary dynamics, Varela, Maturana and others studying autopoiesis have focused on how the minimal organisation of life can be extended to "to construct a bridge between biology and cognitive sciences" (Luisi, 2003, p. 58) —to see how 'life' relates to 'mind'; how the autopoietic organisation can form the basis for a scientific understanding of teleology, adaptive behaviour and cognition. This is related to a primary theme in this thesis: the relationship between the metabolic (autopoietic) organisation and behaviour. What can autopoietic systems do that non-autopoietic systems can not?

In a methodology that is also similar to the techniques used in this thesis, Varela and Maturana developed a computational model of autopoiesis that they used to study the implications of the autopoietic organisation (Varela et al., 1974). This model (depicted in Figure 2.3) consists of abstract particles undergoing stochastic motion on a two-dimensional lattice. Particles that are next to each other can react according to rules that
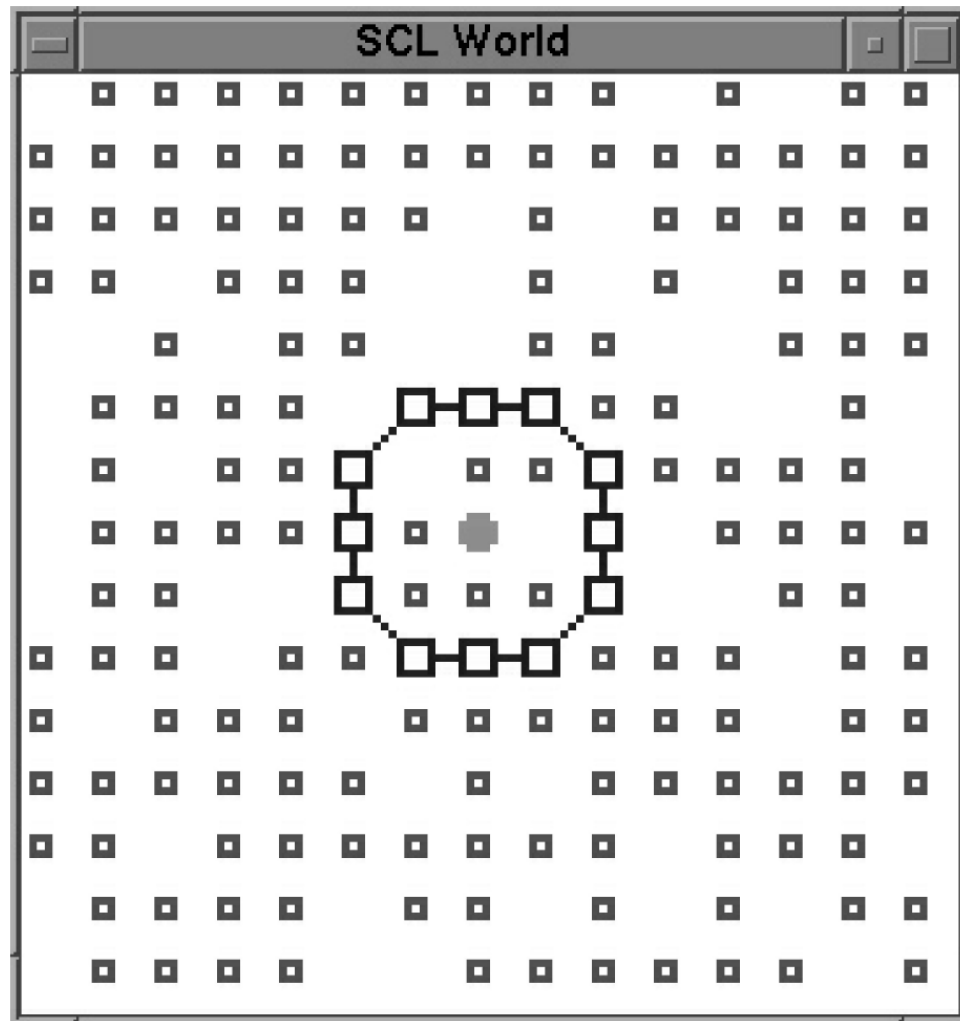
Taken from (McMullin, 2004).

Figure 2.3: Varela, Maturana and Uribe's 1974 computational model of autopoiesis.

were hand-designed by the researchers to demonstrate how an autopoietic individual can emerge from some basic interactants and how it can maintain itself dynamically despite being intrinsically unstable. The reactions are simple: a catalyst particle (the circle in the centre) changes substrate particles (the small grey squares) into link particles (the larger darker squares). The link particles can bind with each other (lines connecting the link particles) and form a membrane that surrounds the catalyst. The membrane does not allow link particles to pass through it, so it maintains a high concentration of free link particles on the inside of the membrane that are ready to repair the holes that spontaneously form when the membrane link particles degrade into substrate.

The model shows how an autopoietic individual (a network of interdependent self-maintaining processes) can emerge from a set of rules and simple reactants. The interdependence is apparent in a cycle of dependence in which the membrane depends upon its own containment of link particles. In the absence of the membrane, the link particles would diffuse away, making the formation of a membrane unlikely. But with

the membrane in place, the reactants are maintained at a high concentration, allowing for the ongoing production and repair of the membrane. The maintenance of the membrane is made possible by the presence of the membrane and so we have a system that depends on itself.

It is worth pointing out that the catalyst particle is not produced by any other part of the system, so it is arguably not part of the network of interdependent processes. It should also be mentioned that the original publication describing the system failed to include the rule of "chain-based bond inhibition" (McMullin, 2004, p. 285) in which the presence of bonded link particles inhibits the linking of two free link particles. This rule is important as it prevents the "clogging-up" of the system with short chains of link particles inside the membrane that prevent the possibility of single-link particles repairing holes in the membrane.

There have been a myriad of extensions of this model with a variety of goals including, for instance, the addition of a third spatial dimension (Madina et al., 2003), facilitating growth (McMullin, 2004), movement (Suzuki and Ikegami, 2009), or self-reproduction by division (Ono and Ikegami, 2000). These and other extensions are reviewed by McMullin (2004).

2.3 LIFE AND MIND

In the previous section, we mentioned that a main contribution of autopoiesis is a possible "bridge between biology and cognitive sciences" (Luisi, 2003, p. 58). In this section we discuss this bridge between metabolism and behaviour, reviewing recent relevant work in this area.

How can the metabolic organisation be relevant to behaviour? Barandiaran and Moreno (2008) summarise nicely this relationship. They argue that the status of the autopoietic system as a precarious and holistically-interdependent system (all of the processes depend upon one another) makes it possible to objectively describe how the system can do things to avoid being destroyed and can, in this sense, act on its own behalf. This provides the foundation for an intrinsic normativity, i. e. a way of describing what is good or bad *for the autopoietic system* instead of for an outside observer, and this leads to stronger concepts of purposeful or teleological behaviour, in which purposeful behaviour is behaviour that the system engages in to maintain itself.

> "Because of [the] circularity in autonomous systems, identity and normative functionality are not observer dependent but intrinsically causal: the network as a whole (the very system) will not exist in the absence or malfunctioning of the component processes (given its FFE [(far-from-equilibrium)] nature and the circular dependency between processes). In other words, in autonomous systems what-the-system-does (the way it functions) and what-the-system-is (its structure) are highly intertwined and they merge together in its organiza-

> tion. [...] the goal state of the system and the organization that instantiates it are one and the same thing. Autonomous systems have an implicit teleology as their internal causal circularity makes each process of the system a contribution to its global self-maintenance." (Barandiaran and Moreno, 2008, pps. 329–330)

This connection between the metabolic organisation and behaviour has been elaborated upon and clarified by many. The primary contributions are reviewed in the remainder of this Section.

2.3.1 Basic autonomy

Ruiz-Mirazo and Moreno (2004) endorse the main message of Varela and Maturana, but object to the "highly abstract" formulation of autopoiesis, and argue that the main concepts must be re-conceptualised in more concrete and specific terms if they are to advance the development of artificial autonomous systems and aid in our understanding of the origin of life (Ruiz-Mirazo and Moreno, 2004, p. 237). They develop the concept of 'basic autonomy', an interpretation of autopoiesis that is more directly related to the biochemistry and thermodynamics of life, and one that they believe will contribute more directly to our understanding of the origins of life and in the creation of artificial life (Ruiz-Mirazo and Moreno, 2004, p. 249).

Where Maturana and Varela were satisfied with the more abstract requirement of operational closure (a more formal way of describing the idea of interdependent processes that we discuss in detail in Chapter 4), Ruiz-Mirazo and Moreno and Barandiaran and Moreno (2008) advocate a more concrete formulation related to the thermodynamics of metabolism, and more generally, far-from-equilibrium (FFE) systems. Organisms are "thermodynamically "hungry," in need of coupling with the environment, which is no longer a mere source of uncomfortable perturbations to be compensated but the source of an essential flow" (Barandiaran and Moreno, 2008, p. 328). In other words, autonomous systems are seen as a subset of far-from-equilibrium systems, that use and need energy and material resources in their environment to maintain their existence. But not just any FFE system is autonomous. "[S]pontaneous FFE" such as candles and whirl-pools are not "real autonomous systems" as, unlike autonomous systems, they do not control the flows of energy and matter that keep the system away from equilibrium (Ruiz-Mirazo and Moreno, 2004, p. 238). For a system to be autonomous, it must "manage the flow of matter and energy through it so that it can, at the same time, regulate, modify, and control: (i) internal self-constructive processes and (ii) processes of exchange with the environment. Thus, the system must be able to generate and regenerate all the constraints—including part of its boundary conditions—that define it as such, together with its own particular way of interacting with the environment." (Ruiz-Mirazo and Moreno, 2004, p. 240)

According to this view, the synthesised protocells (see e. g., Luisi et al. (2006); Rasmussen (2009)) only capture a weak form of autopoiesis (they have a network of interdependent processes involving a membrane constraining the reactions that produce the membrane), but do not count as autonomous systems because these systems have "no control over the material and energetic flow through [them]. The cellular systems that appear in this scenario can change their shape, grow, even reproduce autocatalytically, but do so without any control [and] they have no mechanism to modify or adjust critical boundary conditions" (Ruiz-Mirazo and Moreno, 2004, p. 248).

This view of life as a *particular kind* of dissipative, FFE systems that regulates its interaction with its environment is echoed by Kauffman (2000) who proposes a similar idea in his conception of an autonomous system as something that uses work (in the physics sense of the word) to establish constraints that allow more work to be accomplished. It is also echoed by Bourgine and Stewart (2004, pps. 337–338) who, in the same year as Ruiz-Mirazo and Moreno wrote,

> "purely physical (or physicochemical) dissipative structures (such as the flame of a candle, or a cyclone) are intrinsically ephemeral; they last only so long as certain external conditions, over which they exert no control at all, are maintained. By contrast, there is something special about living organisms considered as dissipative structures: [...] their functioning extends to include a regulation of their own boundary conditions [...] this is a key feature that distinguishes autopoietic systems from dynamical systems in general.

> Cognition, in this perspective, focuses on exactly this aspect: management of the interactions between an organism and its environment." (Bourgine and Stewart, 2004, pps. 337–338)

Bourgine and Stewart claim that "A system that is both autopoietic and cognitive system is a living system." (Bourgine and Stewart, 2004, p. 342). But, in contrast with Varela and Maturana's idea that autopoiesis is a sufficient criterion to call a system living, Ruiz-Mirazo and Moreno's basic-autonomy they propose as "necessary but not sufficient for life" (Ruiz-Mirazo and Moreno, 2004, p. 235). Basic autonomy does not imply an "efficient capacity for open-ended evolution and long-term sustainability" which they argue are required for a system to be called living (Ruiz-Mirazo and Moreno, 2004, p. 251).

This focus upon thermodynamics of life, and the metabolic organisation in the study of adaptive behaviour and cognition may, to those unfamiliar with this line of research, seem odd. What is to be gained by such an approach? There are three primary gains. First, unlike the approaches to adaptive behaviour and teleology and artificial intelligence discussed earlier in this chapter, autopoiesis (or basic autonomy) provides a way to identify a subject. Second, focusing on the biological organisation of autopoiesis makes it possible to provide a scientific conceptualisation of normativity (the creation of value). Third, and this is the primary contribution of this thesis, biological systems can be

sensitive to their own autopoietic dynamics, responding to changes in their own viability rather than to changes in their environment, leading to a variety of forms of adaptation, both at the level of the individual-lifetime adaptation (see especially Chapters 5, 6) and evolutionary adaptation (Chapter 7). We shall now review some of the work that establishes these first two contributions of autopoiesis-related theory and philosophy, as they help to set the scene for the third, which is elaborated upon in later chapters.

2.3.2 *A foundation for identity*

Barandiaran and Moreno (2008) discuss a theme that has appeared repeatedly in this chapter, what they call the "problem of identity": how to distinguish between the subject that is behaving and its environment. A solution to this problem would be a formal definition for a 'subject', or 'individual' or 'agent' such that we would be able, "given a natural (physicochemical or biological) system or a mathematical/simulation model (where variables and parameters do not bear any information about what they are meant to represent) [...] to distinguish between the system and the environment, to define the norms that the agent must satisfy and to determine whether the system is operating as an agent or not." (Barandiaran et al., 2009, p. 381). Or, put more colloquially, "If we are to attribute adaptive capacities to a system, we must first specify *who* is going to be adjusted to what." (Barandiaran and Moreno, 2008, p. 325, emphasis mine). Early in this chapter we discussed how such a definition of a subject is missing from most discussions in cognitive science. We have also described how the idea of autopoiesis allows us to identify an individual as a network of interdependent processes. But, for some, this definition is unsatisfactory, or in need of further clarification. In particular, there has been considerable debate over whether or not a living system must have a physical boundary (such as a membrane) and whether or not that spatial boundary demarcates the limits of the organism.

A spatial or relational boundary?

The intuitive, yet ultimately unhelpful way of identifying an agent has been to refer to its spatial boundary. The membrane plays a prominent role in Varela et al.'s 1974 model. Must a system include this kind of intrinsically-generated *spatial* boundary for it to be an example of autopoiesis, and therefore be discernible as an agent or subject as described above? And, do the limits of the organism coincide with its spatial limits? There has been substantial discussion concerning the role of the physical, spatial boundary ("the membrane") in autopoiesis and we shall return to discuss it in Chapter 4. We find, along with various authors (such as Thompson, 2010; Bourgine and Stewart, 2004; Barandiaran et al., 2009) that a productive way of understanding the role of the membrane in autopoiesis is to see autopoiesis as an instance of a more general and, arguably, more important concept: *autonomy*. Here, as we now explain, the word 'autonomy' is being

used here to connote something more general than Ruiz-Mirazo and Moreno's basic-autonomy.

We follow Varela (1979) and Thompson (2010, p. 44) in their distinction between autonomy (a network of interdependent processes) and autopoiesis (a biological instance of autonomy that involves a spatial membrane). Bourgine and Stewart (2004, p. 337) also make this distinction, but instead of using a different word (as, for instance, Thompson uses autonomy), they provide "a renewed definition of autopoiesis that does not depend on an excessively reified definition of "membrane" or "boundary." [...] a modified, more general definition of autopoiesis: [...] An autopoietic system is a network of processes that produces the components that reproduce the network, and that also regulates the boundary conditions necessary for its ongoing existence as a network. The difference from the preceding definitions [...] is that the emphasis on the semipermeable boundary, which lends itself to over-reification of a membrane, is replaced by a more dynamical emphasis on the regulation of boundary conditions." (Bourgine and Stewart, 2004, p. 337). In the remainder of this thesis, we take the approach of Thompson, Barandiaran and others of using the word 'autopoiesis' to include the idea of the formation of a boundary and 'autonomy' to capture the more encompassing notion of a network of interdependent processes.

As we shall discuss in Chapter 4, the focus of autonomy is not on the spatial structure but instead upon networks of interdependent processes. This means that a variety of systems could be considered autonomous. Examples might include the immune system, social institutions such as religions and nation-states, multi-cellular organisms, "chemical reactions, molecular structures, physiological structures (such as tissues or organs), neurodynamic patterns at the large scale, sensorimotor loops, social habits," (Barandiaran et al., 2009, p. 378) etc. Interestingly, this implies that "agency does not have to be subordinated to biological/metabolic organization but can appear at different scales responding to a variety of autonomous processes." (Barandiaran et al., 2009, p. 378). In other words, it remains theoretically possible (although it has yet to have been accomplished) for a *non*-metabolic robot to be an autonomous agent, acting on its own behalf. We discuss this more below in Section 2.3.5.

Although Ruiz-Mirazo endorses the interdependent network of processes concept of autonomy, perhaps in part because of his interest in the origin of life and synthetic biology, he argues that the spatial membrane is effectively a requirement for basic autonomy. He writes, "The generation of a physical border is a crucial step toward autonomy, because that is the only way, on the one hand, (i) to assure the control of energy flow required for the robust maintenance of the network, and on the other, (ii) to solve the problem of diffusion and dilution (control of concentrations). Furthermore, if the precursor reaction network does not become self-enclosed, [it] [...] will not have any control over the boundary conditions that bring about its distinctive, far-from-equilibrium dynamics (and, thus, it will be extremely fragile)." (Ruiz-Mirazo and Moreno, 2004, p. 243)

In line with this view, Ruiz-Mirazo focuses upon the homeostatic regulation of membrane pores as a minimal form of autonomy, using abstract but realistic simulations of chemical systems developed to explore how such systems could accomplish self-regulation e. g., (Ruiz-Mirazo and Mavelli, 2007, 2008). The concept of self-regulation is visited repeatedly in this thesis, so it has much in common with the work of Ruiz-Mirazo. However, rather than focusing on the regulation of membrane characteristics, we focus on motile behaviour, in which the autonomous system can regulate its boundary conditions by using movement to select its environment. We chose to work with motile behaviour, because we believe there is potential for a richer form of autonomous regulation. With regulation of boundary conditions by modulation of membrane properties an organism can (speaking anthropomorphically) only say 'more please' or 'no thank you' to environmental features. Slightly more powerfully, regulating boundary conditions through motility, lets an organism say 'yes please' or 'give me something else'. As we shall see in Chapters 5 and 6, given a variety of environments, a motility based regulation of boundary conditions can select from these environments, allowing arguably for a richer form of basic self-regulation than that provided through the regulation of membrane properties such as permeability. That being said, we believe that many of the concepts and conclusions could be carried over from Ruiz-Mirazo's work to ours and vice-versa.

Barandiaran et al. (2009) and Thompson (personal communication) call attention to the excessive focus upon the biological form of autonomy (autopoiesis), calling for increased attention on other forms of autonomy.

> "Certainly, the grounding of individuality and normativity conditions in biological organization and minimal models of metabolism has attracted most attention. But it has also distracted attention away from an almost unexplored avenue of research: the possibility for the emergence of a new level of autonomy in the domain of behavior and neurodynamics. The adaptive regulation of behavior needs not be exclusively subordinated to the viability constraints imposed from biological "survival conditions." Instead, it can be equally governed by the need to maintain neurodynamic and behavioral organization in terms of self-maintenance of habits, coherence of behavior, psycho-dynamic stability, and so forth " (Barandiaran et al., 2009, p. 382)

Although we agree with Thompson and Barandiaran et al., the research presented in this thesis focuses exclusively upon the "biological organisation and minimal models of metabolism". We take this focus because it is the most easily understood and visualised form of autonomy that we are aware of. This is an important feature given that this field is young and we are exploring new areas about which we know little. It is intended that advances in understanding this biological autonomy will facilitate later advances in our understanding of other forms of autonomy.

2.3.3 *A foundation for normativity and teleology*

We have just discussed how the metabolic organisation of self-production can form the basis for an identity, allowing us to distinguish between a system and its environment. In this section, we shall discuss the second advantage of this autonomy-based perspective: how the metabolic organisation allows us to formulate a grounded notion of normativity, i. e. of what is objectively 'good' and 'bad' *for an organism* rather than as ascribed by an outside observer.

The most common way to evaluate a behaviour in modern biology is by referring to its evolutionary implications. A good or beneficial behaviour is seen as one that increases the chances of passing on genes to the next generation (Dawkins, 1976). In contrast to that, we (in agreement with Barandiaran et al. (2009) and Thompson (2010)) support the idea of grounding the purpose of a behaviour in the *current organisation* of the organism. After all, we can still distinguish successful or beneficial behaviour in organisms that are incapable of reproduction (e. g. mules).

As Thompson puts it, "living structures have to be comprehended in relation to norms [...] While sucrose is a real and present condition of the physicochemical environment, the status of sucrose as a nutrient is not. Being a nutrient is not intrinsic to the physicochemical structure of the sucrose molecule; it is a relational feature, linked to the bacterium's metabolism." (Thompson, 2010, p. 74). Or, to quote Barandiaran et al., "[I]t is insufficient that we, as observers, make judgements on behalf of the agent about the "adequacy" of its behavior in relation to some of our own norms, standards, or goals (epistemic, artistic, ethical, functional, or otherwise). If we are to adopt a naturalistic approach we must be able to justify this normativity based on the very "nature" of the agent." (Barandiaran et al., 2009, pps. 372–373).

This idea can be made more concrete by referring to the 'viability' of autonomous systems. We elaborate on the idea of viability in Chapters 4 and 5, but it suffices here to describe it as the likelihood of a self-maintaining system continuing to exist. A viable system is one that is 'healthy' in the sense that its self-maintenance is (for the moment) sufficing to counteract its perpetual degradation. It is also possible to discuss trends in viability (Di Paolo, 2005; Egbert and Barandiaran, 2011). The system is either moving toward a situation where it will likely survive, or toward a situation where the interdependent network of processes is likely to cease to exist. Out of this idea of viability trends emerges a normative dimension for the behaviour of autonomous systems. "[I]nteractions that contribute to this natural purpose [of self-maintenance] are seen as intrinsically *good* according to this self-generated norm. And those aspects that challenge this end, are intrinsically *bad*. Thus, a purely dynamical principle of self-continuation engenders an intrinsic relation with the world in terms of values and norms" (Di Paolo, 2010, p. 140). In other words, 'good' or 'beneficial' behaviour *for the organism*, is behaviour that causes positive trends in its viability (i. e. behaviour that increases its chances of continuing to exist).

Unfortunately, like much of the terminology in this area, the word 'viability' has been used to mean different things by different authors. Bourgine and Stewart's use the concept of 'viability constraint' in a way that is "deliberately rather vague". Their viability appears initially to be similar to that described above, in that it applies to the idea of the viability of an autopoietic system. However Bourgine and Stewart wanted the notion of viability to be applicable to situations involving higher-forms of cognition, or situations in which there is no autopoiesis such as a "robot that navigates on the surface of a table may be required neither to simply remain immobile, nor to fall off the edge of the table and crash." (Bourgine and Stewart, 2004, pps. 338–339). This appears to abandon the notion of 'viability' grounded in self-maintenance and therefore also the grounding of normativity in the idea of autonomy, and so we do not agree with their arbitrary or vague definition of viability. There are other ways to extend theory surrounding autonomy to include higher forms of cognition (discussed below in Section 2.3.5), that we believe are more fruitful.

Returning to the issue of teleology, the question of why an organism performs a specific behaviour rather than another can now be addressed by referring to the normative dimension of the behaviour. A behaviour is not teleological or purposeful just because it involves negative feedback as discussed by Rosenblueth et al. (1943), but instead, a behaviour is purposeful / teleological when it satisfies a norm of the behaving system.

Teleological behaviour is thus *not arbitrary*. It is 'normative' – it satisfies the needs of the behaving system, making it possible for it to survive. As Di Paolo puts it, "A real animal [...] can be trained to do lots of things, but never to treat a punishment as a reward. The link between the external situation, the internal dynamics and the overall affective state may be contingent (different species will look on the same piece of fungus with desire or repugnance), but never arbitrary as it is defined by the continuous existence and renovation of the autopoietic organisation and subordinated to it and to the network of interacting tensions and satisfactions it creates in a behaving animal." (Di Paolo, 2010, p. 144)

2.3.4 *Autopoiesis does not imply adaptivity*

As mentioned above, Ruiz-Mirazo and Moreno (2004) observe that spontaneous dissipative systems that do not self-regulate do not qualify as autonomous. They propose that autonomous systems are those that, in addition to being far from equilibrium, regulate their interaction with the environment in such a way as to increase the situations in which they can survive. This self-regulation can be, but is not always, *adaptive*. Self-regulation is non-adaptive when it is "hard-wired", i.e. it is a fixed mechanism that can not change in the presence of a new environment or changes in the needs of the agent. In contrast, an adaptive behavioural mechanism would be able to maintain its beneficial normative dimension even if aspects of the environment or the the needs of the organism were to change radically. To borrow (and slightly modify) an example from Ashby (1952), a
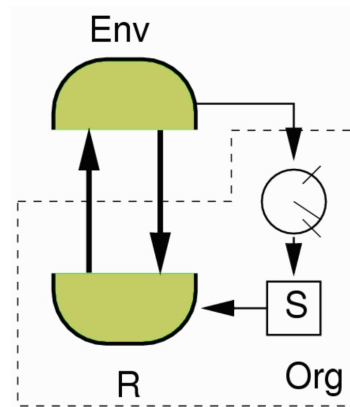
Taken from (Di Paolo, 2003).

Figure 2.4: Di Paolo's Ashbian view of life. Di Paolo argues that adaptivity requires a double-feedback structure, in which the sensory-motor loop is regulated by a second feedback system, that reorganises the sensor-motor loop when essential variables go out of their viability boundaries.

typical autopilot is incapable of adapting to an inversion in the control system such that what formerly rolled the plane to the right would roll it to the left and vice versa. A person, however, given enough practice would be able to adapt to such a situation, and be able to keep the plane level. To accomplish this adaptivity, the system needs to be able to be sensitive to how well it is doing. It needs to be able to recognise that although it is acting, the way that it is acting is not improving the situation.

One of the offshoots of the cybernetics and later embodiment movements was the recognition of 'sensorimotor' loops as useful behavioural mechanisms. The computationalist approach had divided behavioural mechanisms into linear sequences of events; i. e. first sense the environment, then process it to determine an action, then act. The sensorimotor approach drew attention to the rapid and ongoing interaction between sensing and action. It changed the linear sense, process, act structure into the structure of a kind of feedback loop. This recognition that sensory input changes and is changed by motor output in an immediate and ongoing manner led to advances in behavioural science and robotics. But, as Di Paolo points out, "a closed sensorimotor loop is *not* enough for adaptation, but that at least a double feedback structure is needed." (Di Paolo, 2010, p. 142). This double feedback mechanism is depicted in Figure 2.4. Here we see a sensorimotor feedback loop, but part of the loop is under the control of a second feedback loop that evaluates the success of the current behaviour, and reorganises the sensorimotor loop if the current behaviour is unsatisfactory. This architecture is very similar to that used by Ross Ashby in his 'ultrastable' homeostat (Ashby, 1952), and has been employed by Di Paolo in developing a robot capable to adapting to inversions of its visual field (Di Paolo, 2000).

Di Paolo points out that there is no such adaptivity implicit in the definition of autopoietic or autonomous systems. "If Varela and others have managed to see in autopoiesis a natural link with intrinsic teleology and sense-making, it is because they

have complemented the theory with an additional assumption that has guided its interpretation." (Di Paolo, 2005, p. 430). Systems such as the original model of autopoiesis by Varela et al. are autopoietic but show no such adaptivity. It is a defining characteristic of basic autonomies that they regulate their boundary conditions, but for them to do so *adaptively* would require supplementary organisation.

Di Paolo's conception of adaptivity involves two systems of modulation and Barandiaran and Moreno (2008, p. 340) argues that adaptivity "requires a regulatory subsystem within an autonomous system, partially decoupled from the dynamical organization of the constitutive processes of the system.". However, the metabolism-based mechanism discussed in this thesis is highly integrated with the processes of self-construction, and this conception of two systems of modulation, where one modulates the other does not seem to us to be strictly necessary for adaptivity. Instead, as we show in Chapters 5 – 7 it suffices for a system to be directly sensitive to its viability.

It is worth mentioning that implicit in the idea of adaptivity is the "target" of adaptation: i. e. the change(s) to which the system is capable (or incapable) of adapting. In the example above, the change was the inversion of the controls and the autopilot was incapable of adapting while (at least some) people could adapt. Nothing is universally adaptive, for any system there are changes that it can adapt to and others for which it cannot. There are however certain types of adaptivity that are more general than others. In this thesis we discuss metabolism-based behaviour, a very generalistic form of adaptive behaviour that allows organisms to respond to a wide variety of changes, in the environment (Chapters 5 – 7), in the needs of the organism (Chapter 5), and even to changes in the organisation of the organism itself (Chapter 7).

Autopoiesis does not imply adaptivity, but it does provide the basis for a scientific conceptualisation of normativity, teleology and adaptivity. In addition, as we discuss in this thesis, it provides the foundation for some powerful *mechanisms* of adaptivity, allowing organisms to adapt to a variety of environments by being sensitive to changes in their viability, rather than responding directly to environmental phenomena. The adaptive benefits of this "self-sensitivity" mechanism are discussed throughout, but in particular in Chapters 5, 6 and 7.

2.3.5 *Habits and 'higher' cognition*

A common criticism of the philosophy discussed in this section is that the grounding of identity, normativity, adaptivity, etc., is in concepts that are occurring at the chemical level, yet cognitive science is primarily concerned with brains and the behaviour of multicellular organisms that is largely influenced by their nervous systems (NS). It has also been argued that the grounding of normativity and teleology in viability is that people (and other animals) sometimes intentionally act in a way that is detrimental viability (e. g. by smoking or participating in high-risk sports). How can the ideas discussed in

the past few sections connect to 'higher level' cognition and to behaviour that appears to run against the subject's viability?

The dynamics that operate over the NS appear to be highly decoupled from the dynamics of metabolism (Barandiaran and Moreno, 2008, p. 336) and perhaps survival is the "mother value" from which others indirectly emerge (Di Paolo, 2003), but it seems wrong to think that all behaviour is grounded in metabolic-dynamics of self-maintenance. How then, can these formulations of identity, normativity, adaptivity etc., be relevant to the NS?

A nascent area of research is the conceptual extension of biological autonomy (i.e. autopoiesis and the metabolic organisation) into a different domain. Is it possible to have an autonomous system in a domain that is not chemical or spatial in the same way that a cell is autonomous? Can we conceive of autonomous systems in the domain of the nervous system? What would such autonomous behavioural patterns be? What form would they take and can they help us understand the impressive adaptivity and apparent autonomy observed in biology? If they are analogues to the autonomous systems discussed above, then these 'mental-life forms' (Barandiaran, 2008; Barandiaran and Moreno, 2006) or 'habits' (Di Paolo, 2003; Barandiaran et al., 2009) would have to be precarious, self-maintaining organisations. What would their other properties be, and could they help us to explain some properties of the higher cognition we see in nature?

If autonomies can exist in the domain of the mind (involving interaction between brain + body + world), this leads to the interesting idea that organisms with 'higher cognition' are instances of multiple autonomies (biological and mental) operating over the same physical material, as neurons have the metabolic organisation but also form the basis of neural dynamics. "[A]gency does not have to be subordinated to biological/metabolic organization but can appear at different scales responding to a variety of autonomous processes. The possibility is also open for different forms of autonomous organizations to overlap in their material substrates." (Barandiaran et al., 2009, p. 378) These concurrent, superimposed autonomies would be linked in their physical substrate, but interestingly they could also be somewhat independent of each other. As Di Paolo (2003) describes, "habits may indeed die out without implying the death of the system. They may drive the system to situations that are contrary to its own survival or well-being (think for instance of addictive or obsessive behaviour)."

Interestingly, if this view is correct, it might be possible to observe or build artificial autonomous agents that do not have 'biological autonomy'; i.e. that do not have the metabolic organisation. "Minimal life forms already come to satisfy the necessary and sufficient conditions for agency. This does not imply, however, that living organization is necessary for agency, nor that all forms of agency need to trace their normative or individuality conditions back to living organization" (Barandiaran et al., 2009, p. 376). The bodies of these artificial autonomous systems would not be autonomous, but they would provide a domain in which the autonomous 'habits' would exist and in so doing be an autonomous system. In trying to build an autonomous robot, "We may invest

our robots not with *life*, but with the mechanisms for acquiring *a way of life*, that is, with habits. This may be enough for them to generate a natural intentionality, not based now on metabolism, but on the conservation of 'one' way of life as opposed to 'another one'." (Di Paolo, 2003)

Beer observes that "nervous systems significantly enrich the cognitive domains of the animals that possess them. By increasing the internal state that can be maintained and thus the structural changes that can be tolerated, nervous systems expand enormously the range of interactions that an organism can engage in without loss of organization." (Beer, 2004, p. 319). It might be the case that multiple autonomies could 'cooperate' or interact in such a way that increases the situations that they can tolerate. In Chapter 8 we use a computational model to make an early exploration of how superimposed autonomies might cooperate or conflict and how they can benefit from each other.

Understanding other forms of autonomy is a "big-picture" background motivation underlying the work presented in later chapters, but the relation of biological autonomy to 'higher cognition' is not a primary focus of this thesis. We do discuss some of the advantages of decoupling behavioural mechanisms from metabolic dynamics in Chapter 8, but we focus primarily on 'biological autonomy', i.e. the metabolic organisation and how this level of autonomy can produce interesting forms of adaptivity.

2.4 BEHAVIOUR BASED IN METABOLISM

This chapter began by reviewing some recent approaches to the study of the behaviour of organisms. We pointed out problems shared by the early cybernetics and computationalist explanations of behaviour. In particular, we observed how in these orthodox approaches and in the more recent embodied approach, there is no established method for objectively identifying the subject, its goals and its norms. We then introduced ultra-radical embodiment, in which the metabolic or autopoietic organisation of living systems provides a conceptual foundation from which we can develop scientific conceptions of identity, normativity, and adaptivity.

In this thesis, we present computational models of ultra-radically embodied agents that we use to study the relationship between metabolism and behaviour. We consider how metabolism can guide behaviour, how it is possible to model this relationship, and how by doing so, we can develop a better understanding of natural adaptivity. Developing an understanding of these issues will help in the development of theory concerning the foundations of biology, adaptive behaviour and ultimately, biologically grounded theories of cognition. This in turn will aid in the engineering of new kinds of robots, synthetic organisms, and other artifacts. But, before we can begin presenting the research, we must be clear and provide definitions of three key concepts: the metabolic organisation, behaviour, and adaptive behaviour.

metabolic organisation – A network of interdependent processes, each of which depends upon all of the others and each of which enables at least one other. This network is inherently precarious (i. e. unstable and perpetually degrading) but can, to a varying degree, counteract this degradation through processes of self-maintenance, resulting in a dynamical stability.

behaviour – the modulation by a metabolic organisation of its interaction with its environment.

A good example of behaviour is motility; the self-induced motion of an autopoietic organism that changes its environment, thereby changing the interaction between the organism and its environment. A second example is the regulation of membrane permeability by an organism as studied by Ruiz-Mirazo and Mavelli (2007, 2008). For reasons discussed in Section 2.3.2, we focus on motility in the research presented in this thesis.

Behaviour has a normative dimension in that it can have a positive, negative or neutral effect upon the likelihood of the ongoing existence of the metabolic organisation. This normativity allows us to define adaptivity without referring to its evolutionary history.

adaptive behaviour – behaviour that maintains its positive normativity (i. e. its maintenance of the metabolic organisation) despite changes in the environment or in the needs or organisation of the agent.

All of the models presented in later chapters are models of metabolic organisations. Although a spatial boundary (membrane) is not required in the definition of a metabolic organisation, our models do often include such a membrane, as they are a straightforward way of capturing in a model the idea of interdependent processes (a set of reactions that maintain a membrane that maintains the reactants at high enough concentrations to be able to continue to react). As Di Paolo (2005) discusses, the metabolic organisation does not automatically imply adaptive behaviour, but in this thesis we explore how metabolic organisations can be extended to include forms of metabolism-based, 'self-sensitive' mechanisms that can drive an adaptive form of motile behaviour.

Previous work connecting metabolism and behaviour

As we discuss in Section , there are many models of metabolism and many models of adaptive behaviour, but few that target both phenomenon in a rich way that allows for the study of their interaction. There is, therefore, not a great deal of work that is very similar to that presented herein. Perhaps the research that is the most closely related to ours is the work by Suzuki and Ikegami (2004) who extend the original model of autopoiesis (Varela et al., 1974) to include, in an integrated manner, a mechanism of behaviour. By simulating an autopoietic system similar to Varela's original with some minor modifications to the rules, Suzuki and Ikegami show that on a gradient of substrate particles (S), the autopoietic agent will grow more rapidly on the high-concentration

(up-gradient) side than it does on the down-gradient side. This asymmetric growth results in the eventual motion of the centre of the agent up the gradient. This is a first example of a positive coupling between metabolism (autopoiesis) and behaviour. Where the autopoiesis is operating more efficiently, it grows more, and therefore moves in that direction. This direct correlation between the behaviour and what is "good for the metabolism" inspired the models presented in this thesis.

2.5 RECAPITULATION

In this chapter we reviewed contemporary approaches to the study of behaviour. In the process we highlighted a problem shared by these approaches; none have a way of distinguishing between the subject and its environment. We then introduced a possible resolution to this problem: ultra-radical embodiment, an approach that focuses upon the metabolic organisation. We reviewed research related to this recursive, self-maintaining organisation and the body of philosophical work that connects 'life' with 'mind', using the metabolic organisation to identify a subject of action and norms intrinsic to that subject. As we discuss in more detail in Chapter 4, this approach allows for a formulation of behaviour and goals that are intrinsic to the present organisation of system that is behaving rather than a product of its creator or evolutionary history.

The reviewed literature includes definitions of various categories of behaviour, differentiating between the passive undergoing of dynamics of non-autonomous systems and the behaviour of autonomous systems, pointing out the required extensions of autonomy for the existence of *adaptive* behaviour, and discussing briefly how this approach may one day be able to connect to higher cognition.

The chapter has described the context into which our research, which concerns how the metabolic organisation can influence the adaptive behaviour of radically embodied agents, fits. In the next chapter we shall describe and justify the abstract, computational-simulation methodology taken in this work. The following chapter defines the concepts studied in the computational models in more detail and Chapters 4 – 8 present our computational models along with our analysis of them.

PRIMARY POINTS OF CHAPTER

- Contemporary approaches to the study of behaviour were briefly reviewed, and a problem common to them was highlighted: they do not have a strong objective method for identifying the subject that is behaving ('the identity problem'), and the goals that are intrinsic to such an agent ('the problem of normativity').

- The *metabolic organisation* is an organisational property shared by all known individual organisms. In the simplest possible terms, it describes systems that degrade, but rebuild or repair themselves using energy and material from their environment.

The metabolic organisation provides the foundation for a theory of autonomous individuals, and provides a solution to the identity and normativity problems.

- Various categories of behaviour were discussed, and distinctions were drawn between passive dynamics and the behaviour of an autonomous system as well as adaptive and non-adaptive behaviour.

- There is little work involving the computational modelling of *both* the metabolic organisation and behaviour. This thesis is an exploration of the interaction between these phenomena, focusing on how the dynamics of metabolic self-maintenance can drive forms of adaptive behaviour.

3

METHODOLOGY

"In a sense, physics shrinks and biology expands."
(Rosen, 1991, p. xvii)

CHAPTER SUMMARY

This chapter outlines the research methods employed in this thesis. It starts broadly with a brief discussion of the role of computational models in science, and then becomes more specific, discussing how the research presented in this thesis employs both abstract models and models that target specific phenomena observed by empirical biologists. We observe that employing this combination of abstract models and models targeted on specific biological phenomena worked well for our studies, with each type of model feeding back to the development of the other. The abstract models helped us to develop new theories and approaches to studying the more concrete and empirically verifiable models, and vice versa: the work with the more concrete models (those targeted at specific biological phenomena) suggested new abstractions that are most amenable, at least initially, to abstract, conceptual models. Finally, we describe our particular style of computational modelling, in which metabolism and behaviour are modelled in such a way that the interaction between the two is amenable to study. We have labelled this methodology "computational chemo-ethology" as it lies between artificial-chemistry modelling and computational neuro-ethology.

3.1 THE ROLE OF ABSTRACT MODELLING

In studying behaviour and metabolism, this thesis fits within the field of biology. However, there is not any experimentation with macromolecules like DNA or protein, nor are there any new experiments presented that involve animals, plants or fungi as subjects. There are not even any test tubes. Instead, a computational modelling approach is taken, where computer programs are designed, implemented and executed – the output of which is then analysed.

Using computational models to study biological phenomena has been heavily criticised, and passionately defended over the past few decades. Modelling is now an established scientific methodology employed by physicists, chemists and biologists, but debate continues focusing upon which types of modelling are valid contributions to science and what models are "just toys" i.e., irrelevant to science. In particular, the research reported in journals such as *Artificial Life* and *Adaptive Behaviour* have been subject to substantial scrutiny and criticism as e.g. "fact-free science" (attributed to J Maynard Smith by J Horgan (Horgan, 1995)). A primary facet of this criticism is that there is typically insufficient grounding of models in observation of biological phenomena. The argument is that researchers too often develop a model with little connection to biological phenomena, and in this process, draw or imply too strong conclusions from their computational models.

In a recent development of this argument, B Webb criticises models that do not have a specific biological target (Webb, 2009a). Webb criticises models like Beer's models of minimal cognition / categorical perception (Beer, 2003) in which the target of the model is fabricated by its creator instead of developed from observation of biological phenomena. She argues that "as-it-could-be is a rather poor guide to as-it-is" (Webb, 2009b, p. 351) and that "a simpler strategy is to not make things up in the first place" (Webb, 2009b, p. 350). But Webb's arguments are not as extreme as some of her detractors might have you believe – she identifies three ways that hypothesis and model can contribute to each other before empirical validation of the model: (i) clarification of concept(s) involved in the hypothesis, (ii) validation of internal consistency of a hypothesis, and (iii) development of methodological tools (Webb, 2009b, pps. 349-350) . But, she believes that these sorts of contributions are also found in models targeted upon biological phenomena, and therefore sees little reason to fabricate "animats" as targets of study, when "actual biological mechanisms may be more interesting than those conceived by modellers." (Webb 2009b, p. 353, quoting Bechtel 2009, p. 295). Webb observes that a mix of approaches is probably best.

> There has always been a tension in biology between those who believe the best route to general results is to avoid irrelevant details that can be identified with any specific system, versus those who think careful study of specific systems is more likely to produce deep understanding that can be generalized...As

> usual in such things, the "best" approach probably lies between the extremes,
> or in moving back and forth between them. (Webb, 2009b, p. 351)

The role of abstract computational modelling remains a contentious topic. We shall describe and defend the methodology employed in the work presented in this thesis (below), we do not engage the general debate further here. The interested reader may wish to consult Di Paolo et al. (2000) consider the role of models as "opaque thought experiments" – similar to armchair thought experiments but involve situations where the consequences of the assumptions of the thought-experiment are not obvious from the start and require computational simulation for analysis. Also, Barandiaran et al. (2006) who distinguish between and discuss the epistemic uses of four different levels of abstraction in computational model. Finally, the special issue of *Adaptive Behaviour* that contains a target article by Webb (Webb, 2009a) and a selection of counter arguments from various authors is a good and recent summary of the state of the debate.

3.2 THE APPROACH TAKEN IN THIS THESIS

Webb and Bechtel are right to consider that "actual biological mechanisms may be more interesting than those conceived by modellers" (Bechtel, 2009, p. 295), and that "including some constraints derived from actual biology might help us reach the more relevant systems sooner" (Webb, 2009b, p. 353). But, it is also the case that some of the constraints derived from actual biology impede study, through complexification that is the result of how mechanisms formed, but not how they operate or why they are effective.

This seems likely to be the case for metabolism and its relationship with behaviour. Billions of years of evolution have modified, improved and complexified metabolism to the point where it is an unwieldy object of study (see Figure 3.1 which diagrams a small piece of this puzzle). We chose therefore, to study it in an abstract, idealised form – generalised from observations of philosophers and biologists rather than from the specific measurements of a specific metabolism of a specific organism. In making this decision we sacrifice the ability to directly and quantitatively compare the results of our simulation to empirical biological observation, but we maintain the ability to comprehend the system with which we work.

Thus, the work in this thesis is a mixture of biologically grounded models, such as that presented in Chapter 6 in which we recreate phenomena observed in studies of metabolism-dependent chemotaxis in bacteria such as *A. brasilense*, and models that are less directly matched to quantified observation, such as that presented in chapters 5 and 8. In these latter models, the target is indeed biological, but is more abstract and less easily quantified – it is the general organisation of living systems as hypothesised by Maturana and Varela (1980), Gánti (1975, 2003b), Rosen (1991), Morowitz and Smith (2007), Ruiz-Mirazo and Moreno (2004) and others. With these models we have taken ideas discussed predominantly in philosophical literature and endeavoured to make

Figure 3.1: Metabolic networks are very complicated systems that are difficult to analyse. The work presented in this thesis studies metabolism at a higher level of abstraction, limiting the types of questions that can be asked, but allowing us to ignore some of this complexity.

them more concrete and amenable to scientific study. We are in effect, pursuing two of the routes mentioned by Webb that are available even before the model is empirically validated, 1) clarification of concept(s) involved in the hypothesis, and 2) validation of internal consistency of a hypothesis (Webb, 2009b, pps. 349–350).

Interestingly, it was precisely the result of our experimentation with these more philosophical models (i.e., models that target philosophical notions rather than quantifiably empirically observed biological phenomena) that motivated our development of the model that targeted bacterial behaviour. In other words, our work with the abstract and the philosophical helped us to clarify our existing understanding and to develop new theories and approaches to the more concrete and empirically verifiable model (presented in Chapter 6) that led to new explanations for previously unexplained phenomena in bacteria (Section 6.4.6). Working with this more concrete model, led us to new hypotheses concerning the role of metabolism-based behaviour in the early evolution of life (Chapter 7), that (at least initially) are best explored with more abstract models. It is the opinion, and the experience, of the author and his collaborators that as Webb says, a mix of approaches works well; the abstract and the more concrete models operate best in tandem, each benefiting from the other. Figure 3.2 caricatures this relationship. Connections between theory and nature are not established only through experimentation (tendril extending nature towards theory), but also through development of the theory (fingers of theory, some of which are moving towards nature). The abstract models that we have worked with have played this latter role. Instead of building a model to reproduce observations of biological phenomena, we model an idea to better understand it, to be able to improve, refine and manipulate it into a form better suited to connection with empirical biology.

3.3 MODELLING METABOLISM AND BEHAVIOUR

The most important factor to consider when constructing a computational model is the target of the study. For the present work, the principle conceptual targets are adaptive behaviour and autopoiesis. We therefore need to include both of these dynamics in our models in a way that allows us to study how one might influence the other. To accomplish this, we developed a new modelling technique that lies between two established approaches: artificial chemistries (Dittrich et al., 2001) and computational neuro-ethology (Beer and Chiel, 2008): *computational chemo-ethology*.

The 'Computational Neuro-ethology' approach, proposed by Dave Cliff in (Cliff, 1991) advocated the study of cognition through the study of behaviour of embedded, embodied agents. We adopt the dynamical analysis and behavioural analysis aspects of this approach, studying not only the low-level chemical reactions but also the emergent agent-level behavioural dynamics.

This approach is based upon modelling spatially embedded artificial chemistries. A set of abstract artificial chemicals (reactants) are defined along with a set of rules of
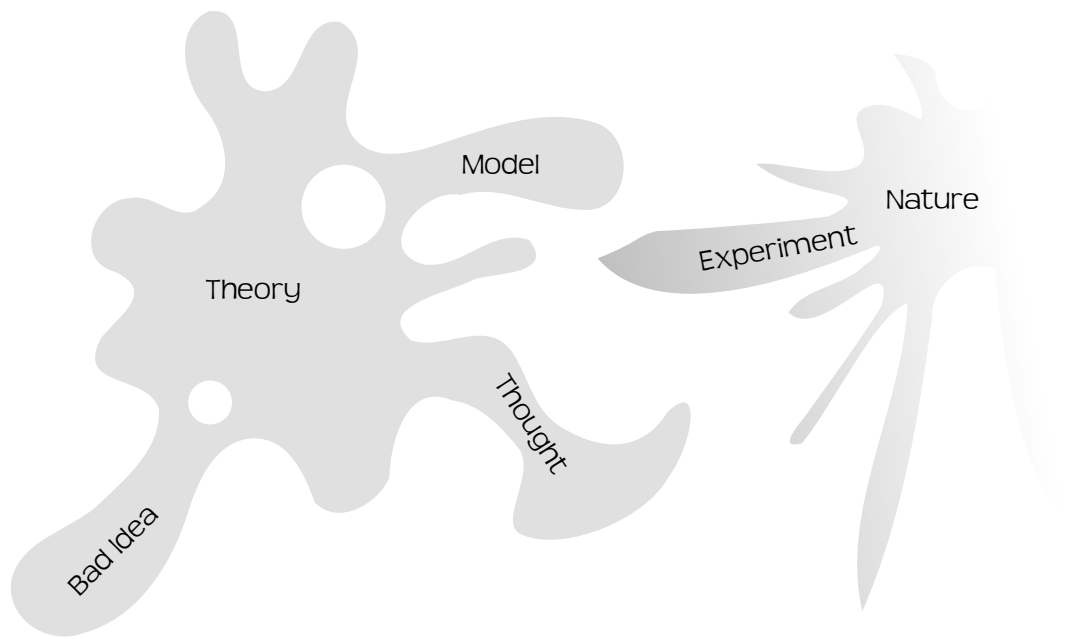
Figure 3.2: Science as the formation of connections between theory and observed nature. Experiments extend the empirically observed while theorising and modelling extend ideas. Here, scientific success is seen as the connection of theory to observation.

interaction and the resulting dynamics are observed. When the chemicals, interactions, and the environment are designed in a particular way, it is possible to observe and analyse different relationships between the processes of autopoiesis and behaviour – a primary goal of this thesis.

The modelling of autopoiesis or behaviour is not new, but the modelling of systems that are simultaneously accomplishing both, and the study of the interaction between these processes is. In the past, models of autopoiesis have only demonstrated autopoiesis *without behaviour*, neither requiring nor performing any behaviour. Similarly, models of behaviour typically ignore or trivialise the role of autopoiesis in their analysis of a behavioural mechanism. By simulating a radically embodied, behaving agent – one that is constructing its behaving self from its environment – as a spatially embedded chemistry, we avoid trivialising self-production/maintenance of a body (autopoiesis) and we avoid trivialising the role of behaviour in autopoiesis. In other words, it is possible in our model to analyse the potentially non-trivial relationship between autopoiesis and behaviour.

PRIMARY POINTS OF CHAPTER

- Computational modelling is an established part of science, but there is an ongoing debate as to what types of models are scientifically valid and which are "just toys".

- The work presented in this thesis involves a variety of types of models, including abstract models of concepts as well as more concrete models that target specific quantifiable biological targets

- Working with the abstract models helped to produce interesting and productive hypotheses that were explored and validated in the more concrete models, which provided new explanations for previously unexplained phenomena and reproduced phenomena observed in empirical studies. This supports the argument that abstract models of ideas are useful tools that can help to refine concepts and theory as argued by others elsewhere.

- Computational chemo-ethology, the methodology employed in models presented in later chapters is introduced. It involves the simulation of artificial chemistry reactants embedded in space, and was designed to allow for the study of non-trivial aspects of the relationship between mechanisms of autopoiesis and behaviour.

<div style="text-align: right">

4

</div>

SELF-SENSITIVE BEHAVIOUR

Our bodies flow like rivulets, their material is renewed like
water in a stream. (Oparin, 1961, p. 9)

CHAPTER SUMMARY

This chapter specifies the theoretical framework for the research presented in subsequent chapters. It begins with an outline of aspects of autopoiesis theory that we consider to be the most interesting and scientifically productive. Core concepts are defined and clarifications are provided as to the role of the 'boundary' in our interpretation of autopoiesis. Section 4.3 discusses how these core concepts can be extended conceptually with the notion of "self-sensitive" *viability modulated behaviour*. This extension forms the basis of a theory of normative behaviour that is strongly intrinsic, unlike the extrinsic basis of behaviour grounded in evolution or weakly intrinsic normativity such as behaviour that satisfies an arbitrary "internal" value such as a battery level. This chapter concludes with a brief discussion of the role of self-sensitive behaviour as a form of adaptive behaviour.

4.1 INTRODUCTION

In Chapter 2 we discussed embodiment, a contemporary movement in philosophy of biology, cognitive science, adaptive behaviour and robotics in which the situatedness (existence in a world) and the interaction between brain, body, and world is seen as playing an important role. This approach has been successful in forging new ground in research, explaining how fish manage to accelerate so quickly, and make such sharp turns; how natural organisms use the passive dynamics of their body to shape their motion (Pfeifer and Bongard, 2006) among other advances.

Clark (1999) differentiates between two approaches to embodiment: 'simple embodiment' where the interactions between the neural system, body and physical world act only as *"constraints upon a theory of inner organization and processing"* (Clark, 1999, p. 348. emphasis his) and 'radical embodiment' in which these same phenomena are seen as *"profoundly altering the subject matter and theoretical framework of cognitive science"* (Clark, 1999, p. 348. emphasis his). Enactivism (Stewart et al., 2011), takes the latter, more radical view, seeing cognition as fundamentally grounded in the sensorimotor dynamics of the interactions between a living organism and its environment.

As mentioned in earlier chapters, this thesis works in parallel with these approaches, but with a different focus on embodiment. Rather than focusing on situatedness and the interaction between brain (neural systems), body and world, the work presented in this thesis considers a different aspect of embodiment – a property shared by all life: the metabolic organisation. All organisms share this organisational property of building themselves from resources available in their environment. This is a fundamental property of the 'bodies' that we observe in biology that is ignored, even by most proponents of the "radically-embodiment" approach. What relevance does this property – a property of the individual organism, not the macro-biology to which it belongs (population evolution), nor the micro-biology from which it is constructed (e. g. enzymes) – have to do with adaptive behaviour and other interesting capabilities that are observed in nature, but so hard to recreate in simulation or in robots? This thesis uses computational models to describe some of the roles that this "ultra-radical embodiment" can have. But before we can delve into the details, it is vital that we define our terms well. This is the goal of this chapter.

4.2 DEFINING A BIOLOGICAL INDIVIDUAL

This thesis is concerned with the relationship between the metabolic organisation and behaviour. The metabolic organisation to which we refer has been discussed by various authors taking different approaches (see Chapter 2). One of the more prominent such approaches is that of *autopoiesis* – the self-production and self-maintenance observed in all life. How can behaviour influence and be influenced by this property? To address this question, it is important to be especially clear what we mean by autopoiesis. This word

has been interpreted in a variety of different ways since its inception in the 1970s. Even the original authors that popularised the term, Maturana and Varela, use the word with subtly different meanings in different publications. It is therefore insufficient for authors to simply refer to autopoiesis. They must provide definitions that make explicit what they mean when they speak of autopoiesis. In this vein, let us now be clear about which aspects of the concept of 'autopoiesis' we find to be interesting and productive concepts.

Autopoiesis is a combination of the Greek words 'auto' meaning 'self' and 'poiesis' meaning 'to make'. In simplest terms, autopoietic systems produce themselves. In contrast, systems that are produced by something other than themselves are called *allo*poietic systems. Cars, for instance, are not the product of cars, but rather of car-factories and are a good example of allopoiesis. The biological cell, actively building, rebuilding, replacing and repairing its constituent parts, is the archetype of autopoiesis. "The bacterial cell is the simplest of living systems because it possesses the capacity to produce, through a network of chemical processes, all the chemical components which lead to the constitution of a distinct, bounded unit." (Varela, 1997, p. 75). With bacteria as inspiration Varela abstracted the property of autopoiesis thus:

> More precisely defined: An autopoietic system is organized (defined as a unity) as a network of processes of production (synthesis and destruction) of components such that these components:
>
> (i) continuously regenerate and realize the network that produces them, and
>
> (ii) constitute the system as a distinguishable unity in the domain in which they exist. (Varela, 1997)

Despite this being one of the later definitions of autopoiesis, it and much of the literature based upon it remain somewhat vague. What is meant by "domain in which they exist" and "regenerate and realize the network that produces them"? That there is a difference between how cells are produced and how cars are produced is relatively intuitive, but much remains undefined and unclear if we are not more formal and explicit about the meaning of autopoiesis. For instance, what is meant by 'self' in the description of autopoietic systems as 'self-producing'? Where does the self start and end? How is it to be distinguished from its environment? To what extent must something be responsible for its own production to be self-producing? Must it be *entirely* responsible for its existence? Does it suffice for it to only coincidentally positively influence its endurability? Where in this spectrum do we draw the line between self-maintaining and not self-maintaining? Why? These are important questions if we are to discuss the properties of self-producing systems.

At the core of the theory, we identify an integration of three clearly expressible, scientifically useful concepts. They are:

1. The flow of material and energy through a dissipative, precarious system.

2. The delimitation of such a system through the relational concept of operational closure.

3. Self-maintenance of that system.

The remainder of this section provides a detailed description of these three core ideas.

4.2.1 *A dissipative, precarious organisation*

Living systems are inherently unstable, degrading systems that counter their degradation with processes of self-repair. Through processes of assimilation and dissimilation, the material that constitutes an organism is in perpetual flux (Oparin, 1961, p. 9). My current structure, a collection of over 10^{27} atoms, will be different by the time I've finished writing this sentence, but my organisation will be more or less the same. I will still have two arms and two legs, a pair of lungs, a circulatory system, etc. In this way, life accomplishes a form of dynamic, organisational stability (Pross, 2008). This can be contrasted with the stability observed in rocks and other non-living forms which is accomplished through a more passive, structure-based stability.

There are other examples of dynamically stable systems that dissipate energy, such as whirlpools and waterfalls. These systems are *dissipative* in that they dissipate energy, but unlike passive dissipative systems such as, for example, a heat sink, dynamically stable dissipative systems are *precarious* – their existence is the result of their dissipation of energy. Should a heat sink be placed in an environment lacking any energy gradient, it would still exist, but the same cannot be said for a whirlpool or a waterfall, etc.

Dynamically stable dissipative systems have a lower state of entropy than their surrounding environment. Their low entropy can only be maintained by increasing entropy elsewhere. Global entropy increases, obeying the 2nd law of thermodynamics, but local entropy of the dissipative system is (temporarily) maintained.

This organisation is universal to living systems. All biological organisms are dynamically stable, dissipative, precarious systems. They require energy to perform work to maintain their highly organised state (see Morowitz, 1979) and they accomplish this by using resources available in their environment. As argued by Maturana and Varela (1980) and discussed below, focusing on these organisational properties provides a method of identifying a biological individual in a way independent of the evolutionary history of the putative organism.

4.2.2 *Delimiting the biological individual*

What is an individual system? If we are to differentiate between systems that are produced by themselves and systems that are produced by other things, we must be able to differentiate between the 'self' and the 'other'. This dichotomy has different guises: organism / its environment, individual / everything else, etc. and there are a

variety of ways in which we can make the distinction. The appropriate way to make the distinction depends upon the context. Here, the context is the notion of self-production, and how we can distinguish between whether a system is responsible for its own production/maintenance or not. We proceed with this as our motivation.

Delimiting a system – identifying where it starts and ends – is not trivial when it exists only as a flux of matter, where the components are constantly changing. If the biological individual has such an organisation, as we argued above, how can we distinguish between it and its environment? This is the subject of an ongoing debate. In the primary literature, Varela and Maturana speak regularly of the 'boundary', but fail to be consistent and clear as to what they mean by this word (Virgo et al., 2009; Luisi, 2003, p. 51). There are two common interpretations.

A spatial or physical boundary

The idea of a spatio-physical boundary is quite straightforward. According to this view, taken by Luisi, Boden and others, the boundary of an autopoietic system is a boundary that surrounds a continuous region of physical space. A cell membrane is often described as playing this role. And, while not autopoietic, a closed cardboard box, a jar or a soap-bubble are boundaries in the same spatio-physical sense. At first, this interpretation may be tempting as it entails the most familiar notion of boundary. Unfortunately, although the spatial boundary is intuitive and easy to identify, it does not function well as a delimiter of a biological individual. Especially, in the present context of self-maintaining systems.

Consider a factory that produces corrugated aluminium. Workers inside this factory use the corrugated aluminium to build and repair the walls of the factory. This system produces a spatio-physical boundary that surrounds the processes that build it. According to the spatio-physical interpretation of autopoiesis, this factory and indeed everything within the boundary of corrugated aluminium would be part of an autopoietic system. But, upon consideration, this is ridiculous. For this corrugated aluminium boundary could surround anything, and then according to the spatio-physical interpretation this thing would be part of the factory. This trivial inclusion of anything simply by surrounding it cannot be the principle by which we identify a self-maintaining system.

Similarly, just because something is inside a cell membrane, does not make it part of the cell or part of the organism. Radiation passing through the cell is not part of it, nor is e. g., a toxin that has been injected into the cell or virus components that have invaded the cell.

Furthermore, when the spatio-physical interpretation of the autopoietic boundary is applied, conceptual difficulties emerge. For instance, Luisi admits difficulty in determining whether or not the self-producing, replicating micelles produced by Luisi et al. (2006) qualifies as to whether or not "the system encompasses reactions that re-generate the components of the system." The "[r]eaction takes place on the boundary, which is part of the system, but a rather restricted one: the internal aqueous core is not part of

the reaction system, and furthermore we are dealing with only one reaction instead of an internalized reaction network...In all the experiments exemplified[...] what is being reproduced is the empty shell. This is not suitable for making models of biological cells." (Luisi, 2003, p. 56)

We ask *Why*, fundamentally, is it important for the membrane to surround other processes? What does this add to the theory? What is the principle that drives Luisi's dissatisfaction with the artificial self-producing, replicating micells? Is there something missing from the self-reproducing micelles or is there something wrong with the approach being used to identify autopoiesis?

Boden (2000) also appears to support a physico-spatial interpretation of autopoiesis, stating that:

> There is no independent ground, other than the wish to save strong A-Life, for deleting reference to "the physical space". Doing so would obscure the fundamental distinction between starfish and societies, or butterflies and businesses...All these systems are autopoietic in the most general, abstract, sense. But only self-organizing systems in the physical space can originate real, metabolising, life. (Boden, 2000, p. 142)

Boden's desire to define life in such a way as to distinguish between "starfish and societies" is understandable, the solution does not lie in the notion of the physical / spatial boundary, as this concept comes with its own problems. Fundamentally, it does not provide a particularly strong or coherent way for identifying the biological individual. A rather more coherent view emerges when instead of considering the boundary of the autopoietic system as spatio-physical, we think of it as a relational boundary.

A relational or organisational boundary

Seeing that the spatio-physical approach to the boundary of an organism appears to be lacking in certain respects, what is an alternative approach? Bourgine and Stewart (2004, p. 337) argue that the "semi-permeable boundary" has been over-emphasised and that it should be "replaced by a more dynamical emphasis on the regulation of boundary conditions." They see autopoiesis as a "network of processes that produce the components that reproduce those processes" (Bourgine and Stewart, 2004, p. 327). In other words, autopoiesis is a property shared by a set of processes whereby they are all interdependent. Only processes that have this relationship with each other are considered part of the system – i.e. part of the identity or the 'self' in self-maintaining (Virgo et al., 2009).

More formally, given a collection of processes **C**, we can identify an operationally closed subset of those processes **S**, such that for every constituent process P, the following conditions are true.

1. Another process P' requires conditions produced by process P
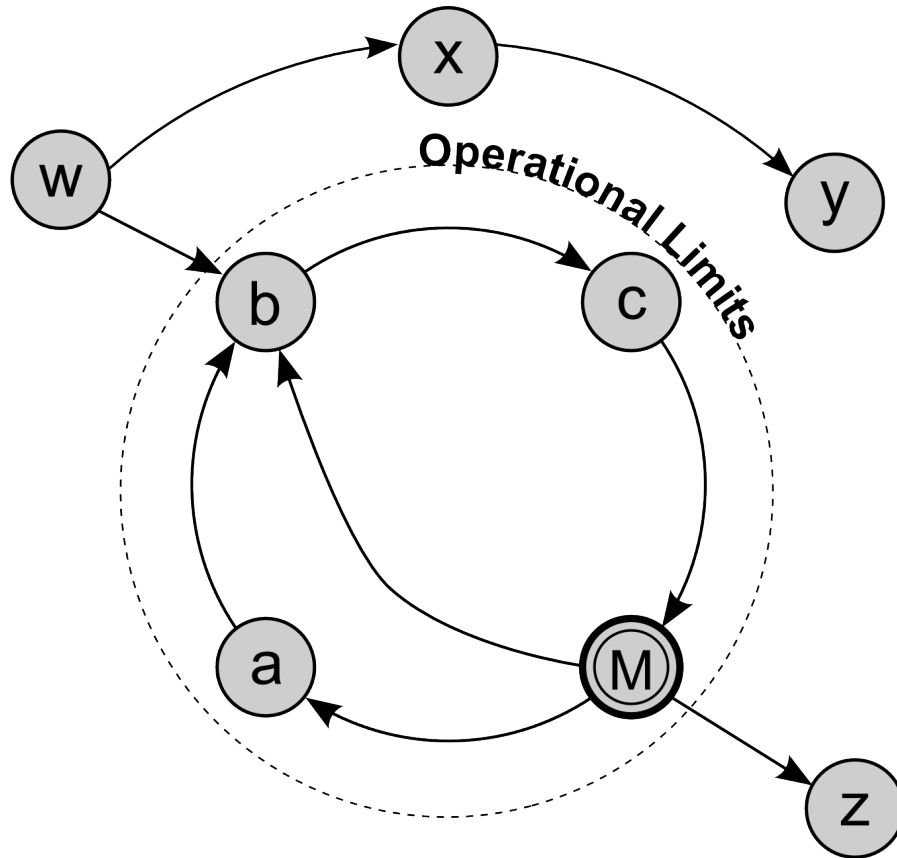
Figure 4.1: A diagram indicating operational closure, the relational interpretation of the boundary of the individual. Nodes are processes and arrows indicate dependence. For instance, process Z depends upon process M. See main text for further explanation.

2. Process *P* is conditioned by another process *P″*

3. *P′* and *P″* ∈ **S**

4. *P′* and *P″* can be (but are not required to be) the same process.

In graph theory, this defines a strongly connected sub-graph of the directed graph of process dependencies. Assuming that all of the processes (vertices) and interdependencies (arrows) are included in Figure 4.1, processes *w*, *x*, *y* and *z* are not part of any operationally closed network. This is the case because each one of these processes does not depend upon another process which is in turn dependent upon the original process. For instance process *x* is dependent upon w which is not dependent upon any process. The absence of cyclical dependence indicates an absence of operational closure. In contrast, processes *a, b, c* and *M* form an operationally closed network. Process *c* depends upon *b* which depends upon *a* which depends upon *M* which depends upon *c*, closing the loop and making the set of four processes operationally closed. A second, smaller operationally closed loop also exists, consisting only of processes *M, b,* and *c* (in

Chapter 8 we discuss nested and other networks of operational closure). These are the only operationally closed loops within this system.

If this relational interpretation is followed, then the spatio-physical boundary remains important in that 1) it enables many processes and 2) it physically separates the individual from the environment (individuation), allowing also for separation from other instances of a similar organisation which allows, among other things, social interaction and evolutionary processes. But, it is less central in defining or delimiting a *self*-maintaining system. To emphasise this, in Figure 4.1, we have indicated the membrane or spatial boundary as node M, one of the component processes of the self-maintaining system.

In the past, because it is easily visualised, because terminology was confusing, and perhaps because of difficulties associated with translation, we have seen the physical boundary of an autopoietic system as playing a special, definitive role. But here we have argued that in the relational domain the spatial boundary should take its place among the other interdependent conditions. It is not of importance where, physically, a process is occurring, but rather how that process is related to other processes, and ultimately how that process is (possibly indirectly) dependent upon the result of its occurrence (Virgo et al., 2009).

The models presented in the next chapters are all of cell-like systems that include a spatial boundary, similar to a membrane. Why, when we have just made the argument that the spatial boundary is not the important boundary in defining an individual, would we then include a spatial boundary in all of our models? The answer is partially pragmatic. The simplest way to capture the idea of a precarious network of interdependent processes capable of self-regulating behaviour was to copy the interdependent relationship between the membrane and the metabolism observed in cells. The interdependence is between the membrane and the chemical reactions that produce the membrane. The membrane depends upon the reactions for its production and maintenance, and the reactions depend upon the membrane to limiting the diffusion of the reactants, thus maintaining their concentration as sufficiently high to continue to react.

It may not be that this is the just the easiest way to capture this idea. Ruiz-Mirazo and Moreno argue that "The generation of a physical border is a crucial step toward autonomy, because that is the *only* way, on the one hand, (i) to assure the control of energy flow required for the robust maintenance of the network, and on the other, (ii) to solve the problem of diffusion and dilution (control of concentrations). Furthermore, if the precursor reaction network does not become self-enclosed, [it] [...] will not have any control over the boundary conditions that bring about its distinctive, far-from-equilibrium dynamics (and, thus, it will be extremely fragile)." (Ruiz-Mirazo and Moreno, 2004, p. 243, emphasis mine) We agree that the spatial membrane likely played a central, important role in the origin of life and indeed, in organisms ever since. However, we want to emphasize that the important *principle* is the network of interdependent processes and that the spatial membrane, the chemicals it contains, and their interdependence is only an *instance* of that principle. Although it might be difficult at this stage to conceive of
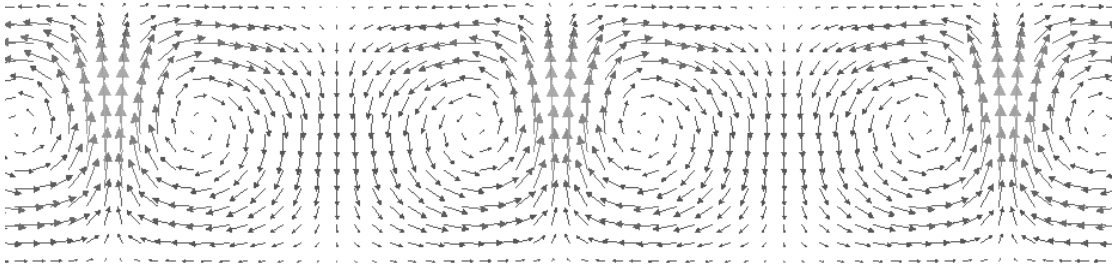
Figure 4.2: Bénard convection cells are dissipative convective structures that spontaneously form in certain conditions. They are most commonly observed when heating oil in a frying pan as an array of dips in the oil. This figure shows the result of a numerical simulation using the CFD software Flotherm indicating the convective flow in a cross-section of three Bénard cells.

an alternative instance of an autonomous system capable of self-regulation, we think it would be a mistake to conflate the theoretical principle with the best (or only known) concrete example of it.

4.2.3 *Self-maintenance of organisation*

Dissipative systems are reliant upon a source of energy and, often, a source of material renewal. Energy and matter cannot be created *ex nihilo*, therefore self-production can only ever be partially accomplished by the self. While a system cannot be fully autonomous, i. e. fully and independently capable of determining the likelihood of its future existence, a systems future existence can fundamentally depend upon its own influence.

If we consider, for instance, a biological cell. In the absence of a membrane, the enzymes and other reactants would rapidly diffuse, becoming too low in concentration for cellular processes to continue to occur. Its membrane makes possible the processes that produce and maintain it. Similarly, food and water are acquired by organisms that use those resources to maintain their organs, limbs, etc. Without the organs and limbs, the acquisition of these resources would be impossible, and without the food acquisition, the limbs cannot continue to exist. These networks of processes are self-dependent – disrupted for too long, and they will collapse. They will die.

A sceptic might argue that an inanimate object, like a chair, depends upon its own existence for its continued existence. The chemical and atomic bonds all must be configured just so *now* for the chair to exist a moment later. But chairs are not particularly dissipative systems, and importantly do not rely upon their dissipation of energy for their ongoing existence.

Should we consider relatively simple dissipative systems like hurricanes or a Bérnard convection cells (see Figure 4.2) to be self-maintaining organisations? Both of these systems positively influence the likelihood of their continued existence. There is, however, a significant difference in degree between this type of self-maintenance and that observed

in biology. We can conceive of this difference as a difference in complexity of the interdependent processes that comprise the system. The intricacy of the Bérnard cell is much less than that for the simplest organism. Simpler dissipative systems comprise fewer processes and require less specific conditions to operate than biological systems. They are therefore vastly more likely to spontaneously emerge than an organism. Certain perturbations to a simple dissipative systems (like Bérnard convection cells) cause them to collapse, but in the right conditions the system can spontaneously reform. Unlike the more intricate network of processes involved in e. g. a cell. Perturb a cell and it dies. Once dead, it will not spontaneously reform.

Thus, it is possible to use operational closure to identify the boundaries of a system that is maintaining itself, but these self-maintaining systems can vary in complexity. Early, simpler life forms must have been simple enough to spontaneously form in abiotic conditions, but modern life is a more complex, tangled network of interdependent processes, which if excessively perturbed will be destroyed and will never spontaneously reform.

4.2.4 *Combining it all*

In this section, we have provided a definition of an autopoietic system. The definition is built upon the integration of three ideas:

1. Precarious, dissipative organisation

2. An individual defined by its operational closure

3. The self-maintenance of that operationally closed, precarious, dissipative organisation.

We can combine these ideas to define a biological individual as a precarious, dissipative, operationally closed system that fundamentally relies upon its influence to exist. The precarious, dissipative organisation and the idea of operational closure allow us to identify an individual organism and to differentiate between it and its environment. This helps us to be clear about what is meant by *self*-maintenance, or 'autopoiesis' – a universal property of living systems.

This property, specifies a way that a system can be inherently needful of certain conditions and therefore capable of acting in such a way as to achieve those conditions. In so doing, these systems can act to satisfy their own demands and not the demands of another. Only if a system has the described organisation can it act *on its own behalf*. The next section starts the examination of how autopoiesis can produce this kind of 'self-motivated' behaviour.

4.3 EXTRINSIC AND INTRINSIC BEHAVIOUR

The widespread approach to conceptualising adaptive behaviour has been to consider it the optimisation of certain parameters (captured by the notion of fitness) or as the maintenance of certain, essential variables (see e.g. (Ashby, 1952)) within viability limits. As a consequence, models of adaptive behaviour generally fall into one of two categories (or a combination of both).

Extrinsic: Optimisation techniques are used to constrain the behaviour of a system to achieve the desired adaptive coupling with its environment in relation to a set of parameters or "fitness" criteria. This category includes different types of supervised learning algorithms for neural networks, simulated annealing or artificial evolution techniques to design control architectures (as used in Evolutionary Robotics, (Nolfi and Floreano, 2004)) or, simply, hand design.

Weakly Intrinsic: Models belonging to this class incorporate a set of internal variables often interpreted as energy sensors, pain or pleasure indicators, etc. These "value modules" are then coupled to other control mechanisms in order to tune the behaviour of the system (as in reinforcement learning) or to choose between competing possibilities for action (acting as an action selector (McFarland and Houston, 1981; Meyer and Guillot, 1990)).

In both cases the parameters or functions to be optimised are explicitly represented either as an external fitness function or as an internal value module, abstractly measuring how well adapted/adapting the system is. There is generally no reference or feedback to the processes from which these criteria emerge. How those boundaries of viability or optimal values come to be there in the first place is rarely addressed and modelled. As Randall Beer recognises, "this explicit separation between an animal's behavioural dynamics and its viability constraints is fundamentally somewhat artificial. (...) However (...) we can assume that its viability constraint is given a priori, and focus instead on the behavioural dynamics necessary to maintain that existence. (Beer, 1997, p. 265)".

In many studies, this abstraction may be reasonable. For instance it is obvious that above a certain temperature value an organism will die or that without a certain quantity of resources it would cease to exist. However, these conditions (or value functions) are often variable and difficult to determine. They show temporal variability and subtle interactions with other processes (e.g. you can survive at a low temperature for some time but not for "too long" and this in turn might depend on your diet, etc. (see Stewart et al., 2011, pps. 45–48 for discussion concerning the limitations of the value-system approach). Critically, the behaviour of organisms might be sensitive to these conditions in many and sophisticated ways that are lost when a priori abstractions are made. For instance, organisms might display a complex dynamic interplay between internal and behavioural adaptive modulations where mechanisms of self-repair, growth, digestion

and maintenance are integrated with behaviour generating mechanisms in many subtle ways (Alexandre and Zhulin, 2001).

What happens when we remove this somewhat artificial and explicit "separation between an animal's behavioural dynamics and its viability constraints"? To address this question requires reference to the more fundamental aspects of biological organisation: the the modelling of energy consumption processes, metabolic organisation, generation of movement, etc. There have been a variety of models addressing protocells (see Rasmussen, 2009 for a collection of recent work in this area) or autopoietic organisations (see McMullin, 2004 for a review) dynamics. However, as mentioned in Chapter 3, these models tend to ignore or dismiss adaptivity and organism-level behaviour. They describe life as a networked set of chemical reactions (metabolism) continuously re-producing the conditions required for their existence, focusing sometimes upon the biochemical processes that make viability conditions and value functions be there in the first place, but fail to consider interaction between organism-level behaviour and metabolism. In these models, the autopoietic system is placed in environments that do not require any system-level regulation of interactions with the environment (i. e. behaviour) to maintain themselves. Therefore, it is clearly the case that behaviour and the interaction between behaviour and autopoiesis are impossible to study in these types of models.

In other words, most contemporary models tend to *either* focus on adaptive behaviour *or* on autopoiesis, but not both. In both cases, due to the design of the models, the possibility of interaction between these two phenomena is either trivialised or omitted completely. A very small number of recent models have begun incorporating mechanisms of system-level behaviour which the autopoietic processes depend upon, such as motion towards resources (Suzuki and Ikegami, 2009), or management of membrane transport systems to avoid osmotic crisis (Ruiz-Mirazo and Mavelli, 2007). Yet, the interplay between behaviour and autopoiesis remains largely unknown. All of the work presented in this thesis focuses upon this underexplored interaction between autopoiesis and behaviour.

4.3.1 *Sensitivity to the viability threshold*

In contrast to extrinsic and weakly-intrinsic behaviour, *strongly intrinsic* adaptive behaviour is the result not of external processes, but of mechanisms that operate within the operationally closed individual. In theory, these processes can benefit or damage the individual, but for the moment, we are concerned with adaptive behaviour and therefore focus upon those strongly intrinsic behaviour mechanisms that produce behaviour that increases the chance of continued existence of the behaving individual.

This *cognitive* ability of autopoietic systems is described by Varela thus: "the autopoietic mechanism will maintain itself as a distinct unity as long as its basic concatenation of processes is kept intact in the face of perturbations and will disappear when confronted with perturbations that go beyond a certain viable range which depends on the specific
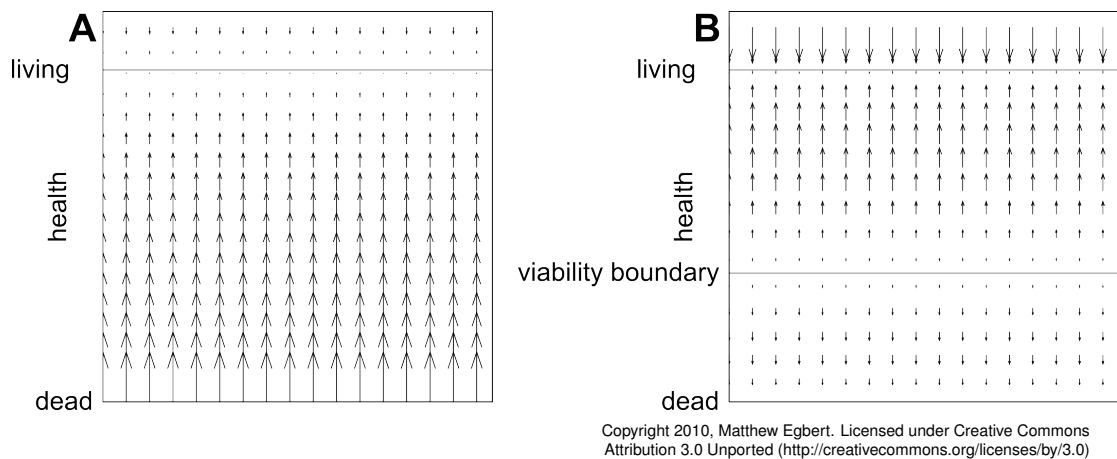
Figure 4.3: A viability boundary is an unstable equilibrium between living and dead states.

system considered." (Varela, 1997, p. 76). The viable range, Varela and Maturana term the "cognitive domain" of the organism, an idea that is made particularly clear in Beer's excellent paper *Autopoiesis and Cognition in the Game of Life* (Beer, 2004) where he provides a didactic example of the cognitive domain of a glider in Conway's game of life.

Bourgine and Stewart (2004), following (Varela, 1997) emphasise that "autopoietic systems can deal with a certain range of perturbations, but that perturbations beyond this range lead to the collapse of the system" and further the analysis by comparing two mathematical models of autopoietic systems. The first has two steady states. A 'zero-state' that is unstable in that a small perturbation causes it to move into the second non-zero state (illustrated in the left of Figure 4.3). To Bourgine and Stewart, this is unsatisfactory as a model of an autopoietic organisation as the system should depend more fundamentally upon its actions. In other words, it should not be a trivial thing to perturb the system from a dead-state into the living state. In their second model, the system has two stable attractors, a zero-state corresponding to death, and a non-zero state corresponding to life (illustrated the right of Figure 4.3). In between these two fixed-point attractors is an unstable equilibrium – the *viability threshold*. A small perturbation here can send the system on a trajectory towards either fixed point. Bourgine and Stewart observe that this 'cognitive' property was present in Varela et al's original model, but difficult to observe due to the stochastic nature of the model and that this motivated their dynamical systems analysis of the two models.

In this thesis we explore various ways that an agent with a metabolic organisation can be self-sensitive to its metabolic viability, i. e. modulate its behaviour according to the state or changes in the state of its viability.

4.3.2 *Is this "minimal cognition"?*

Some might argue that this self-sensitive behaviour is too simple to be considered cognition – that cognition is only accomplished by more complicated "higher" organisms

and that bacteria, or protocells are not cognitive entities. This depends, of course, upon how one defines cognition. Bourgine and Stewart (2004) provide the following definition of cognition:

> A rough preliminary definition of a cognitive system is that of a set of processes in structural coupling with its environment such that the system adapts to its environment and/or transforms that environment in such a way as to adapt the environment to the needs of the system. (Bourgine and Stewart, 2004, p. 327)

But strangely, they go on to suggest that this definition "does not apply in an immediately obvious way to unicellular organisms or multicellular plants" (Bourgine and Stewart, 2004, p. 328). One need not study bacteria or plants for long to observe such adaptation to the environment or transformation of the environment. Simple movement suffices for the latter in that by changing position in the environment, a bacterium (or a plant) changes the environment that they are exposed to. It is in this way that *E. coli* climbs resource gradients and plants perform phototaxis. If we are to use the (quite reasonable) definition of cognition provided by Bourgine and Stewart, then we should include bacteria and plants as cognitive. What about robots?

In defining cognition, Bourgine and Stewart (2004, p. 338) speak of a coupling between senses and actions to satisfy 'viability constraints'. They say that they keep the notion of viability constraint intentionally vague rather than limiting it to the maintenance of autopoiesis. The motivation behind this is the desire to allow for artificial, ungrounded, abstract behaviour (they give the example of a robot that must remain on a table without remaining still) to be classified as cognitive. They have broadened the meaning of a viability condition to include other dynamics. The viability boundary ceases to be a viability boundary and becomes an arbitrary dynamical boundary, determined not by the organisation of the system, but by an external observer. This is throwing the baby out to save the bathwater. A primary contribution of autopoiesis approach to cognition is the provision of a viability boundary that is determined by (i.e. is an intrinsic property of) the system. In other words, Bourgine and Stewart want to keep the externally-determined (ungrounded) behaviour of robots as categorised as "cognitive" and they shouldn't.

This argument raises the question of behaviour that does not serve metabolic viability. Writing this thesis does not improve my metabolic health and behaviour such as smoking, taking drugs or committing suicide accomplishes just the opposite. It would be an odd definition of cognition that did not include these behaviours. One solution to this conundrum is the notion that these behaviours are the result, not of the metabolic autonomy (autopoiesis), but of other autonomies – dynamics that operate over the nervous system and the body and the environment that are self-maintaining. According to this view, cognitive systems are collections of interacting autonomies, "habits" or "mental life"-forms (Barandiaran, 2008) sometimes working synergetically, sometimes in conflict (the smoking example). This thesis focuses upon the simpler, better understood

metabolic autonomy (although we do discuss these ideas more in Chapter 8). Perhaps by understanding metabolic autonomy, we can gain insight into how to study other autonomous dynamics that produce higher forms of cognition.

4.4 RECAPITULATION

In this chapter, we have outlined a definition of self-maintaining, self-sensitive agents. This definition is intended to capture an interesting (i.e. scientifically productive) organisational property of living systems. Taking this approach has proven successful at illuminating new properties and capabilities available to biological organisms. We demonstrate this in the next four chapters by elucidating potential advantages provided by this organisation to living organisms. These include mechanisms of context-sensitive adaptive behaviour at the scale of the biological individual. They also include mechanisms through which self-sensitive behaviour of self-maintaining individuals can facilitate adaptive Darwinian evolution. This definition has been a work in progress since outlined by Maturana and Varela. Over time, it is becoming clearer and better described, but there is work that remains to be done. This chapter has outlined what we believe to be the most concrete and scientifically tractable aspects of this paradigm – the theoretical basis for the research presented in subsequent chapters.

PRIMARY POINTS OF CHAPTER

- This chapter describes the theoretical framework of the research presented in subsequent chapters. Three core concepts of the autopoietic or "metabolic organisation" are defined and explained:
    1. Precariousness / dissipative organisation
    2. Operational closure
    3. Self-maintenance

- Three categories of behaviour are presented: *extrinsic*, *weakly intrinsic*, and *strongly intrinsic*. Unlike the others, strongly intrinsic behaviour is generated by mechanisms that are internal (with respect to the operational closure), and does not include the behaviour of any typical robots, but allows us to consider for the first time, a scientific approach to studying a system that is acting on its own behalf.

- An extension of these core concepts was outlined whereby an autopoietic system, by being sensitive to its viability can behave adaptively employing *strongly intrinsic* behavioural mechanism. This organisation is the focus of the next four chapters of experimentation with computational models.

- A brief discussion is given concerning the role of self-sensitive behaviour in the study of "minimal cognition".

<div style="text-align: right">

5

</div>

IDEAS IN ACTION: A MODEL OF THE IDEAS DEVELOPED THUS FAR

The idea that biological organization is fully determined by molecular structures is popular, seductive, potent and true up to a point — yet fundamentally wrong... It disregards the fact that the cell as a whole is required to create the proper environment for self-assembly to proceed. (Harold, 2003, p. 56)

CHAPTER SUMMARY

This chapter presents a model that we developed and later presented at the 10th European Conference on Artificial Life (Egbert et al., 2009). The model is one of a series of steps we took in developing the concepts presented in this thesis and serves well as an example of many of the ideas outlined in the previous chapter. It simualtes the dynamics of an abstract protocell that adheres to the criteria outlined in Chapter 4, and demonstrates a behavioural mechanism that is integrated with the mechanism of autopoiesis. This relationship makes it possible for the protocell to respond not directly to environmental phenomena, but *to the effects of those phenomena upon processes of self-production*. This chapter serves therefore as an demonstration of how by being a self-producing system, a system can also be sensitive to how well it is accomplishing that self-production and

modulate its behaviour accordingly, demonstrating a simple but powerful mechanism of adaptation.

We describe and discuss four experimental scenarios that demonstrate key-concepts and emergent dynamics of the simulated protocell. Specifically, we show how the self-sensitive behaviour of the protocell provides it with a mechanism of adaptation that is context sensitive and influenced by the history of experience of the system. We also demonstrate how self-sensitive behaviour demonstrates a life-like 'lazy-efficiency' in that the protocell only moves to new resources if doing so improves its situation. This is unlike the more obsessive behaviour seen in many models of adaptive behaviour.

5.1 MODEL OVERVIEW

The model described in this chapter follows the chemo-ethology approach (see Section 3.3) – a combination of artificial chemistry techniques and behavioural modelling and analysis. Within the simulation is an "agent" that can be thought of as a highly simplified model of a protocell (Rasmussen, 2009), simulated at the molecular level. The agent meets the criteria outlined in Chapter 4, in that it is operationally closed, precarious, and self-maintaining (we elaborate upon how the model fits these criteria in Section). It also demonstrates a form of viability-sensitive behaviour in which a mechanism of motility is coupled to the mechanism of self-maintenance. We discuss in Section 5.4 a series of experiments with the model how this coupling of behaviour to metabolism provides a powerful mechanism of adaptability to the agent. The next section describes the model in detail.

5.2 MODEL

The model involves the simulation of three types of interactants: particulate metabolites, gradients of diffusing resources, and a circular membrane that encapsulates the metabolites. These three phenomena are simulated in a two-dimensional environment, 256 units square, where they interact and influence one and other, resulting in a self-maintaining, operationally closed, precarious agent that is capable of a metabolism-based chemotactic behaviour. Figure 5.1 is an image taken from a typical run of the simulation, indicating the membrane (circle) surrounding the metabolites (points and stars within the circle) in an environment containing three different types of resources (diffuse areas of colour). Before describing these higher level emergent system dynamics, we describe how the three types of low-level interactant (metabolites, resources and membrane) are simulated.
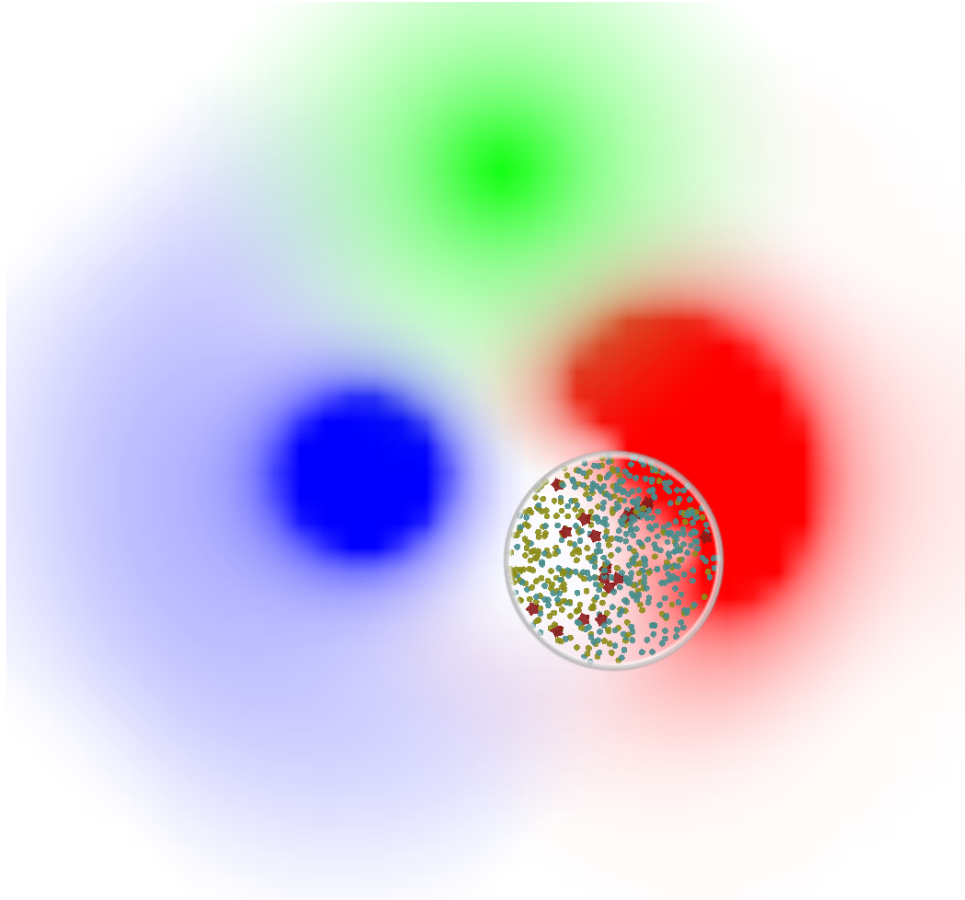
Figure 5.1: A screen-capture our model "chemo-ethology of an adaptive protocell" presented in this chapter.

5.2.1 *Membrane and motion*

A single circular membrane is simulated that surrounds all of the metabolites. The membrane can move. Its motion is governed by a simple simulation of Newtonian physics including drag (fluid resistance). In the equation below, which describes the acceleration of the membrane, the vector **x** represents the two-dimensional location of the centre of the membrane. The first term indicates the linear drag with a drag coefficient, $k_d = 0.01$. The second term indicates the influence of the interaction that takes place when certain metabolites contact the membrane. This interaction is described in detail below in Section 5.2.2.

$$\ddot{\mathbf{x}} = -k_d \dot{\mathbf{x}} + \mathsf{n}(\mathbf{p}, m) \tag{5.1}$$

The membrane is not of a fixed diameter, but can grow and shrink. The size of the membrane is described by the variable ρ which indicates the number of molecules or

"phospholipids" that form the membrane. The radius of the membrane is directly proportional to the number of phospholipids thus: $r = 20\rho/2\pi$. The number of phospholipids in a membrane decays exponentially according to the equation $d\rho/dt = -5\rho \times 10^{-4}$. The number of phospholipids can also grow when certain metabolites contact the membrane (described in the next Section).

5.2.2 *Metabolites*

Metabolites are simulated as dimensionless points that are capable of moving through the two-dimensional simulated environment. There are three different types of metabolite, X, Y, and Z. Each simulated metabolite is one of these types. The type of the metabolite determines its properties. These properties include the metabolite's diffusion rate constant, its stability (the chance of it being removed from the simulation), what reactions it participates in (see Table 5.1) and how it interacts with the membrane should it come into contact with it.

The motion of metabolites is the result of two processes: Brownian motion and the advection of the contents of the protocell. The Brownian motion is captured by simulating a random walk by generating a change in location matrix **g** where the elements are selected from a Gaussian distribution with a mean of 0 and a standard deviation equal to the diffusion rate constant of the metabolite (see Table 5.2). We included in our simulation the advection of the metabolites. We refer to the idea that when the protocell is in motion, it is not only the membrane that is in motion but also its contents. To capture this in a simple manner, we include the second term in the equation below which means that the metabolites take on a fraction ($k_a = 0.75$) of the velocity of the membrane ($\dot{\mathbf{m}}$), and thereby moving with it.

$$\frac{\delta \mathbf{x}}{\delta t} = \mathbf{g} + k_a \dot{\mathbf{m}} \tag{5.2}$$

Each metabolite is inherently unstable. Each iteration, there is a chance ($p = 5d \times 10^{-3}$) that the metabolite disintegrates (is removed from the simulation). Here d indicates the instability of the metabolite, i. e. how likely it is to decay. This parameter depends upon the type of metabolite, and can be found by refering to Table 5.2.

When in proximity, metabolites can react. The only possible reactions are indicated in Table 5.1. Some of these reactions take place not between two metabolits, but between a metabolite and a resource. Resources are described in more detail below. Each reaction can take place in the forward direction or backward direction —although the backward rate constant for most of the reactions is zero, meaning they only proceed in the forward (i. e., left to right) direction. For a reaction to take place, the relevant reactants must be in close proximity (within 2 units of distance from one and other). Additionally, the chance of a reaction taking place is constrained by the rate constants ρ_f and ρ_b (indicated in Table 5.1) that specify the likelihood of the reaction occurring in the forward direction

and backward directions respectively. To simulate these reactions, *N* times each iteration, 2 metabolites within the simulation are selected at random. If the metabolites are within 2 units of distance from each other and if a random number taken from a flat distribution between 0 and 1 is less than the relevant ρ value, then the relevant reaction occurs. *N*, the number of reaction checks per iteration, is calculated each iteration to be equivalent to two percent of the total number of metabolites in the simulation, so that the total number of metabolites does not influence the probability of a reaction occuring between two given metabolites.

In addition to these metabolite-metabolite reactions, metabolites can interact with the membrane. A metabolite-membrane interaction occurs when the motion of the reactant causes it to come into contact with the membrane. For metabolites X and Y, the metabolite-membrane reaction is minimal. These metabolites "bounce off" the membrane, and are relocated to a position slightly closer to the centre of the protocell. Metabolite Z reacts more substantially with the membrane, influencing it in two ways. Firstly, it imparts a force to the membrane, causing an outwards acceleration to the membrane. Secondly, it increases ρ, the number of phospholipids in the membrane, increasing the membranes size. In the process Z is used up, i.e. removed from the simulation as if it were a number of phospholipids that became part of the membrane.

These affects of Z upon the membrane are simulated in the following way. Upon contact with the membrane, a Z metabolite is removed from the simulation, ρ is increased by 0.15, and a force **f** is applied to the membrane. The force vector *f* is calculated by normalising the vector **j** that goes from the centre of the membrane to the last position of the Z metabolite to determine the outward orientation of the force. This vector is then scaled by the force constant of the metabolite $v = 0.1$ (see Table 5.2).

$$\mathbf{f} = v\frac{\mathbf{j}}{|\mathbf{j}|} \tag{5.3}$$

Finally, it is worth pointing out that in this simulation, metabolites never exist outside of the cell as they are created inside the cell and cannot move through the membrane.

5.2.3 *Resources*

The model includes diffusing 'resources', intended to be analagous to sugars or other metabolic resources. The idea is that these resources are transformed *by* the protocell *into* the protocell as part of the metabolic process. There are three types of these resources, R_0, R_1, and R_2. Each is used in a reaction in Table 5.1 in which a metabolite transforms the resource into a metabolite. We wished to study different environments in which the resources could be easily added and consumed and in which the resources underwent a process of diffusion. So, for computational efficiency we modelled resources in a different manner than the metabolites. Instead of modelling the resources as particles, we model a lattice of concentrations of resources. This is a 64×64 lattice of squares 4.0 units wide.

Reactions

#	$R_1+R_2 \leftrightarrow P_1+P_2$	ρ_f	ρ_b	κ
0:	$Z + R_0 \leftrightarrow Z + Z$	1×10^{-2}	0	0.7
1:	$X + R_1 \leftrightarrow X + Y$	1×10^{-2}	0	0.7
2:	$Y + R_2 \leftrightarrow Y + X$	1×10^{-2}	0	0.7
3:	$X + Y \leftrightarrow Z + Z$	5×10^{-3}	1×10^{-3}	n/a

Table 5.1: Chemical Reactions. ρ_f and ρ_b represent the rate of chemical reactions in the forward and backward directions respectively.

Metabolite Types

Name	Diffusion Rate	Stability	$\Delta\rho$	v
X	0.125	0.005	0.00	0.0
Y	0.125	0.005	0.00	0.0
Z	0.2	0.001	0.15	0.1

Table 5.2: Metabolite parameters. $\Delta\rho$ indicates the change in the number of phospholipids should the metabolite contact the membrane. v indicates the magnitude of the force applied to the membrane should the metabolite contact the membrane.

Each square of the lattice is associated with a concentration of each of the three resources. The lattice is updated through Euler integration of the following differential equation which simulates diffusion between adjacent squares.

$$d\phi(\mathbf{r}, t)/dt = D\nabla^2 \phi(\mathbf{r}, t) + q(r) \tag{5.4}$$

In this equation, $\phi(\mathbf{r}, t) \in [0, 3]$ represents the concentration of the resource at location $\mathbf{r}$ at time t, and $q(r)$ represents the addition of resources to the environment at 'resource generators' which are placed in different areas depending on the experimental scenario. The local concentration of resource is increased by 0.01 every iteration. Resource concentrations are constrained such that if they drop below 0, they are immediately set to 0 and if they exceed 3 they are immediately set to 3.

To simulate metabolite-resource interactions, a random sampling of metabolites is taken of the same size N as in the simulation of metabolite-metabolite reactions. For each possible metabolite-resource reaction, a check is performed whereby a random number is selected from a flat distribution between 0 and 1 and the reaction occurs if this number is less than the relevant ρ value. To capture the relationship between resource concentration and the likelihood of a reaction occuring with that resource, a second random number is selected, also from a flat distribution between 0 and 1. This number must be less than the concentration of the relevant resource at the lattice square closest to the location of the metabolite if the reaction is to occur. When a metabolite-resource reaction takes place, the parameter κ indicates the quantity of resource that is consumed by the reaction.

5.3 CONCEPTS FROM CHAPTER 5 IN THE MODEL

This section describes how properties discussed in Chapter 4 are instantiated in this model. We describe how this is a model of an operationally closed, self-producing, dissipative and precarious individual with a strongly intrinsic behaviour.

5.3.1 *A precarious, dissipative organisation*

All of the components of the self-maintaining system (chemicals X, Y, and Z and the membrane) degrade. They are inherently unstable, but can be dynamically stable (Pross, 2008) through a process of production that compensates for their degradation. In the model, the processes of production always consumes resources. Because all of the components of the system are unstable and must consume resources to dynamically maintain non-zero concentrations, the modelled system can be considered dissipative.

Some environments lead trivially to the growth of the protocell (see e.g., the exponentially growing protocells in Figure 5.3 from Experiment 1). In other environments, it is impossible to survive (grey lines in the same figure). In a third type of environment, (such as that in Experiments 2–4), the protocell is neither guaranteed to survive nor to fail; survival depends upon the behaviour of the agents and their health oscillates between improving and worsening trends. The precariousness of the system is most clearly seen in the second two types of environments, although arguably, is present in all of the environments, if only as a counter-factual along the lines of "In the absence of sufficient resources, the protocell will die."

We consider death of the protocell to occur when the number of metabolites falls to zero. In this state, because all metabolites are autocatalytic, no new metabolites can be created. The system has reached a quiescent "zero-state" from which it cannot return and that we associate with protocell death.

5.3.2 *The operationally closed individual*

As discussed in Chapter 4, an operationally closed system consists of a network of interdependent processes. Each process enables another process within the network and also is dependent upon another process within the network.

One example of this relationship within the model is the mutual interdependence between the maintenance of the membrane above a certain size and the autocatalysis of Z. Let us, for the moment, ignore X, Y, R_1 and R_2. If, for some reason, the membrane is below a certain threshold, Z particles will be much more likely to interact with the membrane than to react with each other. Should this be the case for an extended period of time, no new Z particles would be produced and, in the absence of other reactions, the chemical would become extinct. The autocatalytic production of Z is therefore
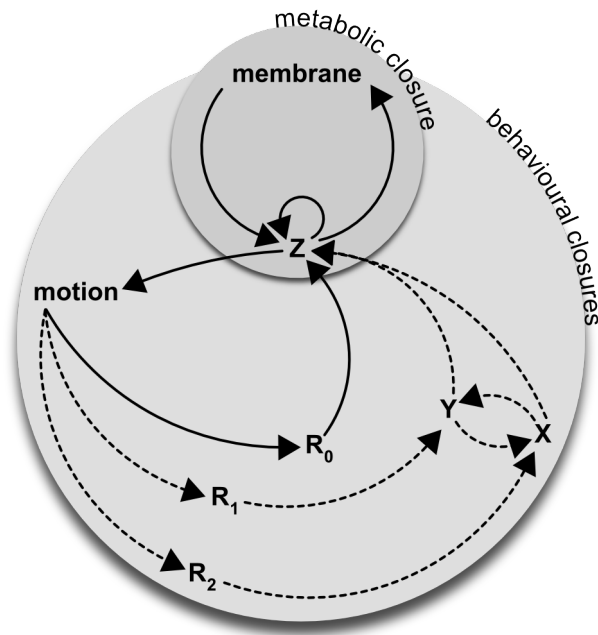
Figure 5.2: Cycles of interdependence in the model, indicating operational closure. Arrows represent dependence, e.g. the production of Z relies upon R_0. Dotted lines are used to distinguish between the R_0 metabolic pathway and the alternative pathway that is possible when R_1 and R_2 are available in the environment.

dependent upon the membrane being of a sufficiently large size. It is also the case that the maintenance of the size of the membrane is dependent upon the production of Z, for the only process that increases the (otherwise degrading / shrinking) membrane is the integration of Z into the membrane. Thus production of Z and membrane maintenance are reflexively interdependent (as indicated in Figure 5.2).

When we include X and Y and the other resources, we can see a larger operationally closed network of interdependence (see Figure 5.2). Unless in an environment that perpetually provides all necessary resources, all of the production reactions in the model rely upon Z production (if indirectly), because they all need resources, and Z is the only chemical that makes possible the further acquisition of new resources (whether it be by growth or movement). If resource R_0 is absent, Z cannot autocatalyze, and the only remaining possible production of Z is through reactions involving X and Y. So, in this situation, Z depends upon the existence of both of these chemicals, for in their absence no process would counter the processes whereby Z is lost.

It is interesting to note that the closure of some of these loops of interdependence is partially determined by the environment of the system. For instance, I mentioned that the system depends upon Z to move around to acquire various resources *unless in an environment that provides all those resources without needing movement*. This is one example of dynamical operational closure, an idea that we return to discuss in detail in Chapter 8.

5.3.3 *Self-production*

As already mentioned, metabolites X, Y, and Z and the membrane all degrade. These processes of degradation are countered by processes of production whereby resources are transformed into constituents. These processes are all made possible by the existence of the operationally closed individual. Therefore the system is self-producing in the sense described in Chapter 4.

If certain constituents were made by something else, then those constituents would not be part of the self-maintaining system, even if the system depended upon them. They would be *a resource* of the system rather than a component of the system, similar to R_1 and R_2.

5.3.4 *Strongly intrinsic behaviour*

The behaviour found in more typical evolutionary robotics models of action selection, minimal perception or cognition tends to be of a different kind than that found in the model described in this chapter. Section 4.3 described how the root cause of behaviour can be extrinsic, weak intrinsic, or strong intrinsic. The behaviour in this model is based upon a sensitivity of the system to its own viability and therefore falls into the category of 'strongly intrinsic' as described in the preceding chapter. To demonstrate this strongly intrinsic behaviour, we perform 4 experiments that are described in the next section. In Section 5.5.2 we return to contrast strongly intrinsic behaviour with the conventional genetic, extrinsic description of behaviour.

5.4 EXPERIMENTAL SCENARIOS

In this section, we expose the organisation and behavioural capabilities of the autopoietic, protocell-like agent through four experimental scenarios.

5.4.1 *E1: System viability*

The first, simplest scenario that we examine with our model is one in which the environment contains a fixed quantity of homogeneously distributed R_0. Inside this environment there is one protocell containing a number of Z metabolites. Z is auto-catalytic in the presence of R_0 (see Table 5.1). Also note that Z contributes phospholipids to the membrane and that this contribution is the only process that counteracts the continual degradation of the membrane. It follows that if R_0 is sufficiently high, the autocatalysis of Z will be sufficient to completely compensate for the degradation of the membrane. If not, the
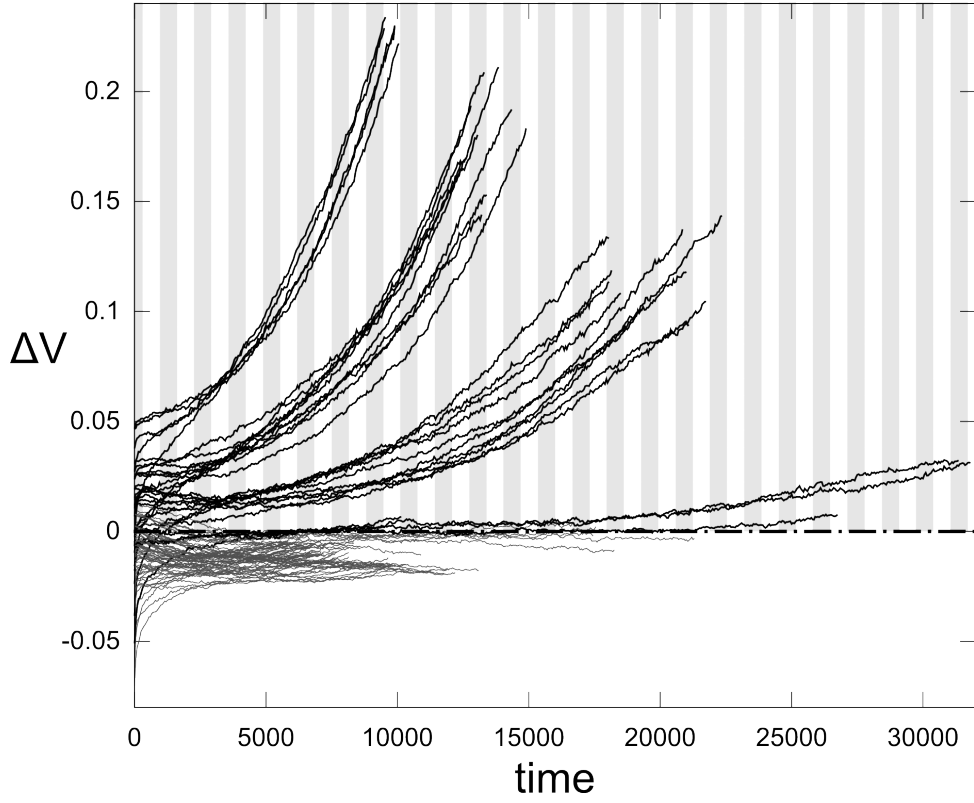
Figure 5.3: System viability in different uniform environments.

membrane will shrink until the cell dies[1]. Therefore, a good indicator of viability across different environmental situations is the rate of production of Z in relation to the rate at which the membrane degrades, $\Delta V \equiv d(Z/\rho)/dt$ (where Z is the number of Z metabolites and ρ is the number of phospholipids in the membrane). Furthermore, $\Delta V = 0$ is an interesting reference as protocells that maintain a negative ΔV for an extended period of time will die, unlike those that maintain a ΔV of 0 or greater. This value, ΔV can be thought of as an indicator of what would happen should the protocell remain in its current situation for a long time. Negative values indicated a propensity towards death and positive values indicate the opposite. Note that this viability condition of the system is not explicitly encoded (unlike classical approaches) but is rather a statistical measure of spatially distributed molecular processes.

Figure 5.3 depicts values of ΔV for agents in the fixed resource environment. To generate this plot, the simulation was initialised with protocells with different starting conditions (#Z = $\{50, 100, 150\}$, $\rho = \{8, 10, 12\}$, $R_0 = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$) and we plotted the mean trajectory of 25 runs in ΔV over time (data was also smoothed using a 250 iteration running-mean low-pass-filter). Thinner trajectories plotted in grey tended

---

1 The relationship between resource availability and membrane size is not as simple as it might first appear. A smaller membrane requires less Z-production to maintain its size, but also has non-linear effects upon the levels of resource that are available to the protocell.
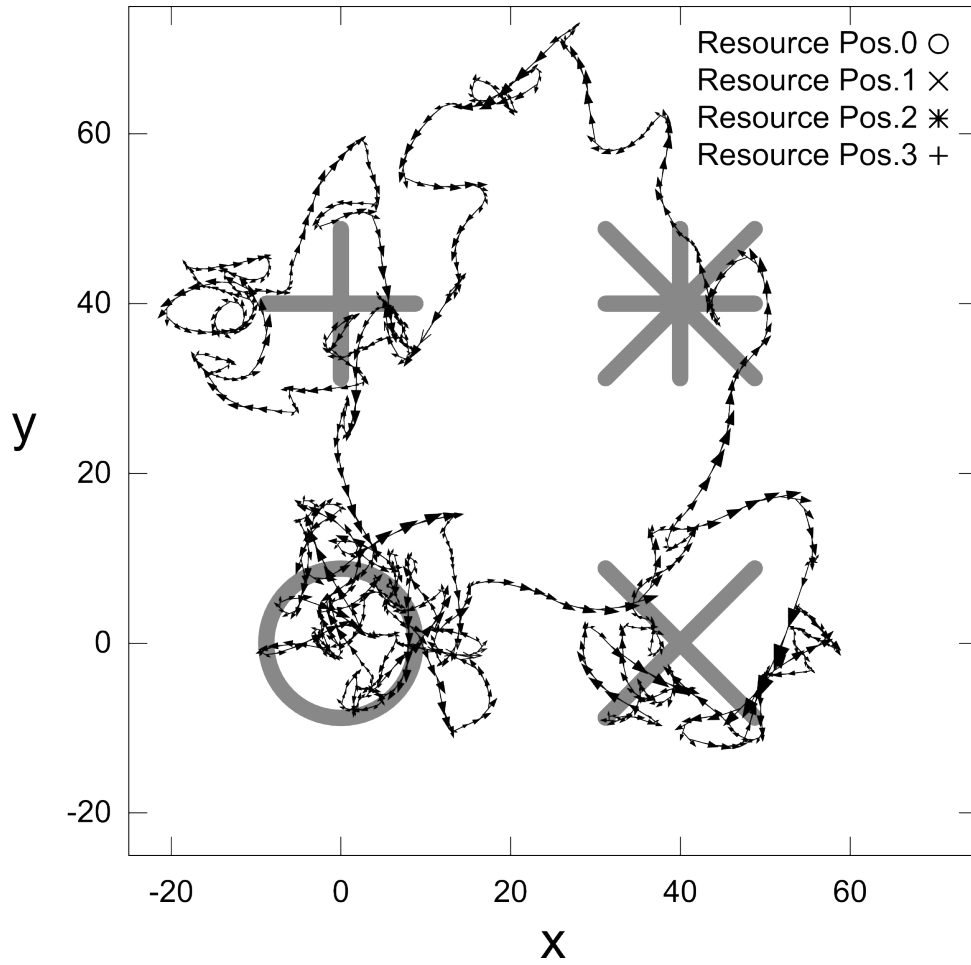
Figure 5.4: Experiment 2: The protocell's response to a moving resource generator.

to die. Note the 'viability boundary' located at $\Delta V = 0$, dividing those trajectories that tend to live (striped background) from those that tend to die.

5.4.2 *E2: Chemotaxis and its effect on viability*

For the second experiment we move into a more complex scenario where rather than having a fixed homogeneous concentration of R_0, we utilise a R_0 generator which rotates through four different locations, moving in discrete jumps once every 5000 iterations. Figure 5.4 shows the behaviour of the protocell, which performs chemotaxis towards the generator as it moves from one location to the next. This motion is the result of the asymmetrical distribution of R_0 within the protocell. The portion of the protocell that has a higher concentration of R_0 will produce more Z. Accordingly, more Z particles will collide with the membrane in this area of the protocell, inducing an overall up-gradient motion. Figure 5.6 describes graphically this process that underlies chemotaxis in this model.
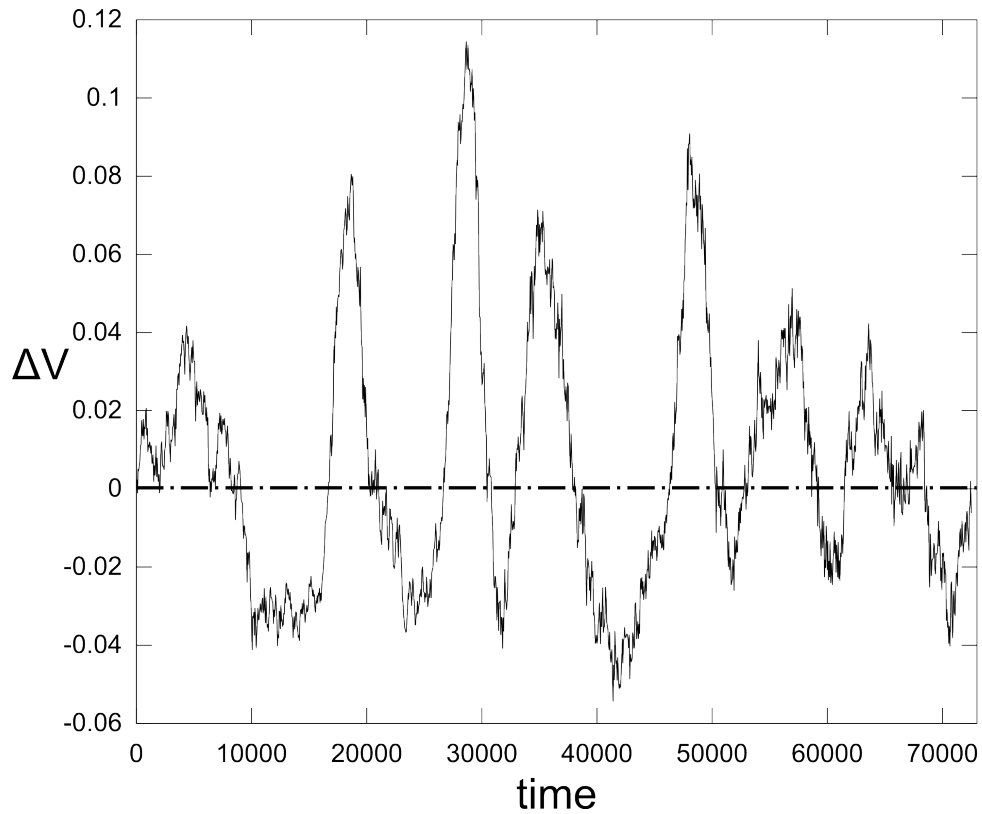
Figure 5.5: Experiment 2: Change in ΔV as the protocell responds to a moving resource generator

Figure 5.5 shows how ΔV oscillates above and below ($\Delta V = 0$). This plot indicates that the system is behaving adaptively; not in relation to an a priori and somewhat artificial parameter, but in relation to the very conditions upon which the system's ongoing survival depends. When the generator disappears, the ΔV becomes increasingly negative. This tendency is inverted by the system as it moves toward the next generator. The behaviour of the protocell is compensating for the negative tendency of ΔV (i.e., it changes the conditions such that the ΔV becomes positive again).

5.4.3   *E3: Oscillation between generators*

In our third experiment, we designed an environment in which the protocell cannot survive without behaving, forcing the simulated protocell into a continuous transient of viability. We accomplished this by simulating an environment absent of any R_0. Instead, two spatially separated stationary resource generators were substituted, one of R_1 and one of R_2. Observe that in the absence of R_0, Z can only be accomplished through the production of X and Y (requiring resources R_1 and R_2). Without any modification to the behavioural mechanism, in this environment, the simulated protocell automatically oscillates back and forth between both generators (see Figure 5.7).
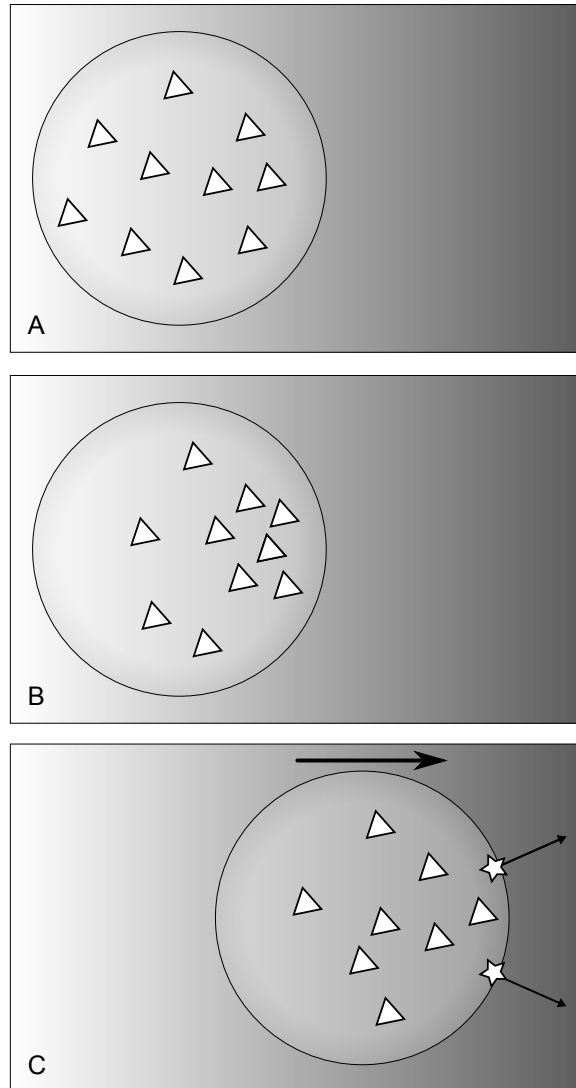
Figure 5.6: The mechanism of viability modulated motility. (A) An initial even distribution of Z becomes asymmetrical due to asymmetrical distribution of R_0 (B). The asymmetrical distribution of Z causes an asymmetrical distribution of contact between Z and the membrane, resulting in an up-gradient motion (C).
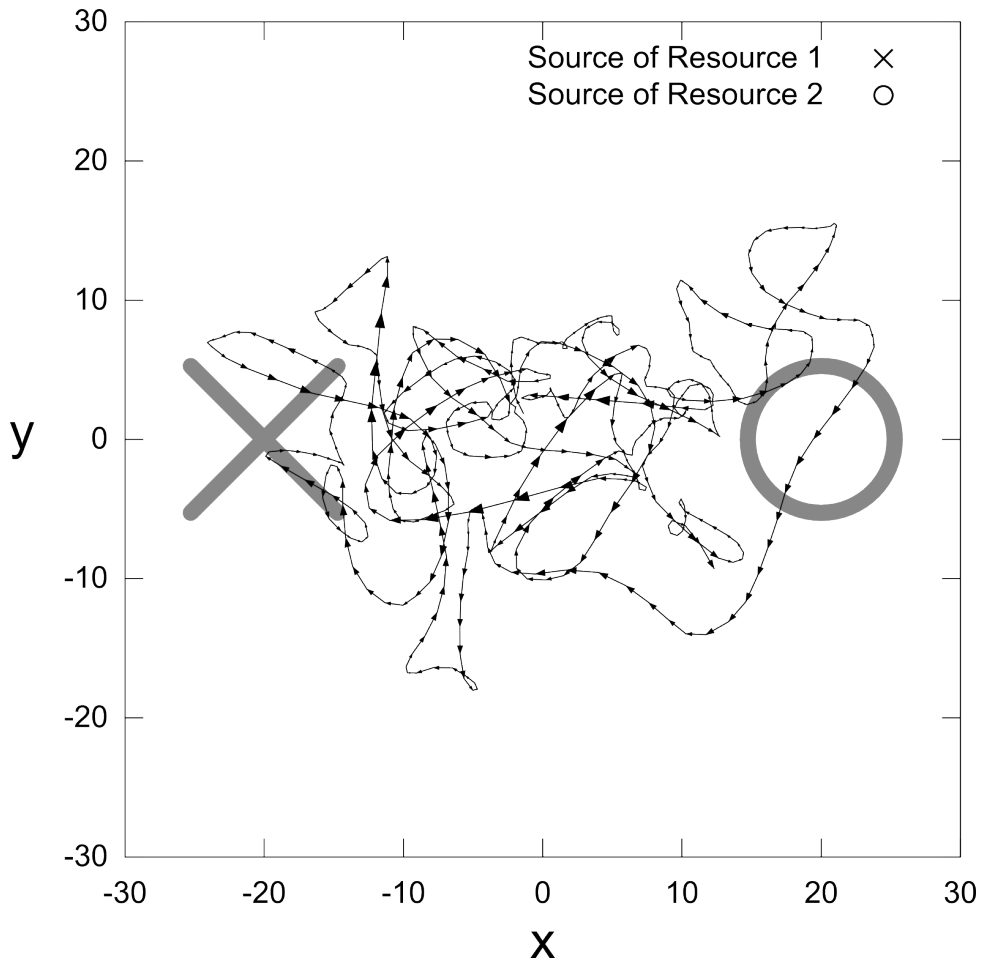
Figure 5.7: Path taken by the protocell in Experiment 3, where it is dependent upon two spatially separated resources.

Again the motion towards the relevant resource-source is produced primarily by the asymmetry within the cell of the production of Z. In this scenario, Z is only produced by an interaction between X and Y, for in the absence of R_0, metabolite Z is the product of only one reaction: $X + Y \rightarrow Z + Z$. Thus, if Z is to be produced, the protocell will require some of X and Y. Metabolites X and Y are reflexively autocatalytic, i.e., X catalyses the production of $R_1 \rightarrow Y$ and Y catalyses the production of $R_2 \rightarrow X$ (reactions 1 and 2 in Table 5.1). As generators of R_1 and R_2 are separated spatially, it is necessary for the cell to move back and forth between the two resources if it is to maintain non-zero concentrations of X and Y. Figure 5.7 indicates that the protocell does indeed move between the resources. How is this accomplished?

Imagine for a moment that the protocell is located at the generator of R_1 (Figure 5.9A). At this location, for a relatively healthy protocell, there would tend to be lots of Y throughout the cell, as R_1 is the resource that is transformed into Y. The rate of Z production by reaction 3 is proportional to the product of the concentrations of X and Y, so given a high concentration of Y, the limiting factor in the production of Z will be
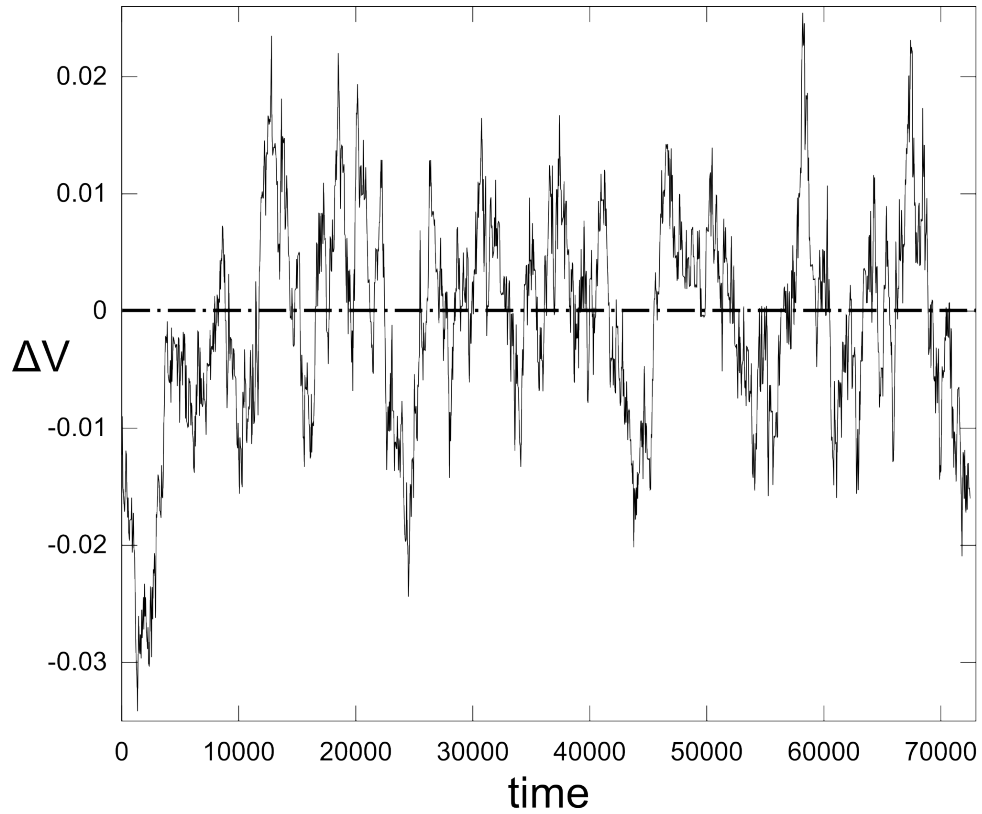
Previously published in Egbert et al. (2009).

Figure 5.8: Change in ΔV in Experiment 3, where it is dependent upon two spatially separated resources.

the concentration of X. In other words, a small amount more Y will make little to no difference in the rate of production of Z, but the same quantity more of metabolite X would make a substantial difference. In other words, in this situation, the protocell is in greater need of X to continue producing Z. It is also the case that in this situation, the protocell is more responsive to – i.e. more likely to climb – the R_2 gradient than the R_1 gradient. This is true because more of the limiting metabolite, X, will be produced in areas higher in concentration of R_2. In the part of the protocell where there is more R_2, there will be more X, and in this situation, where there will be more X, there will be more Z (Figure 5.9B). We have already seen in Experiment 1 and in Figure 5.6, that when Z is asymmetrically distributed within the protocell, it statistically produces a motion in the direction of the higher concentration of Z (Figure 5.9C).

As the cell moves up the R_2 gradient to its generator, the concentration of Y decreases and becomes the limiting factor in the production of Z. A symmetrical process causes the cell to move back towards the original resource generator (Figure 5.9D–F). These two processes result in the oscillation of the cell between the two resource generators. We can again observe how the system behaves adaptively in relation to viability (Figure 5.8): when the ΔV starts to decay, behavioural shifting towards the other generator inverts the tendency.
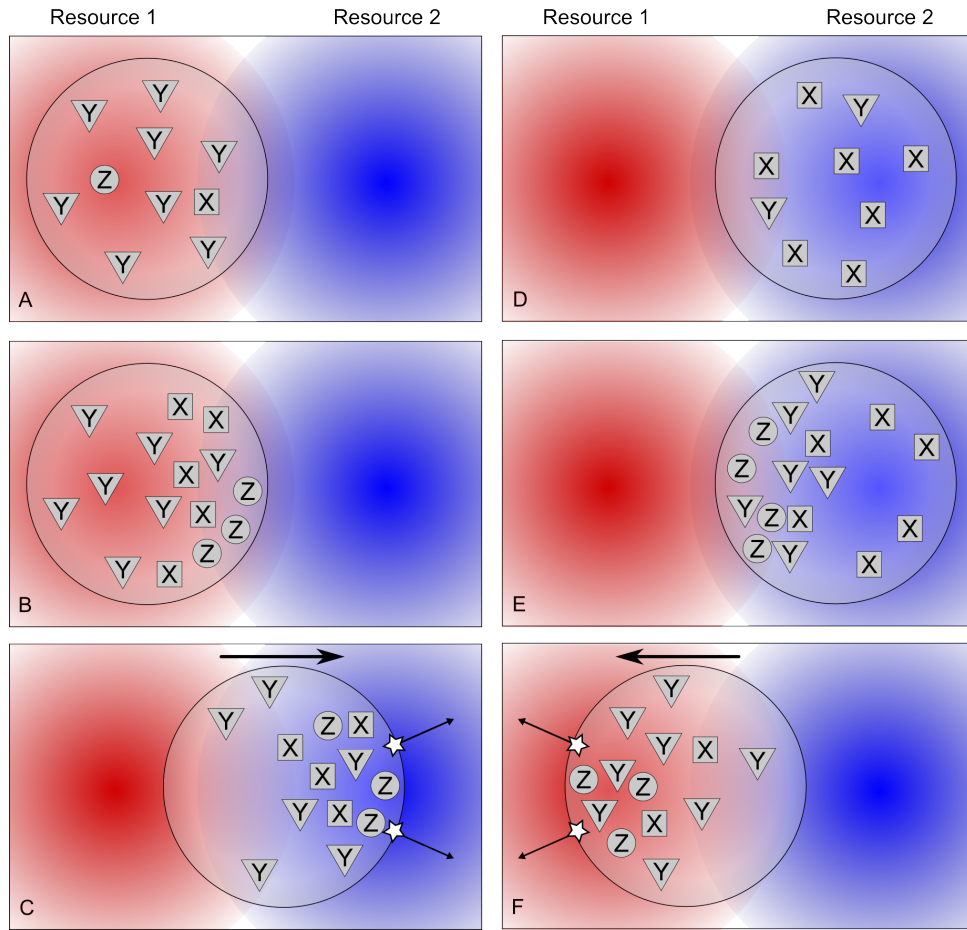
Figure 5.9: Diagram indicating how the protocell oscillates between resources.

5.4.4 *E4: Preference for superior generator*

At this start of this scenario, the two same generators of R_1 and R_2 are presented, as in Experiment 3 (see Figure 5.10). At iteration 5000, a generator of R_0 is added at location $(x = 0, y = 50)$, indicated by the $*$ in Figure 5.10. This generator is "superior" for the protocell in the sense that it suffices, on its own, for the direct production of Z. Soon after its introduction, the protocell moves towards the new resource generator. The protocell remains at the new R_0 resource for some time. Meanwhile, the original resources grow as they are not being consumed by the protocell. As the concentrations of R_1 and R_2 are increased, they diffuse into the environment. This takes place until there is an area between the two generators that is so high in concentration of R_1 and R_2 that it can sustain faster levels of production of Z than that supported by the R_0 generator. At this point, the protocell moves away from the source of R_0, back to in between the sources of R_1 and R_2! This last result was a surprise, and it acts as an excellent demonstration of how the protocell is responding not to a particular gradient, but rather to whatever best suits its metabolism.
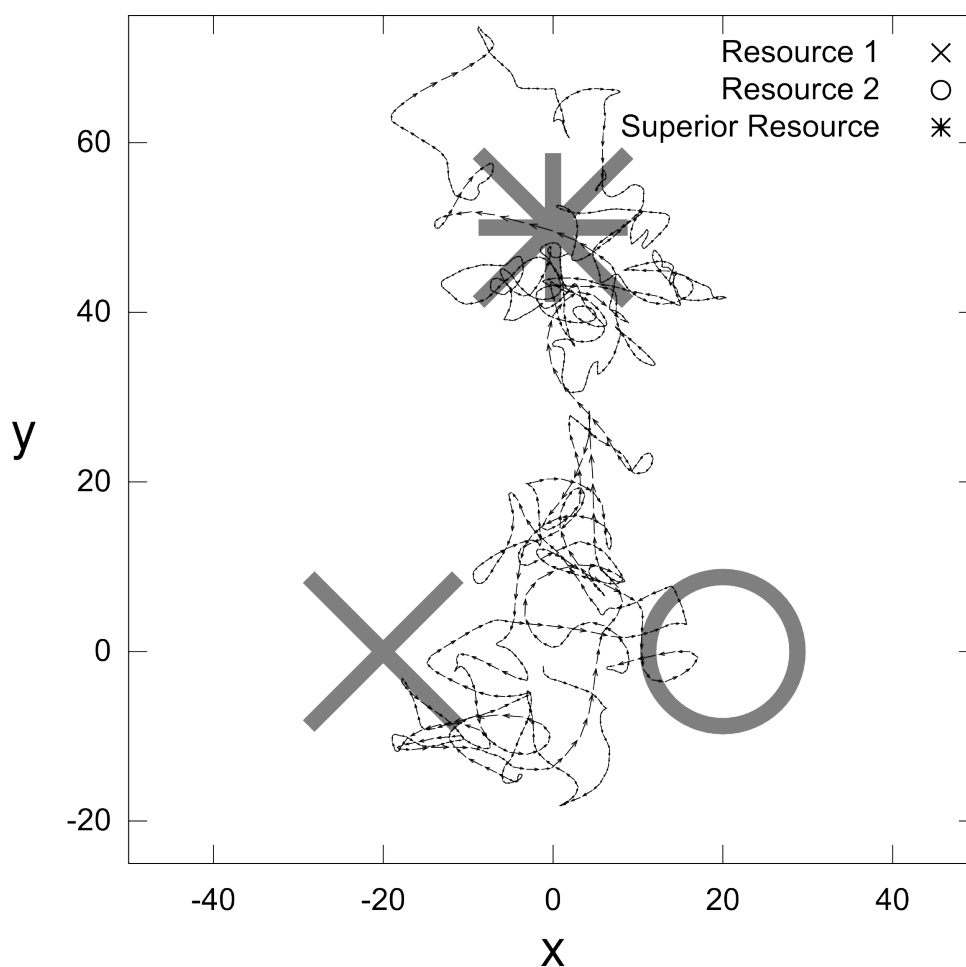
Previously published in Egbert et al. (2009).

Figure 5.10: Experiment 4, the protocell moves to utilise the most profitable resources, maximising its viability.

In this situation and in Experiment 3, the protocell is performing an action selection sensitive to the effects of the environment upon the emergent viability of the protocell. Its behaviour in the moment depends upon the *context of the situation* and the previous *history of the protocell's interactions with its environment*, as well as the present levels of reactants within the protocell.

5.5 DISCUSSION

The results described above show how the simulated protocell behaves in relation to the conditions of long-term viability, generating instances of chemotaxis and simple, but non-trivial examples of action selection. It accomplishes this without explicit sensors and without an explicit encoding of viability conditions as previous models of adaptive behaviour have assumed necessary. The adaptive nature of the behaviour is shown not only by the behavioural inversion of ΔV tendencies but also because the behaviour is not purely reactive nor stimulus driven, but is contextual, i.e. substantially influenced by the
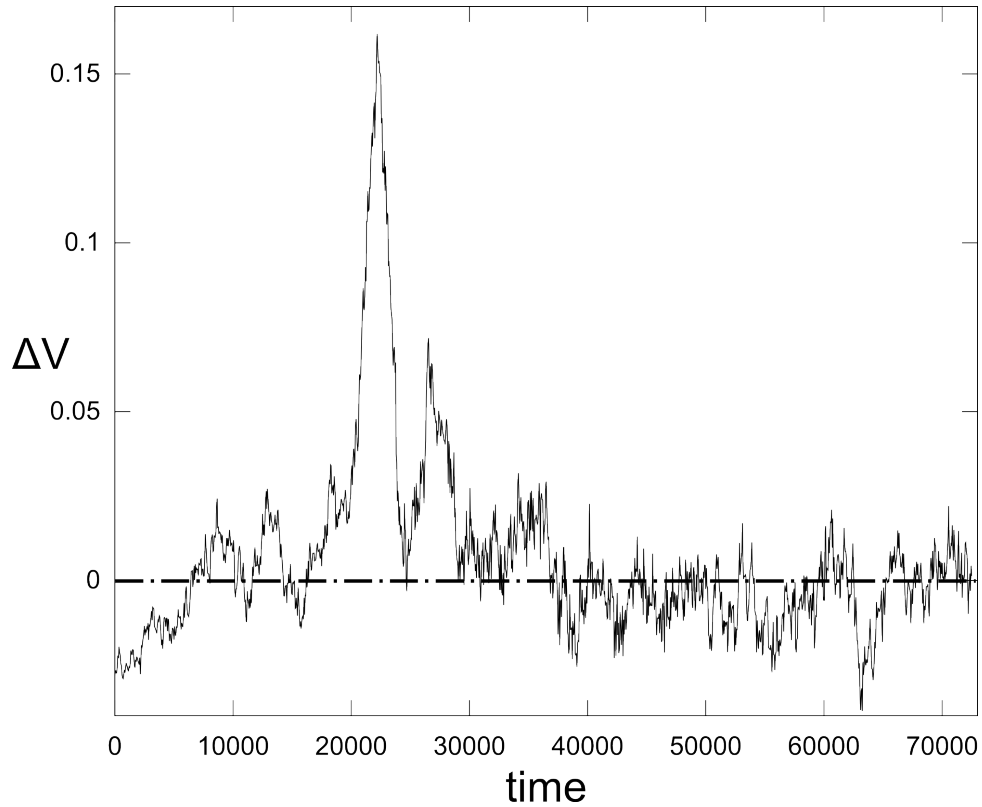
Figure 5.11: Change in ΔV in Experiment 4.

state and history of the protocell. As a case in point, Experiment 3 demonstrated that the protocell's behaviour will be different (it will climb different gradients) if it has just previously visited R_1 or R_2. The protocell can be interpreted as evaluating the value of its interactions with the environment with respect to their effect upon viability. It does not respond directly to environmental phenomena, but rather to the influence of local environmental conditions upon its viability.

One aspect of this behaviour is the protocell's tendency to perform whatever action maximises viability without performing superfluous actions. For instance, when placed in an Environment uniform in R_0 in Experiment 1, the protocell did not move around (results not shown). This 'lazy efficiency' is more like the behaviour observed in nature than the compulsive behaviour that is ignorant of the repercussions of the behaviour, often found in artificial robots. This connects with our motivation of building more biological behaviour and we shall return to this lazy efficiency in Section 6.4.2 where we recreate a form of lazy efficiency observed in bacteria.

In Experiment 1, we described the conditions under which the protocell is non-viable in the long term. Yet, in experiments 2–4, we see it move into those conditions and out of them in transients that are brief enough that the protocell does not die. Here we have a concrete example of a system that is moving towards a viability boundary (i.e., moving

towards death), but is sensitive to that trend and acts to invert it. This is a concrete example of adaptivity as defined by Di Paolo (2005, 2009).

Our model also exemplifies the holistic and dynamic (temporal) nature of viability. Consider Experiment 3, in which the protocell requires two resources that are spatially separate. The long-term tendency of an unmoving cell at any point in space is certain death – no location presents a sufficient level of combined resources. However, in this environment, the cell can survive if it moves. In this situation, the protocell's behaviour inverts the negative ΔV tendencies within the system and is, in fact, a necessary system property for its self maintenance. Thus, adaptive behaviour, typically conceived as something added on top of metabolism and that confers certain advantages to an already stable self-sustaining entity, turns out in this case to be an essential ingredient for the very conditions that keep the system alive. We conclude that we should remain open to seeing agency as implicated in metabolism and metabolism as implicated in agency. Not only does behaviour depend upon metabolism, but so too can metabolism depend upon behaviour. We return to discuss this idea in Chapter 8.

5.5.1 *Capabilities unique to self-producing systems*

Our model is atypical because it includes organismic behaviour and autopoiesis as well as a direct coupling between the two. The coupling comes from the dual role of metabolite Z which both contributes to the stability of the system by maintaining the membrane and imparts velocity to the cell in such a way that it moves towards the area producing the most Z. In having such an organisation, the behaviour of the simulated cell in our model is modulated by a primary dimension of its viability conditions. This direct relationship produces an agent capable of adaptive chemotaxis and an adaptive form of action-selection. In other words, the behaviour of the agent is fundamentally based upon an interaction that is missing in more conventional approaches.

Other approaches, such as evolutionary robotics (ER), have succeeded in identifying mechanisms of behaviour that respond mainly to environmental features. But how could e. g. the ER methodology produce mechanisms that are sensitive to viability such as that described in this paper? Certainly one requirement is for the agents to have viability requirements in the first place! This is not common, though see work on homeostatic adaptation e. g. Harvey et al., 2005. This is one advantage of modelling an autopoietic system in studying adaptive behaviour. Autopoietic systems are precarious in that they require ongoing maintenance to avoid death. In this model, viability is not simplified to just a scalar value (c.f. a battery level), but instead remains a complex, variable process that can interact with behaviour and the environment in a rich manner.

So, what can autopoietic systems do that allopoietic systems cannot? They produce themselves and can therefore modulate their behaviour according to how well they are self-maintaining. This answer is simple, but the implications, as we shall see, are intriguing. We have already seen in the model some first examples of how by being

sensitive to viability, an agent can modulate behaviour to perform simple forms of adaptation and action selection. The next Chapters build upon this, connecting the work to phenomena observed in nature and exposing other implications of viability modulated behaviour including how it may have played an integral role in the origins of life.

5.5.2 *A scalar value representing 'health' is not enough*

The excellent paper "Evolving Action Selection and Selective Attention Without Actions, Selection, or Attention" by Seth, 1998, describes a model that serves well as a demonstration of behaviour that would fall into the category of 'weak internalist'. In the model, two-wheeled, low-inertia agents are simulated in a 2D environment that contains two resources (food and water) and deadly 'traps'. Included in the model are two batteries, the 'food-battery' and the 'water-battery'. Each battery level steadily decreases, but is increased when the agent comes into contact with the associated resource. If one of the battery levels falls to 0, or if the agent touches a trap, it dies. These agents are subjected to an evolutionary algorithm (Nolfi and Floreano, 2004; Harvey et al., 2005) which rewards 'fit' agents. Fitness is proportional to the accumulated sum of the battery levels over lifetime of the agent or the maximum trial duration, whichever is shorter. The genetic algorithm optimises parameters that specify (i) basic mapping of of sensor-input to motor activation and (ii) how this mapping is modulated by the battery levels. The evolutionary algorithm produces agents that demonstrate several properties associated with cognition: action-selection, "a balance between persistence and dithering", "interrupting current behaviour if necessary", opportunism, prioritisation, etc. Seth successfully demonstrates that these phenomena do not require discrete symbols or actions, or 'modules' for selective-attention or action-selection. In this respect it is a successful model. We should not think, however, that because the agents in this model are sensitive to "their" battery level, that their behaviour is intrinsic. The battery levels in the model don't belong to the agent in any strong sense. We only call them "the agent's batteries", but they could, for all intents and purposes, be a quantity external to the agent. Furthermore, the behaviour is determined by the evolution and the fitness function rather than being fundamentally linked to the organisation of the individual. If for instance, we placed a negative sign in front of the fitness function, the agents would seek out traps and avoid resources. In contrast, "A real animal [...] can be trained to do lots of things, but never to treat a punishment as a reward." (Di Paolo, 2010, p. 144)

Some might argue that just as the fitness function determines the behaviour of the agents in Seth's model, it is natural evolution that produces our behaviour. The relationship between evolution and behaviour is not as direct, nor as simple as this. Evolution plays a role in determining the substrate, over which cognition operates. The body, including nervous tissue is indeed partially determined by genetic influence. But I am not writing this thesis because natural selection has produced thesis-writers. I do not prefer Asian food because I have a set of nucleotides in a certain order in every one of

my cells. To base cognition in the genes is to remove lifetime processes from cognition. Cognition is an online, during-the-lifetime dynamic that changes over time-spans far too short to be part of genetic evolution. Our behaviour does not primarily answer to genetic evolution. If that were the case, there would be no suicide, no voluntary celibacy and no homosexuality – for all of these behaviours are strongly disadvantageous with respect to the passing on of genes.

The model described in this chapter is an example of a different organisation of strongly intrinsic behaviour. It shows how even simple systems can act on their own behalf, demonstrating cognitive acts such as action-selection and value-judgement by being sensitive to the repercussions of the environment upon their well-being. This organisation allows for online evaluation of phenomena never before experienced by the individual (nor by any of its ancestors). It also better describes the behaviour of living organisms, from single celled bacteria, to plants, to animals, these systems seem to share this organisation and the benefits it provides. This is not simply a negative statement that evolution does not explain everything, but also contributes the positive assertion of a new organisation that provides new capabilities to organisms.

It is worth noting that the protocell is not sensitive in a holistic sense to its viability. It responds, more specifically, to dynamics related to metabolite Z. Here Z is an example of how simple, low level phenomena can indicate general, emergent system-health and that can accordingly be used to influence the system to adapt in general to novel environments. There are a great variety of "system health indicators" which organisms could respond to. Some of these are more or less directly coupled to viability (a topic we return to in Chapter 8). Most importantly it should be noted that Z is not weakly intrinsic (as the battery level in various behavioural robotics models) but is more fundamentally part of the system producing the behaviour, for without it, there is no system to behave and there is no behaviour.

5.6 RECAPITULATION

In this chapter we have described a model of an autopoietic system that demonstrates a behaviour that is modulated by the rate of one of the processes that contributes to its autopoiesis. The model acts as a didactic example for concepts introduced in Chapter 4. In addition to demonstrating the idea of a dissipative, precarious, self-maintaining system – it shows how such a system, by being sensitive to changes in its viability, can adapt. This adaptation is shown to be context sensitive and influenced by the history of environmental interactions of the agent. It also produces behaviour that operates in a life-like, 'lazily-efficient' manner, performing tasks that maximise viability while avoiding superfluous tasks. The basis of this adaptability stems from the adaptive mechanism operating in response to the value of environmental phenomena with respect to self-production rather than directly to the environmental phenomena themselves.

PRIMARY POINTS OF CHAPTER

- An abstract computational model is presented that demonstrates ideas introduced in Chapter 4, in particular: an operationally closed, precarious, self-maintaining dissipative organisation performing viability-sensitive behaviour.

- An indicator of system viability is examined that compares the rate of self-production to the rate of degradation.

- The simulated agent performs viability modulated behaviour, a *strongly intrinsic* behavioural mechanism. This enables it to respond to the influence of local environmental conditions upon its viability rather than directly to the environmental phenomena.

- The model demonstrates adaptability provided by viability modulated behavioural mechanism, including context sensitivity, sensitivity to history of environmental interactions and a life-like 'lazy efficiency' in which the organism doesn't compulsively perform behaviour, but only does so when it benefits its viability.

6

CONNECTING WITH EMPIRICAL BIOLOGY: METABOLISM & CHEMOTAXIS

CHAPTER SUMMARY

In this chapter we connect the concepts explored abstractly in previous chapters to concrete biological phenomena. We explore the relationship between behaviour and autopoiesis by examining the relationship between bacterial chemotaxis (chemical seeking behaviour) and metabolism. The chapter starts with a historical review of how the role of metabolism as perceived by bacteriologists has changed over the past half century, distinguishing between three different possible relationships between metabolism and chemotaxis: 'metabolism-independent', 'metabolism-dependent', and 'metabolism-based' and outline relevant empirical studies of bacteria. We then present the first, minimal model of metabolism-based chemotaxis. The development of this model required a formulation of a 'minimal metabolism' which is described in Section 6.3.1. The model recreates experimental findings of various bacterial chemotaxis and may provide explanations for previously poorly understood observed phenomena. The results and their implications are discussed in the final sections of the chapter.

6.1 INTRODUCTION

John Maynard Smith, having conceived the theoretical notion of an evolutionarily stable strategy (ESS) had a difficult time convincing empirical biologists of its relevance. Only

after he could explain the idea in the context of an established biological phenomenon, was the idea taken seriously, and eventually accepted as one of the principles of evolutionary biology[1]. The work presented in this chapter was motivated by a desire to make a similar connection between the theory presented elsewhere in this thesis and empirical biology. This chapter was first published as an article in *PLoS Computational Biology* entitled 'A Minimal Model of Metabolism Based Chemotaxis' (Egbert et al., 2010b).

6.2 A REVIEW OF METABOLISM'S INFLUENCE UPON CHEMOTAXIS

Bacterial chemotaxis is one of the best known examples of adaptive unicellular motility. In particular, the mechanisms underlying chemotaxis in *Escherichia coli* have been studied in detail for the last 40 years (for recent, comprehensive reviews see e. g., (Eisenbach, 2004, 2007; Wadhams and Armitage, 2004)). Since the work of pioneers such as Adler (Adler, 1969; Kresge et al., 2006), Berg (Berg and Brown, 1972), Macnab (Macnab and Koshland, 1972), and Spudich (Spudich and Koshland, 1975), considerable advances continue to be made concerning the molecular structure of motors (Stock et al., 1999; Sowa and Berry, 2008), the structure of transmembrane receptors and their collective dynamics (Stock and Levit, 2000; Hazelbauer and Lai, 2010) and the details of a two component signal transduction system (Stock et al., 2000) that mediates between sensors and motors (Falke et al., 1997). Computer simulations of the underlying biochemical processes have helped to support and clarify the current model of chemotaxis mechanisms (Andrews and Bray, 2004; Bray et al., 2007).

In this paper we explore, by means of minimal simulation models, the widespread assumption that the mechanisms of bacterial chemotaxis operate independently of metabolism (Adler, 1969). In this prevailing 'metabolism-independent' view, the behaviour generating mechanisms such as sensors, transduction pathways, flagella, etc., are the product of metabolism, but their ongoing, short-term activity is not subsequently influenced by metabolism. In other words, in the short term, behaviour is not sensitive to changes in the metabolism.

In contrast to this view is metabolism-dependent chemotaxis, where the metabolism has an ongoing influence upon behaviour. The concept dates back at least as far as 1953 (Clayton, 1953), but fell out of favour when Adler demonstrated that in *E. coli*, metabolism of a reactant is neither necessary nor sufficient for taxis (Adler, 1969). Interest was rekindled in a 1983 review of the role of the proton motive force in taxis mechanisms (Taylor, 1983) and there is growing evidence for metabolism-sensitive chemotaxis in *Azospirillum brasilense* (Alexandre et al., 2000), *E. coli* (Taylor and Zhulin, 1998), and other bacteria (Alexandre and Zhulin, 2001), suggesting that *metabolism-dependent* chemotaxis might be more prevalent than previously assumed (see (Alexandre, 2010) for a recent review of metabolism-dependent 'energy taxis').

---

1 Inman Harvey (personal communication)

In this paper, we clarify the distinction between the different relationships between metabolism, chemotaxis and its generative mechanisms and we demonstrate how a metabolism-based chemotaxis mechanism is capable of generating several phenomena observed in bacteria. Our model demonstrates the substantial adaptability provided by the simple metabolism-based mechanism in the form of an ongoing, contextualised and integrative evaluation of the environment. We conclude by discussing this adaptability, the possibility of fumarate playing a role in metabolism-based chemotaxis in bacteria, and some consequences of relaxing the metabolism-independent assumption.

To avoid misunderstanding, we shall clarify two different usages of the term "adaptive" or "adaptation" in this paper. The first usage is that of "organismic or physiological adaptation", meaning the capacity of an organism to homeostatically maintain essential variables (e. g., temperature, pH level, etc.) within viability boundaries, or to maximise or minimise their value (e. g., maximise the amount of available food or minimise exposure to a toxin). Regulated motility is a widespread means for achieving this type of adaptation. For instance, an organism can maintain a stable level of sucrose by moving to sucrose rich environments or moving away from them, or maximise the amount of light by moving to brighter areas, etc. This meaning of adaptivity is well established in biology and complex systems, (see e. g., (Ashby, 1952)). The second sense of adaptation is that of "sensory adaptation", used specifically by the bacterial chemotaxis research community to mean the capacity of transmembrane receptors to maintain the same degree of sensory sensitivity in an extremely wide range of base-stimulus levels (Berg and Tedesco, 1975). Bacterial chemotaxis can be adaptive (in the first, physiological / organismic meaning of adaptation) without any sensory adaptation (second meaning) taking place, provided that mechanisms other than sensory adaptation are capable of guiding behaviour efficiently. So, what matters for a behaviour to be adaptive or conducive to the stability (or maximisation of a given variable) is not the type of behaviour generating mechanisms (e. g., transmembrane signal transduction proteins), nor the dynamics of sub-components of such mechanisms, but just the global resulting pattern of behaviour. Unless stated otherwise, in this paper the word adaptation is used in the first, organismic sense.

We wish to include one additional terminological clarification: we consider a behaviour to be *chemotactic* if there is, in general, an effective overall behaviour that results in an approach towards specific chemical environments, irrespective of the mechanism that generates the behaviour.

6.2.1 *Chemotaxis through running and tumbling*

Behavioural analysis of chemotaxis in *E. coli* has shown that up-gradient or down-gradient directional movement is achieved through the combination of two basic types of movements, tumbling and running (see Figure 6.1). These two behaviours are both achieved through the rotation of flagella. Rotating the flagella in one direction (counter-
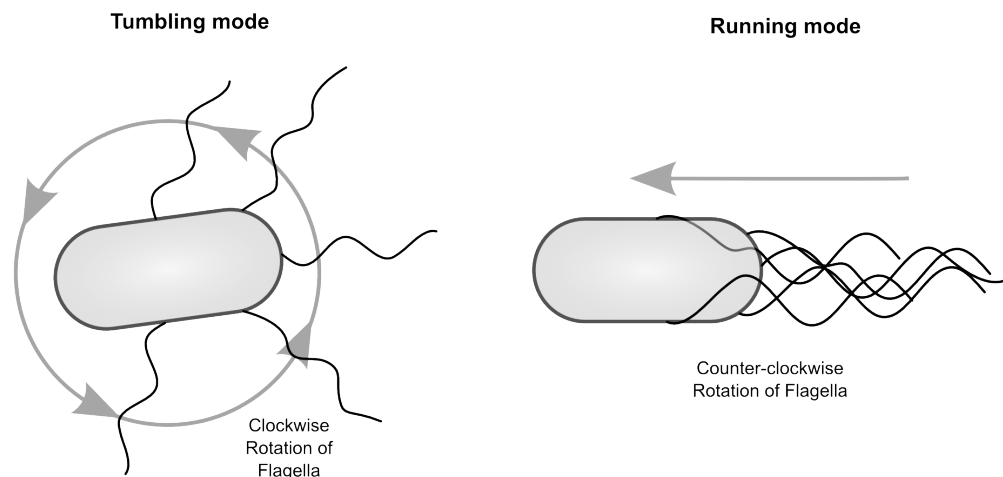
Tumbling mode

Running mode

Counter-clockwise
Rotation of Flagella

Clockwise
Rotation of
Flagella

Figure 6.1: Tumbling and running modes associated with *E. coli* and *S. typhimurium*. CW rotation results in a random re-orientation for the bacterium, but CCW rotation of flagellar motors produces a directed motion.

clockwise) results in a directed motion of the bacterium called 'running' while brief periods of rotation in the other direction cause 'tumbling', the production of a more-or-less random new orientation. Many swimming bacteria make use of similar patterns, alternating between a straight motion mode and a random change of direction (Armitage and Schmitt, 1997). Bacteria combine these movements in such a way as to produce a stochastic chemotactic behaviour. How do they accomplish this? In some species this is relatively well understood. For instance, in *E. coli*, a two component signal transduction system (Stock et al., 2000) compares current and past concentrations of attractants and causes the bacterium to run if the concentration of an attractant is increasing and tumble otherwise (see e.g. (Falke et al., 1997)). This general strategy is sometimes known as 'adaptive gradient climbing' because it is capable of adapting to a wide range of concentrations, climbing gradients whether the local concentration is very low or very high.

This is not the only way that these basic-movements can be modulated to produce chemotaxis. What we have called the *selective-stopping* strategy (also called 'inverted response' elsewhere (Goldstein and Soyer, 2008)) consists of a combination of running and tumbling to perform a random walk until the relevant concentrations are high at which point the tumbling motion dominates and the bacterium more or less tumbles in-place.

6.2.2 *Metabolism-independent chemotaxis*

Regardless of which chemotactic strategy is employed, it is important to identify its *sensitivity*—what is it responding *to*? The commonly accepted view is that chemotaxis
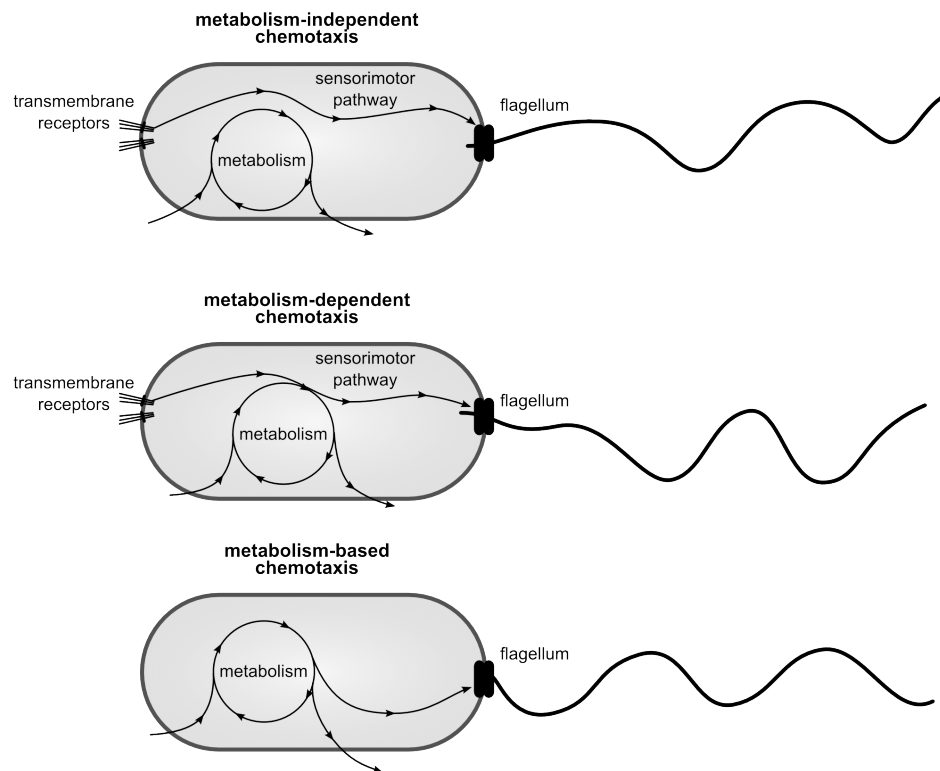
Figure 6.2: Differences between metabolism-dependent, metabolism-independent and metabolism-based chemotaxis. Arrows indicate *only* short-term dynamical influence between processes.

mechanisms are only responsive to the concentration of attractant chemicals in the local environment of the bacterium and are not influenced by the current state of the bacterium's metabolism. This view is known as metabolism-independent chemotaxis, for while the metabolism produces the mechanisms of sensitivity, transduction and response, the behaviour of the bacterium is not influenced by the ongoing dynamics of the metabolism (see Figure 6.2 top). In metabolism-independent chemotaxis, it is not the effect of the attractant upon the metabolism that causes it to move towards it, it is simply the way the attractant excites the sensors. Adler, in his seminal 1969 paper, showed seemingly compelling evidence for this view of bacterial chemotaxis, providing evidence in support of the following (taken from (Adler, 1969)).

1. Some chemicals that are extensively metabolized fail to attract bacteria.

2. Non-metabolizable chemicals act as attractants.

3. Chemicals attract bacteria even in the presence of a metabolizable chemical.

4. Compounds that are closely related in structure compete with each other as attractants but not with structurally unrelated compounds.

5. There are mutants which fail to carry out chemotaxis to certain attractants but are still able to metabolize them.

"[R]esults show" Adler concluded "that extensive metabolism of the attractants is not required, or sufficient, for chemotaxis" (Adler, 1969, p. 1596). After this evidence was presented the metabolism-independent nature of *E. coli* chemotaxis became a generally accepted fact. As Alexandre and Zhulin note: "From that time on, research focused on the metabolism-independent information flow from membrane receptors to flagellar motors." (Alexandre and Zhulin, 2001, p. 4681). But this assumption has recently been challenged and the consequences for the study of chemotaxis are yet to be fully disclosed.

6.2.3 *Evidence for metabolism-dependent chemotaxis*

Despite the predominance of metabolism-independent research on bacterial chemotaxis, there is evidence of different types of metabolism mediated chemotaxis. Metabolism-*dependent* chemotaxis involves an ongoing influence of the metabolism upon the chemotaxis mechanism (see Figure 6.2 middle). This introduces the potential for a sensitivity to the effects of environmental phenomena upon the metabolism. In *A. brasilense*, metabolism-dependent chemotaxis has been systematically studied and is considered as the dominant behavioural strategy (Alexandre et al., 2000; Alexandre and Zhulin, 2001). The following behavioural phenomena have been well-established (Alexandre and Zhulin, 2001):

1. nonmetabolizable analogues of metabolizable attractants are not attractants,

2. inhibition of the metabolism of a chemical attractant completely abolishes chemotaxis to and only to this attractant, and

3. the presence of another metabolizable chemical (exogenous or endogenous) prevents chemotaxis to all attractants studied (in *A. brasilense*, there is a direct correlation between the efficiency of a chemical as a growth substrate and as a chemoeffector).

Similar metabolism-dependent chemotactic phenomena have also been found in other bacteria species like *Pseudomonas putida* (Sarand et al., 2008) and *Rhodobacter sphaeroides* (Jeziore-Sassoon et al., 1998; Greer-Phillips et al., 2003).

Evidence of metabolism-dependent chemotaxis has also been found in the same species that was studied by Adler, *E. coli* (Taylor and Zhulin, 1998). Adler found that although glycerol is extensively metabolised by *E. coli*, it is *not* found to act as an attractant. As such, the case of glycerol seemed to provide supporting evidence in favour of metabolism-independent chemotaxis, and so argued Adler. Interestingly, in apparent contradiction to Adler's findings, Zhulin and collaborators observed chemotaxis towards similar levels of glycerol as studied by Adler (Zhulin et al., 1997, p. 3199).

Evidence of metabolism-dependent chemotaxis in *E. coli* has also been found for proline and succinate metabolic substrates (Alexandre and Zhulin, 2001). In addition, chemotaxis to oxygen (aerotaxis) in *E. coli* (Taylor and Zhulin, 1998) has been shown to depend on metabolism (i.e., reduction of oxygen is required for aerotaxis) together with redox gradient climbing (Edwards et al., 2006). This type of taxis has been termed "energy taxis" meaning that bacterial movement is sensitive to the energy production (generally by the electron transport system having modulatory effects over CheA phosphorilation, but other mechanisms have also been proposed) (for recent reviews of energy taxis, see e. g., (Schweinitzer and Josenhans, 2010; Alexandre, 2010)).

The results described in this section all suggest that metabolism-dependent chemotaxis might be more widespread than previously thought.

6.2.4 *Simulation models of bacterial chemotaxis*

The biochemical organisation of bacteria is astonishingly complex and difficult to model. One approach is to assume some sort of functional decomposition and to study sub-systems separately. Dennis Bray and his group have achieved some impressive results along these lines (Andrews and Bray, 2004; Bray et al., 2007; Bray, 2009). Their model of *E. coli*'s chemotactic behaviour includes molecular level details of the membranes, signal transducers, Brownian motion of molecules within the cytoplasm and a number of genetic details. Their approach has achieved unprecedented levels of predictability and empirical accuracy. However, despite the promising results, some aspects of *E. coli* chemotaxis remain elusive. For instance, how is it possible for a small number of types of sensor (5) to cause an appropriate response to a large number of attractants and repellents ($\sim$50)? It is possible that the elusiveness of these and other aspects of bacterial chemotaxis may not be a question of lack of mechanistic detail in the simulations themselves but may relate to the long-standing assumption that chemotaxis is fundamentally a metabolism-independent phenomenon.

Using a different approach, Goldstein and Soyer artificially evolved metabolism-independent chemotactic pathways (abstracting away the sensory and motor details) (Goldstein and Soyer, 2008). Chemical pathways partially reproducing a gradient climbing response could only evolve under special conditions and for a limited range of basal levels of stimuli. Contrarily, they found that the selective-stopping strategy (tumble in place if resources are high, otherwise perform a random walk) was comparatively easy to evolve. The resulting selective-stopping strategy is simpler, yet robust and efficient. This strategy is also found in different types of bacteria particularly in those performing metabolism-dependent chemotaxis—like that of *Rhodobacter sphaeroides* (Packer and Armitage, 2000)—and in *E. coli* when "normal" transduction pathways are knocked down (Barak and Eisenbach, 1999). Goldstein and Soyer found very simple chemical pathways capable of generating such response patterns. They speculate that "non-adaptive dynamics [i.e. the selective-stopping strategy] could even be achieved without any signaling proteins;

a small molecule, that is a by-product of metabolism or is taken into the cell via a transporter, could directly regulate tumbling probability of the cell", (Goldstein and Soyer, 2008, p. 5).

In order to further develop this hypothesis, we start from an idealised minimal metabolism and make it support and generate chemotactic behaviour. Our model provides a proof of concept of how a minimal metabolism could already support experimentally observed metabolism-dependent chemotactic patterns. The model displays the response patterns shown by Goldstein and Soyer and, in addition, it reproduces some of the metabolism-dependent phenomena described by Alexandre and Zhulin (2001).

6.2.5 *Summary of key concepts*

To summarise, *metabolism-independent* chemotaxis refers to types of chemotaxis where the chemotaxis generating mechanisms operate rather independently from metabolism. Flagellar rotation is only influenced by chemical pathways that are independent from the metabolic network and only modulated by transmembrane receptor activity (see Figure 6.2 top). In *metabolism-dependent* chemotaxis, the chemical pathway that mediates transmembrane receptors and flagellar rotation is influenced by or coupled with metabolic pathways and processes (like the electron transport system) with the result of a metabolism-sensitive behaviour (see Figure 6.2 middle). A third possibility is *metabolism-based* chemotaxis, in which metabolism itself directly modulates behaviour. In this case, there are neither specialised sensors (like transmembrane proteins) nor specialised and dynamically decoupled chemical pathways (see Figure 6.2 bottom).

Most of the available simulation models of chemotaxis are of metabolism-independent scenarios, there are no models of metabolism-dependent or metabolism-based chemotaxis. In the next section, we introduce a first model of metabolism-based chemotaxis to study its chemotactic potential and some theoretical implications. The likelihood of the specific mechanisms in play and experimental support remain out of the scope of this paper.

6.3 METHODS

6.3.1 *A minimal model of metabolism-based behaviour*

The first step to create a minimal metabolism-based chemotactic agent is to distill and justify what counts as a minimal model of metabolism. We do not pretend to settle this issue here but we provide a general context that justifies the assumptions built into the model.

It is generally accepted that life depends upon energetic and material resources in its ongoing self-(re)-production, self-maintenance and growth. Energy and matter flow through living systems, maintaining their biophysical and chemo-dynamic organisation.
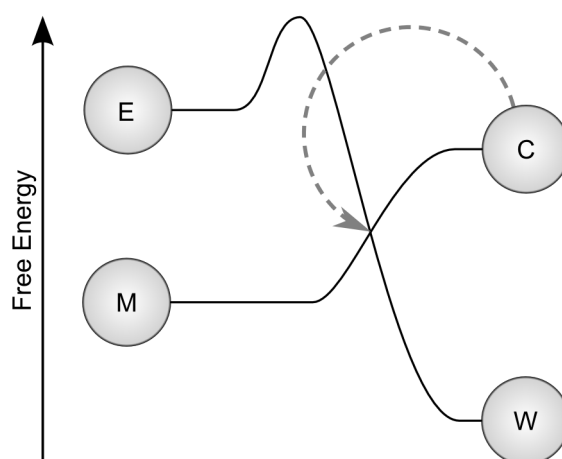
Figure 6.3: Conceptualisation of a minimal metabolism. An exergonic reaction (E → W) is coupled to a material transforming reaction (M → C) such that system components are produced (C) that catalyse those reactions.

These two flows are coupled: energy is used to transform materials into structures that harness energy to perform further material transformations (Kauffman, 2000; Morowitz and Smith, 2007; Rosen, 1991). In concrete terms, the energetic flow through the system depends upon the existence of catalysts which in turn depends upon the flow of energy through the system. This autocatalytic closure of chemical reactions has long been argued to constitute the core or essence of living organisation, exemplified in metabolism (Maturana and Varela, 1980; Kauffman, 1986; Gánti, 2003b; Rosen, 1991)—see (Ruiz-Mirazo et al., 2004; Letelier et al., 2006) for recent reviews and, more generally, see (Barandiaran and Moreno, 2008; Barandiaran et al., 2009) for the central role played by metabolism in grounding biological agency and adaptivity. This metabolic organisation stands far from thermodynamic equilibrium. Energy and matter are lost as heat and waste, requiring the continued acquisition of new resources.

Abstracting away from the particularities of different metabolic networks, the following three key features remain essential to characterise metabolism:

1. A flow of matter through the system.

2. A flow of energy through the system.

3. A dissipative (or degrading) organisation that in the prolonged absence of sufficient resources (energetic or material) ceases to exist.

One of the simplest systems that has these features consists of the following autocatalytic reaction.

$$M + E \xrightarrow{C} C + W$$
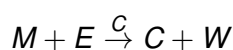
where M and E represent material and energetic resources respectively, C is a constituent or catalyst molecule and W is low-energy waste. The C above the arrow represents catalysis of this reaction by C. This system is illustrated in Figure 6.3, which also indicates the relative free energies of the reactants on the vertical axis. One can conceive of this reaction, according to the free-energies, as an exergonic reaction (in which energy is released as $E \rightarrow W$) that is coupled to (and drives) the endergonic reaction of $M \rightarrow C$. Real metabolisms, of course, have many intermediate steps in the production of enzymes which complicate the system. We assume that such intermediate steps can be justifiably abstracted away to illustrate a minimal instance of metabolism-based chemotaxis, taking the equation above to represent the higher order chemical dynamics of a whole metabolic network.

The $M + E \xrightarrow{C} C + W$ autocatalysis and modulation of flagellar rotation by the concentration of C is indeed a simple mechanism. In theory, this simplicity could be taken even further. For instance, the energetic and material resources could come from a single molecule, $R \xrightarrow{C} C + W$. We chose to keep these resources separate so that we could investigate issues related to integration and to explore more complex environments (see Experiments 2–6). For those experiments where M and E are placed in the same location, the reaction $M + E \xrightarrow{C} C + W$ plays a qualitatively equivalent roles to the simpler reaction $R \xrightarrow{C} C + W$, and behavioural results are identical (results not shown). We chose to maintain the same core reaction of $M + E \xrightarrow{C} C + W$ throughout this paper for simplicity.

The waste particle W also does not play a critical role in the dynamics that we have observed. And in fact, the autocatalysis is not strictly necessary to produce chemotactic behaviour (although it has dynamic consequences). A chemical reaction as simple as $R \rightarrow C$ could have captured a large portion of the behavioural dynamics we have described here. However, we set out with the motivation of exploring metabolism-based chemotaxis. As we described earlier in this section, a metabolism requires the channelling of energy by enzymes into reactions that produce more of those enzymes. We have tried to capture that essential relationship in the minimal reaction $M + E \xrightarrow{C} C + W$.

The following subsections describe the details of the modelling environment, how the above core reaction (and some variations) are implemented to model a minimal metabolism and how a simple coupling of this metabolism to an abstraction of the flagellar machinery generates chemotactic behaviour.

6.3.2 *Model*

The model consists of a two-dimensional environment, containing resource gradients and simulated bacteria. Each bacterium has a position, orientation and velocity as well as a metabolism, which is represented by a set of chemical concentrations. The concentrations of these chemicals are updated each iteration through numerical integration of the differential equations that represent ongoing chemical reactions in the metabolism as well

as degradation of metabolites and transport of 'resource' molecules from the immediate environment of the simulated bacteria into their interior.

Each bacterium is always either 'running' (moving in a straight line) or 'tumbling' (changing its orientation randomly). The probability of tumbling is directly proportional to the concentration of C, the autocatalytic product of the metabolism. This metabolism-based behavioural mechanism causes the bacterium to remain still when the metabolic rate is sufficiently high and to run when its metabolism is not operating above a threshold rate. The simulated environment is 200 units square. Bacteria trying to move out of this area are prevented from doing so as if running into a wall. Details of the reactions, the behavioural mechanisms and the environment are given below.

6.3.3   *Chemical reactions*

The autocatalytic reaction that constitutes metabolism is more explicitly described by the following reaction equations that include the intermediate stage where the catalyst, C, is bound to one of its substrates, M, forming MC.

$$M + C \rightleftharpoons MC \tag{1}$$

$$MC + E \rightleftharpoons 2C + W \tag{2}$$
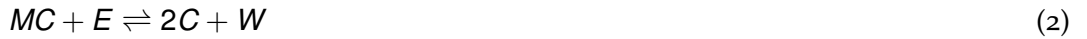
Two other processes influence the concentration of the metabolic reactants. The first is the degradation of reactants W and C into non-reactive products. These chemicals are removed from the simulation at rates specified in Table 6.1. The second is the influence of the local environmental chemical concentrations upon the concentration of reactants within the simulated bacteria (described below).

The metabolism dynamics are simulated by numerical integration of the differential equations in Table 7.2 (we used an Euler time step of 0.01 and typical chemical concentrations ranged between 0 and 2.0). These equations include some reactants that are only used in certain experimental scenarios and are explained later in the text. The rate constants (kf_n and kb_n) in the differential equations were determined by assigning *free-energies* to each reactant and *activation-energies* for each reaction such that the system adhered to the constraints given in our definition of a minimal metabolism. Given chemical free-energies and reaction activation-energies, reaction rates can be calculated by applying the following equations which indicate the reaction rate for a forward (exergonic) reactions and backward (endergonic) reactions respectively.

$$k_f = \exp(A)$$

$$k_b = \exp(A + R - P)$$

Free Energies		E,F	=	200
		M,N	=	40
		C	=	48
		MC,NC	=	91
		W	=	10
		S,SE	=	100
Reaction activation energies	Rxn #1	$M + C \rightarrow MC$	=	1.0
	Rxn #2	$MC + E \rightarrow 2C + W$	=	0.5
	Rxn #3	$N + C \rightarrow NC$	=	1.0
	Rxn #4	$NC + F \rightarrow 2C + W$	=	0.5
	Rxn #5	$S + E \rightarrow SE$	=	0.01
		$C \rightarrow \varnothing$	=	4.0
		$W \rightarrow \varnothing$	=	3.0
Reaction rates		kf_1, kf_3	=	0.0183156
		kb_1, kb_3	=	0.367879
		kf_2, kf_4	=	0.606531
		kb_2, kb_4	$\approx$	0.0
		kf_5	=	0.9905
		kb_5	$\approx$	0.0
Degradation rates		k_{degC}	=	0.0183156
		k_{degW}	=	0.0497871
Resource diffusion rate		k_{in}	=	0.2

Table 6.1: Constant parameters used in the model. Reaction and degradation rate-constants are determined according to formulae applied to the free-energies and activation energies which are hand-designed to adhere to the constraints inherent in our definition of a minimal metabolism and to display the phenomena of interest.

Figure 6.4 indicates why the forward and backward equations are different. This method of determining reaction rates allows the exploration of abstract chemistry while remaining congruent with the 2nd law of thermodynamics.

The environment of a bacterium can affect the concentration of certain chemicals within it, specifically: E, M, F, N, and S. For simplicity, we assume that these resources are actively transported into the bacteria at a rate independent of the concentration of the chemicals inside the membrane. Thus, the internal resource levels are increased by continuous transport from the environment into the bacterium according to the following function:

$$\frac{d[\kappa]}{dt} = k_{in}\epsilon(K, \mathbf{x})$$

where $[\kappa]$ is the concentration of the relevant chemical inside the bacterium, $k_{in} = 0.2$ is the rate of transport across the membrane and $\epsilon(K, \mathbf{x})$ is the concentration of K at the position of the organism in the environment, indicated by the vector $\mathbf{x}$. This influence of the environment is included in the differential equations in Table 7.2 as the last terms of
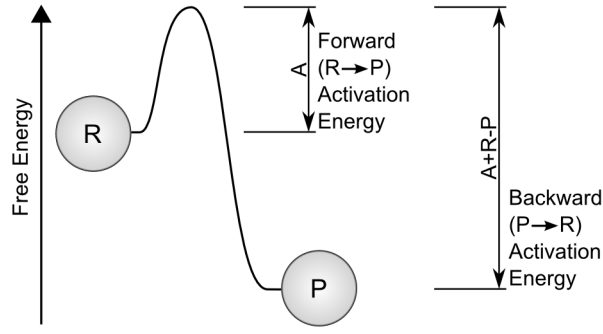
Figure 6.4: Energy required for a reaction to take place.

those equations that update chemicals that are affected by the environment (chemicals E, M, F, N, and S).

$$
\begin{aligned}
d[E]/dt &= -kf_2[MC][E] + kb_2[C]^2[W]/2 + k_{in}\epsilon(E, \mathbf{x}) \\
&\quad -kf_5[S][E] + kf_5[SE] \\
d[M]/dt &= -kf_1[M][C] + kb_1[MC] + k_{in}\epsilon(M, \mathbf{x}) \\
d[MC]/dt &= -kb_1[MC] + kf_1[M][C] \\
&\quad -kf_2[E][MC] + kb_2[C]^2[W]/2 \\
d[F]/dt &= -kf_4[NC][F] + kb_4[C]^2[W]/2 + k_{in}\epsilon(F, \mathbf{x}) \\
d[N]/dt &= -kf_3[N][C] + kb_3[NC] + k_{in}\epsilon(N, \mathbf{x}) \\
d[NC]/dt &= -kb_3[NC] + kf_3[N][C] \\
&\quad -kf_4[F][NC] + kb_4[C]^2[W]/2 \\
d[C]/dt &= -kf_1[M][C] + kb_1[MC] \\
&\quad -2kb_2[C]^2[W]/2 + 2kf_2[E][MC] \\
&\quad -kf_3[N][C] + kb_3[NC] \\
&\quad -2kb_4[C]^2[W]/2 + 2kf_4[F][NC] - k_{\mathrm{degC}}[C] \\
d[W]/dt &= -kb_2[C]^2[W]/2 + kf_2[E][MC] \\
&\quad -kb_4[C]^2[W]/2 + kf_4[F][NC] - k_{\mathrm{degW}}[W] \\
d[S]/dt &= -k_5[S][E] + k_5[SE] \\
d[SE]/dt &= -k_5[SE] + k_5[S][E]
\end{aligned}
$$

Table 6.2: Full list of the differential equations that determine the change in chemical concentrations within each bacterium. $[K]$ indicates the concentration of K within the bacterium. Parameters kf_n and kb_n represent the reaction rate constants for the nth reaction in the forward or backward direction (see Table 6.1). Parameters of the form k_{degK} indicate the rate of degradation of chemical K. k_{in} is the rate of resource diffusion into the bacterium. Finally, the function $\epsilon(K, \mathbf{x})$ returns the concentration of resource K in the environment at the location of the bacterium, indicated by the vector $\mathbf{x}$.

Implementation of movement

The above chemical reactions are simulated as enclosed within a membrane, comprising a simulated bacterium. Although minimal metabolisms have been the subject of simulation models, there have been very few attempts (see e. g., (Egbert and Di Paolo, 2009)) to study the dynamics of metabolism coupled to some form of movement generation
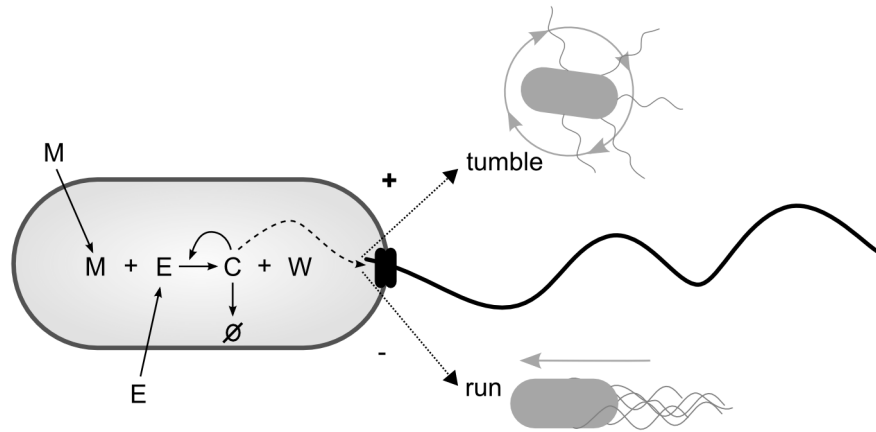
Figure 6.5: Diagram indicating the minimal metabolic reaction and its coupling to tumbling and running modes of behaviour.

mechanism. In this model, inspired by the motion mechanism of *E. coli* and other species, the simulated bacteria are capable of moving in either a directed, 'running' motion or by randomly changing their orientation ('tumbling'). Bacteria are, by default, in a 'running' mode. Each iteration, however, a bacterium has a chance of tumbling that is proportional to the concentration of the product of the metabolism, C: $p_{\text{tumble}} = 1 \times 10^{-5} + ([C] * 0.01)$. Running bacteria move in a straight line in the direction of their orientation (α), $\frac{dx}{dt} = 0.05 \cdot \cos(\alpha)$, $\frac{dy}{dt} = 0.05 \cdot \sin(\alpha)$. Tumbling bacteria remain at the same location, with α changed to a random value selected from a flat distribution between 0 and 2π.

A full schematic diagram of the minimal metabolism and its coupling to behavioural mechanisms can be seen in Figure 6.5.

6.4 RESULTS

We now describe five experimental scenarios where simulated bacteria are placed in environments containing different distributions of different chemical compounds. At the start of each simulation, 100 simulated bacteria are distributed evenly around the 200×200 unit square environment in a 10×10 grid (indicated in the left-most plot of e.g. Figure 6.6B). Each iteration, the metabolism and position of each bacteria is updated according to the equations described earlier. Bacteria are all initiated with a low (0.05) concentration of their metabolites unless otherwise indicated. Except for Experiment 5, all environmental resources have a peak concentration of 1.0 and fall off with distance from their centre according to the following equation where c is the concentration of the relevant resources and d is distance from the centre: $c = \exp(-\frac{d^2}{2000})$. Resources in the environment are always kept constant (i.e., there are no stigmergic effects – bacteria do not affect the concentration or distribution of resources in the environment).

6.4.1 *Experiment 1: Chemotaxis to metabolizable sources*

In this first scenario, a source of E and M is centred at $x = 100, y = 0$. Figure 6.6 shows the distribution of simulated bacteria at the start, middle and end of a 2.5×10^6 iteration simulation. It can be seen how bacteria perform chemotaxis to the area of high-concentration of E and M (which is indicated by the concentric circles).

When a bacterium has access to plenty of resources, it produces significant quantities of C and W. The high concentration of C causes the rate of tumbling to increase to the point where the bacterium is more or less standing in place since the tumbling frequency is so high that it never runs for a significant distance in any direction. If, on the other hand, the bacterium has insufficient available resources to maintain high levels of C, the probability of tumbling will fall and the bacterium will perform a combination of running and tumbling that results in a random walk. This random walk will continue until it comes across a region where it can produce sufficient C to push it "above threshold". In this manner, the simulated bacteria perform a simple form of "selective stopping" chemotaxis whereby they move in a random manner until they are in a resource rich area, at which point they tend to remain where they are.

Statistically, the simulated bacteria show a correlation between final location of the bacteria and high concentration of metabolizable substrates in the environment, i.e., a "chemotactic" movement towards high-concentration of attractants. The individual behaviour of the bacteria may not follow a direct chemotactic path, but the probabilistically directed (i.e., corrected or regulated) behaviour clearly results in an up-gradient movement tendency. Note that even in experimentally observed chemotaxis in *E. coli*, with their more sophisticated gradient-climbing adaptive strategy, the path followed by a single bacteria is difficult to characterise as chemotactic, it is rather the global effect of a population of bacteria that results in a clear chemotactic distribution. This experiment demonstrates that a metabolism-based control of flagellar rotation could potentially perform chemotaxis without dedicated signal transduction pathways, transmembrane receptor proteins nor sensory adaptation.

6.4.2 *Experiment 2: Local presence of metabolizable resources inhibits chemotaxis to other resources*

Alexandre and Zhulin observed that "the presence of another metabolizable chemical (exogenous or endogenous) prevents chemotaxis to all attractants studied" (Alexandre and Zhulin, 2001, p. 4682). We tested our simulated metabolism-based chemotactic bacteria to confirm that they also undergo inhibition of chemotaxis due to the presence of alternative metabolizable resources. To perform this test we used the above chemotaxis scenario as a control. The new experimental scenario is identical to the first except that two new resources, F and N, are included, uniformly distributed throughout the environment at a concentration of 0.5.
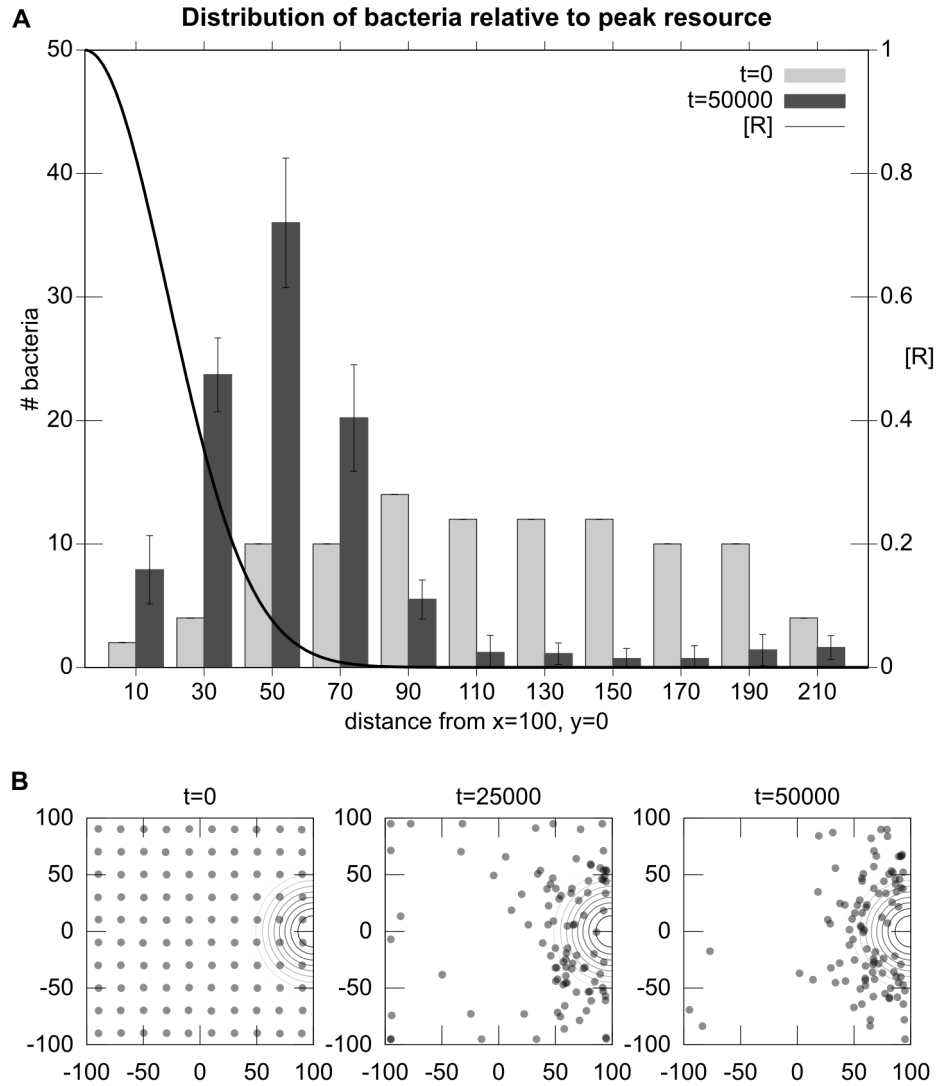
Figure 6.6: Experiment 1: Chemotaxis towards a gradient of M and E. Plot A is a histogram that indicates the distance of bacteria from the location of highest concentration of M and E ($x = 100, y = 0$) at the start and end of trials. Data are averaged from 10 runs of 100 bacteria each. Plot B indicates the spatial distribution of the simulated bacteria as time progresses in a typical trial. The concentric circles indicate the centre of the Gaussian distribution of resources M and E.
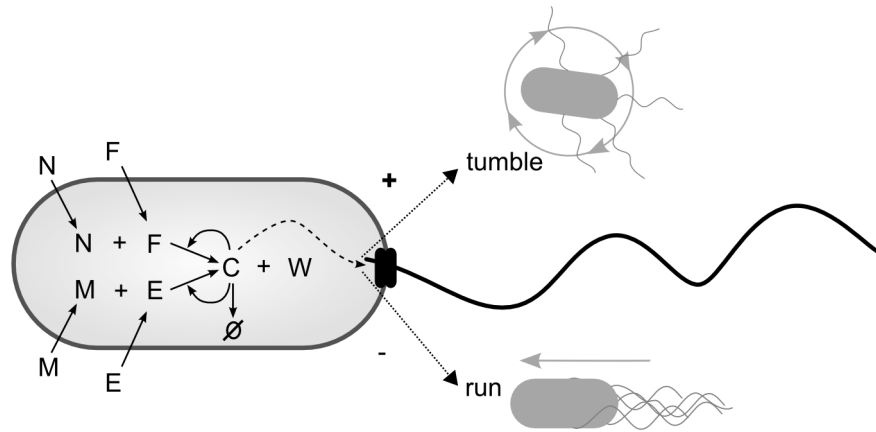
Figure 6.7: Diagram of the metabolic reactions for Experiment 2.

For this experimental scenario (and subsequent ones) alternative resource molecules and reactions were required. For the sake of simplicity we created a duplicate metabolic pathway with identical stoichiometry, reaction rates, etc., the only difference being the chemicals involved (see Figure 6.7): resource F (as analogous to E) and material resource N (analogous to M).

$$N + C \rightleftharpoons NC \tag{3}$$
$$NC + F \rightleftharpoons 2C + W \tag{4}$$

The results can be seen qualitatively by comparing Figure 6.6 (the control) and Figure 6.8. It is clear that chemotaxis has been inhibited. The mean distance from the source ($x = 100, y = 0$) for 10 runs of 100 agents each was 56.03 (std. 36.33) for the control, and 118.17 (std. 50.46) for the experimental abundance of alternative resource condition.

The mechanism for this inhibition is simple. Resources F and N are ubiquitous and sufficient to maintain the concentration of C at the high value necessary to keep the bacteria tumbling. The predominant tumbling keeps the agents stationary, preventing any chemotaxis to the $E + M$ resource.

6.4.3 *Experiment 3: Inhibition of metabolism to a resource inhibits chemotaxis to and only to this resource*

A second result was published by Alexandre and Zhulin in support of energy-taxis as the primary mechanism of chemotaxis in *A. brasilense*: "[The] inhibition of the metabolism of a chemical attractant completely abolishes chemotaxis to and only to this attractant" (Alexandre and Zhulin, 2001, p. 4682).

To test this phenomenon in our simulation of metabolism-based chemotaxis, simulated bacteria are placed in an environment with two resource gradients. The first, consisting
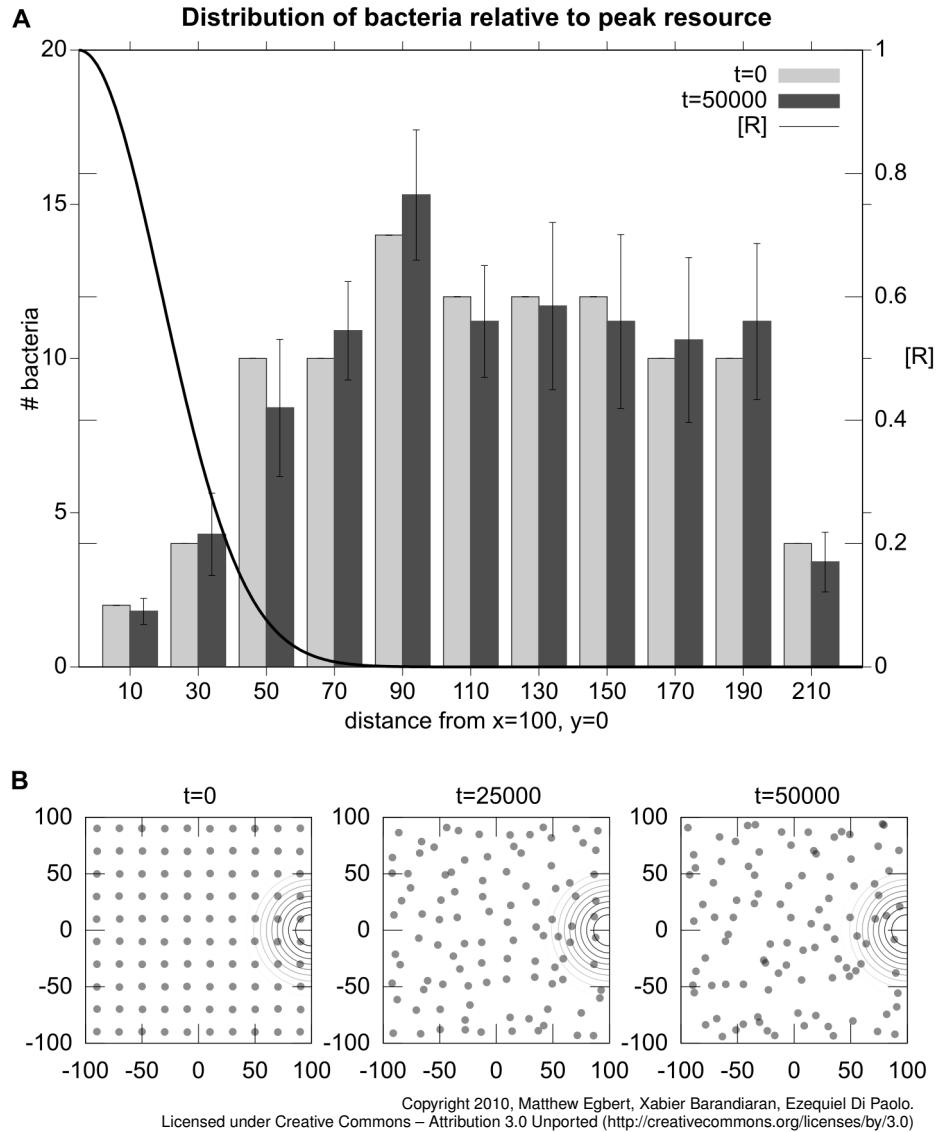
Figure 6.8: Experiment 2: Abundance of sufficient resources inhibits chemotaxis to attractants. In this trial, agents were placed on the same gradients as in Figure 6.6, but with an additional uniform distribution of alternative metabolizable sources [F], [N] = 0.5. The presence of these resources clearly inhibits chemotaxis to the gradients of E and M.
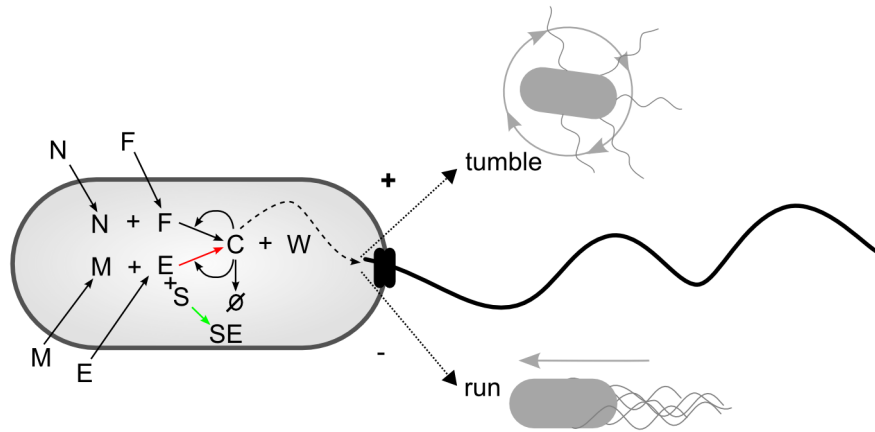
Figure 6.9: Diagram of the metabolic reactions for Experiment 3.

of equal parts of resources *F* and *N* is highest in concentration in the upper-right corner of the simulated environment. The second is equal parts of *E* and *M* and is highest in concentration in the lower-right corner. In the upper-right corner, resources *F* and *N* are sufficient for a healthy metabolism to continue to autocatalyze *C* and maintain its concentration high enough for the bacteria to remain in this corner. The same is the case for *E* and *M* in the lower-right corner. This can be seen in Figure 6.11.

Halfway through this scenario, we add a uniform concentration of $[S] = 1.0$ to the entire simulated environment. This chemical inhibits the *E*/*M* metabolic pathway by exothermically and rapidly bonding to metabolizable substrate *E*, transforming it into a non-reactive chemical, *SE* (see Figure 6.9). This process is described by the following reaction equation:

$$S + E \rightleftharpoons SE \tag{5}$$

After *S* is added to the environment, the simulated bacteria cease to remain in the area high in concentration of *E* and *M*, but continue to be attracted to the high concentrations of *F* and *N*, as shown in the right-most plot of Figure 6.11, demonstrating inhibition of chemotaxis to a reactant by inhibition of metabolization of that reactant.

6.4.4 *Experiment 4: Metabolic inhibitors act as repellents*

Specific metabolic inhibitors such as oxidised quinones or specific electron transport inhibiting molecules such as myxothiazol have been shown to inhibit chemotaxis in bacteria such as *Campylobacter jejuni* (Vegge et al., 2009). It has also been shown that such metabolic inhibitors can act as repellents (Alexandre et al., 2000; Bespalov et al., 1996). We tested to see if our model could also display metabolic inhibitors (or toxins) acting as repellents.
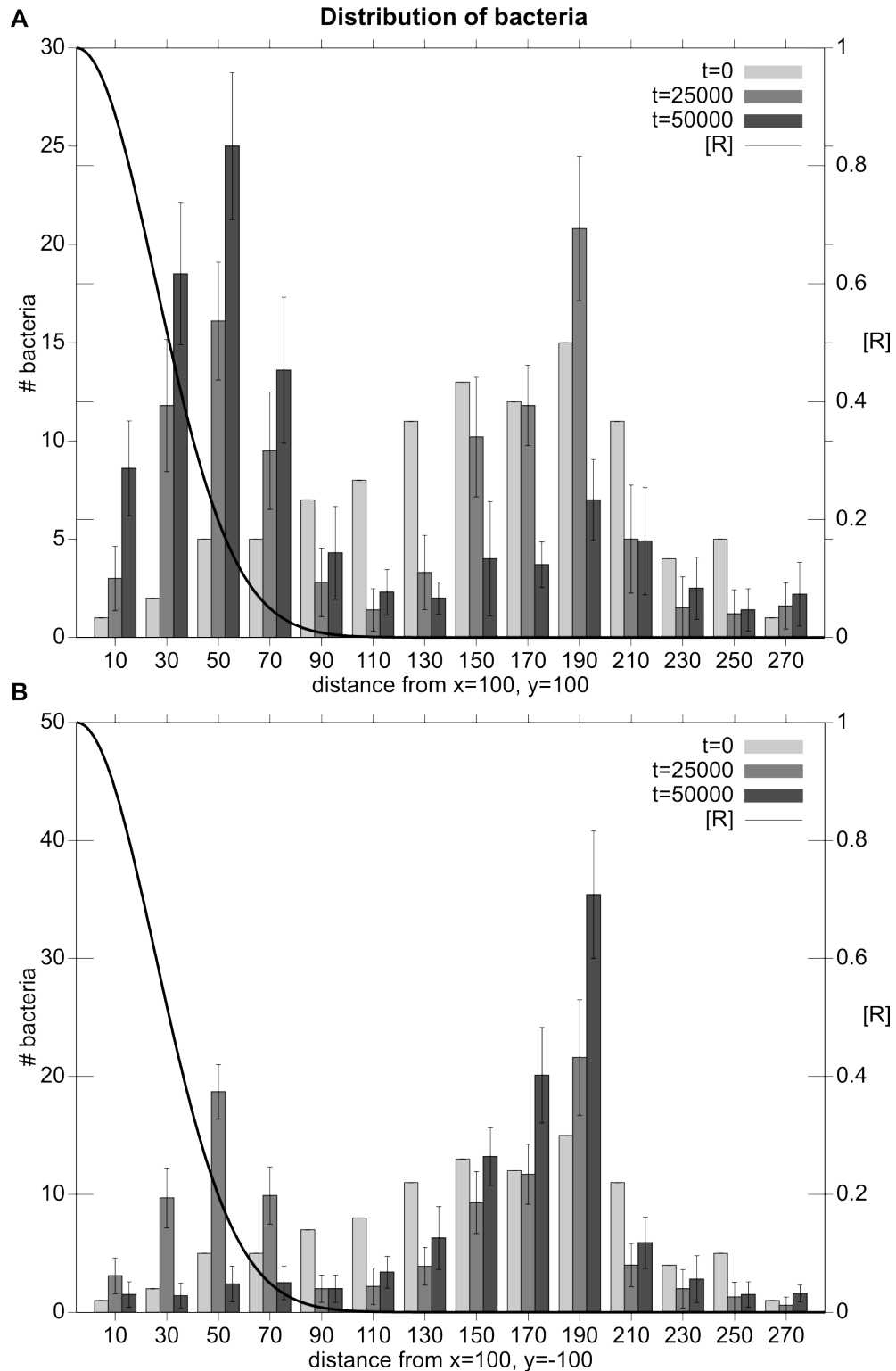
Figure 6.10: Experiment 3, final distribution relative to resource centres. Inhibition of a metabolic pathway inhibits chemotaxis only to the relevant resources. Histogram A illustrates average distance to the non-inhibited (N + F) source. Histogram B shows the distance from the inhibited source (M + E). Data is taken from 10 trials of 100 bacteria each.
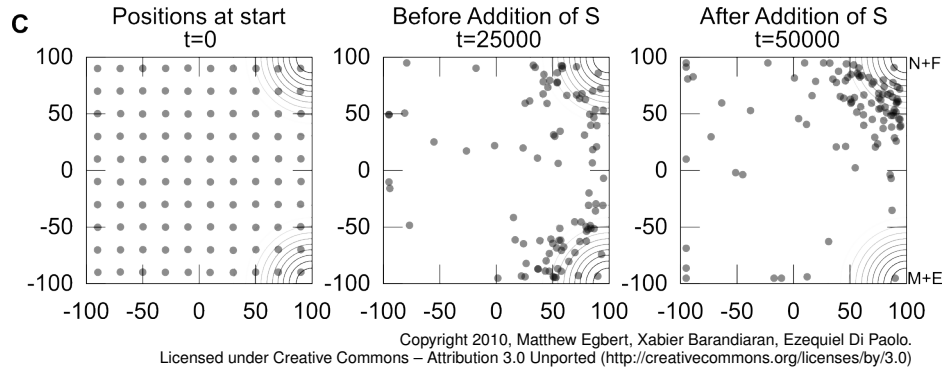
Figure 6.11: Spatial distribution of simulated bacteria in a typical trial of Experiment 3. In an environment with two sets of metabolizable resources, the simulated bacteria move to the areas higher in concentration of either resource-pair. The insertion of S inhibits the metabolism of $M + E$, ending chemotaxis towards the no-longer metabolizable resources without influencing chemotaxis to the other attractants.
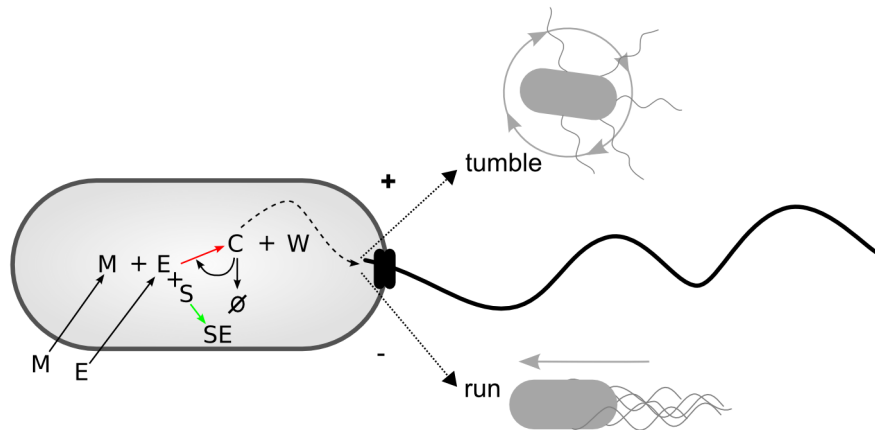
Figure 6.12: Diagram of the metabolic reactions for Experiment 4.

A   **Distribution of bacteria relative to peak concentration of toxin**

B

Positions at start
t=0

Before Addition of S
t=25000

After Addition of S
t=50000
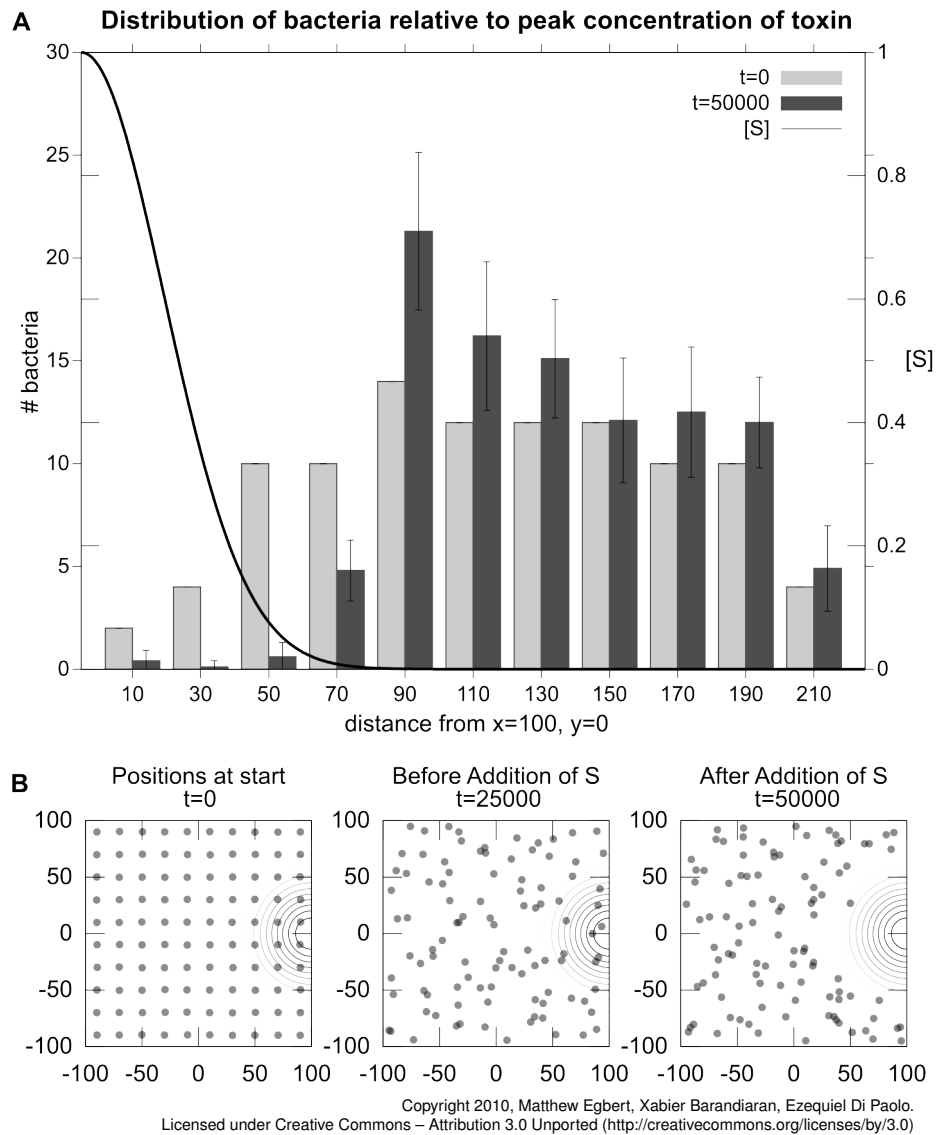
Figure 6.13: Experiment 4: Metabolic inhibitors act as repellents. In a uniform distribution of resources, simulated bacteria move away from high concentrations of the metabolic inhibitor, S.

A repulsion due to a metabolic toxin can be clearly seen in Figure 6.13 which shows a scenario in which bacteria are evenly distributed in an environment of uniform distributions of $[E]$, $[M] = 0.5$. Halfway through the simulation, a gradient of metabolic inhibitor S is added to environment (with a peak concentration 5.0, centred at $x = 100, y = 0$) and the bacteria move away from the higher concentrations of that toxin. Figure 6.12 indicates the reactions that occur in this scenario.

6.4.5   *Experiment 5: Sensitivity to history*

Experiments 1–4 have reproduced empirical observations made by Alexandre and Zhulin. The following experiments explore additional phenomena that could lead to some empirical predictions. This scenario is inspired by Alexandre and Zhulin's observation that bacteria demonstrate a sensitivity to their history of exposure to different resources. Specifically, "[s]tronger chemotaxis responses are observed when cells are grown on the sugar under test as the growth substrate" (Alexandre and Zhulin, 2001, p. 4682). In this experiment, there is no M in the environment except along a strip defined by $x < -90$. Agents are initialised with no M. A gradient of E is placed at the centre of the right side of the environment.

Bacteria have no M, so the resource E is insufficient to produce C. Only after they have encountered the region with M and incorporated M into their metabolism does E act as an attractant. This process can be seen through observation of Figure 6.14 where agents are drawn with $\times$s if they have concentrations of M less than 0.5 and as circles otherwise. Early in the simulation, agents tend to run as none have access to resources sufficient to produce and maintain significant quantities of C. As time passes, the random motions of the bacteria cause some to encounter the M on the left. More time passes, and these agents, now rich in M can produce C while in areas rich in E. At this point, chemotaxis towards high concentrations of E is observed.

The experiment shows how what becomes an attractant for metabolism-based chemotactic bacteria is not an environmental compound *per se*, but rather what, at a given point in the history and internal state of the bacterium, is required for metabolism to occur. Bacteria performing metabolism-based chemotaxis operate according to their current metabolic needs.

6.4.6   *Experiment 6: Integration of environmental phenomena*

By basing chemotaxis in metabolism, the simulated bacteria respond not to specific environmental phenomena, but to the combined effects of all environmental features upon metabolism. In Experiment 6, we demonstrate this ability to integrate environmental phenomena by placing bacteria in a more complicated environment than those of previous experiments. The environment for Experiment 6 consists of perpendicular linear gradients
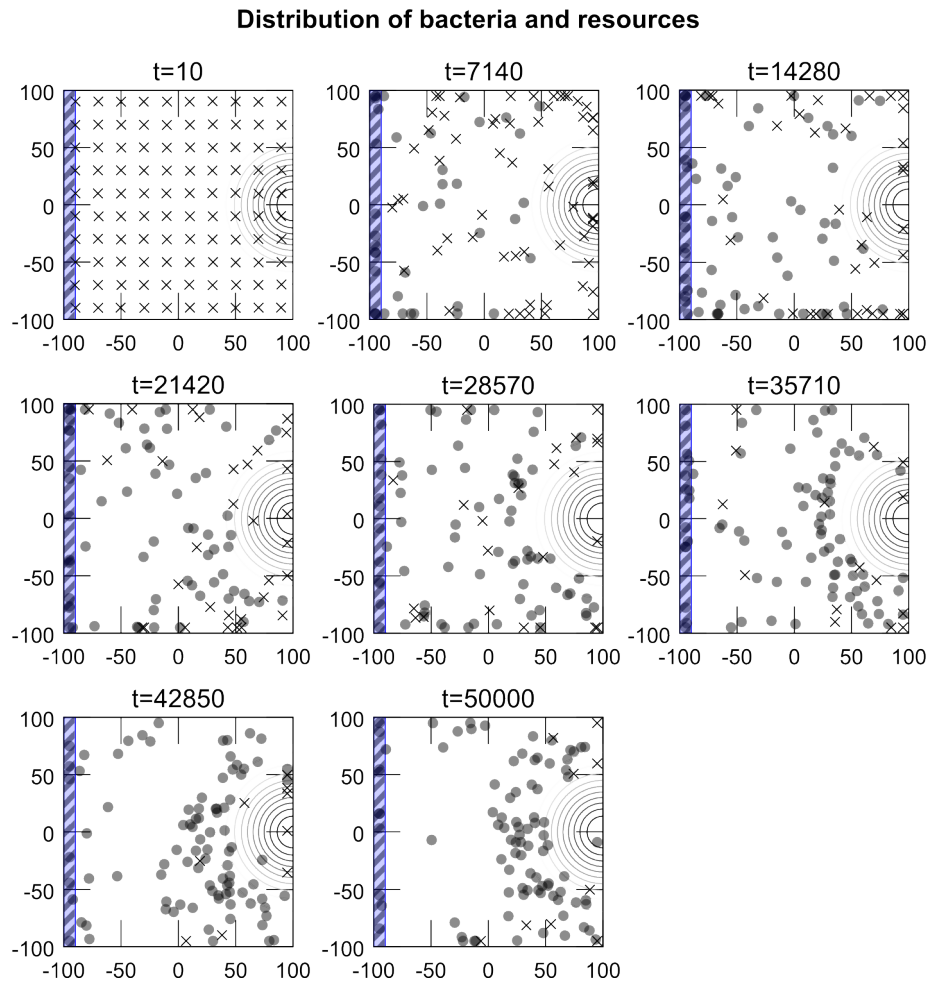
Distribution of bacteria and resources

Figure 6.14: Experiment 5: Sensitivity to history. The metabolism of bacteria is changed by their encounter with M on the left of the simulated area where M can be incorporated into their metabolism. The E rich area (concentric circles on right side of plots) is initially insufficient to support production of C, but for agents that have incorporated M into their metabolism, E becomes an attractant. Bacteria with $[M] < 0.5$ are shown as Xs rather than circles.

of resources M ($[M] = 0.5*(y+100)/200$) and E ($[E] = 0.5*(x+100)/200$) with a Gaussian distribution of toxin S, with a peak concentration of 3.0 centred at ($x = 75, y = 75$). Figure 6.15 indicates the final position of 5000 bacteria (the results of 50 trials, each with 100 evenly distributed bacteria as in the other experiments). It can be clearly seen that the bacteria are neither maximising concentrations of M or E, nor the combination of them, but are performing chemotaxis to the areas where the combined effects of the environmental resources E, M, and S allow the metabolism to operate sufficiently well. The overall bacterial distribution appears correlated with the spatial distribution of the optimal combination of compounds.

One interesting aspect of this plot is the asymmetric distribution of bacteria along the x and y axes, corresponding to the the gradients of E and M respectively. It appears that for our model bacteria, it is more important to have a high concentration of M than of E. This difference may be caused by a high concentration of M transforming 'free' C into MC. MC, unlike C, does not degrade, so a high concentration of M makes the metabolism less likely to degrade than a high concentration of E. An alternative possibility is that the bottleneck in the metabolism lies in the first reaction ($M + C \rightarrow MC$) as this reaction has a slower rate constant than the second reaction ($MC + E \rightarrow 2C + W$). High E has little influence on this bottleneck, but a significant concentration of M can open up the bottleneck, allowing for a more rapid production of C. We confirmed that a third possibility, the asymmetric influence of S upon the reactions, is not responsible for the asymmetric distribution of bacteria (results not shown).

It is interesting to note that in aerotaxis experiments with e. g. *E. coli* and *A. brasilense*, the aerotactic bands can form an asymmetric profile as well (see e. g. (Xie et al., 2010, p. 2238 Figure B, top)). The environmental conditions at the different sides of these bands could be slightly different with respect to the metabolism of the bacteria, and perhaps a mechanism similar to that described here could explain the asymmetric distribution of bacteria in these experiments.

6.5 DISCUSSION

Research in bacterial chemotaxis has operated largely under the assumption that the behaviour is supported by transmembrane receptors and dedicated signalling pathways and that such pathways are metabolism-independent. Despite the growing body of evidence that in many species this might not always be the case (even for those largely thought to be so, such as *E. coli*) available simulation models of bacterial chemotaxis assume metabolism-independence. Here we have presented the first minimal model of metabolism-based chemotaxis. It recreates phenomena observed in bacteria and allows us to explore some potential consequences of metabolism-based behaviour.

The behavioural strategy employed by the simulated bacteria in our model is the "selective stopping strategy" in which bacteria move around in a random walk until they reach a satisfactory area, at which point they tumble in place. Recent artificial evolution
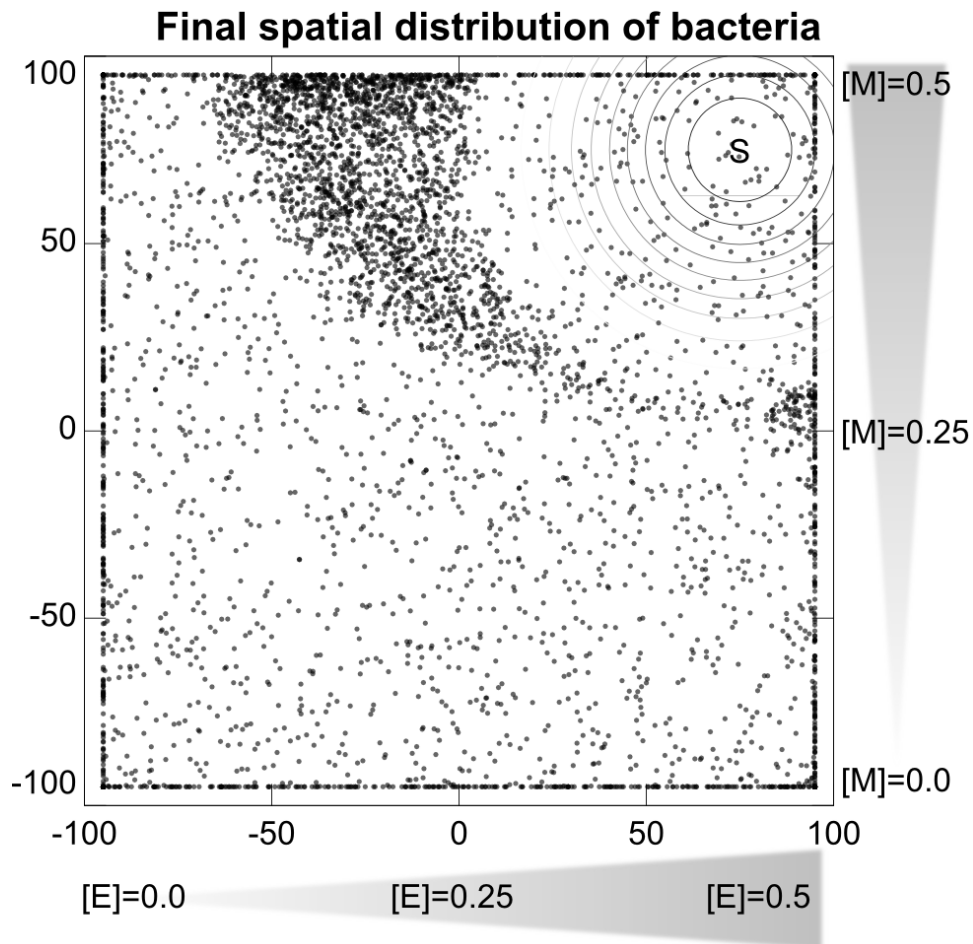
Final spatial distribution of bacteria

Figure 6.15: Experiment 6: Metabolism based behaviour does not respond directly to environmental features, but rather to their combined effect upon metabolism. Here we see simulated bacteria concentrated in the area of the environment where the combined effects of all environmental effects allow for a sustainable metabolism.

of simulated chemotaxis (Goldstein and Soyer, 2008) has shown that this strategy (also referred to as the "non-adaptive or inverted response") is, under certain ecological constraints, the most likely chemotactic strategy. It has also been observed in some cases of metabolism-dependent chemotaxis (Packer and Armitage, 2000). In order to address how metabolism could directly produce such behavioural patterns, we have developed a model of what we have called *metabolism-based chemotaxis* (a simpler case than that of metabolism-dependent chemotaxis). We first identified a minimal metabolic organisation as that of an autocatalytic reaction. We then assigned a probability of running or tumbling to the concentration of the auto-catalyst. The resulting system is very simple, yet capable of instantiating four chemotactic phenomena observed in bacteria.

- Chemotaxis to metabolizable compounds. [Experiment 1]

- A local abundance of metabolic resources inhibits chemotaxis to other resources. [Experiment 2]

- Inhibition of the metabolism of an attractant inhibits chemotaxis to that attractant, and that attractant alone. [Experiment 3]

- Metabolic inhibitors act as repellents [Experiment 4]

The observations of history-dependence reported by Alexandre and Zhulin led to the exploration of Experiment 5 where a compound is incorporated into metabolism and results in a change in chemotaxis pattern according to past experience.

Finally, Experiment 6 demonstrated the capacity of metabolism-based chemotaxis to respond appropriately to combinations of a variety of simultaneous environmental influences. This experiment showed the potential of metabolism as a mechanism for effective chemotactic integration.

The present simulation is not a model of the specific mechanisms supporting metabolism-dependent chemotaxis in bacteria. Yet, it serves as a proof of concept of how a very simple abstraction of metabolism can support, without the addition of specific signalling pathways and even without the need of transmembrane receptors, a wide range of chemotactic phenomena. As a conceptual model it can be further used to explore some theoretical implications of relaxing the metabolism-independent assumption.

Evidence has been found of metabolism-dependent chemotaxis where metabolic processes such as the electron transport system influence flagellar rotation indirectly, by way of the dedicated chemotactic two-component signalling system. There is also evidence for a mechanism through which metabolism directly influences flagellar rotation, i.e., without an intermediate dedicated signalling system, in a manner more similar to the metabolism-based chemotaxis modelled here. Specifically, it has been shown that *E. coli* can perform chemotaxis even when stripped of most of the signalling pathway typically associated with chemotaxis (Barak and Eisenbach, 1999), suggesting that there might be at least two different mechanisms supporting chemotaxis in *E. coli* (Eisenbach, 2007, p. 575). Interestingly, the concentration of fumarate, an intermediate in the citric acid cycle that is part of the "universal metabolism" (Smith and Morowitz, 2004), has been shown to influence the direction of flagellar rotation. A high concentration of this metabolic product increases chance of clockwise, tumble inducing, rotation; the same relationship of metabolic influence upon flagellar rotation that is used in the selective-stopping strategy. Fumarate operates directly upon the flagellar motor switch (Cohen-Ben-Lulu et al., 2008) in a manner that is independent of the protein signalling pathway typically associated with chemotaxis (Prasad et al., 1998). It turns out that fumarate might be currently instantiating mechanisms of metabolism-based chemotaxis; something that still remains to be experimentally tested. This hypothesis was anticipated by (Goldstein and Soyer, 2008, p. 5) and we have shown how fumarate-like intermediate metabolites (C in our model) could not only produce simple chemotaxis but could reproduce a wide spectrum of non-trivial chemotactic phenomena. What is needed to achieve these behavioural patterns is not a complex system of transmembrane receptors influenced by metabolism

in subtle ways but simply a metabolite capable of influencing flagellar rotation in the right manner.

The present model also suggests that it might be time to re-consider part of the terminology and the externalist approach to chemotactic studies. For instance, it is generally assumed that environmental compounds are invariably either *attractants* or *repellents* for bacteria, as if bacteria were simply stimulus-driven systems. The model of metabolism-based chemotaxis shows, however, that environmental compounds are not attractants or repellents purely on the basis of their binding properties and their stereotypically elicited responses. Environmental compounds must instead be categorised within the context of metabolism, which is influenced by the history of the cell and its internal organisation (metabolic rates, active and non-active metabolic pathways, etc.). In other words, the behavioural significance of chemical compounds becomes a relational property that depends on the metabolic dynamics of the cell (which cannot be abstracted away in the study of behaviour). As Experiment 3 shows, if a resource ceases to be metabolized, it ceases to act as an attractant for bacteria. Also, (as shown in Experiment 4) chemical compounds that are toxic for the metabolism of the bacteria can act as repellents without the need of any specific binding of it, or even without the bacteria ever encountering that compound in its evolutionary past. This capacity to be behaviourally sensitive to the effects of environmental compounds on metabolism provides a powerful means of behavioural evaluation and increased adaptive response (at the organismic level). It is not clear how the same adaptation could occur for a metabolism-independent mechanism that requires binding with specific compounds to elicit specific responses.

While advances have been made on the understanding of how receptor complexes integrate sensory information (Thomason et al., 2002; Borrok et al., 2008) the potential *integrative role of metabolism* remains under-explored. The classic view is that integration in metabolism-independent chemotaxis is accomplished in the group dynamics of the transmembrane sensors that all modulate CheA activity (Parkinson et al., 2005). This metabolism-independent mechanism of integration relies upon specific interactions between stimulus chemicals and transmembrane sensors. We can compare this to metabolism-based behaviour, which responds not directly to environmental phenomena but to the *combined effects of environmental phenomena upon the metabolism*. This indirect sensitivity to the environment makes it a good candidate for integrating different stimuli and producing the appropriate response (move toward or away). Goldstein and Soyer (Goldstein and Soyer, 2008) acknowledge this issue but their model does not address any integrative phenomena—their simulation results correspond only to a single attractant gradient scenarios. Despite its simplicity, the model presented here is able to effectively integrate information from multiple gradients in a straightforward manner (Experiment 6). By being sensitive to the production of C, it integrates the effect of all environmental features upon metabolic rate.

As an example of the potential integrative power of metabolism-based chemotaxis, imagine two compounds, T_1 and T_2, each of which acts as a metabolic toxin when

encountered on its own. But, when encountered together, they act as excellent metabolic resources. Metabolism-based chemotaxis would respond appropriately (move towards $T_1 + T_2$ when encountered together and away from T_1 or T_2 when either is encountered on its own), while metabolism-independent chemotaxis would require the evolution of considerable specific machinery to accomplish the same appropriate behaviours.

A further development of the model presented here (see (Egbert et al., 2010a)) has allowed us to explore the potential of metabolism-based mechanisms to produce gradient-climbing strategies, in particular we have shown how a single new reactant could turn a network of metabolic reactions that produces the selective-stopping behaviour into one that produces the more intricate gradient-climbing behaviour. We have also designed a scenario where a simulated protocell incorporates a new attractant from the environment into its metabolism and becomes chemotactic towards it. We have used the above experiments to explore theoretically the potential of metabolism-based chemotaxis, the feedback between metabolism and behaviour, to bootstrap and accelerate early evolutionary processes (Egbert et al., 2010a).

Among the further extensions, a very interesting development would be to study which situations are conducive to which relationships between metabolism and chemotaxis and how transitions from one form to the other could occur. In particular, artificial co-evolution of metabolic networks and behavioural mechanisms could help address questions regarding a) the likelihood of metabolism-dependent or metabolism-independent chemotaxis under various environmental conditions, b) the possibility of metabolism-independent chemotaxis arising from metabolism-dependent precursors and c) how both types of chemotaxis might co-exist with a varying degree of influence.

Despite the considerable advances that the segregated one-compound-one-response approaches to chemotaxis have provided so far, it is perhaps time to start integrating not only metabolism into the picture but richer and varying environments where metabolic modulation might be playing a more relevant behavioural role. We hope to have shown that such an integrative move does not necessarily require the inclusion of an overwhelming level of detail, but might instead be effectively dealt with by metabolism-based forms of regulation. Many aspects of metabolism can potentially be abstracted away to reproduce the higher order dynamics of complex metabolic networks and then coupled to a behavioural mechanism. Moving in this direction opens the space for interactions between internal and environmental chemical dynamics that are not reducible to the influence of environmental compounds upon transmembrane receptors.

From the reported experiments we can generalise that, despite its simplicity, metabolism-based chemotaxis allows for an *ongoing* evaluation of environmental conditions. This evaluation is *indirect* in that behaviour is not in response to the environment, but rather to the influence of the environment upon the metabolism. The ongoing and indirect nature of metabolism-based chemotaxis makes possible an automatic and appropriate response to a variety of encounters with environmental conditions that have never been experienced by the bacterium, nor even by its evolutionary ancestors, for it is not

necessary to evolve trans-membrane sensors that interact in specific ways with each environmental influence. The evaluation of the environment is accomplished by the influence of the metabolism.

These generalisations should be further examined both by empirical studies and elaborations of the current model. We would like to stress that the current model plays the role of a proof of concept by allowing us to see the possibility of metabolism-based chemotaxis at work and unveil some implications. As variations of the model start to address more specific issues, they will have to incorporate more realistic assumptions such as energetic requirements for movement, biomechanics, differences in timescales between behaviour and metabolism, and potential interactions between optimal behavioural control, metabolic dynamics and stochasticity. Also required is a study of the parametrical robustness of the phenomena reported here.

PRIMARY POINTS OF CHAPTER

- This chapter presents the first ever minimal model of metabolism based chemotaxis. It recreates chemotaxis patterns observed in bacteria such as *Azospirillum brasilense* and *Campylobacter jejuni*, supporting the idea that metabolism based behaviour may operate in nature and that it can explain behavioural phenomena.

- The model also provides a new explanation for the asymmetric distribution of bacteria observed in various studies such as Xie et al. (2010).

- Some benefits of metabolism based behaviour were discussed, specifically, (i) the integration of environmental phenomena into a coherent and appropriate response and (ii) the ability to adapt ontogenetically (during the life-time) to phenomena never previously experienced rather than relying upon multiple generations of exposure as in evolutionary adaptation.

- Fumarate part of the universal metabolism is hypothesised to play a role in accomplishing metabolism-based chemotaxis in *E. coli*.

BEHAVIOURAL METABOLUTION

*"What is it that enables living things, apparently so moist,
fragile, and evanescent, to persist while towering mountains
dissolve into dust, and the very continents and oceans dance
into oblivion and back?" (Rosen, 1991, p. 11)*

CHAPTER SUMMARY

In Chapter 6, we presented and analysed our computational model of metabolism-based chemotaxis. Our experimentation with this model and extended consideration of metabolism-based chemotaxis prompted new theories concerning how such mechanisms could interact with evolutionary processes in such a way as to accelerate them in a variety of ways.

In this chapter we outline the fundamentals of how this could occur and demonstrate some of its potential through extensions of the computational simulation presented in Chapter 6. Specifically, we demonstrate in computational simulation how metabolism-based behaviour can adapt not only to changes in the environment of the organism, but also to changes in the organisation of the organism itself (e. g., genetic mutations, or chemical 'avalanches' caused by environmental encounters). We show how this kind of adaptation can produce a change in the environment of the organism, and

how an encounter with new environmental phenomena can qualitatively transform a metabolism-based behaviour.

We use these demonstrations to develop the notion of behavioural metabolution, a circle of mutual influence between metabolism-based behaviour, metabolism and environment, in which metabolism-based behaviour adapts to accommodate changes in the metabolism, while also increasing the chances of further changes in the metabolism. This is a combination of directed and non-directed processes that results in adaptive-evolution facilitating dynamics.

We then develop the hypothesis that the simplicity of the mechanisms underlying behavioural metabolution and the broad adaptability that behavioural metabolution provides suggest that it could have played an interesting role in the early history of life.

Most of the text of this chapter was originally submitted to the 12th International Conference on the Synthesis and Simulation of Living Systems, where it received excellent reviews and was deemed of broad enough interest that the organisers invited us to present it in one of the four plenary talks of the conference.

7.1 INTRODUCTION

This chapter presents a third simulation model of protocells capable of metabolism-based chemotaxis. Similar to the model presented in Chapter 6, a minimal metabolic system capable of modulating behaviour by influencing the probability of flagellar rotation (like in *E. coli* chemotaxis). We perform two illustrative experiments. In the first, the incorporation of a chemical compound into metabolism qualitatively improves the chemotactic strategy. In the second, an encounter with a specific chemical compound leads to a reaction that opens up a new metabolic pathway while automatically regulating chemotaxis towards that same compound. Both experiments illustrate the adaptive potential of metabolism-based behaviour and can be used to explore the idea of "Behavioural Metabolution," a co-evolutionary synergy between behaviour and metabolism. We abstract some principles of behavioural metabolution and discuss its application to early prebiotic evolution.

More specifically, the two experiments provide further evidence that metabolism can modulate behaviour in an adaptive manner. They also demonstrate that behaviour can change the metabolism by changing the environment in which it exists and finally that changes in metabolism can produce new types of behavioural patterns. The chapter concludes with discussion of the evolutionary dimension of metabolism-based chemotaxis, what we term "behavioural metabolution", and its potential application to the question of early evolution of life.
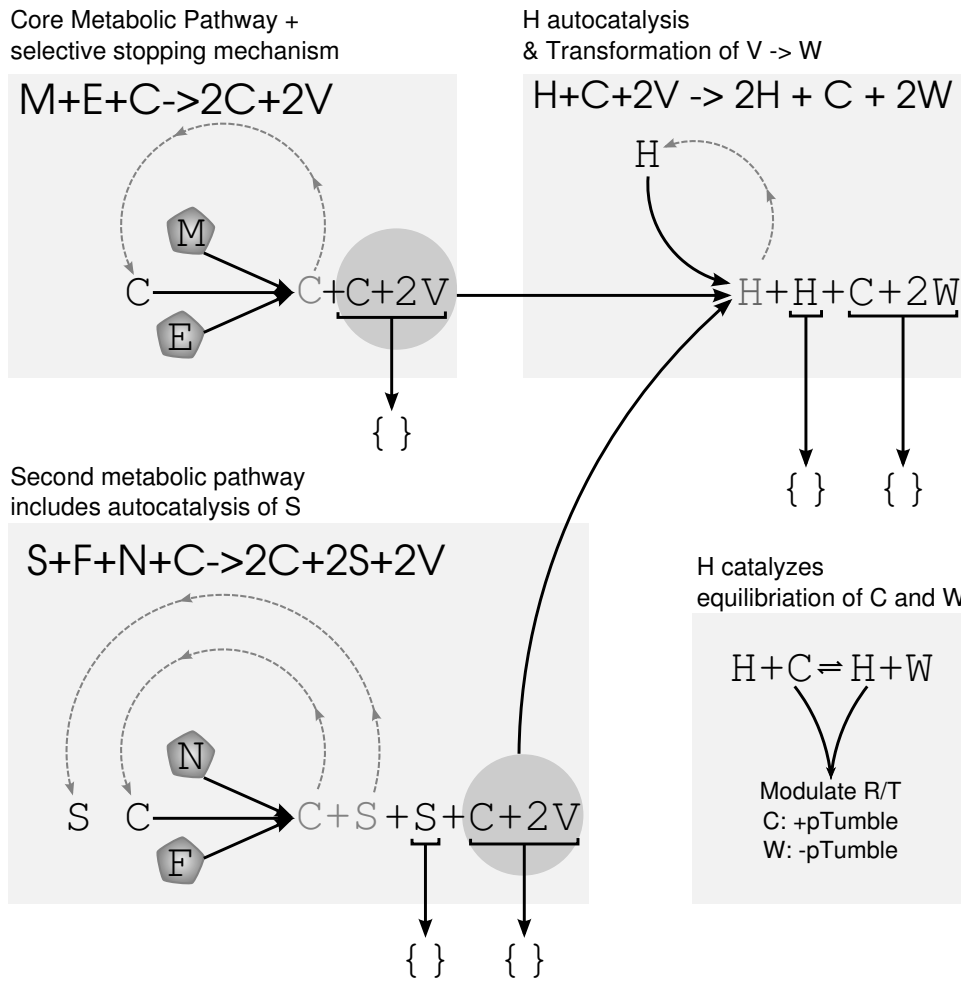
Core Metabolic Pathway +
selective stopping mechanism

$M+E+C->2C+2V$

H autocatalysis
& Transformation of V -> W

$H+C+2V -> 2H + C + 2W$

Second metabolic pathway
includes autocatalysis of S

$S+F+N+C->2C+2S+2V$

H catalyzes
equilibriation of C and W

$H+C \rightleftharpoons H+W$

Modulate R/T
C: +pTumble
W: -pTumble

Figure 7.1: Reactions grouped conceptually by their 'role' in the model. Resources are surrounded by pentagons. Auto-catalytic reactions are indicated by circular paths. Degradation of reactants is indicated by an arrow to the empty set.

7.2 METABOLISM-BASED CHEMOTAXIS, THE MODEL

As before, we consider metabolism as the self-production of a chemical network through the transformation (by the network) of available energetic and material resources into constituents of the network. This process is realised in this model through an auto-catalytic reaction whereby energetic and material resources (E and M respectively) are transformed by network constituent C into more C and a low energy waste V thus: $M + E \xrightarrow{C} C + 2V$. As with the previous model, this single reaction may be understood as a higher level abstract representation of a whole network of processes, considering that the essence of metabolism is that of an auto-catalytic network. To capture the requirement of far-from-thermodynamic equilibrium, C and V are considered thermodynamically unstable and degrade rapidly. Their continued presence is therefore only possible through a *dynamic* equilibrium of degradation countered by production. We label this reaction the "core metabolism" and expose it to various other reactants in different experiments. Table

#	reactants		products	k_f	k_b
0:	$M + E + C$	$\rightleftharpoons$	$2C + 2V$	0.61	≈ 0
1:	$H + C$	$\rightleftharpoons$	$H + W$	0.006	0.006
2:	$H + C + 2V$	$\rightleftharpoons$	$2H + C + 2W$	0.37	≈ 0
3:	$C + 2V$	$\rightarrow$	$\{\}$	0.006	n/a
4:	$C + 2W$	$\rightarrow$	$\{\}$	0.006	n/a
5:	H	$\rightarrow$	$\{\}$	0.02	n/a
6:	S	$\rightarrow$	$\{\}$	0.0001	n/a
7:	$S + F + N + C$	$\rightleftharpoons$	$2C + 2S + 2V$	0.99	≈ 0

Table 7.1: A list of the chemical reactions in each simulated metabolism. Also indicated are the reaction rates (forward and backward). These rates are referred to in Table 7.2.

7.1 and Figure 7.1 show all of the chemical reactions that can be active in the protocells simulated in our model. The upper-left square indicates the core metabolism described in this section. The other pathways are described in the experiments and results section.

The metabolic dynamics are described by the differential equations in Table 7.2. These equations include some reactants that are only used in some of our experimental scenarios and are explained later in the text. The rate constants (k_{fn} and k_{bn}) in the differential equations were determined using the same methods as described in Chapter 6. *Free-energies* were assigned to each reactant and *activation-energies* were assigned to each reaction such that the system adhered to the constraints defined in our definition of a minimal metabolism. Given these values, we could derive rate constants for the reactions (see Section 6.3.3 for the details of this derivation).

Resources encountered in the environment diffuse into protocells at a rate proportional to the local concentration of the environmental resource. The rate constant for this diffusion, $k_d = 0.04$, is the same for all resources.

As in the previous model, the chemical reactions are simulated as occurring in a compartment surrounded by a membrane that includes a set of flagella. The clockwise and counter-clockwise flagellar rotation is determined by the relation between the concentrations of C and W compounds. In analogy to the working of flagellar rotation in *E. coli* chemotaxis, when the overall movement of flagellar rotation is counter-clockwise the protocell is propelled in straight direction (what is generally called the "running mode"), whereas when flagella rotate clockwise, the protocell rotates on its axis changing direction randomly ("tumbling mode"). Protocells are simulated in a 2D square 'petri-dish' of 200 units. By default, protocells are always running, i.e., moving in a straight line in the direction of their orientation, α, thus: $\frac{dx}{dt} = 0.05 \cdot \cos(\alpha)$, $\frac{dy}{dt} = 0.05 \cdot \sin(\alpha)$. A baseline probability of tumbling allows for the random direction to be changed occasionally, even in running protocells. Tumbling protocells remain at the same location, with α changed to a random value selected from a flat distribution between 0 and 2π. The effect of the influence of C and W concentrations on flagellar rotation is abstracted

$$
\begin{aligned}
dE/dt &= -k_{f0}EMC + k_{b0}C^2V^2/4 + k_d[E](\mathbf{x}) \\
dM/dt &= -k_{f0}EMC + k_{b0}C^2V^2/4 + k_d[M](\mathbf{x}) \\
dC/dt &= -k_{f0}EMC + k_{b0}C^2V^2/4 \\
&\quad -2k_{b0}C^2V^2/4 + 2k_{f0}EMC \\
&\quad -k_{f1}CH + k_{b1}HW \\
&\quad -k_{f3}CV^2/2 - k_{f4}CW^2/2 \\
&\quad -k_{f7}CFNS + k_{b7}C^2V^2S^2/6 \\
&\quad -2k_{b7}C^2V^2S^2/6 + 2k_{f7}CFNS \\
dV/dt &= -2k_{b0}C^2V^2/4 + 2k_{f0}EMC \\
&\quad -2k_{f2}CHV^2/2 + 2k_{b2}CH^2W^2/4 \\
&\quad -2k_{f3}CV^2/2 \\
&\quad -2k_{b7}C^2V^2S^2/6 + 2k_{f7}CFNS \\
dW/dt &= -k_{b1}HW + k_{f1}CH \\
&\quad -2k_{b2}CH^2W^2/4 + 2k_{f2}CHV^2/2 \\
&\quad -2k_{f4}CW^2/2 \\
dH/dt &= -k_{f2}CHV^2/2 + k_{b2}CH^2W^2/4 \\
&\quad -2k_{b2}CH^2W^2/4 + 2k_{f2}CHV^2/2 - k_{f5}H \\
dF/dt &= -k_{f7}CFNS + k_{b7}C^2V^2S^2/6 + k_d[F](\mathbf{x}) \\
dN/dt &= -k_{f7}CFNS + k_{b7}C^2V^2S^2/6 + k_d[N](\mathbf{x}) \\
dS/dt &= -k_{f6}S - k_{f7}CFNS + k_{b7}C^2V^2S^2/6 \\
&\quad -2k_{b7}C^2V^2S^2/6 + 2k_{f7}CFNS + k_d[S](\mathbf{x})
\end{aligned}
$$

Table 7.2: Differential equations specifying how chemical concentrations change within each simulated protocell (excluding influence of the environment). k_{fn} and k_{bn} represent the reaction rate constants for the nth reaction in the forward or backward direction. $[\rho](\mathbf{x})$ represents the local environmental concentration of the resource ρ.

and summarised in the following equation governing the probability of tumbling of the protocells: $P_{\text{tumble}} = 0.001 * \max(-0.1 + [C]^2 - 0.9[W]^2, 0.01)$.

7.3 EXPERIMENTS AND RESULTS

The goal of these two experiments we now present is to provide a proof of concept of how, in metabolism-based chemotaxis, small changes in metabolism can lead to qualitative changes in behaviour (Experiment 1) and how behaviour can lead to fixation of new metabolic pathways (Experiment 2).

7.3.1 *Experiment 1: A change in metabolism can change behaviour*

In this experiment, we demonstrate how a small change in metabolism can lead to a substantial, qualitative difference in behaviour. In this scenario, one form of chemotaxis (selective-stopping) is transformed into a more sophisticated form (gradient-climbing) through exposure to a new reactant. To do this, we compare two different types of protocells, placing 100 of each type evenly distributed on a petri dish containing at its centre a resource of $M + E$; the concentration of which decays with distance following a Gaussian distribution (indicated in the histograms). The control group starts with only reactant $[C] = 0.5$ which provides a functioning core metabolic pathway. The experimental group is the same as the control group except that it starts with an additional

reactant, $[H] = 1.0$. The presence of this chemical produces a self-maintaining gradient-climbing mechanism by enabling reactions 1 and 2 (see Table 7.1 and Figure 7.1 top-right and lower-right). These two conditions allow us to examine the differences between protocells that have not encountered H (control group) and those that have (experimental group).

Figure 7.2 indicates the behaviour of the control group which demonstrates the selective-stopping mechanism accomplishing a simple form of chemotaxis. The histogram at the top indicates the number of protocells at different distances from the peak resource at the end of the trial, (data taken from 10 trials, each of 100 protocells). The three plots at the bottom of the figure indicate the spatial distribution of the protocells in the petri dish at the start, halfway through, and end of a typical trial. The behaviour of these protocells is a simple result of the metabolism and its influence on motion. In the absence of W, the concentration of C will drive the behaviour of the protocell: if the metabolic activity (i.e., the production of C) is high the probability of tumbling will increase and the protocell will remain in the local area. If C is low the probability of tumbling will decrease and the protocells will move, still in a random walk, but with increasingly long durations of directional movement until C is produced again (e. g., when the protocell finds a place where M and E are abundant). The mechanisms resemble the Ashbian principles for adaptation (Ashby, 1952) except that the system is altering its relation to the environment, instead of reconfiguring itself internally. In this way, behaviour is directly modulated by the rate of metabolic production in a "selective stopping" manner that is beneficial for metabolism: "stay where you are if the metabolism is running sufficiently well, otherwise run". This is the simplest example of what we call *metabolism-based chemotaxis* where the "sensorimotor" pathway is the metabolism itself.

Protocells with $[H] > 0$ are capable of the the more sophisticated "gradient climbing" strategy (widely found in bacterial chemotaxis) whereby the protocells are capable of comparing, as they move, the current concentration of a chemical compound with its concentration the protocell experienced earlier. To explain how this is accomplished, we must describe the dynamics of the new reactant, H. H is auto-catalytic in the presence of C and V, so once a functioning metabolism encounters H, its concentration will be maintained above 0. In this simulation, H performs two roles. It catalyses an equilibration between C and W, ($H + C \rightleftharpoons H + W$) and additionally, in its auto-catalysis, transforms V into W which inhibits tumbling. These equations produce a system that is described conceptually in Figure 7.4 whereby 1) stoichiometry and reaction rates cause W to change more rapidly than C, 2) W and C tend to equilibriate to equal concentrations, and 3) W inhibits the probability of tumbling and C enhances it. These properties produce an adaptive gradient climbing mechanism (adaptive in the sense used by bacteriologists to describe the ability to adapt to a wide range of base levels of stimulus). It can be seen how in both conditions protocells approach the resource centre but H produces a more efficient result due to its adaptation; as is evident when comparing Figures 7.2 and 7.3 where the gradient-climbing protocells move to the highest concentration of resource,

unlike the selective-stoppers that stop when the resources are above a threshold. (In both cases, a secondary peak around a distance of 190 can be observed due to the effect of the petri dish wall).

The net effect of incorporating H into the metabolism is that the protocell is capable of performing *gradient climbing* towards those resources that increase its metabolism. Figure 7.4 illustrates the mechanism: Since W degrades faster than C and is produced twice as fast, the difference between the concentration of both chemicals is a reliable indicator of the rate of metabolism. If metabolism is increasing there will be more W than C, if metabolism is kept constant $[W] = [C]$ and if the metabolic rate is decreasing $[W] < [C]$. Since W inhibits the probability of tumbling and C enhances it, an increasing metabolic rate ($[W] < [C]$) will increase the running probability. If, on the contrary, the protocell is moving down a gradient of a metabolizable compound, its metabolic rate will decrease, $[W] < [C]$, the probability of tumbling will increase and it will change direction of movement.

The experiment shows how changes in the metabolic network of a metabolism-based chemotactic agent can lead to qualitative adaptive changes and improvement on its behaviour, through relatively simple means. While moving through its environment, a protocell can potentially encounter a new component H that is incorporated into the metabolism through its self-catalytic activity and through its capacity to improve the adaptive behaviour of the protocell. The chances of this event happening are enhanced by the self-movement of the protocell. Note that the specific changes that have occurred here have been designed to make the system as simple to understand as possible. The details are, therefore, quite contrived and we are not suggesting that the transformations described have occurred in this way in biology.

7.3.2 *Experiment 2: A change in behaviour can change the metabolism*

In this new experiment we include a second metabolic pathway. In this pathway, energetic and material resources (F and N respectively) are converted into C and V. Like the core metabolic pathway, this is an auto-catalytic production requiring C to be present to occur. However, unlike the core metabolic pathway this reaction is also auto-catalytic with respect to S. This means that S is both produced by the reaction and required for the reaction to occur (it helps at this point to look again at Figure 7.1, especially the bottom-left).

Protocells, (initialised with $C = 0.5$, $H = 1.0$ and $S = 0.0$) are evenly distributed around a petri dish containing two sources of E and M, located at ($x = -75, y = 0$) and ($x = 75, y = 0$). One source of F and N is located at ($x = 0, y = 0$). There is no S in the environment except within a circle of radius 0.5 around the left peak of resource E and M ($x = -75, y = 0$), where $[S] = 1.0$ (indicated by the star).

Figure 7.5 indicates the distribution of the protocells over the course of the simulation. Figure 7.5B is the same as in Figures 7.2 and 7.3, but the histogram now indicates the

A

Distribution of "Selective Stoppers" (H=0.0) relative to peak resource

B
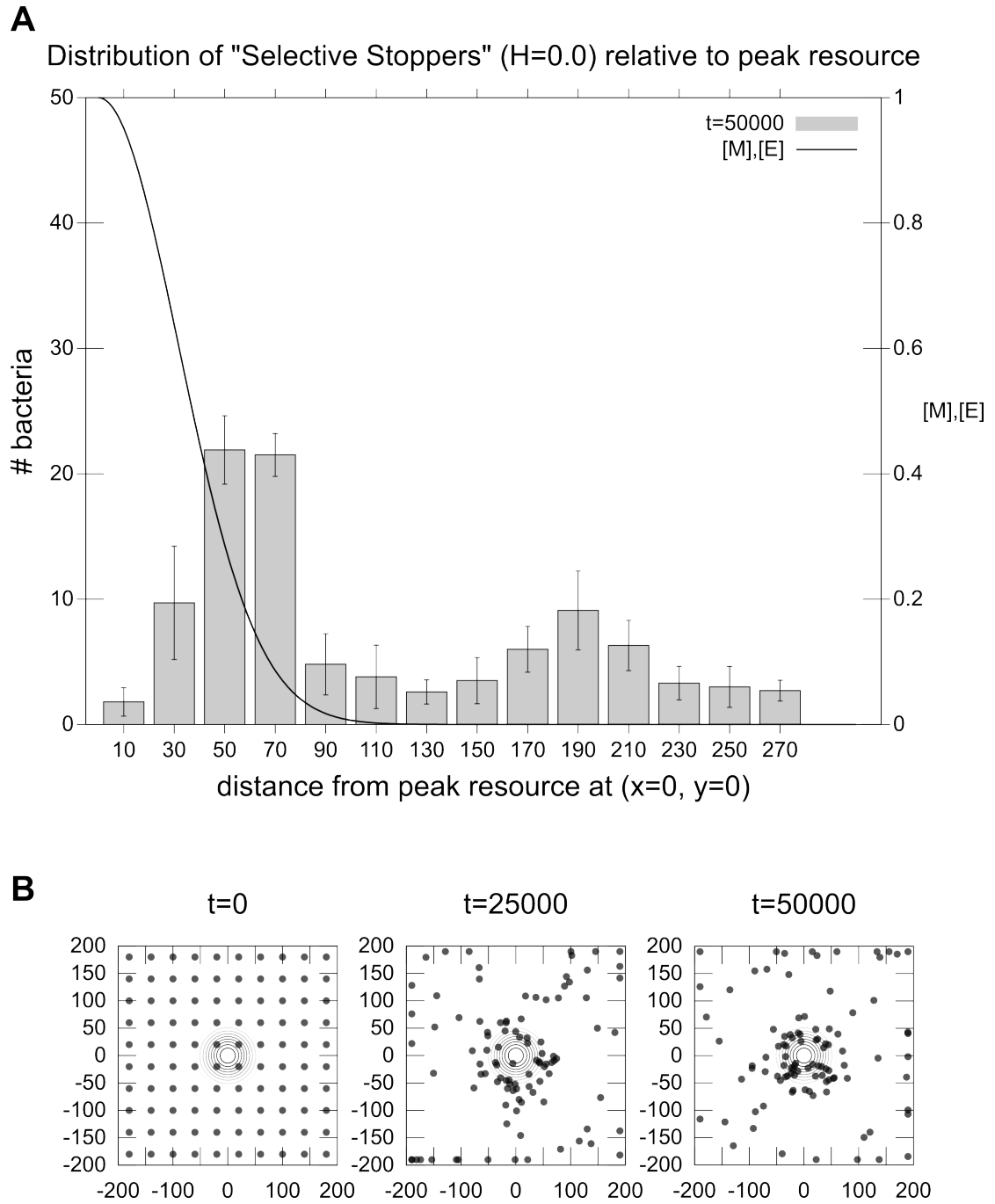
Figure 7.2: Selective-stopping protocells distance from peak resource (A) and spatial distribution (B).

A

Distribution of "Gradient Climbers" (H>0.0) relative to peak resource

B
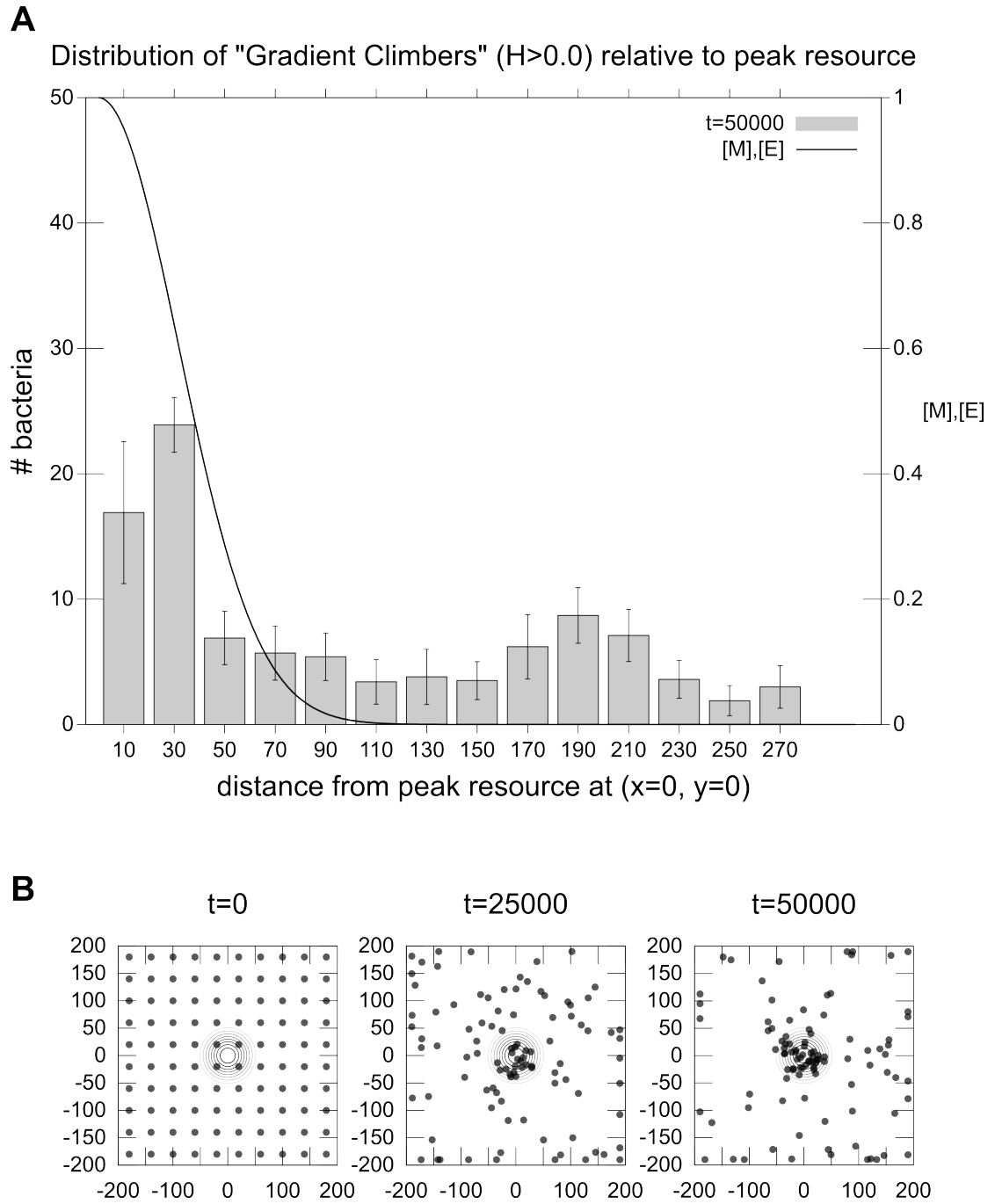
Figure 7.3: Gradient-climbing protocells distance from peak resource (A) and spatial distribution (B).
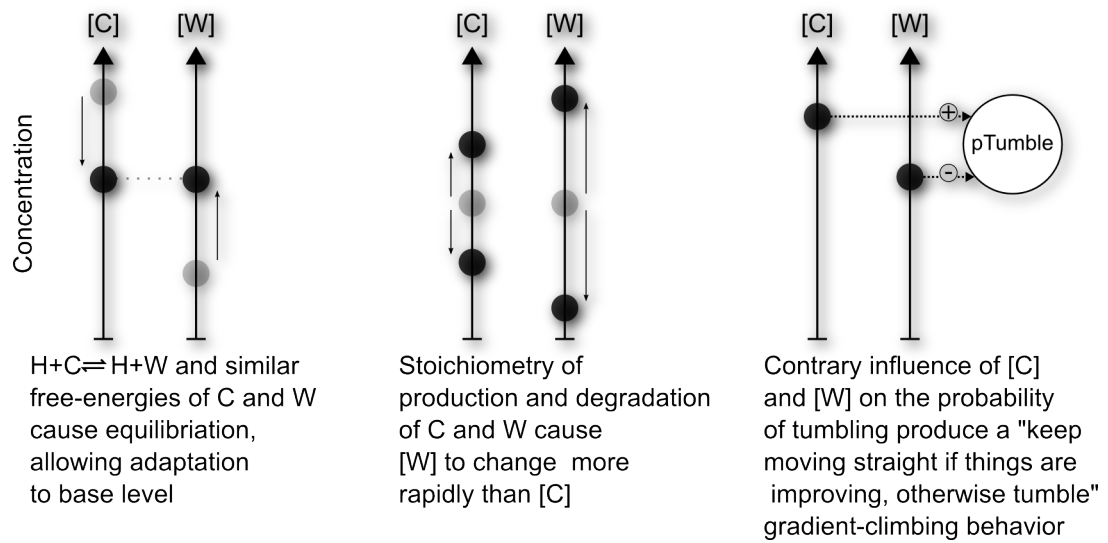
Figure 7.4: Diagram describing the gradient climbing mechanism.

distribution of protocells along the x-axis, comparing the distributions of protocells that have zero and non-zero concentrations of S. Data have been collected at the end of 10 different trials, each of 100 simulated protocells. As before, at the start of the simulation, the protocells are evenly distributed around the arena. The gradient climbing mechanism attracts the protocells to one of the sources of E/M. At this stage, none of the protocells have any S, so F/N is not metabolizable and has no effect on the behaviour of the protocells as the metabolism-based mechanism automatically ignores resources that are irrelevant to the metabolism. As time progresses, protocells tend to gravitate towards the higher concentrations of E/M, and those that are at the left source have an increasingly high chance of encountering the pocket of S. Those protocells that come into contact with S become capable of auto-catalysing S. Their metabolism has been changed and the odds of this change occurring have been significantly influenced by their behaviour. Those protocells with $[S] > 0$ have gained a new metabolic pathway. They are now capable of metabolising F/N and as time progresses, those protocells that through their random walk are brought close enough to "taste" F/N, now also climb that gradient. Protocells that were initially attracted to the right-most source of E/M never encounter S and accordingly never are drawn away from their initial F/N resource source and at the end of the simulation there are in a certain respect two 'species' of protocells – one that consumes and is attracted to both pairs of resources and one that is only attracted to, and only consumes the original pair.

Distribution of bacteria and resources at t=500000
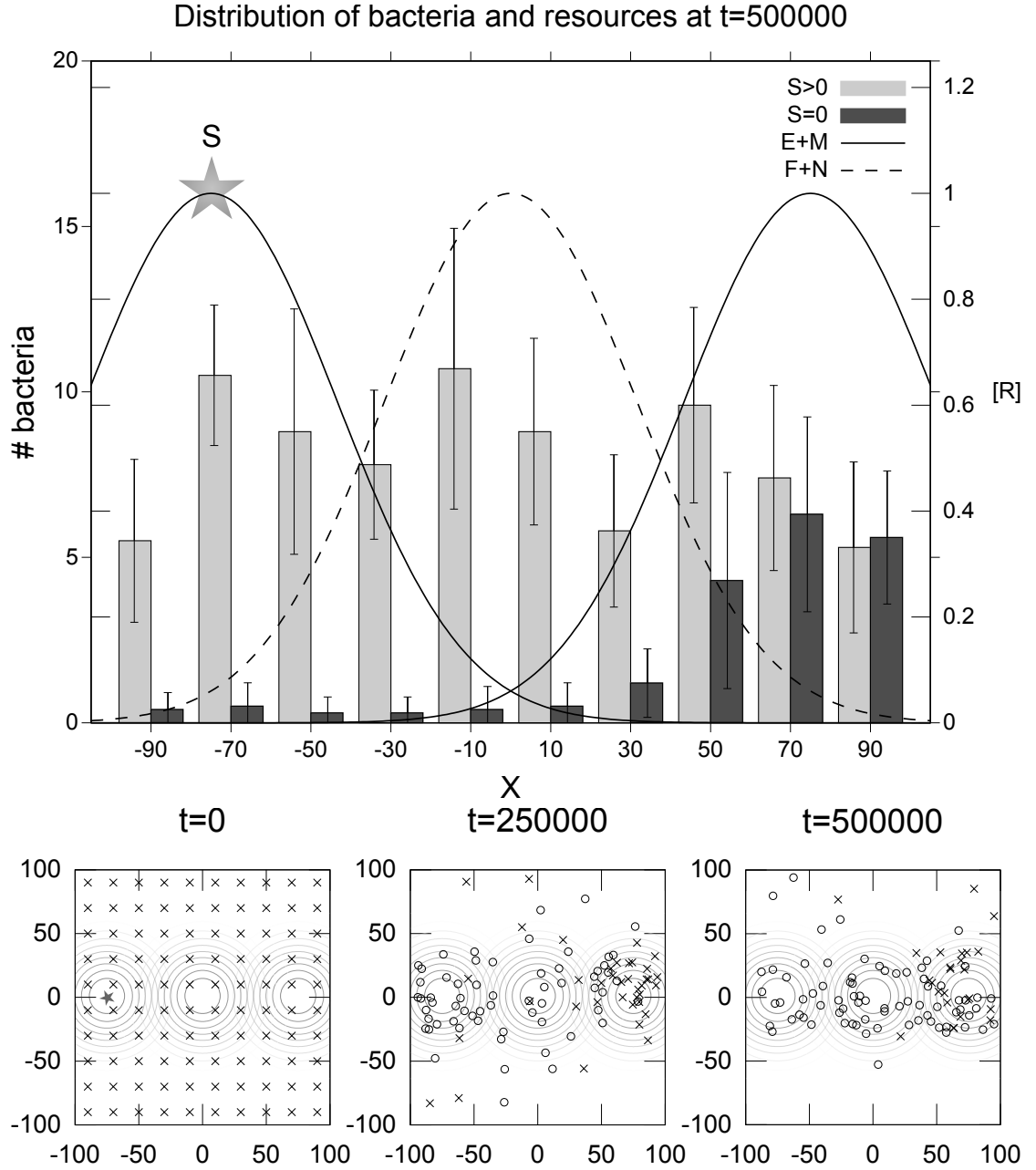
Figure 7.5: Experiment 2. Protocells are initially attracted to sources of $M + E$, but those that encounter the metabolic-path-opening reactant S, automatically become also attracted to new resources $N + F$. $\times$ indicate protocells with $[S] = 0$ and O indicate protocells with $[S] > 0$.

In the experiments described above, we demonstrated how in an organism with metabolism-based behaviour, a change in metabolism can cause an adaptive change in behaviour and conversely, a change in behaviour can instantiate adaptive changes in metabolism through influencing the environment to which the organism is exposed. In this section, we elaborate on how, through this adaptive cycle, an organism with a metabolism-based behavioural system is more "evolvable" – more pliable to adaptive evolution – than an organism that is identical except with metabolism-independent behaviour. There are a few ways that this *facilitation of evolution by metabolism-based behaviour* can occur.

1. Metabolism-based behaviour provides a phenotypic plasticity that facilitates the slower, but more plastic mechanisms of neo-Darwinian evolution – an instantiation of the Baldwin effect (Crispo, 2007) where the phenotypic plasticity is provided by metabolism-based behaviour.

2. Metabolism-based behaviour allows for adaptation not only to external changes (changes in the environment), but also to internal changes (changes in the organisation of the organism itself. This effectively transforms a variety of what would otherwise be deleterious or neutral mutations into beneficial mutations.

3. Metabolism-based behaviour can expose an organism to a greater variety of environments and can increase the likelihood of speciation events.

To provide an example of the Baldwin-effect mechanism, imagine the appearance of a metabolic toxin, never previously experienced by a population of protocells, nor by its ancestors. Metabolism-independent protocells would most likely be insensitive to this toxin, having had no chance to evolve a sensitivity, but metabolism-dependent protocells would automatically avoid the toxin as it damages the metabolism. This aversion could reduce, but not eliminate exposure to the toxin, allowing, over longer time scales for the evolution of an immunity, or more sensitive metabolism-independent toxin-aversion mechanism.

To describe the second two mechanisms, it helps to first recapitulate the adaptive benefits of metabolism-based behaviour at the level of the individual.

7.4.1 *The adaptive power of metabolism-based chemotaxis*

As discussed in Section 4.3.1, adaptive behaviour is generally understood and modelled as optimising some value function or as maintaining essential variables under viability constraints. The dynamics of biology that produce these viability constraints are rarely considered in any depth or included in a model of adaptive behaviour. But, when they are included, as the metabolic dynamics have been included here, there is the potential for

interactions between behavioural mechanisms and the viability-constraint determining dynamics. We have shown here and earlier in the thesis how when metabolic dynamics are coupled to behavioural mechanisms, a number of adaptive phenomena come to the surface that generally pass unnoticed due to the typical abstractions made in adaptive-behaviour models. From the previous experiments we can generalise that, despite its simplicity (or perhaps thanks to it), metabolism-based behaviour can enable a number of powerful adaptive capacities:

1. The metabolic consequences of behaviour can be evaluated online (i.e., in onto-genetic time and in relatively short timescales) and behaviour can be modulated accordingly.

2. Organisms can adapt not only to the presence of specific chemicals but also to other environmental conditions (e.g., temperature) that might influence metabolism.

3. Organisms can adapt not only to changes in the environment, but to changes in their own metabolic organisation by modulating their behaviour accordingly.

4. Organisms can integrate information from the environment and from within, which means that behavioural and metabolic processes of adaptation can feed back to each other.

As a consequence, organisms can adapt (respond appropriately) to various environmental and internal chemical compounds and conditions that were never previously experienced by the individual nor even by any of its ancestors. The system will be attracted to any compound or condition that increases metabolic rate and will be repelled by those that decrease or inhibit metabolism. This does not, of course, rule out potential cases of maladaptation such as parasitic interactions that override the behavioural mechanism or interactions that increase the short-term metabolic production, but damage metabolism in the long-term by e.g., destroying the membrane – though see discussion in Section 8.5.5 about metabolism-based behaviour being less susceptible to being 'fooled' than metabolism-independent mechanisms.

7.4.2 *A system that adapts to* itself

Now let us look at the implications of these and other outcomes of metabolism-based behaviour at the level of the evolution of populations. Figure 7.6 illustrates the case of a mutation (genetic or otherwise inheritable) on metabolic pathways that permits one protocell to exploit and metabolise a new environmental resource. Metabolism-independent chemotactic protocells (left) will remain in place and the benefits of the mutation will pass unnoticed; unless there is an extra-ordinarily unlikely coincident mutation that makes transmembrane receptors sensitive to the new metabolic source *and* generates attraction to it. Genetic drift dictates that, most probably, the *potentially*
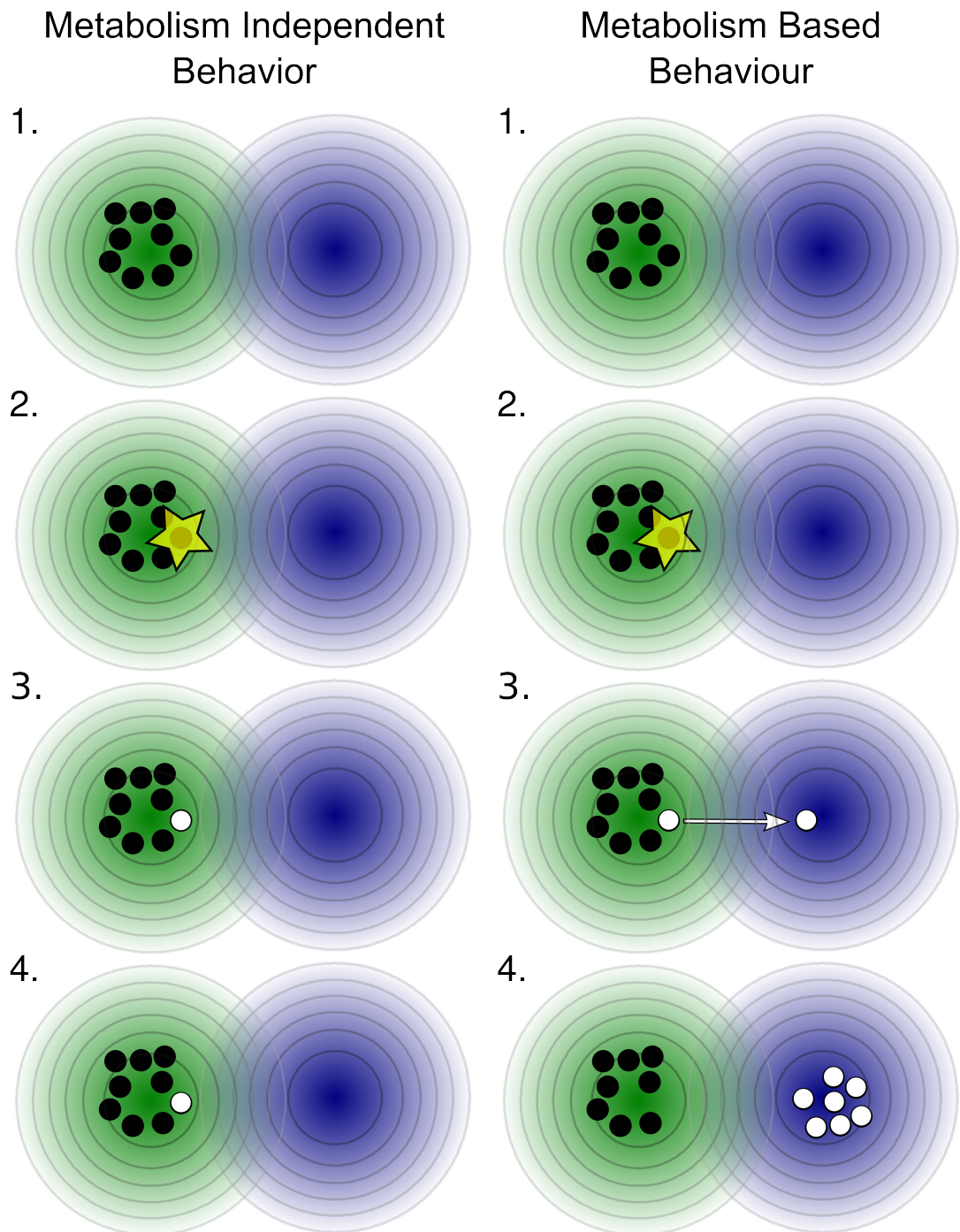
Metabolism Independent
Behavior

Metabolism Based
Behaviour

1.

1.

2.

2.

3.

3.

4.

4.

Figure 7.6: A comparison of metabolism-independent and metabolism-dependent responses to a change in organisation (represented by a star in frame 2) that allows them to consume a new resource (blue gradient on right).

Examples for each type of change
1) A mutation or organization modifying peturbation
2) Chemotaxis to newly metabolizable resource
3) The maintenance of a high concentration of environmental resource X that indirectly influences the concentration of the metabolism's component chemicals.

1) A change in organization

3) That can change organization to suit the new behaviour

2) Automatically produces a corresponding change in behaviour

Time

Behavior   Organization

Figure 7.7: Behavioural metabolution, a cycle of mechanisms contributing to adaptation.

beneficial mutation (ability to metabolise the new resource) will be lost since it has no beneficial effect on the protocell (the protocell cannot use the new ability). Metabolism-based chemotactic agents (right), contrarily, will immediately and automatically be attracted to the new resource (for it benefits metabolism) if they are exposed to it. They will benefit from the mutation by incorporating a new metabolizable resource into their organisation; the mutation is much more likely to be retained and a new population could emerge in the new resource-rich environment, leading potentially to speciation. The abilities of the organism changed, and the metabolism-based behaviour automatically adapted to that change.

We have just discussed how metabolism-based behaviour can produce an in-the-moment automatic adaptation to change. This change that is adapted to could be *external* (e.g. an environmental change), or *internal* — when the organisation or the needs of the behaving system have changed. This adaptation is directed; in the majority of cases, it is likely to produce changes that are beneficial to the organism. This is because the behaviour is driven by a sensitivity to what is beneficial for the organism's metabolism. This, as we mentioned, can make some otherwise neutral mutations more likely to be beneficial mutations. It can also increase the chance of further metabolism-changing events by increasing the variety of environmental encounters experienced by the organism or by influencing the local concentrations of reactants in the environment of the organism. These changes in metabolism could be helpful or damaging, but they seem more likely to be helpful as if they are damaging, the metabolism-based behaviour organism (if it can respond in time) would move away from such metabolism-damaging encounters.

This all culminates in the idea of *behavioural metabolution*: A cycle of interactions between behaviour, metabolism and evolution in which a change in the metabolism of an organism can automatically produce an adaptive change in behaviour that can

cause further changes in the metabolism, possibly inducing further adaptive changes in behaviour etc. This cycle is depicted in Figure 7.7. A change to the organisation of an agent causes it to automatically behave differently, in a way appropriate to its change in organisation. The new behaviour brings the system to a new environment where new mutations (or old mutations) and/or new environmental conditions might be beneficial for metabolism, or as demonstrated in Experiment 1, can produce a new (possibly improved) behavioural mechanism. In this way, a push-me/pull-you dynamic interplay can be established between changes in behaviour and changes in metabolism, influencing evolutionary processes in ways that remain mostly unexplored.

The goal of the above experiments is not to provide evidence for this phenomenon but to show the very possibility and some potential dynamics. Further extensions of the present work shall include an open artificial chemistry with moving protocellular systems that could be used to determine whether the presence of self-movement largely increases the probability of chemical-evolutionary adaptation.

7.4.3 *Behavioural metabolution as proto-evolution*

The adaptability provided by metabolism-based behaviour is powerful, yet extraordinarily simple. Not only does it enable mechanisms of adaptation at the level of the individual, but also, has the potential to facilitate evolution. There is even the possibility that at the very early stages of life, the coupling between metabolism and behaviour could have played a particularly powerful role by instantiating, on its own (and without the presence of a genetic code or even without reproduction), a form of proto-evolution. Let us see how this is possible.

Assuming an origins-of-life scenario where membrane compartments or oil-droplets enclose proto-metabolic reaction networks undergoing natural selection (Shenhav et al., 2005; Fernando and Rowe, 2008; Shapiro, 2006) it is evident how any tendency to move (even randomly) would become beneficial to such systems: local metabolic resources would soon be consumed and random movement would lessen competition for local resources. Any bias of random movement towards metabolically more beneficial environments would rapidly be selected. The selective-stopping chemotactic strategy has been shown to be easily evolvable (Goldstein and Soyer, 2008) and it seems simple enough that it could, sooner or later, appear and be a metabolism-based rather than metabolism independent. Metabolism-dependence seems likely as it is simpler than the metabolism-independent mechanisms and early metabolic networks would tend to be highly integrated and simple — certainly not highly specialised as is necessary for metabolism-independent modes of chemotaxis.

The motion mechanism in our models is an abstract version of a flagellar movement, which is far too complicated to be found at an early stage of evolution. However, at such early stages, movement could be implemented through a wide variety of metabolism-controllable ways – simple reaction-diffusion spots have been shown to be capable of

movement (Krischer and Mikhailov, 1994), and more recent work on self-propelled oil droplets (Toyota et al., 2009) or simpler mechanisms such as modulation of protocell buoyancy or membrane permeability are potential early prebiotic environment modulating behavioural mechanisms.

In any of its possible instantiations, the coupling between metabolism and behaviour could enable behavioural metabolution as an alternative or overlapping evolutionary mechanism. Under this scenario the adaptation enabled through interactions between metabolism and behaviour (described above) precede or supplement early neo-Darwinian evolution. Internal and external variation is selected behaviourally, and behaviour increases the diversity of environmental encounters, which increases the chances for beneficial metabolic chemical avalanches, and the likelihood of speciation.

The combination of both behavioural and metabolic variation and selection offer unexplored possibilities. Behaviour that causes directed exploration of chemistry-rich environment can facilitate the exploration of the 'internal' metabolic space, discovering new metabolic pathways, etc. Meanwhile the 'evolving' metabolism is exploring the space of behavioural strategies.

We have outlined some possible interactions here, but the flexibility of metabolism-based behaviour demands further exploration to establish the likelihood of these dynamics as well as the potential significance of their impact. In this vein, the present model should be extended to include a wide (or even open-ended) set of chemical reactions available as environmental resources and as metabolites. A model of this form (similar in design, but different in focus to the work developed by Fernando and Rowe, 2008) would allow us to explore the capabilities of behavioural metabolution in the early, possibly pre-genetic era scenarios.

Central to the idea of behavioural metabolution and its relevance to early forms of life is the simple yet powerful potential of coupled metabolic and behavioural systems. Metabolism-based behaviour provides the ability to adapt in a variety of ways – not only to changes in the environment, but to changes in its own organisation. These forms of adaptation at the level of the individual have the potential to substantially facilitate, accelerate or perhaps even instantiate mechanisms of population-level evolution-like adaptation.

7.5 CONCLUSIONS

Despite the central role that both metabolism and adaptive behaviour play in artificial life and theoretical biology, very little attention has been paid to the interplay between the two, especially at the ontogenetic and evolutionary scales. When behaviour is not controlled by a subsystem that maximises some function (generally external to the subsystem itself, in the form of selected adaptations or satisfaction of internal "needs") but is, instead, directly modulated by metabolism, then a wide range of adaptive phenomena come to the surface. We have shown, through a model of metabolism-based chemotaxis, how changes

to metabolic pathways can qualitatively improve behavioural strategies (e. g., from a selective-stopping to a gradient-climbing strategy; experiment 1) and how behaviour might serve to explore and fixate new metabolic pathways (experiment 2). These two examples have been used to reveal some fascinating roles that the behaviour-metabolism interplay can play in the context of evolution: forms of metabolism-behaviour coupling could have bootstrapped or driven the evolution of early (pre-genetic) life and could be currently instantiating forms of non-genetic inheritance or genetic assimilation of phenotypic plasticity.

Incorporating this type of connection between behaviour and metabolism opens up a promising line of artificial life research where the long term (evolutionary) consequences of interactions between behaviour, system organisation and environment and can be systematically studied in simulation.

7.6 RECAPITULATION

In this chapter, we have shown that in addition to the adaptability provided by self-sensitive behaviour at the level of the individual, forms of self-sensitivity can facilitate, or perhaps even instantiate forms of population adaptation, supplementing or perhaps even preceding neo-Darwinian evolution.

This argument was supported by a computational model in which we demonstrated how metabolism-based behaviour can increase the variety of encounters with the environment, and how it is possible for metabolism to be changed through such encounters such that behaviour is qualitatively improved. These act as demonstrations of aspects of the idea of *behavioural metabolution* in which behaviour can automatically adapt to changes in the metabolic organisation of the behaver, and changes in the behaviour of the organism can increase the variety of environmental interactions, increasing the likelihood of further metabolic changes, some of which will be beneficial.

PRIMARY POINTS OF CHAPTER

- Benefits to the individual provided by metabolism-modulated behaviour, some of which were introduced or demonstrated in previous chapters are recapitulated and discussed.

- These properties of an individual performing metabolism-modulated behaviour can facilitate adaptive evolution in a variety of ways.

- The idea of 'behavioural metabolution', a synergistic interaction between processes of behaviour, metabolism and evolution is introduced and discussed. Behavioural metabolution may have played a role in bootstrapping the early evolution.

- Aspects of behavioural metabolution were demonstrated in simulation, but work remains to evaluate the likelihood of these processes occurring in nature.

8

FRACTALS OF AUTONOMY

"Clearly, form and function in amoeboid cells do not depend on a particular set of molecular players...Any suite of molecules will do, so long as they can be articulated into cellular structures that support the function at hand." (Harold, 2003, p. 142)

CHAPTER SUMMARY

Chapter 6 discussed in detail the notions of metabolism-independent, metabolism-dependent, and metabolism-based chemotaxis. Also discussed in that chapter were the advantages of metabolism-based behaviour. The reader may recall, however, that there is strong evidence of behavioural mechanisms that are independent of metabolism (e. g., Adler, 1969). Why should there be such metabolism-independent mechanisms? What is gained from metabolism-independence? This chapter describes some benefits of metabolism-independence of behavioural mechanisms.

The chapter starts by introducing the idea of *decoupled* behavioural systems and the dynamic nature of operational closure. Then, a model is introduced that provides an example of behaviour partially decoupled from metabolism. The model acts as an didactic example, helping us to understand how a system can maintain a 'core' operationally closed identity, while being capable of temporarily becoming part of or separating from larger operationally closed networks of processes, a form of *dynamic operational*

closure. The model demonstrates some of the variety of possible relationships between mechanisms of autopoiesis and behaviour. Other possible relationships are discussed, considering i) the degree to which the behaviour benefits/facilitates autopoiesis and ii) the degree to which the organisation of the mechanisms is integrated, i. e. are the processes that produce the behaviour the very same processes that perform autopoiesis, or are they completely distinct, or somewhere in between?

This discussion leads to the idea of *autonomous behavioural patterns* whereby a set of processes that causes a behaviour are self-maintaining, but their self-maintenance is independent of the self-maintenance of the organism that is performing the behaviour. Such behavioural mechanisms can be in conflict with self-production, providing a positive (i. e. without resorting to the system being broken or 'short circuited') account for unhealthy behavioural patterns.

8.1 DECOUPLING BEHAVIOUR AND METABOLISM

Excluding the work presented in this thesis, most computational models of autopoiesis demonstrate only self-maintenance but no behaviour. Generally the simulated autopoietic entities exist in an environment which requires no organism-scale action to continue to exist (e. g. Varela et al. 1974; McMullin 2004). A few more recent models have demonstrated agents performing a slightly extended autopoiesis; extensions such as incorporating a behaviour such as osmotic crisis avoidance (Ruiz-Mirazo and Mavelli, 2007) or chemotaxis (Suzuki and Ikegami, 2009). In these cases, the majority if not all of the added behaviours are actually *extensions of the mechanisms of autopoiesis* – they are inseparable from the autopoiesis. To stop the mechanism of behaviour is to stop the mechanism of autopoiesis. However, this is not the case for the majority of behaviours observed in nature. Most behaviours can stop, perhaps starting again later while autopoiesis continues.

Moreno and Etxeberria (2005) and Barandiaran and Moreno (2008) have argued that the mechanisms underlying the production of behaviour (regulated coupling with the environment) have acquired complexity in the history of life through a succession of 'decouplings' from underlying metabolic levels.

How can mechanisms of behaviour can be related to, but somewhat independent of the mechanism of autopoiesis? How can behaviour be integrated with and yet "decoupled" from mechanisms of autopoiesis?

8.2 DYNAMIC OPERATIONAL CLOSURE

We described the idea of operational closure in some detail in Section 4.2.2. It plays an important role in parts of this chapter, so we shall briefly redescribe the idea here.
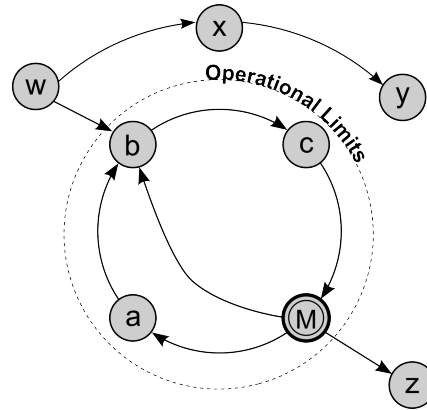
Figure 8.1: A schematic of an operationally closed system and its environment. Filled circles represent processes and arrows represent dependencies. The processes within the dashed circle fit the criteria for operational closure and are thus part of an operationally closed system.

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. The aim of this definition is to provide a framework for the notion of a self-sustaining system. Other expressions for this concept can be found in (Varela, 1979; Thompson, 2010; Di Paolo, 2009) and elsewhere.

Given a collection of processes **C**, we can identify an operationally closed subset of those processes, **S** such that for every constituent process *P*, the following conditions are true.

1. Another process P' requires conditions produced by process *P*

2. Process *P* is conditioned by another process P''

3. P' and $P'' \in$ **S**

4. P' and P'' can be (but are not required to be) the same process.

With these criteria, given a set of processes and their relationships of condition, it is possible to identify operationally closed networks of processes.

So, given a collection of processes and their inter-dependencies, it is possible to identify operationally closed sets. However, the situation becomes less well defined when we recognise the dynamic nature of inter-dependencies. Process A might depend upon process B *now*, but not later (or vice versa). The point we wish to underline here is that system inter-dependencies can change, not only because the processes are dynamic but also because relationships of contingency change. It is already established that the realisation of an operationally closed system can also change over time but, in addition to these structural changes, the relations between the constituent processes of an operationally

closed system may also vary in important ways. As we shall show in this chapter, an operationally closed system can recruit behavioural processes, while conserving operational closure. Perhaps understanding this dynamic aspect of operational closure can help us understand the relationship between autopoiesis and behaviour, why behaviour is often decoupled from metabolism and how to build systems with behaviour that is partially decoupled from metabolism.

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 and 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 organisation of processes which may include physical, chemical, mechanical and/or behavioural 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.

8.3 MODEL

The model presented in this chapter was, in fact, the first model that we developed in this collection of work. It is, accordingly, quite exploratory, and not the minimal model necessary for the conclusions presented. Nevertheless, we present the full model here, as conclusions drawn from our experimentation with it (outlined below and described in detail later in this chapter) are interesting and the model suggests directions in which this research could be extended – e.g. through exploring more complex membrane and metabolism processes such as phagocytosis.

As with other models presented in this thesis, a primary motivation in the construction of this model was to produce a simulation that incorporates behaviour and autopoiesis in a single, unified model. As before, the autopoiesis and the chemotaxis of the protocell are both the result of the same type of dynamics: 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 as there is in other types of models such as typical evolutionary robotics (see discussion of this in Chapter 2).

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 protocell' (see Figure 8.2). We use the model to demonstrate the following four ideas concerning the potentially complex and dynamic relationship of inter-dependence between metabolism and behaviour.
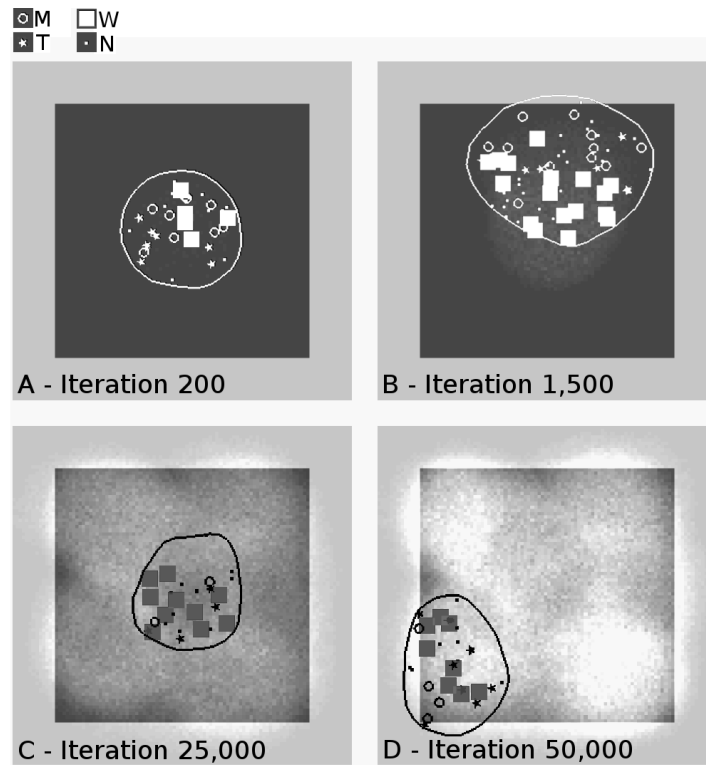
Figure 8.2: A time series of a non-lesioned protocell. To improve readability, only 10% of the enzymes are drawn. See text for a detailed description of the processes involved. A) Soon after the start of a trial, stochastic processes have already produced some small, random motion in parts of the membrane. B) The protocell has begun moving upwards. Note the asymmetry in concentrations of M and W. Also note the increased size of the protocell due to the high local levels of R. C) After having visited most parts of the arena, the protocell returns to the centre where the R is now high again relative to the protocell's previous location. Note the protocell has shrunk due to the decrease in availability of R. D) The end of a typical run for a non-lesioned protocell.

1. Behaviour is dependent upon metabolism, but also metabolism can be strongly dependent upon behaviour.

2. *Dynamical operational closure*: how an organism can maintain an operational identity, while forming (and breaking) larger operationally closed systems with the environment.

3. How behaviour and autopoiesis can be *partially* decoupled through dynamic operational closure.

4. Some advantages and disadvantages of partial decoupling.

Before describing the various reactants and reactions within our model in more detail, we should point out that the modelling of chemical reactions in this model is more *ad hoc* than in the previous models. For instance, we use some short-cuts for chemical reaction rates rather than including the details of all of the chemical reactants and reactions that would be necessary to accomplish the dynamics demonstrated. We feel these abstractions are justified as this model is primarily an exploration of the general concepts and not of the particulars of the reaction network. To include the full reaction networks would only obfuscate the model.

8.3.1 *Resource*

The protocell exists in a 2D, infinite environment initialised with a patch of 'resource' (R). Reactions 1a and 2a in Table 8.1 show the use of R in the production of M, an enzyme that is fundamentally important in the maintenance of the membrane. Thus, the protocell 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 protocell.

The initial distribution of R is a square of concentration $1000R/unit^2$, with dimensions 1200 x 1200 units (about 3 times the diameter of a protocell) with a border of lower concentration (500) of width 200. This distribution of resource was selected to be large enough to observe the protocells behaviour in a healthy environment, but limited enough to determine how the protocell responds as it consumes R and finds itself in an increasingly R impoverished environment. As the protocell consumes R in the production of M, the distribution of R changes. This can be seen in the background values of Figure 8.2 which indicate higher concentrations of R in darker colours. Other than consumption by the protocell, nothing affects the distribution of R.

R can be thought of as a high energy resource that is consumed in the transformation of precursors (P_1 and P_2 discussed below) into enzymes. This is similar to the coupled material and energetic flows discussed in our requirements for a minimal metabolism Section 6.3.1. R plays the role of the energetic resource and the precursors play the role of material resources.

Rate	Reactants		Products	
$2 \times 10^{-4}[R]$	$M + P_1 + R$	$\Longrightarrow$	$M + M$	(1a)
0.15	$M + M$	$\Longrightarrow$	$M + P_1$	(1b)
$2 \times 10^{-4}[R]$	$W + P_1 + R$	$\Longrightarrow$	$W + M$	(2a)
0.15	$W + W$	$\Longrightarrow$	$W + P_1$	(2b)
0.4	$T + P_2$	$\Longrightarrow$	$T + T$	(3a)
0.15	$T + T$	$\Longrightarrow$	$T + P_2$	(3b)
0.4	$N + P_2$	$\Longrightarrow$	$N + N$	(4a)
0.15	$N + N$	$\Longrightarrow$	$N + P_2$	(4b)
$f([R], [A])$	$T + M$	$\Longrightarrow$	$T + W$	(5)

Table 8.1: Metabolic reactions. In the rate column, $[X]$ represents the local concentration of reactant X. See main text for further details.

8.3.2 *Enzyme properties and chemical reactions*

There are four particulate enzymes, M, W, T, and N and two particulate precursors, P_1 and P_2. These are all simulated as 2D points in Brownian motion. Each iteration, a value selected from a Gaussian distribution (mean $= 0$, std $= 20$) is added to each of the spatial coordinates of each particle. In addition to these enzymes that are modelled as spatial particles, we have included A, which is assumed to diffuse very rapidly, and is thus simulated in the less computationally expensive method of representing global concentration within the membrane by a scalar value.

Precursors P_1 and P_2 are reactants that are transformed into enzymes by autocatalytic reactions (reactions 1a, 2a, 3a, and 4a in Table 8.1). We included two types of precursor model to facilitate our analysis of the model. Specifically, part of our analysis involves lesion studies, in which we systematically remove reactants to study their influence on the system. By using two precursors instead of one, we can study the influence of these enzymes without the indirect influence of competition for resources in their autocatalysis. We can remove 'behavioural' enzymes such as T and N, and study their influence upon the concentration of 'metabolism' enzymes such as M and W and know that this influence is not due to a change in the resources available for autocatalysis. We can be confident of this because the behavioural and metabolic enzymes are produced from different precursors (behavioural enzymes are produced from precursor P_1 and the metabolic enzymes are produced from precursor P_2).

Each of the autocatalytic reactions can run in reverse (reactions 1b, 2b, 3b, and 4b). In other words, the enzymes can all degrade back into their material precursors. Note however that the energy resource R is not produced by the backwards reaction. The relative rates of reaction were selected to adhere to the laws of thermodynamics (i.e. not produce the chemistry equivalent of a perpetual motion machine) and to produce an equilibrium for each reaction pair such that typically a small quantity of precursor would be available for the autocatalytic reactions at any given location within the protocell.

Chemical	Name	Description
R	Resource:	An energetic resource that is used up in enzyme-producing reactions and never replaced.
M	membrane Motion:	Adds phospholipids to the membrane. Produces a local outward acceleration
W	membrane inverse-Motion:	Adds phospholipids to the membrane. Produces a local inward acceleration
A	Averager:	Stores a representation of the mean quantity of resource experienced by the protocell as a whole
N	seNsor:	Maintains A through local interactions with R
T	Transformer:	Transforms M into W when the local concentration of R is lower than the concentration of A
P_1	Presursor 1:	A precursor particle that can be transformed into M
P_2	Precursor 2:	A precursor particle that can be transformed into T or N

Table 8.2: Summary of enzyme roles.

Additionally, all non-precursor particulate enzymes can spontaneously degrade into precursor. Every iteration, there is a $p = 0.001$ chance that an enzyme will degrade into its precursor (M and W degrade into P_1 and the other non-precursor enzymes degrade into P_2).

Particle roles

The reactions between A, R, T, M, W and N produce a comparison mechanism. Summarised loosely, the protocell keeps track of the average concentration in its local area of "resource" R by varying the concentration of A ("average"). It uses this value to transform the membrane outward-motion-inducing particle M into a form that causes inward-motion, W in areas of the protocell that are low in concentration of R relative to the average concentration of R throughout the protocell. This produces an asymmetrical distribution of M and W that induces chemotaxis. Table 8.2 provides a reference for the roles of each chemical in the model. We now describe the mechanism in detail.

Particles M and W interact with the membrane, causing it to grow and to move. M causes a local motion of the membrane toward the outside of the protocell and W causes a motion toward the inside. Chemical reactions involving N, T, and A result in an asymmetrical distribution of M and W that result in various chemotactic and other behavioural phenomena. How does this occur? The concentration of A, a rapidly defusing reactant that is therefore distributed evenly throughout the protocell, correlates with the average concentration of R recently encountered by the protocell. This correlation is maintained by N which increases the concentration of A in areas that have greater concentration of R than A and decreases the concentration of A in areas where the

concentration of R is lower than A. N is distributed evenly throughout the protocell. Its influence upon A therefore produces a correlation between the concentration of A and the average concentration of R throughout the protocell. This influence of N upon A operates according to the equation below where $[R]_n$ is the local concentration of R at the location of the nth N enzyme. A rapidly diffuses and is therefore modelled as a scalar value representing its concentration throughout the protocell. This is the only reaction that affects the concentration of A.

$$\frac{dA}{dt} = \sum_n 1 \times 10^{-3}([R]_n - A) \tag{8.6}$$

T transforms M particles into W particles. The rate of this reaction is determined by the relative concentrations of R and A according to the function $f(R, A) = 0.9 \cdot H(A - R)$ where $H()$ is the unit step function. This reaction can be thought of as a comparison mechanism. Where the local concentration of R is lower than the concentration of A, there is a high probability of T transforming M into W. This mechanism is inspired by chemotactic biochemical pathways found in bacteria but is different from those discussed in previous chapters, in that it compares the concentration of a chemical in one part of the protocell to its concentration in other parts of the protocell. This is in general not an effective strategy for bacteria as their minuscule size and the effects of thermal motion makes it difficult to discern a gradient by comparing such proximal chemical concentrations, but the underlying principle of a differential reaction based upon a comparison of one chemical concentration to another is an established phenomenon in bacteria.

Computational simulation of particulate reactions

To simulate the enzyme-enzyme reactions described in Table 8.1, every iteration, the arena is divided up into a grid of 'pockets', each containing the particulate reactants 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 8.1) that the reaction will occur, i. e. that the reactants will be replaced by the products.

This method of simulating 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 1000 T (and sufficient P_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, in conjunction with the bi-directionality of the autocatalytic reactions, means that if one of the autocatalytic enzymes reaches a high concentration, the 'backward' reaction (breakdown of the autocatalyst into precursor) becomes more likely than the 'forward'
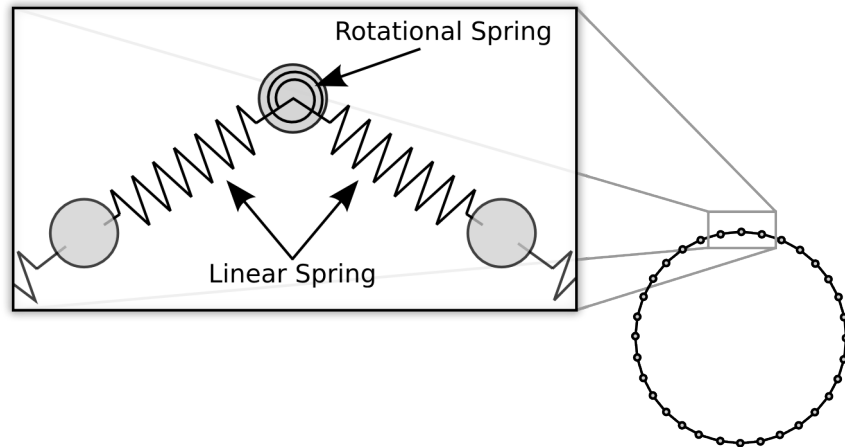
Previously printed in Egbert and Di Paolo (2009).

Figure 8.3: Schematic of a membrane-section.

reaction. In this way, the system does not transform all precursor to enzymes, but instead finds an equilibrium state that includes precursor and product.

8.3.3 *Membrane dynamics*

A flexible, semi-permeable membrane is modelled as a circle of 32 membrane-sections. Each section is modelled as a mass-point with associated linear and rotational springs connecting it to its neighbouring points (see Figure 8.3). The rest-length (l) of the linear springs and the mass of the membrane-section is proportional to the number of phospholipids in the membrane-section (ϕ): $l = \phi/100$. The linear springs apply a force $F = -kx$ to the mass-point and its neighbours (where k is the spring constant and x is the distance that the spring has been displaced from its rest-length). Similarly, the rotational spring applies torque $\tau = -\kappa\theta$ (where κ is the spring constant and θ is the displacement of the spring from its rest angle) to the associated mass-point and its two neighbours. The rest-angle for each of the rotational springs is $\pi \cdot 15/16$, i. e. a slightly smaller angle than $180°$ between membrane-sections, resulting in an overall circular membrane at rest. These forces can be thought of as an approximation of membrane rigidity.

Each membrane-section degrades at a rate proportional to the number of phospholipids in the membrane-section. The particular rate of degradation was selected to produce a system that degrades rapidly in the absence of the system's metabolism, but is stable when the protocell is has sufficient resources to healthily self-maintain. Left unchecked, the degradation of the membrane causes the protocell to eventually shrink to a size too small to maintain the populations of enzymes. This is the equivalent of death within our model.

Phospholipids diffuse between neighbouring membrane sections. The following equation describes how the diffusion and degradation cause the number of phospholipids, ϕ,

to change over time. The first term captures degradation, and the second term captures diffusion between neighbouring membrane sections. In this equation, $k = 0.01$ is the rate of diffusion and ϕ_x represents the number of phospholipids at mass point x.

$$\frac{d\phi_x}{dt} = -\phi_x \cdot 10^{-6} + k\big((\phi_{x-1} - \phi_x) + (\phi_{x+1} - \phi_x)\big) \tag{8.7}$$

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 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 impart a small force to the membrane section, perpendicular to the tangent of the surface of the membrane. M imparts a force towards the outside of the membrane and W imparts a force is towards the inside of the protocell. These interactions between enzymes and membrane cause the deformations of the membrane as well as the motion of the protocell.

In addition to the forces applied by impacts with M and W enzymes, membrane-sections are subjected to a drag force, proportional to the square of their velocity, approximating the drag that would be present for a cell in a viscous medium. The following equation indicates how the position of membrane-sections changes according to drag force and the influence of particles M and W (indicated by I_M and I_W).

$$x'' = \frac{dv}{dt} = \frac{-0.5v^2 + I_M + I_W}{\phi \cdot 10^{-4}} \tag{8.8}$$

The membrane is permeable to all other particulate enzymes, but the enzymes degrade immediately to their precursors when outside of the membrane.

8.3.4 *Initial conditions*

We initialised the model described above with a circular membrane of radius 300, placed in the centre of a square arena, with 10 particles each of T, N, and M placed inside the membrane. Also, P_1 and P_2 particles were randomly distributed around the arena at a density of 0.001 particles per unit square for each type. The arena is initialised 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 this border, the concentration of R is zero.

To expose the interactions between the mechanisms of behaviour and of autopoiesis, we performed three lesion studies of our protocell. In the first lesion study, we completely remove N, ('seNse' – the chemical that maintains the correlation between A and the average concentration of R). In the second study we remove N and T ('Transformer' – the chemical that changes M into W). These removals consist of transforming all of the pertinent chemicals into their precursors after an initial settling period of 1000 iterations which was identified, by observing non-lesioned trials, as sufficient time for the system to relax from the initial conditions. Both N and T are only produced by autocatalysis, so when they are removed from the simulation, their population remains zero for the duration of the trial.

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

We provide the following names for the lesioned protocells. The protocell lacking both T and N is called 'metabolism-only', as the removal of these enzymes make all reactions irrelevant except those involved in the autocatalysis of M, i.e. those that act as a metabolism, maintaining a non-zero concentration of enzymes and the membrane. The protocell lacking N we call 'senseless' as by removing N, we remove the system through which the protocell "senses" the change in the average concentration of R in its environment.

8.4.1 *Dynamics of non-lesioned protocells*

Initially the protocell remains momentarily stationary (see Figure 8.2A). 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 protocell extending into an area higher in concentration of R than those areas that have been depleted by the protocell. N particles in this area increase the concentration of A. Then, T transforms M into W in the parts of the protocell 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 protocell away from the area low in concentration of R. Figure 8.2B shows an protocell 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 8.4.

As the protocell moves, it continues to reduce the local concentration of R. Areas that the protocell has occupied for a longer period of time are more depleted of R. This causes areas towards 'the rear' of the protocell to be lower in R which causes further production of W which propagates the motion of the protocell. In this way, the protocell moves

Path travelled

Single Trial   10 Trials
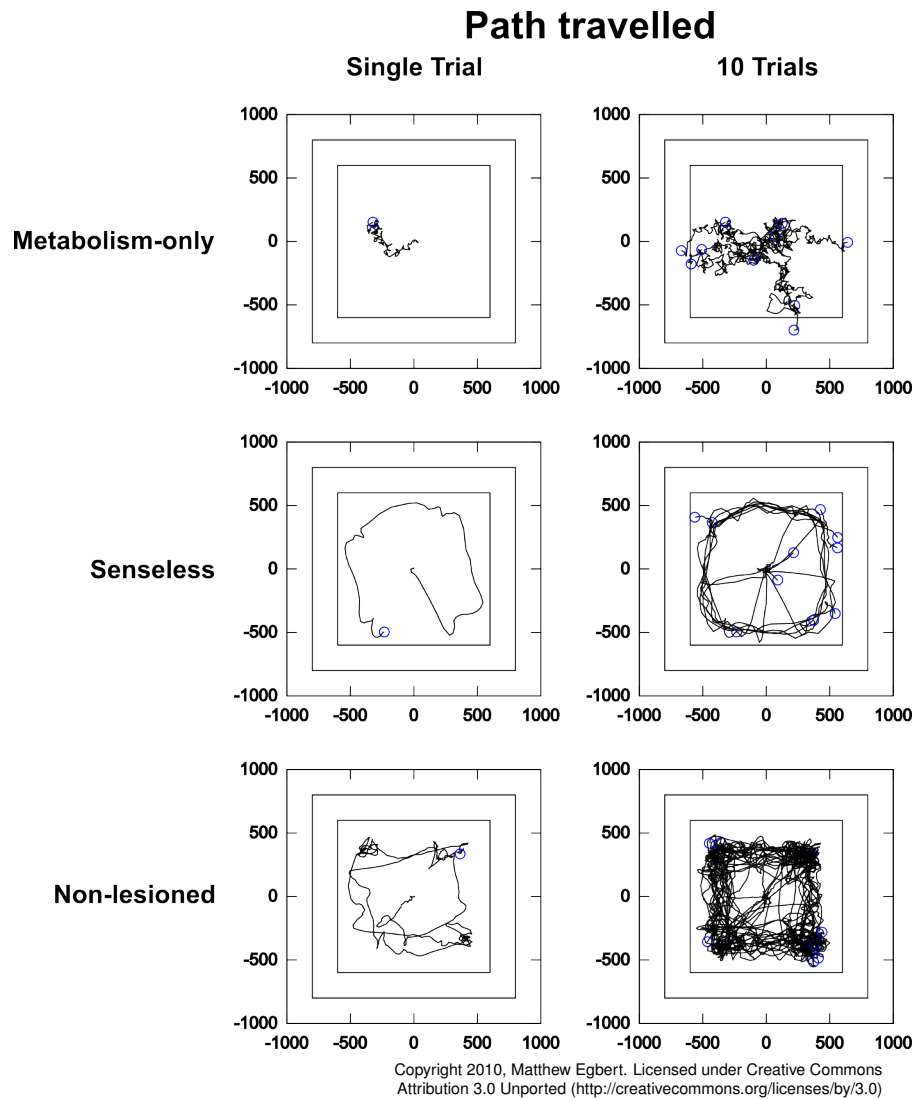
**Figure 8.4:** Paths of non-lesioned (bottom) and lesioned protocells (middle and top). Paths are indicated for a single trial (left column) and for 10 different trials (right column). Circles indicate the location of the centre of the protocell at the end of the trial. Note the doubling back in the non-lesioned protocell and the avoidance of previously visited areas by the protocell that has T but not N.
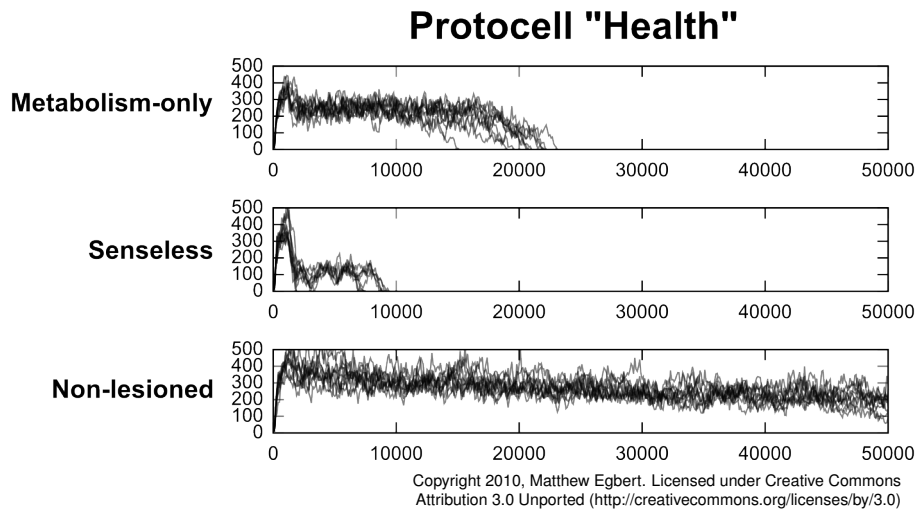
Protocell "Health"

Figure 8.5: Health, as indicated by the sum of the number of M and W particles. Plotted against time for non-lesioned (bottom) and lesioned protocells (middle and top). Data is taken from the same 10 trials as in Figure 8.4.

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.

This directed motion is a good demonstration of how local molecular interactions can acquire a global coherence, creating a spatial asymmetry that influences local conditions, enabling the continuation of the protocell's motion. This is an example of many local (small-scale) interactions producing a more global (i. e. larger scale) dynamic that feeds back into the the local interactions maintaining the conditions necessary for the processes to continue. Although the underlying mechanisms are quite different, these kinds of local-to-global / global-to-local reflexive interactions have been found in other minimal systems capable of generating self-movement in a homogeneous spatial situation, e. g. self-moving oil droplets studied by Hanczyc et al. (2007); Hanczyc and Ikegami (2010). In their physical system and in our model, the conditions that cause movement are maintained by the movement itself.

The lower left plot in Figure 8.4 shows the motion of the centre of a non-lesioned protocell 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 returning to areas where it has been before. This behaviour results in an efficient consumption of resource. The protocell tends to move towards whatever proximal area has the greatest density of R. As the resource in the environment is consumed, the total amount available decreases, and the protocell will return to areas that were previously not as resource rich, but are now better than other available options. The results of this behaviour can be seen in Figures 8.2 and 8.6. The background of Figure 8.2 shows how the concentration of resource changes during a typical trial of the non-lesioned protocell, with darker tones indicating higher concentration of resource. By iteration $25,000$, the agent has moved around the environment consuming resource everywhere, reducing levels in the

Distribution of resource
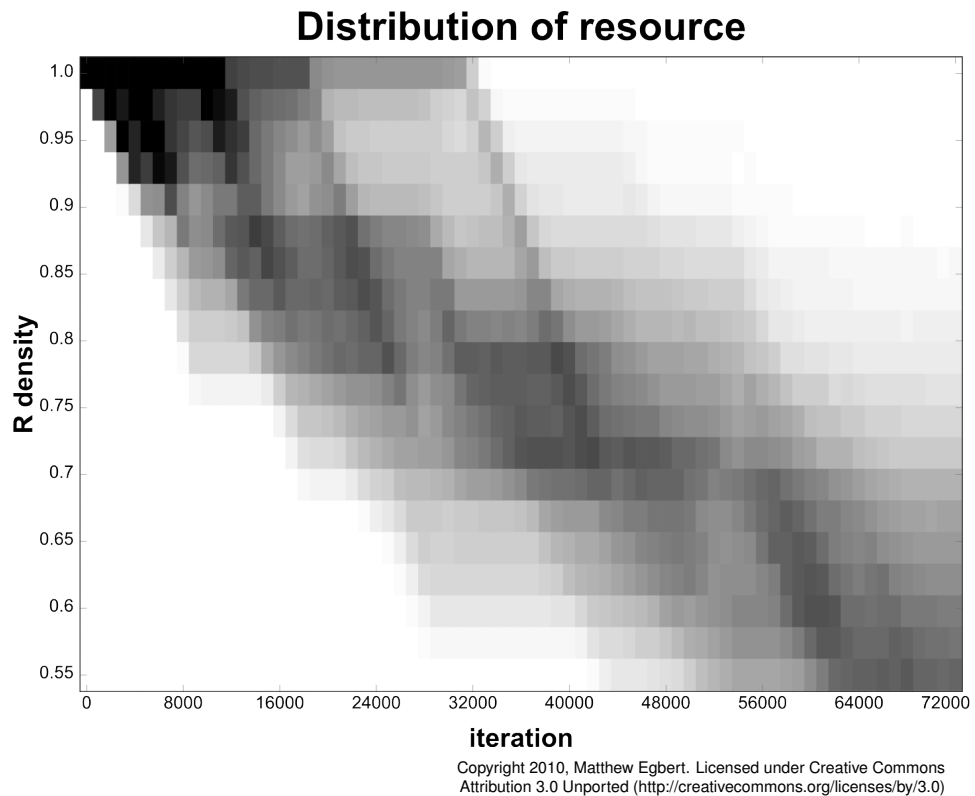
Figure 8.6: The distribution of R density plotted against time for the non-lesioned protocell. The darker the square, the greater proportion of the area has the indicated R density.

central high-concentration area to roughly even amounts. A further $25,000$ iterations later, the agent has continued to consume resources throughout the environment, again in a roughly even manner. This even consumption of resources can be seen more easily in Figure 8.6 which shows how the concentration of resources in the centre 1200×1200 changes over time. At the start, all of this area has resource concentration of 1.0 (dark square in upper left corner) and over time the agent consumes resource. The agent tends to consume areas of higher resource before moving on to areas of lower concentration; by iteration $33,000$ all of the area of concentration 1.0 has been consumed, before the agent begins to consume areas of $[R] < 0.9$. This pattern can be seen a second time: after the agent consumes the majority of areas of concentration ≈ 0.9, it begins consuming areas of $[R] < 0.75$. This data corresponds to the path taken by the non-lesioned agent indicated in the lower left plot of Figure 8.4

What produces this behaviour? The chemical N updates the concentration of A, maintaining a correlation between it and the average concentration of R recently encountered by the protocell. The concentration of A is used as a basis of comparison for the transformation of $M \rightarrow W$ by T. If T is in an area of the protocell where the local concentration of R is lower than that of A, then it transforms M into W. Thus, areas that are lower than average in concentration of R will have more W, and regions higher than average in concentration of R will have more M. W moves the membrane inwards and M moves the membrane outwards, and the asymmetrical distribution of these chemicals results in

Distribution of resource
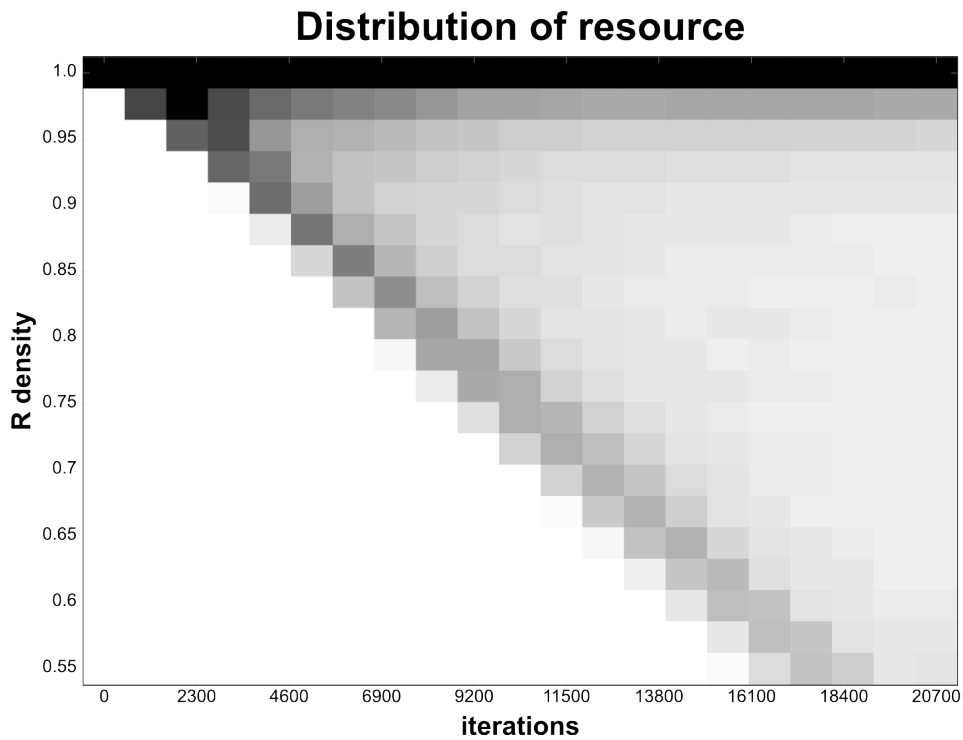
Figure 8.7: R density plotted against time for a protocell lacking N and T. Brighter values indicate that a larger portion of the arena has the indicated R-concentration.

more outward motion in areas high in concentration of R and more inward motion in areas low in concentration of R. In this process, N is responsible for the protocells ability to detect the average concentration of R recently encountered. Let us now look at what happens when we remove N from the protocell.

8.4.2 *Dynamics of 'senseless' protocells*

Removing all N enzymes produces what we call a 'senseless' protocell as removal of this enzyme makes it impossible for the protocell to maintain the correlation between A and R– it makes it impossible for the protocell to "sense" the change in the average R that the protocell is exposed to. In other words, lesioning the N enzymes causes the concentration of A to become fixed, which speaking loosely is like the protocells losing their ability to change their standards. Their local environment is no longer compared to the average of the recent exposure to R, but is instead compared to their environment as experienced at the time of lesioning, when R is still at a relatively high concentration. With A fixed at a relatively high value, the protocells move around the arena, consuming some R at each location until the R-concentration falls below the standard of comparison A.

  After moving around the arena once, the entire area has lower concentration of R than the protocell's inappropriately high concentration of A (inappropriate when compared to a non-lesioned protocell). This will cause T to change *all* M to W preventing any

asymmetry. At this point, the protocell remains where it is, eventually consuming all local resource and dying.

8.4.3 *Dynamics of 'metabolism-only' protocells*

If, in addition to removing N, we remove T, the protocell is neither capable of updating A, nor of producing W. The only autocatalyst remaining is M. Unable to create W, one might expect this protocell to be incapable of any behaviour at all. However, rather than standing still, investigation of the top right plot in Figure 8.4 shows that these agents are capable of movement.

How does this movement occur? The motion is caused by an asymmetry in the distribution of M particles. The protocell depletes the local R concentration to levels where further production of M becomes unlikely. In areas of the protocell where this has occurred, the production of M decreases substantially, causing an asymmetrical distribution of M within the protocell, inducing motion. This mechanism is reminiscent of one of the extensions of the original model of autopoiesis made by Suzuki and Ikegami (Suzuki and Ikegami, 2009), in that the lesioned protocells only begin to move when the metabolism starts to fail in a sub-region of the organism. That is to say, the protocells only move after they have come substantially closer to 'death', i.e. close to their viability boundary.

This mechanism of motility is clearly not as sophisticated as that of a non-lesioned agent. The behaviour only kicks in once the system is in a desperate, fragile state of "partially dying". This required fragility means that the random variation in e.g. the particle trajectories can be sufficient to inhibit motion for long enough as to cause death and that these protocells are largely incapable of turning corners – likely because this process involves too lengthily a system reorganisation without addition of resource for the fragile protocells to survive.

The top of Figure 8.5 shows the number of M particles plus the number of W particles plotted against time for 10 trials of the metabolism-only protocell. We use this count as an approximation system health they must be produced for the membrane to maintain sufficient size to enclose enzymes. Note the lower typical health, i.e. more fragile state of the metabolism-only protocells.

It is clear when comparing metabolism-only to non-lesioned that the lesioned entities tend to have a shorter life-span. The longest surviving metabolism-only protocell dies shortly after iteration 23000 with a mean survival-time of 20462 iterations (std. 2274.5) whereas almost all of the non-lesioned protocells survive past the end of the trial (past 50000 iterations). Interestingly, the senseless protocells die the fastest. In this case, having part of the behavioural mechanism appears to be more deadly than not having it at all.

8.5 DECOUPLING BEHAVIOUR AND METABOLISM

With the model in mind, in this section, we discuss the idea of weakening the relationship between behaviour and metabolism through 'partial decoupling' where the metabolism does not completely determine the behaviour, nor is it entirely irrelevant to it. Specifically, we discuss how:

1. Complete integration of metabolism and behaviour has various limitations.

2. Partial decoupling, where behaviour is not completely integrated with metabolism can overcome some of these disadvantages.

3. One way that this partial decoupling can occur is through *dynamical operational closure* in which the inter-dependencies of behavioural and metabolic processes change over time.

4. Metabolism-independent and metabolism-dependent behaviour have complementary advantages and are not mutually exclusive. It seems most likely that natural organisms would employ a combination of both types of mechanism.

8.5.1 *Limitations of metabolism-based behaviour*

This thesis has focused upon models in which there is a very close coupling of mechanisms of behaviour and autopoiesis. We have described several advantages to having such an organisation. But what are the disadvantages? Much of behaviour is not directly determined by metabolism. Why should this be the case? What are the advantages of decoupling behavioural mechanisms from metabolism?

Some disadvantages of close coupling of metabolic and behavioural mechanisms can be seen by examining the metabolism-only protocells, for which the chemotaxis of the metabolism-only protocells is modulated quite directly by the metabolism. This direct relationship between behaviour and viability is a fragile organisation in two ways. First, the response to unhealthy phenomena requires a decrease in metabolic health – a portion of the protocell ceases to produce enzymes. This is a decrease in viability that correlates with an increase in fragility of the protocell making the protocell less capable of survival. Note the lower typical 'health' of these agents compared to the non-lesioned agents (Figure 8.5). Observation of Figure 8.4 shows that these agents seem to be less capable of turning corners. When they encounter the boundary of high R concentration area, they tend to die. Perhaps this difficulty is because turning requires a system reorganisation that takes time while in a relatively low resource location, and the already weakened protocells (remember they only move once weakened) are incapable of surviving this reorganisation. Similarly, their shortened life span is likely the result of their reduced ability (due to their weakened condition) to endure unlucky stochastic dynamics such as brief unlucky motion *down* the R-gradient.

A second weakness of direct coupling between metabolism and behaviour is that it can trap the protocell in local viability maxima. Imagine a locally good (with respect to metabolism) area that is surrounded by a less optimal region. In this scenario, direct metabolic modulation of behaviour keeps the protocell in the local maximum, even if just on the other side of the less healthy areas there might be more advantageous situations. A protocell whose behaviour is determined directly by the efficacy of the metabolic system could never find these superior conditions, but an protocell whose behavioural mechanisms are more decoupled from the metabolism could be capable of forging out into a less safe environment in search of a superior environment. A variation of this sort of limitation can be seen in our model where we have observed agents "painting themselves into a corner" (results not presented). In comparison, the non-lesioned protocells (see Figure 8.4), with their partially decoupled behavioural mechanism, clearly make better use of resource, tending to feed on the higher resource areas before consuming areas down to dangerous levels, returning to areas they have already visited, etc. (see Figures 8.6 and 8.7).

There are, of course, other advantages to decoupling metabolism and behaviour. Decoupled behavioural mechanisms can respond to phenomena that do not *directly* influence metabolism. Such a mechanism could cause an agent to move towards an indicator of a metabolic resource. This allows for delayed reward feedback. As an example, people can respond to the bell of an ice cream truck, and perform "ice cream-taxis" without benefiting from the ice cream until rather late in the process. Such metabolism independent indicators could allow organisms to respond to more spatially distal resources. For example, a metabolism independent mechanism could in theory respond to much lower concentrations of resource than the concentrations that are necessary to have a significant impact upon the metabolism.

8.5.2 *Behaviour and metabolism: related, yet distinct*

This model is different from most other models of metabolism and behaviour, because the behaviour and the metabolism, while emergent from the same level of interaction (abstract chemistry), are not one and the same process. We can see how distinct they are by taking a closer look at the lesion studies. In the long term, the senseless and metabolism-only protocells have a lower average health and shorter life-span than non-lesioned protocells (see right column in Figure 8.4). However, in the short term the metabolism continues. If these lesioned protocells were placed in a different environment, such as one in which R is steadily renewed, the motion of the protocell would be unnecessary for the metabolism to continue to operate. It would be possible for the behaviour to cease without the metabolism also stopping.

This demonstrates that in this model, the more sophisticated behavioural mechanism demonstrated in the non-lesioned agent (involving N, T, A, etc.) is related to its metabolism but is also distinct from it. In the lesion studies, we have removed components of
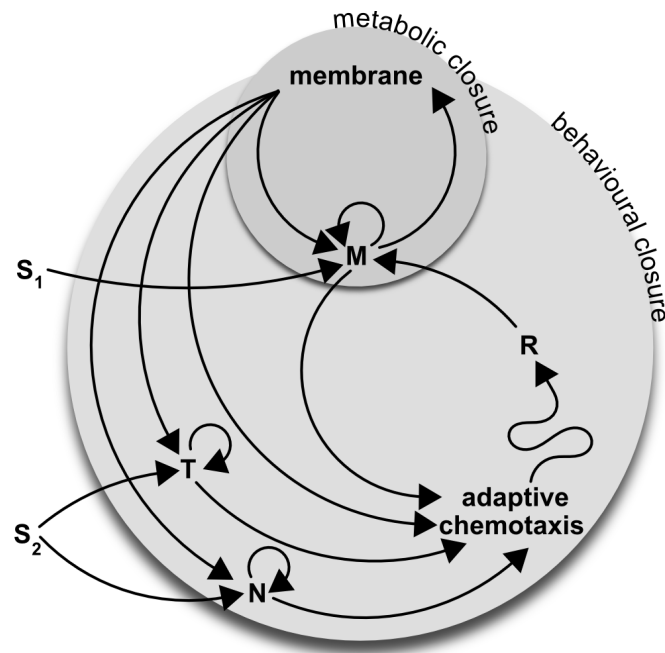
Previously printed in Egbert and Di Paolo (2009).

Figure 8.8: Inter-dependencies within our model. Arrows represent inter-dependencies as in Figure 8.1. M maintains the degrading membrane, which makes possible the continued existence of M, T, and N, etc.

this behavioural mechanism without removing a component of the metabolic mechanism. The behaviour affects the autopoiesis indirectly, through influencing the environment that the protocell (and its metabolism) is exposed to rather than being directly part of the autopoietic process. It is because of this decoupling of behavioural and metabolic mechanisms that it is possible in certain environments for the entity to continue to exist *without* behaving. This is unlike previous models of behaving, autopoietic systems (such as that by Suzuki and Ikegami, 2009) in which the agents metabolic and behavioural mechanisms are so intertwined that it is impossible to prevent its behaviour without also preventing its autopoiesis.

8.5.3 *Dynamical operational closure*

The relationship between metabolism and behaviour is asymmetrical. Without metabolism there is no protocell to perform the behaviour but the metabolism is not necessarily dependent upon the behaviour. In the case of our model, the behaviour depends upon the metabolism but the metabolism can operate (in the short term or in certain environments) in the absence of chemotaxis. Of course, the behaviour is not irrelevant to metabolism. The health plots in Figure 8.4 clearly shows the behaviour tends to have an effect on metabolism (e. g. its longevity), but the relationship is more indirect and variable than the dependence of the behaviour upon autopoiesis.

To understand the relationship between these mechanisms it is helpful to view the system from a perspective of operational closure. This can be most easily realised through study of Figure 8.8 which depicts the inter-dependencies present within our model. The clearest example of operational closure in our model is the "metabolically closed" system, which is highlighted by the smaller grey-filled circle at the top of the diagram. For the M particles to be produced, the membrane must be large enough and for the sufficiently large membrane size to be maintained, M particles must be made.

Just as the autocatalysis of M relies upon the membrane encircling sufficient precursor particles, so does the autocatalysis of T and N. These relationships are indicated by the arrows from 'membrane' to T and N. T, N, M, W (not included in the figure for simplicity) and the membrane are all integral to the reactions that cause the more sophisticated (non-lesioned) chemotaxis. Less simple is the dependence of R upon chemotaxis. This depends upon the environment in which the protocell 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 chemotaxis is unnecessary for the protocell. If however, the protocell depletes R in its local environment, if it is to survive it must move to another area to find more R. This dependence changes depending upon the environment in which the protocell exists.

This is an example of operational closure changing over time. In this case, depending upon the environmental conditions, the operationally closed system can include only the metabolic cycle (the filled circle) or it can include both the autopoietic cycle and the behavioural mechanisms (dashed circle). In this model, the metabolic system is sometimes and sometimes not part of a larger operationally closed behavioural system.

This relation between behaviour 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 inter-dependencies 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 behaviour is an organisational, not merely a structural, change to the system. We suspect that this possibility of entering temporarily different modes of system organisation can play an essential role in developing the theory of autopoiesis towards an account of biological transformations of organisation (e.g. development from embryo to adult), which is at the moment one of its blind-spots.

8.5.4 *Partial decoupling through dynamical operational closure*

As mentioned above, the dependence of metabolism upon behaviour can change. That is to say that while behavioural mechanisms are always dependent upon autopoiesis, autopoiesis is not always dependent upon behaviour and this can change in the short term. It can, for instance, change when the agent enters a different environment. An

environment of steadily replenished resources may not require the same behaviour as a more impoverished environment etc.

If we consider this changing relationship in the context of operational closure, it becomes evident that as the relationships of condition change, as does the boundaries of the operationally closed system. The behavioural mechanism (B) always depends upon the mechanism of metabolism (M). When M also depends upon B, then M+B is operationally closed. When M does *not* depend upon B, then M+B is not operationally closed. In this way, the form of operationally closed systems can change over time.

In some models (Suzuki and Ikegami, 2009), the coupling is so tight that the behavioural and metabolic mechanisms can be thought of as actually one and the same mechanism. In this case, the behaviour is distinct from metabolism only in that the extended dynamics occur in the more noticeable spatio-mechanical domain (e. g. simulated osmotic pressure or membrane spatial location). In this model, we have decoupled the mechanisms enough to explore some the variety of possible relationships between behaviour and metabolism.

8.5.5 *Complementary mechanisms*

Some of the advantages of metabolism-based behaviour stem from its simplicity. By being indirectly sensitive to the environment by way of the metabolism, an organism's response is generalised in the sense that it is not specific to the environmental influence, but just responds to changes in the metabolism, regardless of what caused them. This is a simple, but powerful, general mechanism. As discussed in earlier chapters, it makes it possible for the organism to integrate a variety of influences, or respond to phenomena never previously experienced by the organism nor by any of its ancestors. But, as they say, there is no free lunch. The generality comes at the cost of being incapable of identifying the cause of a recent change. Metabolism dependent behaviour can respond to a decrease in viability, but it cannot, on its own, directly respond to the cause of the decrease in viability. It can "see that things have gotten worse", but it cannot tell if it was a change in temperature or pH or something else that has made things worse. Forms of metabolism independent behaviour do not need to suffer from the same limitations and could supplement metabolism driven behaviour with more specific sensitivities and response mechanisms.

This is one of the advantages provided by more specific metabolism-independent behaviour and we have discussed other advantages of metabolism independence elsewhere in this chapter. But, it would be wrong to draw the conclusion that metabolism independence is the superior mechanism. Just as metabolism independent mechanisms provide benefits unavailable to metabolism dependent behaviour, the converse is also true – metabolism dependent behaviour can supplement metabolism independent behaviour. For example, metabolism independent behaviour is more susceptible than metabolism driven behaviour to being 'fooled' into performing unhealthy behaviour. As a case in point, consider the metabolism independent chemotaxis of *E. coli* to the non-metabolisable

saccharin. Because metabolism independent mechanisms are by definition sensitive to something *other* than metabolic health, it is possible that even when functioning correctly, they can produce non-healthy behaviour (e.g. by being in a strange environment).

What we wish to highlight, is that the two types of mechanism are not mutually exclusive, in fact they seem well complemented to one and other. The simplicity and generality of metabolism-based adaptive behaviour is an excellent starting point – one that can be improved over time through the addition of supplementary metabolism independent mechanisms by e.g. Darwinian evolution. The pattern we would expect to see in biology is therefore a combination of both types of mechanism, where metabolism independent behaviour is the driving force in many situations, but where metabolism-based behaviour acts as an override or "fall-back plan" where if the metabolism-independent mechanisms are failing, it takes over.

This organisation matches that presented in 6.5, where we discussed how there appear to be at least two different mechanisms supporting chemotaxis in *E. coli* (Eisenbach, 2007, p. 575). Evidence shows that these bacteria can perform chemotaxis even when stripped of most of the signalling pathway typically associated with chemotaxis (Barak and Eisenbach, 1999). The well studied two-component transduction system is established as being a metabolism-independent mechanism, while fumarate, a metabolic product, modulates the flagellar rotation (Cohen-Ben-Lulu et al., 2008) in such a way as to operate as a metabolism-based chemotaxis mechanism.

Perhaps there is also evidence of this organisation in higher forms of cognition – hunger can make people "less rational", evidence, perhaps, of more primitive behavioural mechanisms overriding the more intricate or delicate behavioural mechanisms that we are more familiar with. More research should be done in this area.

8.6 a fractal of autonomy

If cognitive acts, such as writing your thesis, are irrelevant to your metabolic needs, or to your continued existence, is there anything that operational closure or the approach taken in this thesis has to say about higher cognition? We mentioned breifly (in Section 2.3.5) the idea of autonomous processes operating in the domain of the nervous system, the body and the environment, and that these kinds of autonomies could be relevant to higher forms cognition. This thesis does not cover this idea in detail (see discussion of 'mental life' (Barandiaran, 2008) and also 'habits' (Di Paolo, 2009, 2003), but there are a few points of contribution that can be made here with the present model in mind.

Let us consider for a moment an alternative hypothetical system organisation. What if the production of T and N depended more directly upon the chemotaxis (i.e., not indirectly through the metabolic processes)? This would produce two new operationally closed loops (one consisting of say 'T' and 'chemotaxis', the other consisting of 'N' and 'chemotaxis'). These loops could have their own optimal operating conditions and it would be possible for such behavioural loops to have norms different or even

in conflict with the norms of the primary autopoietic system. In this hypothetical situation, the behavioural mechanism is an autonomous (operationally closed) system. The constitutional autonomy is in a sense part of the "environment" for this autonomous mechanism.

Consider the asymmetry of dependence in the relationship between the autonomies. A behavioural autonomy is in a sense more fundamentally dependent upon the constitutional or metabolic autonomy than vice versa, for to behave, there must be a behaver! It is possible to imagine a non autonomous (i. e. not operationally closed or precarious) body such as a robot that acts as a substrate upon which behavioural autonomy operates, but for now behaviour is limited to life and all life comprises a metabolic autonomy. It is, conversely, possible to imagine or perhaps even provide examples of organisms that are metabolically closed but that demonstrate *no* autonomous behavioural systems whatsoever. We can also consider early models of autopoiesis (representing metabolic closure) that lack any sort of behaviour, demonstrating the non-dependence of metabolic autonomy upon behavioural autonomy. That being said, there are metabolic autonomies that are almost as fundamentally dependent upon their behavioural autonomies as the behavioural autonomies are dependent upon the metabolic autonomy.

A behavioural mechanism could, for instance, continually move the protocell 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, in fact, the situation modelled in Experiment 3 in Chapter 5 (Section 5.4.3) in which the simulated protocell cannot survive in any one location in its environment but must move back and forth between resources. The metabolism relies upon the behaviour, and the behaviour relies upon the metabolism.

Interestingly, when we consider behaviours as having their own autonomy, we begin to be able to provide new explanations for previously difficult to explain phenomena. Take as an example, smoking. Smoking and other addictions have been described as a broken behavioural system – a "short-circuiting" of the reward system or something similar. But here we can see smoking not as a broken system, but one that has its own positive influence upon its own continued existence, a loop whereby smoking increases the chance that smoking will occur in the future. Unfortunately, in this case, the autonomous behaviour runs in conflict with the metabolic or constitutional autonomy of the smoker, but in other situations autonomous behaviours may operate cooperatively with metabolic constitution, for instance leading the organism out of a local healthy area through a short unhealthy area into a more global maximally healthy area. Or, as mentioned before, transforming through movement a combination of unsurvivable environments into a collection of survivable environments.

8.7 RECAPITULATION

This chapter reports on a computational model that incorporates aspects of computational artificial chemistry and neuro-ethology and has helped us to explore how mechanisms of behaviour and autopoiesis can be related, yet distinct. Specifically, the following clarifications have been made through analysis of the model.

PRIMARY POINTS OF CHAPTER

- Behaviour modulation based upon autopoietic efficacy has limitations, including:
  - A metabolism driven aversion response to dangerous phenomena requires a decrease in metabolic health (viability)
  - Metabolism driven behaviour can become trapped in local health maxima
  - Metabolism independent behaviour can respond to temporally delayed reward
  - Metabolism independent behaviour can be more sensitive to phenomena, responding before they have any significant impact upon metabolic health.
  - Metabolism driven behaviour cannot, on its own, identify (directly respond to) the cause of a change in viability conditions.

- Metabolism-independent mechanisms do not need to suffer from the same limitations, but have their own limitations, in particular:
  - Metabolism-independent behaviour acts in response *not* to the effect of environmental phenomena, but directly to environmental phenomena themselves, so are more prone to being fooled or confused into performing unhealthy behaviour (c.f., bacteria being attracted to non-metabolisable resources such as saccharin)

- Metabolism-independent and metabolism-dependent mechanisms are not mutually exclusive. In fact, they seem like mechanisms that would supplement each other well, each covering for the limitations of the other. We hypothesise that most organisms would utilise both forms of behavioural mechanism and recapitulate some evidence in support of this theory.

- One way that metabolism-independent and metabolism-dependent systems might be related, yet distinct involves the new notion of *dynamical operational closure*, in which the organisation of inter-dependencies that determine an operationally closed system can change over time while aspects remain operationally closed.

9

SUMMARY OF CONCLUSIONS AND
FUTURE WORK

9.1 SUMMARY

In this thesis we have presented and analysed models in which behaviour can interact with the viability-constraint determining dynamics of metabolism. These models have allowed us to study how a system can be sensitive to its own viability, and have an organisation that allows it to respond to trends in that viability. The behaviour of such systems might sub-serve a long-term evolutionary advantage, but it more directly subserves a more immediate and intrinsic 'self-preserving' dynamic through which a variety of adaptive mechanisms are provided. These adaptive mechanisms have been described throughout the thesis and are recapitulated below.

By taking an approach that considers interactions between metabolism and behaviour, we have exposed a variety of phenomena that can help us understand how life manages to be so adaptable. We have developed a better understanding of phenomena observed in bacteria, and we have discovered a new, hypothetical mechanism of evolution-like adaptation, that we have called behavioural metabolution. All of these advances deserve further exploration. They may help us design more adaptable technology, fight disease, or help us better understand the mechanisms of evolution that resulted in our existence. We hope that these advances presented in this thesis will encourage others to consider the metabolic organisation when studying adaptive or cognitive systems, or the evolution of such systems.

9.2 PRIMARY CONCLUSIONS

In this section, we recapitulate the primary conclusions drawn over the course of this thesis.

9.2.1 *Viability-sensitive behaviour*

In Chapter 4, we provided a definition of self-sensitive behaviour, building upon notions of an operational closed, self-maintaining dissipative organisation to capture the notion of 'the metabolic organisation'. We observed that paying attention to this organisational property not only allows us to differentiate between an agent and its environment, but also reveals mechanisms that are only available to systems with this organisation. Mechanisms such as self-sensitivity, in which systems with the metabolic organisation are sensitive to their *viability*. By being sensitive to the state of this viability (or changes in it), an organism can respond to the quality of its environment with respect to its well-being. This makes possible simple, but powerful 'self-sensitive' adaptive behaviours such as *If I am healthy now, keep doing the same as I have been doing – otherwise do something else.*

9.2.2 *Self-sensitivity provides powerful mechanisms of adaptation*

In Chapters 5 – 8 we described and demonstrated using computational simulations the diverse adaptive benefits of self-sensitive behavioural strategies such as that described above. These benefits include:

(i) A mechanism of adaptive behaviour that can be sensitive to the organism's history of experience (Chapters 5 and 7), and sensitive to the current context of the organism (Chapter 5);

(ii) The ability to integrate different environmental influences to produce an appropriate response, even across different domains (e. g. chemical encounters and temperature) (Chapter 6);

(iii) The ability to respond appropriately (e. g. flee from toxins / seek out metabolic resources) to phenomena never previously experienced by the agent nor by any of its ancestors (Chapters 6 and 7);

(iv) The ability to adapt not only in response to changes in the environment but also to changes in the organism's *own organisation* (Chapter 7);

(v) Behaviour is in response to the metabolism rather than directly to the environment, and is therefore less easily "fooled" into producing damaging behaviour (unlike metabolism-independent mechanisms in which, for example, bacteria can be attracted to non-metabolisable reactants such as saccharin) (Chapter 8);

There are also disadvantages of viability-sensitive behaviour. For instance, without supplementary mechanisms, simple viability-sensitive mechanisms cannot determine *the cause* of a change in viability, but can only respond to its occurrence. Also, viability-sensitive behaviour requires a change in viability to occur for a change in behaviour to be produced. This entails certain limitations. For example, unlike a viability-sensitive mechanism, a viability-independent mechanism could respond to resource levels that are too low to have an influence on viability, allowing for more sensitive or indirect sensitivities (discussed in Chapter 8).

These are not problems for the general concept of metabolism-based behaviour for, as we argue in Chapter 8, metabolism-dependent and metabolism-independent mechanisms are not mutually exclusive. They, in fact, supplement each other well, each covering the blind spots of the other. We therefore proposed the hypothesis that in life, one would expect to find a combination of both metabolism-based (i.e., viability-sensitive) and metabolism-independent (environment-sensitive) behavioural mechanisms, each category of mechanism supplementing the other. We discuss this and other relationships between behaviour and metabolism, such as partial-decoupling, in Chapter 8.

9.2.3 *Self-sensitivity can facilitate adaptive evolution*

In Chapter 7, we discussed and demonstrated aspects of how self-sensitive behaviour can facilitate population level adaptation (evolution) in a number of ways:

(a) Adaptation produced through metabolism-based behaviour results in phenotypic plasticity enabling Baldwin-effect-like processes, thereby facilitating Darwinian evolution;

(b) Self-sensitive behaviour can cause agents to move towards new advantageous environments, separating populations, increasing the chance for speciation events;

(c) Behaviour based upon self-sensitivity allows for adaptation to changes in the very organisation of the organism itself which can transform otherwise neutral mutations into beneficial mutations;

(d) The movement of agents into new advantageous environments can increase the likelihood of an encounter with new environmental reactants that can cause a possibly advantageous chemical avalanche in the metabolism;

(e) There is a possibility of a circle of influence through which (c) causes (d), causing (c) again, causing (d), etc. This cycle of directed adaptation in which behaviour adapts to metabolism, possibly causing further changes in metabolism is what we have termed behavioural metabolution;

9.2.4 *Behavioural metabolution could have played a role in the early evolution of life*

Self-sensitive behaviour can be accomplished with a very simple coupling between behaviour and metabolism and as listed above it can provide or facilitate a wide range of adaptive mechanisms. In this respect, self-sensitive behaviour is simple and powerful enough that we argued that it is worthy of further study as having played a role facilitating the early evolution of life on Earth (Chapter 7).

9.3 MORE GENERAL CONCLUSIONS

In addition to the conclusions listed above, the research presented in this thesis led to a few broader, more philosophical conclusions. These are presented in this section.

9.3.1 *An approach to autonomous agents*

One idea that lies underneath all of the work presented in this thesis is that of a system "acting on its own behalf". The work presented in this thesis is the result of thinking about life in these terms. Considering life not as a population of replicating molecules "inside gigantic lumbering robots" (Dawkins, 1976, pps. 19–20), but as individual organisms that act fundamentally for their own survival. Thinking of cognition not as computation, or disembodied (in the radical embodiment sense) dynamics, but as a system dependent upon its own effects. These alternative approaches, as outlined in Chapter 2, motivated this research and led us to the conclusions presented here.

By considering the metabolic organisation, we have been able to define and study systems that are, in a relatively strong sense, intrinsically motivated. The idea of intrinsic motivation is intuitive, but it has proven difficult over the history of philosophy and science to give it, and related concepts such as agency, autonomy and teleology, scientific meaning. One broad motivation for our work is that it and future work in this area will help to make these terms less mysterious through formalising their scientific meaning.

By taking an approach that considers the interaction between behaviour and the viability-constraining dynamics of metabolism, we have exposed mechanisms of adaptability available to living systems, but interestingly are not available to contemporary typical robots. We have come to better understand the behaviour of metabolism-based chemotactic bacteria and we have proposed a new hypothetical mechanism of population evolution that is based upon interactions between behaviour and metabolism. This is all encouraging. It validates the approach taken, and supports an argument for further study in this direction.

9.3.2 *The role of abstract models*

It also should be emphasised that abstract computational models of concepts played a central role in our analysis and in the development of our understanding of these viability-sensitive systems. Self-producing, self-maintaining, precarious yet dynamically stable systems are complicated, recursive systems that cannot be understood simply through thought experiments. Designing, and implementing, computational models of these ideas helped us to be concise and precise about our terms. Analysis of the models helped us to understand the systems and what they are capable of. This benefited the development of subsequent models and theory. Observation of the adaptability of the agent in the model presented in Chapter 5 led to the work connecting the ideas to bacterial chemotaxis in Chapter 6, which eventually led us to develop the hypotheses and demonstrations presented concerning behavioural metabolution presented in Chapter 7. It is hard to see how the presented research could have been accomplished without using abstract computational models of ideas. Accordingly, we believe this work acts as an example in support of the argument for continued use of abstract computational models of concepts in science.

9.4 FUTURE WORK

There are several ways in which the research presented here could be extended. This section presents two extensions that seem particularly likely to be productive.

9.4.1 *Establishing the impact of behavioural metabolution:*

As discussed, self-sensitive behaviour is simple enough to be one of the first behavioural mechanisms. It provides an inherent adaptability and that adaptability can facilitate adaptive evolution in several ways. Self-sensitive behaviour may therefore have played a critical role in facilitating early major transitions in the history of life. But, very little is known about this role of self-sensitive behaviour. In this thesis, we have presented some pilot studies, demonstrating the plausibility of some of the concepts involved, but it is time for a thorough research programme in which we evaluate how self-sensitive behaviour may have been implemented in the early history of life, and the role it played in facilitating early evolutionary transitions. How frequently are the various steps in behavioural metabolution likely to occur, often or hardly ever? How strong an influence are these mechanisms likely to have incurred upon the evolution of life on Earth, fundamental or insignificant? This is the subject of a draft grant proposal, assembled by the author with a collection of international project partners.

9.4.2 *Autonomies in cooperation, conflict, and indifference*

When one considers life from a perspective of autonomy, it is apparent that for an individual organism there is not a single, logical system that produces its behaviour, but a variety of influences. The behaviour of organisms is the result of a variety of influences, sometimes in conflict. A male grasshopper acts to prolong its life, and is clearly motivated to copulate, yet this can lead to its head being bitten off by its mate. Richard Dawkins became famous partially through his promotion of the idea of genetic systems having a sort of autonomy. Siblings, parents or cousins can sacrifice themselves, he argues, if that sacrifice will tend to prolong the existence of the genes carried by that family. In this view, the autonomy of the individual is at times in conflict with the autonomy of the genetic inheritance system. There can be additional conflicts. At times, it appears that the autonomy of the individual can be in conflict with itself. A smoker can simultaneously want a cigarette and want to *not* want the cigarette. I want this thesis to be finished as soon as possible, but I also want it to be as good as possible. Organisms, especially perhaps humans, are in perpetual self-conflict, yet this conflict is rarely a focus in the study of adaptive behaviour or cognition.

The model presented in Chapter 5 demonstrated a dynamic of "being torn in two directions" similar in some respects to the conflicts mentioned above. The data is not reported in detail in this thesis, but during observation of the agent, it was not uncommon to see the agent briefly standing still between two resource generators – attracted to both, yet moving towards neither. In this model, conflict occurred within the same mechanism, but the conflict could stem from different autonomous processes. The model described in Chapter 8 presented two autonomies (one behavioural and one metabolic) that operated in cooperation. What if these autonomies did not always have the same viability constraints or sensitivities. What if they were operated sometimes in conflict and sometimes in cooperation? A fascinating piece of future work would be to explore cognition, not as a monolithic dynamical or computational system, but as a collection of autonomies, sometimes in cooperation, sometimes in conflict, and sometimes indifferent to one another. This too seems like a promising route of research.

PRIMARY POINTS OF CHAPTER

- Viability-sensitive behaviour can produce simple, yet powerful mechanisms of adaptive behaviour at the scale of the individual organism.

- The adaptability provided by viability-sensitive behaviour can facilitate evolutionary processes through a variety of mechanisms.

- Self-sensitive behaviour is simple enough to have possibly facilitated the very early evolution of life, conceivably even pre-genetic life.

- There is more to embodiment than situatedness. The metabolic organisation of an interdependent precarious dissipative network of processes is universal to life and worthy of further attention – it may play a key role in developing an understanding of notions related to autonomy and agency.

- Computational modelling of abstract concepts can be productive, especially when trying to study systems that are difficult to imagine (e.g., highly recursive complex systems) and when biological targets that demonstrate the phenomena of interest are much more complicated than its minimal form.

A

TERMINOLOGY

autonomy a system that consists of a precarious, operationally closed network of interdependent processes, i. e., autopoiesis but without the requirement for a spatial membrane. Different from 'basic autonomy' (see below).

autopoiesis from the Greek 'auto' (self) and 'poiesis' (construction), used to refer to systems that build themselves and / or maintain themselves. In addition to being self-maintaining, these systems are typically considered to be precarious and operationally closed. Some interpretaions require the system to include a spatial membrane that forms a boundary around the system and is produced by the system. A biological cell is often given as an example of an autopoietic system.

adaptive behaviour behaviour that maintains its positive normativity (i. e. its maintenance of the metabolic organisation) despite changes in the environment or in the needs or organisation of the agent.

basic autonomy Ruiz-Mirazo and Moreno (2004)'s interpretation of autopoiesis that draws attention to the idea of autopoietic systems as far-from-equilibrium systems that must be coupled with their environment, to acquire the energetic and material resources necessary to maintain their dissipative organisation.

behaviour following (Bourgine and Stewart, 2004), this thesis considers behaviour to be the actions of a metabolic (or autonomous) organisation that modify its interaction with its environment. Examples include motility and modulation of membrane permeability.

cognition We use this word to capture the general idea of what cognitive scientists study, i. e. thought, intentional behaviour, adaptivity, normative behaviour, etc. These phenomena remain poorly understood, so we use this word to capture them in an intentionally non-specific and informal way.

ontogenetic in reference to the lifetime development of an individual

operational closure A property of a system such that each member component or process is dependent (indirectly or directly) upon every other component or process in the system. If one component fails for a long enough period of time, the whole system will fail.

metabolic organisation A network of interdependent processes, each of which depends upon all of the others and each of which enables at least one other. This network is inherently precarious (i. e. unstable and perpetually degrading) but can, to a varying degree, counteract this degradation through processes of self-maintenance, resulting in a dynamical stability.

precariousness used to describe systems that, in the absence of their own self-influence (self-production or self-maintenance) would cease to exist.

viability a measurement of the health of a metabolic organisation. In the absence of other differences, a system with a high viability is more likely to persist than one with low viability.

BIBLIOGRAPHY

Adler, J. (1969). Chemoreceptors in bacteria. *Science*, 166(3913):1588–1597.

Alexandre, G. (2010). Coupling metabolism and chemotaxis-dependent behaviours by energy taxis receptors. *Microbiology*, 156(8):2283–2293.

Alexandre, G., Greer, S. E., and Zhulin, I. B. (2000). Energy taxis is the dominant behavior in azospirillum brasilense. *J. Bacteriol.*, 182(21):6042–6048.

Alexandre, G. and Zhulin, I. B. (2001). More than one way to sense chemicals. *Journal of Bacteriology*, 183:p. 4681—4686.

Andrews, S. S. and Bray, D. (2004). Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical Biology*, 1(3):137–151.

Armitage, J. P. and Schmitt, R. (1997). Bacterial chemotaxis: Rhodobacter sphaeroides and sinorhizobium meliloti—variations on a theme. *Microbiology*, 143(Pt 12):3671—3682.

Ashby, W. R. (1952). *Design for a Brain: the origin of adaptative behaviour*. J. Wiley, London, 2nd edition.

Barak, R. and Eisenbach, M. (1999). Chemotactic-like response of escherichia coli cells lacking the known chemotaxis machinery but containing overexpressed CheY. *Molecular Microbiology*, 31(4):1125–1137.

Barandiaran, X. and Moreno, A. (2006). On what makes certain dynamical systems cognitive: A minimally cognitive organization program. *Adaptive Behavior*, 14(2):171 –185.

Barandiaran, X. and Moreno, A. (2008). Adaptivity: From metabolism to behavior. *Adaptive Behavior*, 16(5):325 –344.

Barandiaran, X., Moreno, A., Rocha, L., Yaeger, L., Bedau, M., Floreano, D., Goldstone, R., and Vespignani, A. (2006). ALife models as epistemic artefacts. In *Artificial Life X : Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems*, pages 513–519. The MIT Press (Bradford Books).

Barandiaran, X. E. (2008). *Mental Life. A naturalized approach to the autonomy of cognitive agents*. PhD thesis, University of the Basque Country, Spain.

Barandiaran, X. E., Di Paolo, E., and Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5):367–386.

Bechtel, W. (2007). Biological mechanisms: Organized to maintain autonomy. In Boogerd, F. C., editor, *Systems Biology: Philosophical Foundations*, page 269–302. Elsevier.

Bechtel, W. (2009). Some virtues of modeling with both hands. *Adaptive Behavior*, 17(4):293.

Beer, R. D. (1995). On the dynamics of small Continuous-Time recurrent neural networks. *Adaptive Behavior*, 3(4):469–509.

Beer, R. D. (1997). The dynamics of adaptive behavior: A research program. *Robotics and Autonomous Systems*, 20:257—289.

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4):209–243.

Beer, R. D. (2004). Autopoiesis and cognition in the game of life. *Artificial Life*, 10(3):309–326.

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

Berg, H. C. and Brown, D. A. (1972). Chemotaxis in escherichia coli analysed by three-dimensional tracking. *Nature*, 239(5374):500–504.

Berg, H. C. and Tedesco, P. M. (1975). Transient response to chemotactic stimuli in escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 72(8):3235 –3239.

Bespalov, V. A., Zhulin, I. B., and Taylor, B. L. (1996). Behavioral responses of escherichia coli to changes in redox potential. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19):10084–10089. PMID: 8816755.

Boden, M. (1999). Is metabolism necessary? *Br J Philos Sci*, 50(2):231–248.

Boden, M. A. (2000). Autopoiesis and life. *COGNITIVE SCIENCE QUARTERLY*, 1:117—145.

Borrok, M. J., Kolonko, E. M., and Kiessling, L. L. (2008). Chemical probes of bacterial signal transduction reveal that repellents stabilize and attractants destabilize the chemoreceptor array. *ACS Chemical Biology*, 3(2):101–109.

Bourgine, P. and Stewart, J. (2004). Autopoiesis and cognition. *Artificial life*, 10(3):327–345.

Bray, D. (2009). *Wetware: A Computer in Every Living Cell*. Yale University Press.

Bray, D., Levin, M., and Lipkow, K. (2007). The chemotactic behavior of Computer-Based surrogate bacteria. *Current Biology*, 17(1):12–19.

Brooks, R. A. (1991). Intelligence without reason. *Artificial intelligence: critical concepts*, 3.

Cairns-Smith, A. G. (1990). *Seven Clues to the Origin of Life: A Scientific Detective Story*. Cambridge University Press.

Ciciriello, F., Costanzo, G., Pino, S., Crestini, C., Saladino, R., and Di Mauro, E. (2008). Molecular complexity favors the evolution of ribopolymers†. *Biochemistry*, 47(9):2732–2742.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, 3(9):345–351.

Clayton, R. K. (1953). Studies in the photataxis of rhodospirillum rubrum. *Archiv for Mikrobiologie*, 19(2):125–140.

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

Cohen-Ben-Lulu, G. N., Francis, N. R., Shimoni, E., Noy, D., Davidov, Y., Prasad, K., Sagi, Y., Cecchini, G., Johnstone, R. M., and Eisenbach, M. (2008). The bacterial flagellar switch complex is getting more complex. *EMBO J*, 27(7):1134–1144.

Crispo, E. (2007). The baldwin effect and genetic assimilation: Revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution*, 61(11):2469–2479.

Dawkins, R. (1976). *The Selfish Gene - First Edition*. Oxford.

Di Paolo, E. A. (2000). Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. In *From animals to animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior*, page 440–449.

Di Paolo, E. A. (2003). Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop. In Murase, K. and Asakura, T., editors, *Dynamical systems approach to embodiment and sociality*, pages 19—42. Advanced Knowledge International, Adelaide, Australia.

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(1):9—21.

Di Paolo, E. A. (2010). Robotics inspired in the organism. *Intellectica*, 53–54:129—162.

Di Paolo, E. A., Noble, J., and Bullock, S. (2000). Simulation models as opaque thought experiments. In *Seventh International Conference on Artificial Life*, pages 497—506, Cambridge, MA. MIT Press.

Dittrich, P. and di Fenizio, P. (2007). Chemical organisation theory. *Bulletin of Mathematical Biology*, 69(4):1199–1231.

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

Dreyfus, H. (1965). Alchemy and artificial intelligence. *RAND Coprporation*.

Dreyfus, H. L. (1979). *What Computers Can't Do*. Harper & Row.

Edwards, J. C., Johnson, M. S., and Taylor, B. L. (2006). Differentiation between electron transport sensing and proton motive force sensing by the aer and tsr receptors for aerotaxis. *Molecular microbiology*, 62(3):823–837. PMID: 16995896 PMCID: 1858650.

Egbert, M. D. and Barandiaran, X. E. (2011). Quantifying normative behaviour and precariousness in adaptive agency. In Lenaerts, T., Giacobini, M., Bersini, H., Bourgine, P., Dorigo, M., and Doursat, R., editors, *Proceedings of the 11th European Conference on Artificial Life*, pages 210—218. MIT Press.

Egbert, M. D., Barandiaran, X. E., and Di Paolo, E. A. (2010a). Behavioral metabolution: Metabolism based behavior enables new forms of adaptation and evolution. In *Artificial Life XII Proceedings of the Twelfth International Conference on the Simulation and Synthesis of Living Systems*, pages 213—220. MIT Press.

Egbert, M. D., Barandiaran, X. E., and Di Paolo, E. A. (2010b). A minimal model of Metabolism-Based chemotaxis. *PLoS Comput Biol*, 6(12):e1001004.

Egbert, M. D. and Di Paolo, E. A. (2009). Integrating autopoiesis and behavior: An exploration in computational chemo-ethology. *Adaptive Behavior*, 17(5):387–401.

Egbert, M. D., Di Paolo, E. A., and Barandiaran, X. E. (2009). Chemo-ethology of an adaptive protocell: Sensorless sensitivity to implicit viability conditions. In *Advances in Artificial Life, Proceedings of the 10th European Conference on Artificial Life, ECAL.*, pages 242—250. Springer.

Eisenbach, M. (2004). *Chemotaxis*. Imperial College Press, 1 edition.

Eisenbach, M. (2007). A hitchhiker's guide through advances and conceptual changes in chemotaxis. *Journal of Cellular Physiology*, 213(3):574–580.

Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A., and 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. PMID: 9442881.

Feldman, J. and Ballard, D. (1982). Connectionist models and their properties. *Cognitive Science*, 6(3):205–254.

Fernando, C. and Rowe, J. (2008). The origin of autonomous agents by natural selection. *Biosystems*, 91(2):355–373.

Fontana, W. and Buss, L. W. (1996). The barrier of objects: From dynamical systems to bounded organizations. *Boundaries and barriers*, pages 56—116.

Gánti, T. (1975). Organization of chemical reactions into dividing and metabolizing units: The chemotons. *Biosystems*, 7(1):15–21.

Gánti, T. (2003a). *Chemoton Theory: Theory of Living Systems*. Springer.

Gánti, T. (2003b). *The Principles of Life*. Oxford University Press, USA.

Goldstein, R. A. and Soyer, O. S. (2008). Evolution of taxis responses in virtual bacteria: Non-Adaptive dynamics. *PLoS Comput Biol*, 4(5):e1000084.

Greer-Phillips, S. E., Alexandre, G., Taylor, B. L., and Zhulin, I. B. (2003). Aer and tsr guide escherichia coli in spatial gradients of oxidizable substrates. *Microbiology*, 149(9):2661–2667.

Hanczyc, M. and Ikegami, T. (2010). Chemical basis for minimal cognition. *Artificial Life*, 16(3):pps. 1–12. in press.

Hanczyc, M., Toyota, T., Ikegami, T., Packard, N., and Sugawara, T. (2007). Fatty acid chemistry at the oil- water interface: Self-Propelled oil droplets. *J. Am. Chem. Soc*, 129(30):9386—9391.

Harold, F. M. (2003). *The Way of the Cell: Molecules, Organisms, and the Order of Life*. Oxford University Press, USA.

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

Hazelbauer, G. L. and Lai, W. (2010). Bacterial chemoreceptors: providing enhanced features to two-component signaling. *Current Opinion in Microbiology*, 13(2):124–132.

Horgan, J. (1995). From complexity to perplexity. *Scientific American*, pages 74—79.

Izquierdo, E. J. and Lockery, S. R. (2010). Evolution and analysis of minimal neural circuits for klinotaxis in caenorhabditis elegans. *The Journal of Neuroscience*, 30(39):12908 –12917.

Jeziore-Sassoon, Y., Hamblin, P. A., Bootle-Wilbraham, C. A., Poole, P. S., and Armitage, J. P. (1998). Metabolism is required for chemotaxis to sugars in rhodobacter sphaeroides. *Microbiology*, 144(1):229–239.

Kacian, D. L., Mills, D. R., Kramer, F. R., and Spiegelman, S. (1972). A replicating RNA molecule suitable for a detailed analysis of extracellular evolution and replication. *Proceedings of the National Academy of Sciences of the United States of America*, 69(10):3038–3042.

Kampis, G. (1991). *Self-Modifying Systems: A New Framework for Dynamics, Information, and Complexity*. Pergamon, Oxford-New York.

Kauffman, S. A. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1):1–24.

Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. OUP USA.

Kauffman, S. A. (2000). *Investigations*. Oxford University Press, USA, 1 edition.

Kresge, N., Simoni, R. D., and Hill, R. L. (2006). Julius adler's contributions to understanding bacterial chemotaxis. *Journal of Biological Chemistry*, 281(41):e33.

Krischer, K. and Mikhailov, A. (1994). Bifurcation to traveling spots in Reaction-Diffusion systems. *Phys. Rev. Lett.*, 73(23):3165.

Letelier, J., Soto-Andrade, J., Guinez Abarzua, F., Cornish-Bowden, A., and Luz Cardenas, M. (2006). Organizational invariance and metabolic closure: Analysis in terms of (M,R) systems. *Journal of Theoretical Biology*, 238(4):949–961.

Luisi, P., Ferri, F., and Stano, P. (2006). Approaches to semi-synthetic minimal cells: a review. *Naturwissenschaften*, 93(1):1–13.

Luisi, P. L. (2003). Autopoiesis: a review and a reappraisal. *Naturwissenschaften*, 90(2):49–59.

Luisi, P. L. and Varela, F. J. (1989). Self-replicating micelles – a chemical version of a minimal autopoietic system. *Origins of Life and Evolution of Biospheres*, 19(6):633–643.

Macnab, R. M. and Koshland, D. E. (1972). The Gradient-Sensing mechanism in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America*, 69(9):2509 –2512.

Madina, D., Ono, N., and Ikegami, T. (2003). Cellular evolution in a 3D lattice artificial chemistry. *Advances in artificial life*, pages 59—68.

Maturana, H. R. and Varela, F. (1992). *Tree of Knowledge*. Shambhala.

Maturana, H. R. and Varela, F. J. (1980). *Autopoiesis and cognition: the realization of the living*. Springer.

Maynard Smith, J. and Szathmáry, E. (1998). *The Major Transitions in Evolution*. Oxford University Press, USA.

Mayr, E. (1982). *The growth of biological thought: diversity, evolution, and inheritance*. Harvard University Press.

McFarland, D. and Houston, A. (1981). *Quantitative Ethology*. Pitman Advanced Pub. Program.

McMullin, B. (2004). 30 years of computational autopoiesis: A review. *Artificial Life*, 10:277–295.

Meyer, J. and Guillot, A. (1990). Simulation of adaptive behavior in animats: review and prospect. In *Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats*, pages 2–14, Paris, France. MIT Press.

Moreno, A. and Etxeberria, A. (2005). Agency in natural and artificial systems. *Artificial Life*, 11(1-2):161–175.

Moreno, A. and Ruiz-Mirazo, K. (2009). The problem of the emergence of functional diversity in prebiotic evolution. *Biology and Philosophy*, 24(5):585–605.

Morowitz, H. and Smith, E. (2007). Energy flow and the organization of life. *Complexity*, 13(1):51—59.

Morowitz, H. J. (1979). *Energy Flow in Biology*. Ox Bow Press.

Morowitz, H. J. (1999). A theory of biochemical organization, metabolic pathways, and evolution. *Complexity*, 4(6):39–53.

Nolfi, S. and Floreano, D. (2004). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. The MIT Press.

Ono, N. and Ikegami, T. (2000). Self-maintenance and self-reproduction in an abstract cell model. *Journal of Theoretical Biology*, 206(2):243–253.

Oparin, A. I. (1961). *Life, its nature, origin and development.* Oliver and Boyd, (Edinburgh).

Packer, H. L. and Armitage, J. P. (2000). Inverted behavioural responses in wild-type rhodobacter sphaeroides to temporal stimuli. *FEMS Microbiology Letters*, 189(2):299–304.

Parkinson, J. S., Ames, P., and Studdert, C. A. (2005). Collaborative signaling by bacterial chemoreceptors. *Current Opinion in Microbiology*, 8(2):116–121.

Pfeifer, R. and Bongard, J. C. (2006). *How the Body Shapes the Way We Think: A New View of Intelligence*. The MIT Press.

Piantadosi, S. T. and Crutchfield, J. P. (2010). How the dimension of space affects the products of Pre-Biotic evolution: The spatial population dynamics of structural complexity and the emergence of membranes. *1010.5019*.

Prasad, K., Caplan, S., and Eisenbach, M. (1998). Fumarate modulates bacterial flagellar rotation by lowering the free energy difference between the clockwise and counter-clockwise states of the motor. *Journal of Molecular Biology*, 280(5):821–828.

Pross, A. (2008). How can a chemical system act purposefully? bridging between life and non-life. *Journal of Physical Organic Chemistry*, 21(7-8):724–730.

Rasmussen, S. (2009). *Protocells : bridging nonliving and living matter*. MIT Press, Cambridge Mass.

Romano, A. H. and Conway, T. (1996). Evolution of carbohydrate metabolic pathways. *Research in Microbiology*, 147(6-7):448–455.

Rosen, R. (1991). *Life Itself*. Columbia University Press, 1 edition.

Rosenblueth, A., Wiener, N., and Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of science*, 10(1):18–24.

Ruiz-Mirazo, K. and Mavelli, F. (2007). Simulation model for functionalized vesicles: Lipid-Peptide integration in minimal protocells. In *Advances in Artificial Life*, pages 32–41. Springer Berlin / Heidelberg.

Ruiz-Mirazo, K. and Mavelli, F. (2008). On the way towards [']basic autonomous agents': Stochastic simulations of minimal lipid-peptide cells. *Biosystems*, 91(2):374–387.

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

Ruiz-Mirazo, K., Pereto, J., and Moreno, A. (2004). A universal definition of life: Autonomy and Open-Ended evolution. *Origins of Life and Evolution of Biospheres*, 34(3):323–346.

Sarand, I., Osterberg, S., Holmqvist, S., Holmfeldt, P., Skarfstad, E., Parales, R. E., and Shingler, V. (2008). Metabolism-dependent taxis towards (methyl)phenols is coupled through the most abundant of three polar localized aer-like proteins of pseudomonas putida. *Environmental Microbiology*, 10(5):1320–1334.

Schweinitzer, T. and Josenhans, C. (2010). Bacterial energy taxis: a global strategy? *Archives of Microbiology*, 192(7):507–520.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(03):417–424.

Segre, D., Ben-Eli, D., Deamer, D. W., and Lancet, D. (2001). The lipid world. *Origins of Life and Evolution of Biospheres*, 31(1):119–145.

Seth, A. K. (1998). Evolving action selection and selective attention without actions, attention, or selection. In *From animals to animats 5: Proceedings of the Fifth International Conference on the Simulation of Adaptive Behavior*, volume 5, pages 139—147, Cambridge, MA. MIT Press.

Shapiro, R. (2006). Small molecule interactions were central to the origin of life. *The Quarterly Review of Biology*, 81(2):105–126.

Shenhav, B., Solomon, A., Lancet, D., and Kafri, R. (2005). Early systems biology and prebiotic networks. In *Transactions on Computational Systems Biology I*, pages 14–27. Springer Berlin / Heidelberg.

Smith, E. and Morowitz, H. J. (2004). Universality in intermediary metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13168–13173.

Sowa, Y. and Berry, R. M. (2008). Bacterial flagellar motor. *Quarterly Reviews of Biophysics*, 41(02):103–132.

Spudich, J. L. and Koshland, D. E. (1975). Quantitation of the sensory response in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America*, 72(2):710 –713.

Stewart, J., Gapenne, O., and Di Paolo, E. A. (2011). *Enaction: Toward a New Paradigm for Cognitive Science*. The MIT Press.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-Component signal transduction. *Annual Review of Biochemistry*, 69(1):183–215.

Stock, D., Leslie, A. G. W., and Walker, J. E. (1999). Molecular architecture of the rotary motor in ATP synthase. *Science*, 286(5445):1700–1705.

Stock, J. and Levit, M. (2000). Signal transduction: Hair brains in bacterial chemotaxis. *Current Biology*, 10(1):R11–R14.

Suzuki, K. and Ikegami, T. (2004). Self-repairing and mobility of simple cell model. In *Artificial life IX: Proceedings of the 9th International Conference on the Simulation and Synthesis of Living Systems*, pages 421—426.

Suzuki, K. and Ikegami, T. (2009). Shapes and Self-Movement in protocell systems. *Artificial Life*, 15(1):59–70.

Taylor, B. L. (1983). Role of proton motive force in sensory transduction in bacteria. *Annual Review of Microbiology*, 37(1):551–573.

Taylor, B. L. and Zhulin, I. B. (1998). In search of higher energy: metabolism-dependent behaviour in bacteria. *Molecular Microbiology*, 28(4):683–690.

Taylor, R. (1950). Comments on a mechanistic conception of purposefulness. *Philosophy of Science*, 17(4):310–317. ArticleType: research-article / Full publication date: Oct., 1950 / Copyright © 1950 Philosophy of Science Association.

Thomason, P. A., Wolanin, P. M., and Stock, J. B. (2002). Signal transduction: Receptor clusters as information processing arrays. *Current Biology*, 12(11):R399–R401.

Thompson, E. (2010). *Mind in Life*. Harvard University Press, London, England.

Toyota, T., Maru, N., Hanczyc, M. M., Ikegami, T., and Sugawara, T. (2009). Self-Propelled oil droplets consuming "Fuel" surfactant. *Journal of the American Chemical Society*, 131(14):5012–5013.

Varela, F., Maturana, H., and Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5(4):187–196.

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

Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34(1):72–87.

Vegge, C. S., Brondsted, L., Li, Y., Bang, D. D., and Ingmer, H. (2009). Energy taxis drives campylobacter jejuni toward the most favorable conditions for growth. *Appl. Environ. Microbiol.*, 75(16):5308–5314.

Virgo, N., Egbert, M. D., and 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.*, pages 234—242, Berlin. Springer.

Wadhams, G. H. and Armitage, J. P. (2004). Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol*, 5(12):1024–1037.

Webb, B. (2009a). Animals versus animats: Or why not model the real iguana? *Adaptive Behavior*, 17(4):269 –286.

Webb, B. (2009b). Response: The power of selection. *Adaptive Behavior*, 17(4):346.

Williams, P. L., Beer, R. D., and Gasser, M. (2008). An embodied dynamical approach to relational categorization. *Cognitive Science*, 1:223–228.

Xie, Z., Ulrich, L. E., Zhulin, I. B., and Alexandre, G. (2010). PAS domain containing chemoreceptor couples dynamic changes in metabolism with chemotaxis. *Proceedings of the National Academy of Sciences*, 107(5):2235–2240.

Zhulin, I., Rowsell, E., Johnson, M., and Taylor, B. (1997). Glycerol elicits energy taxis of escherichia coli and salmonella typhimurium. *J. Bacteriol.*, 179(10):3196–3201.

Ziemke, T. (2003). What's that thing called embodiment? In Alterman and Kirsch, editors, *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, pages 1305—1310, Mahwah, NJ. Lawrence Erlbaum.